Visually guided flight in Harris' hawks: a computer vision approach



Sofía Miñano

Lady Margaret Hall

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2022

Abstract

Animals with well-developed visual systems often use vision as an active sense. This means they can manipulate the position and orientation of their viewpoint to interrogate their visual environment. In this thesis, I investigate active vision in a particularly challenging scenario: birds in manoeuvring flight. I focus on Harris' hawks (*Parabuteo unicinctus*), which strongly rely on vision to coordinate their manoeuvres. I cover three key challenges of studying birds' gaze strategies in flight. First, I explore the problem of accurately tracking an animal's gaze direction. Marker-based motion capture is the most accurate species-agnostic method to track 3D movements in humans and animals, but commercial systems often underperform when applied to animals. I present a marker-labelling method that solves many of the existing problems, based on an arrangement of markers that is optimal for a custom per-frame labelling algorithm. Second, I explore how to capture the visual scene an animal experiences. I solve this computationally, using a 3D model of the flight environment, and a model of the birds' view during flight. I demonstrate how the proposed method allows us to define alternative gaze strategies for hypothesis testing, and explore different methods of modelling the environment, that adapt to varying levels of geometric complexity. Third, I investigate how we can use the generated synthetic data to analyse the hawks' behaviour. For $N = 4$ birds, I generate semantic maps of the visual scene the animals experience in flight, for a total of $n = 97$ perching and obstacle avoidance flights. I characterise the birds' gaze strategies around the landing perch and the obstacles, and discuss possible explanations as to why these strategies may be advantageous. Overall, my thesis demonstrates a novel and reliable method of investigating avian vision in flight and opens up new avenues for data-driven models of animal behaviour.

Author contributions

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where stated otherwise by reference or acknowledgement, the work presented is entirely my own. Specifically, my contributions towards the data chapters are as follows:

Chapter 2. Optimal marker pack design for a time-independent labelling algorithm.

This chapter is presented as a manuscript *in prep* for submission to a peer-reviewed journal. I designed the experiments, collected and postprocessed the data, analysed the results and wrote the manuscript, with the supervision of Graham Taylor and Stuart Golodetz.

Chapter 3. Through hawks' eyes: synthetically reconstructing the visual field of a bird in flight

This chapter is based on a submitted manuscript of which I am first author. The work was conceived by Stuart Golodetz, Tommaso Cavallari, Graham Taylor and myself. Stuart Golodetz and Tommaso Cavallari provided the SLAM application *SemanticPaint*, and I helped them set up the integration with the motion capture system. I designed the experiments, collected and postprocessed the data, analysed the results and wrote the initial manuscript. All authors contributed to the final manuscript. The single pursuit flight analysed in this chapter is part of a dataset I collected with fellow DPhil student James Kempton, which is described in **Appendix D**.

Chapter 4. Analysis of Harris' hawks' gaze strategies in perching and obstacle avoidance.

This chapter is presented as a traditional thesis chapter. I designed the experiments, collected and postprocessed the data, analysed the results and wrote the manuscript, with the supervision of Graham Taylor.

The dataset described in **Appendix D** is part of a set of experiments designed and collected with fellow DPhil student James Kempton, and supervised by Graham Taylor. I postprocessed the collected data using a custom labelling algorithm developed by me. The model described in **Appendix E** was developed by me, supervised by Graham Taylor.

Derived publications

- Part of the work presented in Chapter 3 was accepted as a conference workshop paper in CVPR 2021 *CV4Animals*; the pre-print is available at bioRxiv:

  > Sofía Miñano, Graham K. Taylor. *Through hawks' eyes: reconstructing a bird's visual field in flight to study gaze strategy and attention during perching and obstacle avoidance*. bioRxiv 2021.06.16.446415; doi: https://doi.org/10.1101/2021.06.16.446415.

- The work presented in Chapter 3 is an extended version of the manuscript submitted to *International Journal of Computer Vision: Special Issue on Computer Vision Approach for Animal Tracking and Modeling*, currently under review. The submitted version is included in **Appendix F**

My work was supported by funding from the Biotechnology and Biological Sciences Research Council (BBSRC) [grant number BB/M011224/1], via the Interdisciplinary Bioscience Doctoral Training Partnership.

Acknowledgements

I would like to thank my supervisor Graham Taylor, for his continual support throughout this project. His positive outlook and his enthusiasm about science have been a strong encouragement these past four and a half years. Thanks for all the help in the highs and lows, and for being supportive of my research interests and ideas, and even of my Spanglish terms. It has been a privilege to learn more about animal flight in your group.

I would also like to thank my co-supervisor, Stuart Golodetz. Thanks for your attention to detail, for your commitment to your students and for your fantastic sense of humour. It was really fun to work on the SLAM bit early on, and I'm very glad your presence sandwiched the thesis the way it did, thanks for all the help. Thanks to Tommaso too, for many fun chats about birds, drones and SLAM, and a few afternoons at the flight lab setting everything up. Thanks to my viva examiners as well, Adrian Thomas and Shane Windsor, for a fantastic discussion and for your enthusiasm in my work.

I am also very grateful to the DTP team for all their support during the pandemic and also afterwards. Thanks to all the team at LMH too, for providing an amazing working environment this final year.

To the current and past members of the Oxford Flight group, a big thank you. Marco, thanks for always being willing to help, and for all those shared battles against the markers and their labels (I think we tamed them). Thanks to Caroline too, for introducing me to all things mocap with birds. Special thanks to James K: a friendship formed in a stressful experimental campaign can never be broken. Thanks for all the help, especially with top-quality feedback on many applications and pieces of writing. Thanks to Natalia for being the best outlet of Spanish humour, and to James S for bringing the chillness to the lab. Thanks to Robin too, for many chats trying to find a way around neural networks and birds. Thanks as well to Indira, Lydia, Inés, Leo, Alix, Emma and Kyu Min for making this a fantastic group to work in. I can't wait to see the cool science that will come next. Thanks also to Helen, Lucy and Mark for their amazing help with the birds during the experiments, and to everyone at the field station for making it a great working environment.

The Oxford experience™ would not have been the same without the DTPeeps. Anna and Lea, thanks for being amazing flatmates and friends. Thanks to Tati,

for always lending an ear and for fuelling the foozball addiction, and to Tika, for sharing de-stressing table-tennis sessions. Thanks as well to Lois, Stefi, Riccardo and Paloma for being there in this final stretch (often with delicious food). Thanks too to Stephen, Signe and Vero, and special thanks to Nick for the stream of bird memes, sustained at a great rate for four years. On the LMH side, I need to thak super-problem-solver Vania, and Suri for many many laughs and a lot of encouragement (and linspecer!). Thanks to Claire and her fantastic London plans, Kao for great chats and walks, Alasdair for being the best host, and Isla and Chris for being a fantastic very dynamic duo. Special thanks to Michael for all the help, the encouragement, the plasma physics and the fleece jokes. Thanks to Ben too for being the best partner for last-minute plans.

The internship at DeepMind was one of the best bits of this DPhil: thanks to Ulrich for being an amazing mentor and for all the help, even after the internship. Thanks to Sanjana, Eden, Roman and Brian for making it so fun, it has been lovely to keep in touch, and thanks to Brendan too. On the Spanish side, I need to thank Julia and the amazing fiestas del audio, and also Eva, Jorge, Martalb, Luis, Laura and Marta for their cariño, ongoing support and encouragement. I would also like to thank Dan Quinn, a fantastic role-model scientist who got me started into this research thing.

Most of all I would like to thank my family. Thanks Isa, por el cariño hermanil and for always offering understanding in anxious times. And thanks to my parents, for their amazing support, for always making me laugh and for all their help, especially these past few years. Sois los mejores, gracias por estar siempre a mi lado.

Contents

List of Figures xii

List of Tables xviii

List of Supplementary Videos xx

1 Introduction 1

1.1 Gaze movements in animals . . . 1

1.2 Challenges of studying gaze in animals . . . 4

1.3 Visual abilities of Harris' hawks . . . 8

1.4 Thesis outline . . . 11

References . . . 14

2 Optimal marker pack design for a time-independent labelling algorithm 21

2.1 Abstract . . . 21

2.2 Introduction . . . 22

2.3 Methods . . . 27

2.3.1 Derivation of an optimal marker pack . . . 28

2.3.2 Construction of a bird's headpack . . . 38

2.3.3 Validation experiments . . . 38

2.4 Results . . . 47

2.4.1 Choice of $l_{min,res}$ . . . 47

2.4.2 Choice of $l_{max,design}$ . . . 48

2.4.3 Comparison between EE and temporal labelling . . . 56

2.5 Discussion . . . 62

References . . . 67

3 Through hawks' eyes: synthetically reconstructing the visual field of a bird in flight **72**
3.1 Abstract . . . 73
3.2 Introduction . . . 73
3.2.1 Related work . . . 75
3.2.2 Contribution . . . 81
3.3 Methods . . . 83
3.3.1 Bird flight experiments . . . 83
3.3.2 Motion capture data postprocessing . . . 84
3.3.3 Modelling the bird's view . . . 85
3.3.4 Hybrid model of the lab environment . . . 93
3.4 Results . . . 98
3.4.1 Pursuit flight . . . 98
3.4.2 Obstacle avoidance flights . . . 104
3.5 Discussion . . . 106
3.5.1 Key features of rendering method . . . 106
3.5.2 Behavioural analyses: pursuit flight . . . 109
3.5.3 Behavioural analyses: obstacle avoidance . . . 110
3.5.4 Limitations and future work . . . 111
3.6 Supplementary videos . . . 113
References . . . 114

4 Analysis of Harris' hawks' gaze strategies in perching and obstacle avoidance **120**
4.1 Abstract . . . 120
4.2 Introduction . . . 122
4.3 Methods . . . 124
4.3.1 Motion capture experiments . . . 125
4.3.2 Visual coordinate system . . . 128
4.3.3 Computational model in Blender . . . 136
4.4 Results . . . 136
4.4.1 Stabilisation of the landing perch and the obstacles in the visual field . . . 137
4.4.2 Longitude angle subtended by the landing perch and the obstacles along the trajectory . . . 141
4.4.3 Gaze strategy in perching . . . 148

4.4.4 Gaze strategy around obstacles . . . 157
4.5 Discussion . . . 164
References . . . 171

5 Conclusions and future work **175**
5.1 Main conclusions . . . 175
5.2 Limitations and recommendations . . . 178
5.3 Future directions . . . 182
References . . . 185

Appendices

A Appendices for Chapter 2: *Optimal marker pack design for a time-independent labelling algorithm* **190**
A.1 Triangle inequality condition . . . 190
A.2 Cayley-Menger determinant condition . . . 191
A.3 Optimisation procedure . . . 192
A.4 Golden solutions . . . 193
A.5 Design plane l_{min}, l_{max} . . . 193
A.6 Computing 3D coordinates of vertices from edges . . . 195
A.7 Design considerations for 3D-printed headpack for birds . . . 195
A.8 Design considerations for 3D-printed two-markers supports . . . 197
A.9 Drill speed measurements with tachometer . . . 197
A.10 Computation of ground-truth labels . . . 198
A.11 Definition of templates in Nexus . . . 202
A.12 Histograms of residuals per trial . . . 202
References . . . 202

B Appendices for Chapter 3: *Through hawks' eyes: synthetically reconstructing the visual field of a bird in flight* **204**
B.1 Bird flight experiments . . . 205
B.1.1 Motion capture system . . . 205
B.1.2 Placement of markers . . . 205
B.1.3 Experimental procedure . . . 206
B.2 Motion capture data postprocessing . . . 207
B.2.1 Markers' reconstruction and labelling . . . 207
B.2.2 Headpack transform per frame . . . 209

B.2.3 Target transform per frame . . . 210
B.2.4 Walls, floor and ceiling geometry . . . 211
B.3 Modelling the bird's view . . . 211
B.3.1 Fixation phase in gaze calibration trial . . . 211
B.3.2 Distance between the origin and the midpoint between the eyes 212
B.3.3 Gaze direction estimate . . . 214
B.3.4 Normal to the sagittal plane estimate . . . 216
B.3.5 Retinal margins . . . 217
B.3.6 Trajectory coordinate system . . . 218
B.4 Hybrid model of the lab environment . . . 218
B.4.1 Geometry of the pursuit flight . . . 218
B.4.2 Geometry of the obstacle avoidance flights . . . 218
B.4.3 Dense 3D map of the curtain . . . 219
B.5 Supplementary videos . . . 220
References . . . 221

C Appendices for Chapter 4: *Analysis of Harris' hawks' gaze strategies in perching and obstacle avoidance* 224
C.1 Motion capture experiments . . . 224
C.2 Motion capture data postprocessing . . . 225
C.3 Headpack coordinate system . . . 226
C.4 Visual coordinate system . . . 227
C.5 Spherical histograms . . . 229
C.6 Latitudinal and longitudinal extension of the landing perch and the obstacles . . . 238
C.7 Alignment with landing perch centre or bisector direction . . . 238
C.8 Deviation from an ideal symmetric expansion of the landing perch . . . 247
C.9 Mean position of the landing perch centre in the visual field . . . 249
C.10 Mean position of the obstacle's edge in the visual field . . . 249
C.11 Location of the minima of optic flow in the bird's visual field . . . 252
References . . . 252

D A dataset of obstacle avoidance flights in Harris' hawks, with static and moving goals 256
D.1 Introduction . . . 256
D.2 Materials . . . 257

D.2.1 Curtain . . . 258
D.2.2 Perches . . . 259
D.2.3 Pulleys and linear motor . . . 259
D.2.4 Obstacles . . . 259
D.3 Experimental procedure . . . 260
D.3.1 Static goal trials . . . 263
D.3.2 Moving goal trials . . . 264
D.4 Summary of recorded flights . . . 265
D.5 Labelling procedure . . . 272
D.6 Randomisation based on stations . . . 272
D.7 Randomisation based on grid . . . 278
D.7.1 Static goal trials . . . 278
D.7.2 Moving goal trials . . . 282

E An attractor-repeller model for the obstacle avoidance behaviour of Harris' hawks **291**
E.1 Introduction . . . 291
E.2 Model overview . . . 292
E.3 Steering commands: normal acceleration . . . 292
E.3.1 Scalar field s . . . 293
E.3.2 $K_{s,i}$ and $K_{\phi,i}$ functions . . . 294
E.4 Thrust commands: tangential acceleration . . . 295
References . . . 295

F Submitted manuscript version of Chapter 3 **297**
F.1 Through hawks' eyes: synthetically reconstructing the visual field of a bird in flight . . . 297

List of Figures

1.1 Gaze movements in birds 3

2.1 A four-marker pack as a tetrahedron 29
2.2 Design plane l_{min}, l_{max} for the selected solution 37
2.3 General workflow for the design of an optimal marker pack 39
2.4 Construction of bird headpack 40
2.5 Experiments with two-markers supports 41
2.6 Design points for optimal headpacks 43
2.7 Experiments with headpacks 45
2.8 Types of bird trials 46
2.9 Percentage of correct frames for each two-markers support 48
2.10 Violin plots for ρ per headpack type 51
2.11 ρ in drill trials 53
2.12 ρ in bird trials 54
2.13 Labelling performance of EE algorithm across headpack types . . . 55
2.14 Labelling performance of EE and temporal algorithm in drill trials - headpack type and rotation axis effect 57
2.15 Labelling performance of EE and temporal algorithm in drill trials - sampling frequency and drill speed effect 60
2.16 Labellings comparison in selected drill trials 61
2.17 Labelling performance of EE and temporal algorithm in bird trials . 62
2.18 Examples of label switching in the temporal algorithm 63

3.1 Related work 76
3.2 Summary of method for reconstructing the visual information contained within the visual field of a bird in flight 81
3.3 Pursuit and obstacle avoidance flights 86
3.4 Estimation of gaze direction for the pursuit flight 88
3.5 Estimation of the normal to the sagittal plane for the pursuit flight 89

3.6 Estimation of the gaze direction for the obstacle avoidance flights . 91
3.7 Estimation of the normal to the sagittal plane for the obstacle avoidance flights . . . 92
3.8 Retinal margins of Harris' hawks eyes . . . 93
3.9 The two coordinate systems used for rendering . . . 94
3.10 Model of the lab environment for the obstacle avoidance flights . . . 96
3.11 Capturing the 3D dense map in the motion capture coordinate system 97
3.12 Snapshots of the rendered pursuit flight . . . 101
3.13 Semantic heatmap for the target . . . 102
3.14 Trajectory of the target in the bird's visual field . . . 103
3.15 Trajectory of the target in the trajectory coordinate system . . . 104
3.16 Trajectory of the obstacles in the visual field sphere . . . 107
3.17 Extension in longitude of the obstacles set and the landing perch per frame, in the obstacle avoidance flights . . . 108
3.18 3D model of a forest environment in the lab . . . 112

4.1 Experimental setup for the perching and obstacle avoidance trials . 126
4.2 Perching flights per bird . . . 127
4.3 Obstacle avoidance flights per bird . . . 127
4.4 Gaze direction estimate per bird . . . 132
4.5 Normal to the sagittal plane estimate per bird . . . 134
4.6 Comparison of spherical histograms in the visual coordinate system and the trajectory coordinate system . . . 139
4.7 Spherical histograms showing the stabilisation of the landing perch in the visual field . . . 140
4.8 Spherical histograms showing the stabilisation of the obstacles in the visual field . . . 141
4.9 Visibility of the landing perch and the obstacles in an obstacle avoidance trial . . . 142
4.10 Longitude angle subtended by the landing perch on a bird's visual field along the trajectory - Drogon and Charmander flights . . . 145
4.11 Longitude angle subtended by the landing perch on a bird's visual field along the trajectory - Ruby and Toothless flights . . . 146
4.12 Longitude angle subtended by the obstacles on a bird's visual field along the trajectory . . . 147
4.13 Deviation from an ideal symmetric expansion of the landing perch - Perching flights . . . 151

4.14 Deviation from an ideal symmetric expansion of the landing perch - Obstacle avoidance flights 152
4.15 Fraction of frames aligned with one of the perch's edges per flight, against the RMSE relative to the ideal symmetric expansion of the perch 154
4.16 Histograms for the fraction of frames aligned with the landing perch's centre and the landing perch's edge per flight 155
4.17 Deviation from an ideal alignment with the obstacles' inside edge . 159
4.18 Mean longitude angle of the MSAC inliers' across all flights (in absolute value) 161
4.19 Estimated loss of fixation point per flight 162
4.20 Loss of fixation point on the birds' trajectories 163

A.1 Examples of feasible regions in the l_{min}, l_{max} plane for three solutions. 196
A.2 Diagram for the computation of the ground-truth labels for the drill trials 199
A.3 Overview of the computation of the geometry and the ground-truth labels for the headpack markers in the bird trials 200
A.4 Diagram for the computation of the ground-truth labels of the headpack markers in the bird trials 201
A.5 Histograms of residuals for bird and drill trials 203

B.1 Reference motion capture markers used in the pursuit trial 206
B.2 Histogram of 3D point residuals 207
B.3 Labelled markers per trial in a local coordinate system 209
B.4 Estimated location of the walls 211
B.5 Examples of the bird's behaviour in the fixation phase of the gaze calibration trial 212
B.6 Distance between selected origin and midpoint between the eyes for the pursuit flight, in the direction perpendicular to the sagittal plane 213
B.7 Distance between selected origin and midpoint between the eyes for the pursuit flight, in the direction perpendicular to the headpack's plate 213
B.8 Distance between selected origin and midpoint between the eyes for the obstacle avoidance flights, in the direction perpendicular to the sagittal plane 214

B.9 Distance between selected origin and midpoint between the eyes for the obstacle avoidance flights, in the direction perpendicular to the headpack's plate . . . 215
B.10 Visual field of Harris' hawks . . . 217
B.11 Correction of the captured curtain mesh . . . 222
B.12 Orthographic projection of the rendered output . . . 223

C.1 Birds and headpacks used . . . 225
C.2 Headpack coordinate system for each design . . . 227
C.3 Perch fixation and eyes-level assumption . . . 228
C.4 Trajectories of the landing perch midpoint in the headpack coordinate system . . . 228
C.5 Spherical histograms computation . . . 230
C.6 Spherical histograms in the trajectory coordinate system reflect the geometry of the trajectory . . . 230
C.7 Spherical histograms for the landing perch in the perching flights - Drogon and Charmander . . . 231
C.8 Spherical histograms for the landing perch in the perching flights - Ruby and Toothless . . . 232
C.9 Spherical histograms for the landing perch in the obstacle avoidance flights - Drogon and Charmander . . . 233
C.10 Spherical histograms for the landing perch in the obstacle avoidance flights - Ruby and Toothless . . . 234
C.11 Spherical histograms for the obstacles in the obstacle avoidance flights - Drogon and Charmander . . . 235
C.12 Spherical histograms for the obstacles in the obstacle avoidance flights - Ruby and Toothless . . . 236
C.13 Spherical histograms for the obstacles and the landing perch, for the flights over the obstacles . . . 237
C.14 Latitudinal extension of the landing perch in the bird's visual field along the trajectory - Drogon and Charmander perching flights . . . 239
C.15 Latitudinal extension of the landing perch in the bird's visual field along the trajectory - Ruby and Toothless perching flights . . . 240
C.16 Latitudinal extension of the landing perch in the bird's visual field along the trajectory - Drogon and Charmander obstacle avoidance flights . . . 241
C.17 Latitudinal extension of the landing perch in the bird's visual field along the trajectory - Ruby and Toothless obstacle avoidance flights 242

C.18 Longitudinal extension of the landing perch in the bird's visual field along the trajectory - Obstacle avoidance leg 1 flights 243
C.19 Longitudinal extension of the obstacles in the bird's visual field along the trajectory (not mirrored) - Drogon and Charmander flights . . . 244
C.20 Longitudinal extension of the obstacles in the bird's visual field along the trajectory (not mirrored) - Ruby and Toothless flights 245
C.21 Median and bisector of triangle defined by the bird's head and the landing perch's edges 246
C.22 Values of β angle per frame across all flights against the normalised distance to the landing perch $\hat{y}_{perch}$ 246
C.23 Deviation from an ideal symmetric expansion of the landing perch in longitudinal direction in the bird's visual field - Obstacle avoidance leg 1 flights 248
C.24 Mean position of the landing perch's centre in the bird's visual field per flight - Perching flights 250
C.25 Mean position of the landing perch's centre in the bird's visual field per flight - Obstacle avoidance flights 251
C.26 Mean position of the obstacles' inside edge midpoint in the bird's visual field per flight 254
C.27 Location of minima of optic flow in the visual field sphere, for two of Drogon's flights 255

D.1 Schematic of the flight lab, with the curtain and obstacles 258
D.2 Flight lab with the two astroturf covers 258
D.3 Curtain 259
D.4 Monoliths 260
D.5 Individual pillars 260
D.6 Pairs of pillars 261
D.7 Monoliths (front) and pairs of pillars (back) 262
D.8 Three pillars (front and back) 262
D.9 High pillars (front) and three pillars (back) 262
D.10 Pillars in 2+1 configuration (front and back) 262
D.11 Initial and end perches (static goal trials) 266
D.12 Randomisation approaches (top-down view) 267
D.13 Restrictions for static goal trials 268
D.14 Pictures of the pulleys' configuration (moving goal trials) 268
D.15 Schematic of the pulleys' configuration (moving goal trials) 269

D.16 Restrictions for moving goal trials 270
D.17 Labelling of markers for a static goal trial 273
D.18 Labelling of markers for a moving goal trial 274
D.19 Labelling approach followed for the individual markers on the headpack 275
D.20 Headpack pose for a sample static goal flight 276
D.21 Grid definitions for static and moving goal trials 279
D.22 Area covered by the obstacles for static goal trials using the randomisation based on stations 284
D.23 Area covered by the obstacles in the static goal trials using the randomisation based on grid - Individual pillars 285
D.24 Area covered by obstacles in the static goal trials using the randomisation based on grid - Pair of pillars (back only) 286
D.25 Area covered by the obstacles in the static goal trials using randomisation based on grid - Obstacles front and back 287
D.26 Area covered by the obstacles in the moving goal trials - Diamond . 288
D.27 Area covered by the obstacles in the moving goal trials - Zigzag . . 289
D.28 Area covered by the obstacles in the moving goal trials - Zigzag and combined 290

List of Tables

1 List of Supplementary Videos. xxi

2.1 Results for all pairings between opposite edges 34
2.2 Solutions per pairing of opposite edges 35
2.3 Headpacks used in drill and bird experiments 42
2.4 Drill trials conditions 44
2.5 Number of trials analysed for each bird and headpack 45
2.6 Distance measured for each two-markers support in the static trials 48
2.7 Percentage of correct frames in lasso trials, and mean distance measured 49
2.8 Mean, median and standard deviation ρ for each headpack type . . 50
2.9 Mean and median labelling performance of the EE algorithm for each headpack type 56

3.1 Description of supplementary videos 113

4.1 Number of flights analysed per bird, trial type and trial leg 128
4.2 Estimated gaze direction vectors $\vec{v}_{gaze}$ per bird 130
4.3 Estimated normals to the sagittal plane $\vec{n}_{sagittal}$ per bird 135
4.4 Mean root-mean-square error (RMSE) of the landing perch's extension relative to an ideal symmetric case, across all flights per bird 150
4.5 Median and interquartile range for the fraction of frames (%) aligned with the landing perch's centre, and the fraction of frames (%) aligned with a landing perch's edge, over all flights 154
4.6 Median fraction of frames (%) aligned with the landing perch's centre, across all flights per bird 156
4.7 Median fraction of frames (%) aligned with the landing perch's edge, across all flights per bird 156
4.8 Mean root-mean-square error (RMSE) between the longitude angle of the obstacles' inside edge and the ideal case, over all flights per bird 160

4.9 Mean perpendicular distance to the obstacles, and mean time period before passing the obstacles at the estimated loss of fixation point, over all flights per bird 164

A.1 Drill speed measurements with optical tachometer 198

B.1 Maximum distance d_{max} between labelled markers and corresponding markers in headpack template 209
B.2 Results of the gaze direction fitting, for the pursuit and the obstacle avoidance flights 215
B.3 Results of the gaze direction fitting per trial, for all trials with the same bird and headpack placement as the rendered obstacle avoidance trial 216
B.4 Results of the normal to the sagittal plane fitting, for the pursuit and the obstacle avoidance trial 217
B.5 Mean distances between the markers fixed to the target 219
B.6 Distances between the motion capture markers fixed to the curtain and the subset of red vertices in the curtain's mesh 220

C.1 Results of the gaze direction fits, considering all trials per bird . . . 229

D.1 Obstacle arrangements used in static and moving goal trials 261
D.2 Types of obstacles and arrangements for the static goal trials 271
D.3 Types of obstacles and arrangements for the moving goal trials . . . 271
D.4 Randomisation based on stations 277
D.5 Randomisation based on grid for static goal trials: individual and pair of pillars 280
D.6 Randomisation based on grid for static goal trials: obstacles in front and back 281
D.7 Randomisation based on grid for moving goal trials 283

List of Supplementary Videos

The set of videos described in Table 1 is provided as supplementary material. For each video the following information is specified in the table: type of manoeuvre recorded, data shown, coordinate system for the virtual camera's axes, projection applied to the 360° data, and chapter in which the videos are mainly discussed.

All videos are reproduced at 20Hz (1/10 of the real speed), except the optic flow data video (`ESM_5.avi`, `ESM_16.avi`) that is reproduced at 5Hz (1/40 of the real speed). The RGB outputs are presented using two projections: equirectangular, in which the geometry appears distorted, but which shows the complete field of view of the bird; and orthographic, in which the distortion is reduced but doesn't include the most peripheral regions of the bird's visual field. The rest of the rendered outputs are only represented in orthographic projection.

In all videos, a red contour around the figure indicates the bird's head transform for that frame was interpolated. The retinal margins and the blind areas of the bird's visual field are overlaid on the rendered output. In the optic flow video, the colormap represents for each frame the instantaneous angular speed per pixel, in degrees per second. The colourbar is in logarithmic scale and capped at $10°/s$ in the lower bound and $1000°/s$ in the upper bound. Further details on the computation of the videos are included in Appendix B.5.

Table 1: List of Supplementary Videos.

Filename	Manoeuvre	Data	Coordinate system	Projection	Chapter
`ESM_1.avi`	Pursuit	RGB	Visual	Orthographic	3
`ESM_2.avi`	Pursuit	RGB	Visual	Equirectangular	3
`ESM_3.avi`	Pursuit	Depth	Visual	Orthographic	3
`ESM_4.avi`	Pursuit	Semantic	Visual	Orthographic	3
`ESM_5.avi`	Pursuit	Optic flow	Visual	Orthographic	3
`ESM_6.avi`	Pursuit	RGB	Trajectory	Orthographic	3
`ESM_7.avi`	Pursuit	RGB	Trajectory	Equirectangular	3
`ESM_8.avi`	Obstacle avoidance (leg 1)	RGB	Visual	Orthographic	3
`ESM_9.avi`	Obstacle avoidance (leg 1)	RGB	Visual	Equirectangular	3
`ESM_10.avi`	Obstacle avoidance (leg 1)	RGB	Trajectory	Orthographic	3
`ESM_11.avi`	Obstacle avoidance (leg 1)	RGB	Trajectory	Equirectangular	3
`ESM_12.avi`	Obstacle avoidance (leg 2)	RGB	Visual	Orthographic	3
`ESM_13.avi`	Obstacle avoidance (leg 2)	RGB	Visual	Equirectangular	3
`ESM_14.avi`	Obstacle avoidance (leg 2)	RGB	Trajectory	Orthographic	3
`ESM_15.avi`	Obstacle avoidance (leg 2)	RGB	Trajectory	Equirectangular	3
`ESM_16.avi`	Obstacle avoidance (leg 1)	Optic flow	Visual	Orthographic	4 (Appendix C.11)

1 Introduction

Animals with well-developed visual systems rely on vision for a variety of behaviours, from locomotion (Rivers et al., 2014) to foraging (Land, 1999b), and predator detection (Yorzinski et al., 2014) to courtship displays (Yorzinski et al., 2013). In most of these examples, vision acts as an active sense, meaning that the animals can manipulate the position and orientation of their viewpoint to explore the environment, or reduce the uncertainty of the required information to execute a task (Hayhoe et al., 2018). As a result, analysing the motion patterns of an animal's gaze direction can be very informative about its cognitive processing (Hayhoe et al., 2005), and provide unique insights into the animal's behaviour (Billington et al., 2020).

1.1 Gaze movements in animals

Animals move their gaze across their environment using different combinations of body, head and eye movements (Land et al., 2012). Humans make use of all of these kinds of motion to move their gaze across a visual scene. Insects rely on head and body movements, since their eyes are fixed relative to their heads (Land et al., 2012; Egelhaaf et al., 2012; Cellini et al., 2020; Ravi et al., 2022). In other animals, such as rock crabs, the head is fixed relative to the body, and so gaze control is achieved through body motion and eye movements relative to the head

(Land et al., 2012; Land, 1999a). In birds, the large size of their eyes limits their motion relative to their skull, so their gaze movements across the environment are largely accomplished by head movements relative to their bodies (Gioanni, 1988; Yorzinski et al., 2013), and by their own body locomotion.

Some types of gaze movement are only observed in certain animals, while others are widespread. For example, smooth pursuit, or the ability to continuously track a target with gaze, is thought to be exclusive to primates (Krauzlis, 2004), while the vestibulo-ocular reflex (VOR) and the opto-kinetic reflex (OKR) are present in all vertebrates, as they evolved early in the lineage (Land et al., 2012). The saccade-fixation pattern is another example of a type of gaze motion that is widespread across animals with good vision: it consists of alternating periods of gaze fixations with fast gaze shifts, also called saccades (Land et al., 2012; Land, 1999a). Humans show this type of motion in their eye movements relative to their heads (Land, 1999a). In insects, it appears in both head and body motions (Egelhaaf et al., 2012; Boeddeker et al., 2010), and birds exhibit it in their head movements (Eckmeier et al., 2008; Kress et al., 2015; Frost, 1978). In fact, birds' lightweight and powerfully articulated heads enable them to saccade at speeds that are comparable to the eye motions of mammals (Kress et al., 2015).

Although humans and other animals use the saccade-fixation strategy in visual searches, to redirect the high-resolution areas of their visual fields (or foveas) to the relevant information, this is likely not the fundamental reason why this motion pattern is so widespread (Land et al., 2012). Most vertebrates do not have specialised high-resolution areas in their visual fields, and still exhibit this type of motion. Instead, the saccade-fixation strategy appears to primarily serve a gaze stabilisation function during locomotion. An animal moving across its environment and aiming to stabilise the image projected on its retina, entirely or partly, will likely reach a point at which its gaze position or orientation needs to be reset due to anatomical limitations (Land et al., 2012; Land, 1999a). With a saccadic movement, this reorientation is executed as fast as possible, minimising the time

during which the image of the environment is blurred on the retina, or suppressed by the visual system (Land et al., 2012).

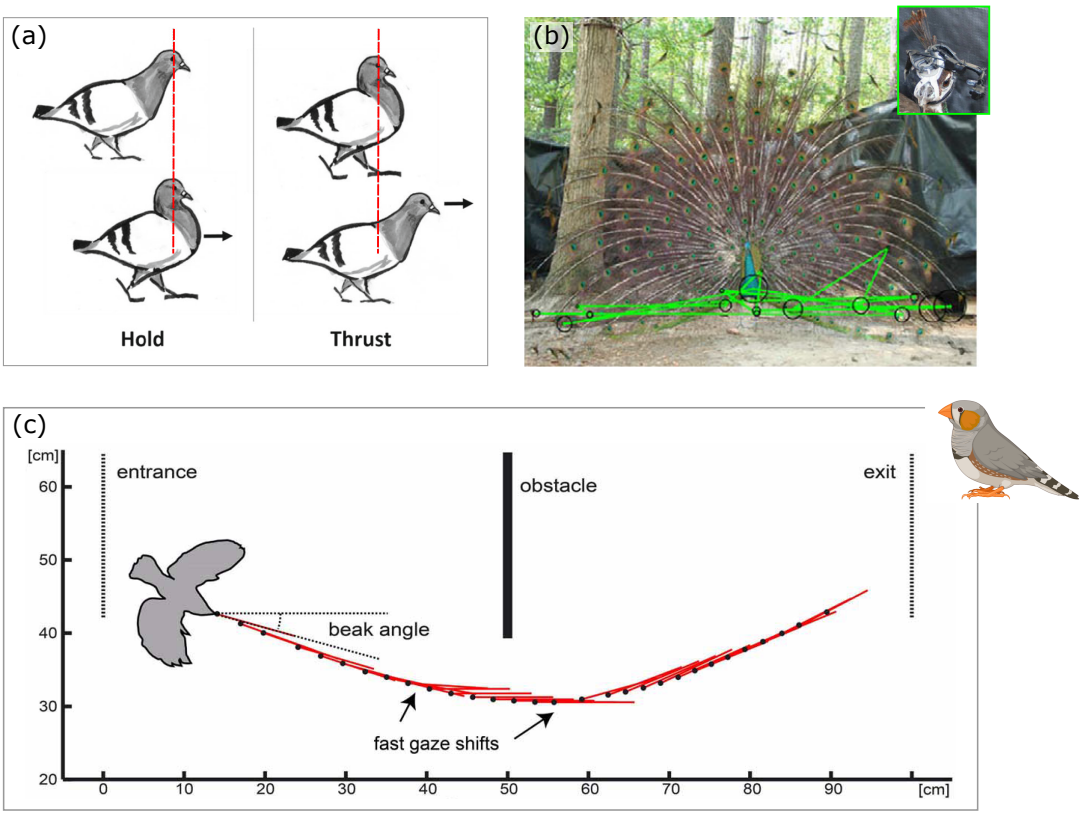


Figure 1.1: Gaze movements in birds. Three examples of gaze movements in birds are shown. Figure (a) shows a schematic of head-bobbing in walking pigeons, in which the time sequence follows the panels from top to bottom. The animal's head is thrust forwards, and then fixated in space while the body catches up underneath (modified from Pritchard et al., 2018). In (b), scanpaths are shown of the eye movements of peahens during courtship; the size of the markers denotes the amount of time fixating at each location. The inset (top right) shows the eye-tracking setup on the bird (Yorzinski et al., 2013). Figure (c) shows the estimated gaze direction of a zebra finch during a turning flight (red lines). The fast gaze shifts are marked with arrows. The bird contour is approximately at the correct scale (Eckmeier et al., 2013). Zebra finch icon created with BioRender.com.

Birds' heads follow a saccade-fixation motion pattern in a variety of behaviours. For example, some ground-feeding birds show a head-bobbing motion while walking, in which the animal's head is first thrust forwards, and then fixated in space while the body catches up underneath (see Figure 1.1a; Frost, 1978; Land et al., 2012). In this way, they maximise the time in which their lateral visual field is stationary relative to their environment, which is thought to support foraging (Land et al., 2012). The saccade-fixation pattern has also been reported in birds performing

visual searches during courtship (Figure 1.1b; Yorzinski et al., 2013) or hunting (Land, 1999b; Ochs et al., 2016), and in turning flight (Figure 1.1c; Eckmeier et al., 2008; Kress et al., 2015; Ros et al., 2017b). In the latter case, it is thought to enable the disambiguation of motion flow on the retina (also called optic flow), by extracting during the fixation phases the translational component of the flow field, which contains the depth information (Eckmeier et al., 2008). This behaviour is also exhibited by flying insects (Schilstra et al., 1998). All of these are good examples of how active vision plays a vital role in a variety of behaviours in birds.

1.2 Challenges of studying gaze in animals

In humans, measurements of gaze movements have been analysed to investigate behaviour since the 1900s, but the field received a particular boost when head-mounted eye-trackers became widely available in the 1980s (Land, 2006). Studies have analysed human gaze movements for a variety of motor tasks, such as walking in a complex natural environment (Matthis et al., 2018; Matthis et al., 2022), preparing food (Land, 2006; Hayhoe et al., 2003), batting (Hayhoe et al., 2012; Land et al., 2000), steering while driving (Land et al., 1994; Mole et al., 2021; Lappi, 2022) and cycling (Wilkie et al., 2008). Despite important differences between human and artificial vision (Hayhoe et al., 2018), the way in which humans control their gaze for the execution of these tasks has inspired active vision approaches in robotic humanoids, both in experimental setups (Seara et al., 2001; Seara et al., 2004; Seara et al., 2002) and in simulation (Merel et al., 2020). Within the field of image understanding in computer vision, several examples exist of neural network architectures inspired by the mechanisms of human visual attention (Mnih et al., 2014; Gregor et al., 2015; Ba et al., 2015; Chen, 2021). However, a similar understanding of gaze strategies in non-human animals, and consequently the derivation of similarly bio-inspired applications, remains elusive.

There are several challenges involved in the study of gaze control in non-human animals that likely explain the relatively limited research in this area. One of these

challenges is accurately tracking an animal's gaze position and orientation. For some animals, such as primates or mice, this may require tracking eye movements, whereas for others like insects or birds, head tracking may suffice (Land et al., 2012). Eye-tracking setups from human research have been successfully adapted for primate studies in the lab (Kano et al., 2018; Damon et al., 2017; Ryan et al., 2019); this may explain why their gaze strategies are the most studied within non-human animals (Billington et al., 2020). For non-primates, adapted or custom eye-tracking setups have been used in a small number of studies, although examples exist for dogs (Somppi et al., 2012), cats (Rivers et al., 2014), peahens (Yorzinski et al., 2013; Yorzinski et al., 2014; Yorzinski et al., 2015) and especially for rodents (Holmgren et al., 2021; Sattler et al., 2021; Meyer et al., 2020; Michaiel et al., 2020; Meyer et al., 2018; Payne et al., 2017; Wallace et al., 2013). These studies typically involve restricting the animal's motion (Somppi et al., 2012), or using invasive surgical procedures (Rivers et al., 2014; Holmgren et al., 2021; Sattler et al., 2021; Meyer et al., 2020; Michaiel et al., 2020; Meyer et al., 2018; Payne et al., 2017; Wallace et al., 2013). From the examples mentioned, only the set of studies on peahens by Yorzinski and colleagues involved free-moving animals in their natural habitat; unfortunately, the weight of their equipment currently restricts its application to large birds and to terrestrial locomotion (Yorzinski et al., 2013; Yorzinski et al., 2014; Yorzinski et al., 2015).

Studies investigating gaze in birds or insects typically track head movements, most frequently using video-based tools (Ros et al., 2017b; Ros et al., 2017a; Lin et al., 2014; Eckmeier et al., 2008; Kress et al., 2015; Ravi et al., 2022; Boeddeker et al., 2010), but also marker-based motion capture (Minano et al., 2021; Quinn et al., 2019; Zeiler et al., 2009; Theunissen et al., 2017; Lin et al., 2017; Mischiati et al., 2015). Video-based methods have the advantage of collecting additional information on the animal's behaviour (Naik et al., 2020) and can be deployed in the field (Brighton et al., 2019). Semi-automatic tracking tools for video data, such as the one developed by Hedrick (2008), are still very widely used in the field of animal behaviour and biomechanics (see for example the recent paper by Ravi

et al., 2022). Advances in computer vision and deep learning in recent years have boosted the performance of fully automated markerless methods, especially for 2D video data. Although similar species-agnostic solutions for tracking 3D movement from video are emerging, these are not yet well-established (Dunn et al., 2021; Karashchuk et al., 2021; Nath et al., 2019). For tracking 3D human and animal movement, marker-based motion capture is still considered the 'gold standard', owing to its achievable sub-millimetre accuracy (Dunn et al., 2021; Naik et al., 2020). Commercially available systems provide the 3D positions of markers attached to the subjects almost straight out of the box, with streamlined camera calibration procedures. However, most systems are optimised to track human motion, and perfom poorly when applied to animals (Naik et al., 2020).

Another challenge to tackle when investigating gaze control in animals is the need to realistically capture the animal's visual scene. One option is to use head-mounted cameras, which allow us to directly capture the visual environment the animal experiences as a result of its own behaviour. This approach additionally enables studies of the animal in its natural habitat; it has been successfully applied to investigate the gaze strategy of falcons and hawks in the field (Kane et al., 2015; Kane et al., 2014). However, head-mounted cameras have some important limitations. First, the available camera options are significantly restricted by the maximum payload requirements determined by welfare considerations. It is thus difficult to approximate the animal's full field of view, or its optical or sampling resolution, with a suitable camera. For example, the 31° vertical field of view of the head-mounted camera used in Kane et al. (2015), and Kane et al. (2014), is far from the 100° vertical extent of the binocular area in Harris' hawks (Potier et al., 2016). Additionally, these cameras' typically low frame rates (usually around 30 Hz, as per Kane et al., 2015; Kane et al., 2014; Lev-Ari et al., 2018; Hazan et al., 2015) lead to motion blur and other artefacts that limit the possibilities of posterior data analyses (Kane et al., 2014). Furthermore, the large mass of head-mounted cameras makes them prone to wobble, unless surgically attached to the animal's head (Lev-Ari et al., 2018; Hazan et al., 2015), which is ideally avoided on welfare grounds.

An alternative approach is to reconstruct the environment in which the animal is moving synthetically. This usually requires us to track the animal's gaze with an independent method, but provides more flexibility than a head-mounted camera solution. For example, recreating the scene in a computational environment facilitates the definition of alternative hypotheses regarding how the animal moved (Eckmeier et al., 2013; Ravi et al., 2022; Ardin et al., 2015) or how the environment looked (Eckmeier et al., 2013; Bian et al., 2021; Bian et al., 2019). Additionally, a variety of 3D modelling approaches enable reconstructions with varying levels of detail, from simple geometric shapes (Minano et al., 2021; Eckmeier et al., 2013), to detailed high-resolution ones using laser scanners and photogrammetry (Holmgren et al., 2021). However, reconstructing in detail an animal's natural environment usually involves expensive and sophisticated equipment, and the postprocessing work is often difficult to automate (Risse et al., 2018; Stürzl et al., 2015). In the literature of animal behaviour, synthetic reconstructions of an animal's visual environment have mostly been used to study the behaviour of small animals, such as insects (Zeil et al., 2014; Ardin et al., 2016; Ardin et al., 2015; Ravi et al., 2022; Ravi et al., 2019; Stuerzl et al., 2016), mice (Holmgren et al., 2021), zebra finches (Eckmeier et al., 2013), or lizards (Bian et al., 2019; Bian et al., 2021). For the case of zebra finches, the synthetic reconstruction was carried out for a single flight.

A third challenge of studies of gaze in animals relates to the methods of analysis. In some cases, methods borrowed from human studies are almost directly applicable, whereas in others the adaptation is more challenging. For example, analyses of fixations in the environment are considerably simplified if these fall on a planar surface, such as a screen in the dog experiments by Somppi et al. (2012), the floor in the walking cats study by Rivers et al. (2014), or the peacock's train in the work by Yorzinski et al. (2013). This facilitates the computation of popular metrics from visual search studies in humans, such as scan paths, which represent the trajectory of the animal's gaze (see Figure 1.1b adapted from Yorzinski et al., 2013), or gaze heatmaps, which represent the density of fixations within a time period (Rivers et al., 2014) or their average duration (Somppi et al., 2012). However,

other scenarios often introduce complications. A good example is the time headway metric used in gaze studies of humans walking or driving, defined as the time it takes to locomote to the point of gaze fixation (Mole et al., 2021; Matthis et al., 2018); this metric is difficult to translate to flying animals or to other non-terrestrial modes of locomotion. Additionally, many of the approaches derived from human studies reduce the individual's visual field to a single high-resolution area. For many animals, such as most raptors that pursue their prey (Mitkus et al., 2018), the high-resolution area may not be a single point. For most vertebrates, a specialised area is not present at all (Land et al., 2012).

1.3 Visual abilities of Harris' hawks

Raptors are renowned in the animal kingdom for their visual abilities (Mitkus et al., 2018). We consider raptors all birds from the orders Falconiformes (falcons), Accipitriformes (hawks and relatives), Cathartiformes (New World vultures and condors), Strigiformes (owls) and Cariamiformes (seriemas), following the definition from Potier et al. (2020b), McClure et al. (2019), and Jarvis et al. (2014). However, within raptors there are notable differences in their visual systems, especially between diurnal and nocturnal ones, and between raptors of different foraging styles (predators versus scavengers, or aerial predators versus ground hunters; see Potier et al., 2020b; Potier, 2020; Mitkus et al., 2018).

In this thesis, I investigate active vision in Harris' hawks (*Parabuteo unicinctus*), which, as diurnal raptors, strongly rely on visual information to coordinate their flight manoeuvres (Mitkus et al., 2018). Additionally, they hunt fast-moving prey very close to the ground (Mitkus et al., 2018; Brighton et al., 2017), which makes them exceptionally agile at avoiding obstacles and manoeuvring through clutter.

The importance of vision for these birds is evident from their eye size: compared to other birds, raptors have very large eyes, both in absolute terms and relative to their body mass (Mitkus et al., 2018; Potier, 2020). Larger eyes enable higher visual acuity (Mitkus et al., 2018), but also limit how much the eyes can move relative

to the skull (Potier et al., 2020b). Owls present an extreme case in point: their range of eye motion is under 1.5°, mainly due to their eyes being partly outside their skulls (Steinbach et al., 1973). Diurnal raptors can move their eyes more than owls, but less than other birds such as corvids (up to 39°, as per Troscianko et al., 2012). However, data is relatively scarce (Potier et al., 2020b). One of the few detailed studies of eye movements in raptors (Wallman et al., 1985) found that a head-restrained little eagle (*Hieraaetus morphnoides*) could move its eyes up to 24° horizontally, but that the vast majority of eye saccades were under 5° in amplitude.

While Harris' hawks and little eagles belong to the same Family (Accipitridae) they have different hunting styles (Wallman et al., 1985), and the visual systems of raptors often show large variation across species (and even across individuals), especially if their foraging habits differ (Mitkus et al., 2018). The range of eye motion in Harris' hawks has not been reported, but estimates exist for other Accipitridae species (apart from the little eagle): $\leq 6°$ is estimated for goshawks (*Accipiter gentilis*) in Kane et al. (2015); $\sim 5°$ is estimated for red-tailed hawks (*Buteo jamaicensis*), in O'Rourke et al. (2010); and $\sim 8°$ is estimated for Cooper's hawks (*Accipiter cooperi*), in O'Rourke et al. (2010). Note however that most of these studies are carried out on restrained birds, and to the best of the author's knowledge, eye movements of birds in flight have not been reported in the literature.

Another factor contributing to the enhanced visual acuity of raptors is their specialised retinas (Mitkus et al., 2018). The raptor retina has three different types of photoreceptors (rods, double cones and single cones), and most diurnal raptors are tetrachromatic (Potier et al., 2020b; Mitkus et al., 2018). The distribution of photoreceptors across the retina is also not uniform. Harris' hawks and most other diurnal raptors that pursue their prey (predators) have two foveas in each retina, that is, two depressed regions in the retinal surface were single cones are concentrated (Mitkus et al., 2018; Potier et al., 2016). One of the foveas is deeper and lateral-facing (deep or central fovea) and another one is shallower and forward-facing (shallow or temporal fovea). The lateral-facing fovea presents the highest visual acuity and is thought to be used to detect prey at distance, whereas the

frontal-facing fovea is thought to support fixation on the prey just before interception (Mitkus et al., 2018; Potier, 2020; Potier et al., 2016).

These and other specialisations confer high visual acuity to diurnal raptors, but popular claims of their visual acuity being 8 times that of humans are inaccurate (Potier et al., 2020b). Some large raptors do have the highest spatial resolution among extant animals (the maximal visual acuity of the wedge-tailed eagle, *Aquila audax*, is close to the theoretical limit imposed by waveguide optics; Potier et al., 2020b), but most medium-sized raptors like Harris' hawks present values that are close to humans (27.4-62.3 cycles/deg; Potier et al., 2016; Potier et al., 2018). Additionally, in diurnal raptors visual acuity decreases significantly with decreasing light intensity, and with decreasing achromatic contrast of the stimuli (Potier et al., 2020b). For example, Harris' hawks' visual acuity drops to 5-8 cycles/deg at 8-9% contrast, showing a lower performance than humans (Potier et al., 2018; Potier et al., 2020b). However, they can resolve purely chromatic patterns (specifically, isoluminant red-green gratings) up to 20 cycles/deg, which is twice the human maximum value (Potier et al., 2018)

To inspect how fast these birds can see, the flicker fusion frequency (FFF) is often used as a proxy metric for temporal resolution (Potier et al., 2020b). In Harris' hawks this value is higher than in humans. Under optimal conditions, these birds can distinguish a light flickering up to 81 Hz, which contrasts with the 40-60 Hz human threshold (Potier et al., 2020a). Diurnal raptors that are also aerial hunters present even higher FFF values than Harris' hawks, which hunt close to the ground: the Peregrine falcon can resolve up to 129 Hz, and the Saker falcon up to 102 Hz, suggesting again an adaptation to their hunting styles (Potier et al., 2020a; Potier et al., 2020b).

The visual field configuration of raptors is also highly tuned to their foraging habits (Mitkus et al., 2018; Potier, 2020). Harris' hawks and other predatory raptors have a larger binocular field than scavengers (for Harris' hawks, 47° horizontally at its maximum; Potier et al., 2016) which is thought to mainly support the control of their feet during prey capture (Potier, 2020). Additionally, Harris' hawks and

other ground predators have a large blind sector above their heads. This blind area is larger than that of most aerial predators or scavengers, and is partly due to an eyebrow ridge that protects them from sun-dazzling (Mitkus et al., 2018; Potier, 2020; Potier et al., 2016).

The projections of the four foveal regions (two in each retina) of Harris' hawks on their visual field sphere has not been determined experimentally, although methods used in other birds should be applicable (Tyrrell et al., 2018; Wallman et al., 1985). For most diurnal raptors, the frontal-facing foveas are usually estimated to project between 9° and 16° longitude from the forward direction of the head, and the lateral-facing fovea somewhere above 30° longitude (Wallman et al., 1985; Frost et al., 1990; Kane et al., 2014; Tucker, 2000). It is however unclear whether the foward-facing foveas of Harris' hawks usually project to a single point in the visual field. In Anna's hummingbirds this is not the case, even when the eyes converge (Tyrrell et al., 2018), whereas in the little eagle it does seem possible but it was rarely observed in the experiments, being their primary gaze position with their frontal-facing foveas at around 13° from the head sagittal plane (Wallman et al., 1985). The effect of eye movements or their convergent/divergent positions on the visual fields of Harris' hawks has not been assessed either; in Potier et al. (2016), measurements were only taken with the eyes at rest.

1.4 Thesis outline

In this thesis, I present a method that overcomes many of the aforementioned limitations of gaze studies in animals, focusing on a particularly challenging scenario: hawks in manoeuvring flight. Combining high-speed motion capture data from experiments with birds and computer vision tools, I generate a detailed description of the visual scene the birds experience in flight, and use it to analyse their gaze behaviours.

In **Chapter 2**, I present a novel method that deals with the problems of accurately tracking 3D movements in animals. As mentioned, animal movements

are often challenging to track with commercially available motion capture systems, as these are largely optimsed for human motion. In this chapter, I describe a solution based on the design of an optimal arrangement of markers that maximises the performance of a custom marker-labelling algorithm. I validate the performance of the proposed method experimentally, show that it outperforms the commercially available approach for scenarios of interest, and demonstrate its use for the case of tracking head movements of birds in flight. The method is readily applicable to other species, and to other marker-based tracking approaches.

In **Chapter 3**, I present a method to synthetically reconstruct the visual scene experienced by a bird in flight, and demonstrate its use for a pursuit flight and two obstacle avoidance flights. In a computational environment, I model the 3D geometry of the lab and the view from the bird's perspective during flight, as a 360° virtual camera representative of the animal's visual field. Using a rendering engine, I generate a set of synthetic visual inputs per frame over the bird's full visual field, which characterise its visual experience in flight. I demonstrate how the proposed method allows us to define alternative gaze strategies for hypothesis testing, and how the generated data may be used to answer behavioural questions. Additionally, I explore different ways of modelling the 3D environment of the lab that adapt to varying levels of geometric complexity. The method includes key aspects to support the collection of large amounts of data across many individuals, such as non-invasive tracking of the animal's head pose, and automated integration of the captured 3D dense maps. This chapter is based on a submitted manuscript currently under review in a peer-reviewed journal.

In **Chapter 4**, I use the method described in Chapter 3 to investigate the gaze strategy of Harris' hawks in perching and obstacle avoidance manoeuvres. I used a subset of 97 flights across four birds to carry out this study. Specifically, I used semantic descriptions of the birds' complete visual fields at each timestep in their trajectories to inspect their gaze strategies around the landing perch and the obstacles. I find the birds use distinct behaviours around these two elements: while the landing perch is largely centred relative to the bird's sagittal plane throughout

the perching manoeuvre, the edge of the obstacles is fixated frontally when they fly around them. I explore patterns in which the birds deviate from these general behaviours and discuss possible explanations as to why the observed strategies may be advantageous for manoeuvring flight. This work provides a solid basis for future studies, by (1) pinpointing specific hypotheses to test with the full dataset, and (2) validating analytical approaches at a small scale, which is necessary before they can be extended to a larger scale.

In **Chapter 5**, I summarise and discuss the main conclusions from the previous chapters, and set the findings in a wider context. Within the limitations of the proposed approach, I consider the possibilities that it opens up in the field of sensory ecology and animal behaviour, and provide recommendations for future directions.

The data chapters in this thesis (Chapters 2 to Chapter 4) are written as self-contained manuscripts, and as such present independent introduction, methods, results and discussion sections. Since Chapter 3 and Chapter 4 share a common methodology, the self-containment requirement means there is some degree of overlap between the two chapters; I aimed to minimise this by focusing on the methodological aspects in Chapter 3 and the behavioural analysis in Chapter 4.

I additionally include two appendices that cover research work I carried out during my DPhil: in **Appendix D**, I describe a dataset of obstacle avoidance flights of Harris' hawks flying to static targets (perches) and moving ones (lures). In **Appendix E**, I describe the rationale of an obstacle avoidance model to characterise the flight trajectories of Harris' hawks, based on the attractor-repeller paradigm from robotics and inspired in the behavioural model by Fajen et al. (2003). The remaining appendices (**A** to **C**) include supplementary information for each of the data chapters, which has been separated from the main text for readability.

Plural pronouns are used in the data chapters, as is customary in manuscripts with several authors. The precise contributions of other authors are described in the preamble of the thesis.

References

Ardin, P., Mangan, M., Wystrach, A., and Webb, B. (2015). "How variation in head pitch could affect image matching algorithms for ant navigation". In: *J. Comp. Physiol. A* 201.6, pp. 585–597. DOI: 10.1007/s00359-015-1005-8.

Ardin, P., Peng, F., Mangan, M., Lagogiannis, K., and Webb, B. (2016). "Using an insect mushroom body circuit to encode route memory in complex natural environments". In: *PLoS Comput. Biol.* 12.2. DOI: 10.1371/journal.pcbi.1004683.

Ba, J. L., Mnih, V., and Kavukcuoglu, K. (2015). "Multiple object recognition with visual attention". In: *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–10. arXiv: 1412.7755.

Bian, X., Chandler, T., Pinilla, A., and Peters, R. A. (2019). "Now you see me, now you don't: Environmental conditions, signaler behavior, and receiver response thresholds interact to determine the efficacy of a movement-based animal signal". In: *Front. Ecol. Evol.* 7.APR, pp. 1–16. DOI: 10.3389/fevo.2019.00130.

Bian, X., Pinilla, A., Chandler, T., and Peters, R. (2021). "Simulations with Australian dragon lizards suggest movement-based signal effectiveness is dependent on display structure and environmental conditions". In: *Sci. Rep.* 11.1, pp. 1–11. DOI: 10.1038/s41598-021-85793-3.

Billington, J., Webster, R. J., Sherratt, T. N., Wilkie, R. M., and Hassall, C. (2020). "The (under)use of eye-tracking in evolutionary ecology". In: *Trends Ecol. Evol.* 35.6, pp. 495–502. DOI: 10.1016/j.tree.2020.01.003.

Boeddeker, N., Dittmar, L., Stürzl, W., and Egelhaaf, M. (2010). "The fine structure of honeybee head and body yaw movements in a homing task". In: *Proc. R. Soc. B Biol. Sci.* 277.1689, pp. 1899–1906. DOI: 10.1098/rspb.2009.2326.

Brighton, C. H. and Taylor, G. K. (2019). "Hawks steer attacks using a guidance system tuned for close pursuit of erratically manoeuvring targets". In: *Nat. Commun.* 10.1, pp. 1–28. DOI: 10.1038/s41467-019-10454-z.

Brighton, C. H., Thomas, A. L., and Taylor, G. K. (2017). "Terminal attack trajectories of peregrine falcons are described by the proportional navigation guidance law of missiles". In: *Proc. Natl. Acad. Sci.* 114.51, p. 201714532. DOI: 10.1073/pnas.1714532114.

Cellini, B. and Mongeau, J. M. (2020). "Active vision shapes and coordinates flight motor responses in flies". In: *Proc. Natl. Acad. Sci. U. S. A.* 117.37, pp. 23085–23095. DOI: 10.1073/pnas.1920846117.

Chen, G. (2021). "Where to Look: A Unified Attention Model for Visual Recognition with Reinforcement Learning". In: pp. 1–11. arXiv: 2111.07169.

Damon, F., Méary, D., Quinn, P. C., Lee, K., Simpson, E. A., Paukner, A., Suomi, S. J., and Pascalis, O. (2017). "Preference for facial averageness: Evidence for a common mechanism in human and macaque infants". In: *Sci. Rep.* 7.March, pp. 1–11. DOI: 10.1038/srep46303.

Dunn, T. W., Marshall, J. D., Severson, K. S., Aldarondo, D. E., Hildebrand, D. G., Chettih, S. N., Wang, W. L., Gellis, A. J., Carlson, D. E., Aronov, D., Freiwald, W. A., Wang, F., and Ölveczky, B. P. (2021). "Geometric deep learning enables 3D kinematic profiling across species and environments". In: *Nat. Methods* 18.5, pp. 564–573. DOI: 10.1038/s41592-021-01106-6.

Eckmeier, D., Geurten, B. R., Kress, D., Mertes, M., Kern, R., Egelhaaf, M., and Bischof, H. J. (2008). "Gaze strategy in the free flying zebra finch (Taeniopygia guttata)". In: *PLoS One* 3.12. DOI: 10.1371/journal.pone.0003956.

Eckmeier, D., Kern, R., Egelhaaf, M., and Bischof, H. J. (2013). "Encoding of naturalistic optic flow by motion sensitive neurons of nucleus rotundus in the zebra finch (Taeniopygia guttata)". In: *Front. Integr. Neurosci.* 7.SEP, pp. 1–17. DOI: 10.3389/fnint.2013.00068.

Egelhaaf, M., Boeddeker, N., Kern, R., Kurtz, R., and Lindemann, J. P. (2012). "Spatial vision in insects is facilitated by shaping the dynamics of visual input through behavioural action". In: *Front. Neural Circuits* 6.December, pp. 1–23. DOI: 10.3389/fncir.2012.00108.

Fajen, B. R. and Warren, W. H. (2003). "Behavioral Dynamics of Steering, Obstacle Avoidance, and Route Selection". In: *J. Exp. Psychol. Hum. Percept. Perform.* 29.2, pp. 343–362. DOI: 10.1037/0096-1523.29.2.343.

Frost, B. J. (1978). "The Optokinetic Basis of Head-bobbing in the Pigeon". In: *J. Exp. Biol.* 74.1, pp. 187–195. DOI: 10.1242/jeb.74.1.187.

Frost, B. J., Wise, L. Z., Morgan, B., and Bird, D. (1990). "Retinotopic representation of the bifoveate eye of the kestrel (Falco sparverius) on the optic tectum". In: *Vis. Neurosci.* 5.3, pp. 231–239. DOI: 10.1017/S0952523800000304.

Gioanni, H. (1988). "Stabilizing gaze reflexes in the pigeon (Columba livia) - I. Horizontbal and vertical optokinetic eye (OKN) and head (OCR) reflexes". In: *Exp. Brain Res.* 69.3, pp. 567–582. DOI: 10.1007/BF00247310.

Gregor, K., Danihelka, I., Graves, A., Rezende, D. J., and Wierstra, D. (2015). "DRAW: A recurrent neural network for image generation". In: *32nd Int. Conf. Mach. Learn. ICML 2015* 2, pp. 1462–1471. arXiv: 1502.04623.

Hayhoe, M. and Ballard, D. (2005). "Eye movements in natural behavior". In: *Trends Cogn. Sci.* 9.4, pp. 188–194. DOI: 10.1016/j.tics.2005.02.009.

Hayhoe, M. M. and Matthis, J. S. (2018). "Control of gaze in natural environments: effects of rewards and costs, uncertainty and memory in target selection". In: *Interface Focus* 8.4, pp. 1–7. DOI: 10.1098/rsfs.2018.0009.

Hayhoe, M. M., McKinney, T., Chajka, K., and Pelz, J. B. (2012). "Predictive eye movements in natural vision". In: *Exp. Brain Res.* 217.1, pp. 125–136. DOI: 10.1007/s00221-011-2979-2.

Hayhoe, M. M., Shrivastava, A., Mruczek, R., and Pelz, J. B. (2003). "Visual memory and motor planning in a natural task". In: *J. Vis.* 3.1, pp. 49–63. DOI: 10.1167/3.1.6.

Hazan, Y., Kra, Y., Yarin, I., Wagner, H., and Gutfreund, Y. (2015). "Visual-auditory integration for visual search: a behavioral study in barn owls". In: *Front. Integr. Neurosci.* 9, pp. 1–12. DOI: 10.3389/fnint.2015.00011.

Hedrick, T. L. (2008). "Software techniques for two- and three-dimensional kinematic measurements of biological and biomimetic systems". In: *Bioinspiration and Biomimetics* 3.3. DOI: 10.1088/1748-3182/3/3/034001.

Holmgren, C. D., Stahr, P., Wallace, D. J., Voit, K.-M., Matheson, E. J., Sawinski, J., Bassetto, G., and Kerr, J. N. D. (2021). "Visual pursuit behavior in mice maintains the pursued prey on the retinal region with least optic flow". In: *Elife* 10, pp. 1–34. DOI: 10.7554/eLife.70838.

Jarvis, E. D. et al. (2014). "Whole-genome analyses resolve early branches in the tree of life of modern birds". In: *Science (80-.).* 346.6215, pp. 1320–1331. DOI: 10.1126/science.1253451.

Kane, S. A., Fulton, A. H., and Rosenthal, L. J. (2015). "When hawks attack: animal-borne video studies of goshawk pursuit and prey-evasion strategies". In: *J. Exp. Biol.* 218.2, pp. 212–222. DOI: 10.1242/jeb.108597.

Kane, S. A. and Zamani, M. (2014). "Falcons pursue prey using visual motion cues: new perspectives from animal-borne cameras". In: *J. Exp. Biol.* 217.2, pp. 225–234. DOI: 10.1242/jeb.092403.

Kano, F., Shepherd, S. V., Hirata, S., and Call, J. (2018). "Primate social attention: Species differences and effects of individual experience in humans, great apes, and macaques". In: *PLoS One* 13.2, pp. 1–25. DOI: 10.1371/journal.pone.0193283.

Karashchuk, P., Rupp, K. L., Dickinson, E. S., Walling-Bell, S., Sanders, E., Azim, E., Brunton, B. W., and Tuthill, J. C. (2021). "Anipose: a toolkit for robust markerless 3D pose estimation". In: *Cell Rep.* 36.13, p. 109730. DOI: 10.1016/j.celrep.2021.109730.

Krauzlis, R. J. (2004). "Recasting the Smooth Pursuit Eye Movement System". In: *J. Neurophysiol.* 91.2, pp. 591–603. DOI: 10.1152/jn.00801.2003.

Kress, D., Van Bokhorst, E., and Lentink, D. (2015). "How lovebirds maneuver rapidly using super-fast head saccades and image feature stabilization". In: *PLoS One* 10.6, pp. 1–24. DOI: 10.1371/journal.pone.0129287.

Land, M. F. (1999a). "Motion and vision: why animals move their eyes". In: *J. Comp. Physiol. - A Sensory, Neural, Behav. Physiol.* 185.4, pp. 341–352. DOI: 10.1007/s003590050393.

Land, M. F. (1999b). "The roles of head movements in the search and capture strategy of a tern (Aves, Laridae)". In: *J. Comp. Physiol. - A Sensory, Neural, Behav. Physiol.* 184.3, pp. 265–272. DOI: 10.1007/s003590050324.

Land, M. F. and Lee, D. N. (1994). "Where we look when we steer". In: *Nature* 369.6483, pp. 742–744. DOI: 10.1038/369742a0.

Land, M. F. (2006). "Eye movements and the control of actions in everyday life". In: *Prog. Retin. Eye Res.* 25.3, pp. 296–324. DOI: 10.1016/j.preteyeres.2006.01.002.

Land, M. F. and McLeod, P. (2000). "From eye movement to actions: how batsman hit the ball". In: *Nat. Neurosci.* 3.12, pp. 1340–1345. DOI: 10.1038/81887.

Land, M. F. and Nilsson, D. E. (2012). *Animal eyes*. 2nd ed. Oxford: Oxford University Press, p. 288. DOI: 10.1093/acprof:oso/9780199581139.001.0001.

Lappi, O. (2022). "Gaze strategies in driving–an ecological approach". In: *Front. Psychol.* 13.March. DOI: 10.3389/fpsyg.2022.821440.

Lev-Ari, T. and Gutfreund, Y. (2018). "Interactions between top-down and bottom-up attention in barn owls (Tyto alba)". In: *Anim. Cogn.* 21.2, pp. 197–205. DOI: 10.1007/s10071-017-1150-2.

Lin, H. T. and Leonardo, A. (2017). "Heuristic Rules Underlying Dragonfly Prey Selection and Interception". In: *Curr. Biol.* 27.8, pp. 1124–1137. DOI: 10.1016/j.cub.2017.03.010.

Lin, H. T., Ros, I. G., and Biewener, A. A. (2014). "Through the eyes of a bird: modelling visually guided obstacle flight". In: *J. R. Soc. Interface* 11.96, pp. 1–12. DOI: 10.1098/rsif.2014.0239.

Matthis, J. S., Muller, K. S., Bonnen, K. L., and Hayhoe, M. M. (2022). "Retinal optic flow during natural locomotion". In: *PLOS Comput. Biol.* 18.2, pp. 1–30. DOI: 10.1371/journal.pcbi.1009575.

Matthis, J. S., Yates, J. L., and Hayhoe, M. M. (2018). "Gaze and the control of foot placement when walking in natural terrain". In: *Curr. Biol.* 28.8, 1224–1233.e5. DOI: 10.1016/j.cub.2018.03.008.

McClure, C. J., Schulwitz, S. E., Anderson, D. L., Robinson, B. W., Mojica, E. K., Therrien, J. F., Oleyar, M. D., and Johnson, J. (2019). "Commentary: Defining Raptors and Birds of Prey". In: *J. Raptor Res.* 53.4, pp. 419–430. DOI: 10.3356/0892-1016-53.4.419.

Merel, J., Tunyasuvunakool, S., Ahuja, A., Tassa, Y., Hasenclever, L., Pham, V. U., Erez, T., Wayne, G., and Heess, N. (2020). "Catch & carry: reusable neural controllers for vision-guided whole-body tasks". In: *ACM Trans. Graph.* 39.4, pp. 1–14. DOI: 10.1145/3386569.3392474.

Meyer, A. F., O'Keefe, J., and Poort, J. (2020). "Two Distinct Types of Eye-Head Coupling in Freely Moving Mice". In: *Curr. Biol.* 30.11, 2116–2130.e6. DOI: 10.1016/j.cub.2020.04.042.

Meyer, A. F., Poort, J., O'Keefe, J., Sahani, M., and Linden, J. F. (2018). "A Head-Mounted Camera System Integrates Detailed Behavioral Monitoring with Multichannel Electrophysiology in Freely Moving Mice". In: *Neuron* 100.1, 46–60.e7. DOI: 10.1016/j.neuron.2018.09.020.

Michaiel, A. M., Abe, E. T., and Niell, C. M. (2020). "Dynamics of gaze control during prey capture in freely moving mice". In: *Elife* 9, pp. 1–27. DOI: 10.7554/eLife.57458.

Minano, S. and Taylor, G. K. (2021). "Through hawks' eyes: reconstructing a bird's visual field in flight to study gaze strategy and attention during perching and obstacle avoidance". In: *bioRxiv*. DOI: 10.1101/2021.06.16.446415.

Mischiati, M., Lin, H. T., Herold, P., Imler, E., Olberg, R., and Leonardo, A. (2015). "Internal models direct dragonfly interception steering". In: *Nature* 517.7534, pp. 333–338. DOI: 10.1038/nature14045.

Mitkus, M., Potier, S., Martin, G. R., Duriez, O., and Kelber, A. (2018). "Raptor vision". In: *Oxford Res. Encycl. Neurosci.* March. Oxford University Press, pp. 1–38. DOI: 10.1093/acrefore/9780190264086.013.232.

Mnih, V., Heess, N., Graves, A., and Kavukcuoglu, K. (2014). "Recurrent models of visual attention". In: *Adv. Neural Inf. Process. Syst.* 3.January, pp. 2204–2212. arXiv: 1406.6247.

Mole, C., Pekkanen, J., Sheppard, W. E., Markkula, G., and Wilkie, R. M. (2021). "Drivers use active gaze to monitor waypoints during automated driving". In: *Sci. Rep.* 11.1, pp. 1–18. DOI: 10.1038/s41598-020-80126-2.

Naik, H., Bastien, R., Navab, N., and Couzin, I. D. (2020). "Animals in virtual environments". In: *IEEE Trans. Vis. Comput. Graph.* 26.5, pp. 2073–2083. DOI: 10.1109/TVCG.2020.2973063.

Nath, T., Mathis, A., Chen, A. C., Patel, A., Bethge, M., and Mathis, M. W. (2019). "Using DeepLabCut for 3D markerless pose estimation across species and behaviors". In: *Nat. Protoc.* 14.7, pp. 2152–2176. DOI: 10.1038/s41596-019-0176-0.

O'Rourke, C. T., Hall, M. I., Pitlik, T., and Fernández-Juricic, E. (2010). "Hawk eyes I: Diurnal raptors differ in visual fields and degree of eye movement". In: *PLoS One* 5.9, pp. 1–8. DOI: 10.1371/journal.pone.0012802.

Ochs, M. F., Zamani, M., Gomes, G. M. R., de Oliveira Neto, R. C., and Kane, S. A. (2016). "Sneak peek: Raptors search for prey using stochastic head turns". In: *Auk* 134.1, pp. 104–115. DOI: 10.1642/auk-15-230.1.

Payne, H. L. and Raymond, J. L. (2017). "Magnetic eye tracking in mice". In: *Elife* 6, pp. 1–24. DOI: 10.7554/eLife.29222.
Potier, S. (2020). "Visual adaptations in predatory and scavenging diurnal raptors". In: *Diversity* 12.10, pp. 1–14. DOI: 10.3390/d12100400.
Potier, S., Bonadonna, F., Kelber, A., Martin, G. R., Isard, P.-F., Dulaurent, T., and Duriez, O. (2016). "Visual abilities in two raptors with different ecology". In: *J. Exp. Biol.* 219.17, pp. 2639–2649. DOI: 10.1242/jeb.142083.
Potier, S., Lieuvin, M., Pfaff, M., and Kelber, A. (2020a). "How fast can raptors see?" In: *J. Exp. Biol.* 223.1. DOI: 10.1242/jeb.209031.
Potier, S., Mitkus, M., and Kelber, A. (2018). "High resolution of colour vision, but low contrast sensitivity in a diurnal raptor". In: *Proc. R. Soc. London B Biol. Sci.* 285.1885. DOI: 10.1098/rspb.2018.1036.
Potier, S., Mitkus, M., and Kelber, A. (2020b). "Visual adaptations of diurnal and nocturnal raptors". In: *Semin. Cell Dev. Biol.* 106.May, pp. 116–126. DOI: 10.1016/j.semcdb.2020.05.004.
Pritchard, D. J. and Healy, S. D. (2018). "Taking an insect-inspired approach to bird navigation". In: *Learn. Behav.* 46.1, pp. 7–22. DOI: 10.3758/s13420-018-0314-5.
Quinn, D., Kress, D., Chang, E., Stein, A., Wegrzynski, M., and Lentink, D. (2019). "How lovebirds maneuver through lateral gusts with minimal visual information". In: *Proc. Natl. Acad. Sci. U. S. A.* 116.30, pp. 15033–15041. DOI: 10.1073/pnas.1903422116.
Ravi, S., Bertrand, O., Siesenop, T., Manz, L. S., Doussot, C., Fisher, A., and Egelhaaf, M. (2019). "Gap perception in bumblebees". In: *J. Exp. Biol.* 222.2, pp. 1–10. DOI: 10.1242/jeb.184135.
Ravi, S., Siesenop, T., Bertrand, O. J., Li, L., Doussot, C., Fisher, A., Warren, W. H., and Egelhaaf, M. (2022). "Bumblebees display characteristics of active vision during robust obstacle avoidance flight". In: *J. Exp. Biol.* 225.4. DOI: 10.1242/jeb.243021.
Risse, B., Mangan, M., Stürzl, W., and Webb, B. (2018). "Software to convert terrestrial LiDAR scans of natural environments into photorealistic meshes". In: *Environ. Model. Softw.* 99, pp. 88–100. DOI: 10.1016/j.envsoft.2017.09.018.
Rivers, T. J., Sirota, M. G., Guttentag, A. I., Ogorodnikov, D. A., Shah, N. A., and Beloozerova, I. N. (2014). "Gaze shifts and fixations dominate gaze behavior of walking cats". In: *Neuroscience* 275, pp. 477–499. DOI: 10.1016/j.neuroscience.2014.06.034.
Ros, I. G., Bhagavatula, P. S., Lin, H. T., and Biewener, A. A. (2017a). "Rules to fly by: pigeons navigating horizontal obstacles limit steering by selecting gaps most aligned to their flight direction". In: *Interface Focus* 7.1, p. 20160093. DOI: 10.1098/rsfs.2016.0093.
Ros, I. G. and Biewener, A. A. (2017b). "Pigeons (C. livia) follow their head during turning flight: Head stabilization underlies the visual control of flight". In: *Front. Neurosci.* 11.DEC, pp. 1–12. DOI: 10.3389/fnins.2017.00655.
Ryan, A. M., Freeman, S. M., Murai, T., Lau, A. R., Palumbo, M. C., Hogrefe, C. E., Bales, K. L., and Bauman, M. D. (2019). "Non-invasive eye tracking methods for new world and old world monkeys". In: *Front. Behav. Neurosci.* 13.March. DOI: 10.3389/fnbeh.2019.00039.
Sattler, N. J. and Wehr, M. (2021). "A Head-Mounted Multi-Camera System for Electrophysiology and Behavior in Freely-Moving Mice". In: *Front. Neurosci.* 14.January, pp. 1–14. DOI: 10.3389/fnins.2020.592417.

Schilstra, C. and van Hateren, J. (1998). "Stabilizing gaze in flying blowflies". In: *Nature* 395.October, p. 1998. DOI: 10.1038/27114.

Seara, J. F., Lorch, O., and Schmidt., G. (2001). "Gaze Control for Goal-Oriented Humanoid Walking". In: *Proc. IEEE/RAS Int. Conf. Humanoid Robot. (Humanoids), S. 187–195, Tokio, Japan.*

Seara, J. F. and Schmidt, G. (2004). "Intelligent gaze control for vision-guided humanoid walking: Methodological aspects". In: *Rob. Auton. Syst.* 48.4 SPEC. ISS. Pp. 231–248. DOI: 10.1016/j.robot.2004.07.003.

Seara, J. F., Strobl, K. H., and Schmidt, G. (2002). "Information management for gaze control in vision guided biped walking". In: *IEEE Int. Conf. Intell. Robot. Syst.* 1.October, pp. 31–36. DOI: 10.1109/IRDS.2002.1041357.

Somppi, S., Törnqvist, H., Hänninen, L., Krause, C., and Vainio, O. (2012). "Dogs do look at images: Eye tracking in canine cognition research". In: *Anim. Cogn.* 15.2, pp. 163–174. DOI: 10.1007/s10071-011-0442-1.

Steinbach, M. J. and Money, K. E. (1973). "Eye movements of the owl". In: *Vision Res.* 13.4, pp. 889–891. DOI: 10.1016/0042-6989(73)90055-2.

Stuerzl, W., Zeil, J., Boeddeker, N., and Hemmi, J. M. (2016). "How wasps acquire and use views for homing". In: *Curr. Biol.* 26, pp. 470–482. DOI: 10.1016/j.cub.2015.12.052.

Stürzl, W., Grixa, I., Mair, E., Narendra, A., and Zeil, J. (2015). "Three-dimensional models of natural environments and the mapping of navigational information". In: *J. Comp. Physiol. A Neuroethol. Sensory, Neural, Behav. Physiol.* 201.6, pp. 563–584. DOI: 10.1007/s00359-015-1002-y.

Theunissen, L. M. and Troje, N. F. (2017). "Head stabilization in the pigeon: Role of vision to correct for translational and rotational disturbances". In: *Front. Neurosci.* 11.OCT, pp. 1–12. DOI: 10.3389/fnins.2017.00551.

Troscianko, J., Von Bayern, A. M., Chappell, J., Rutz, C., and Martin, G. R. (2012). "Extreme binocular vision and a straight bill facilitate tool use in New Caledonian crows". In: *Nat. Commun.* 3. DOI: 10.1038/ncomms2111.

Tucker, V. A. (2000). "The deep fovea, sideways vision and spiral flight paths in raptors". In: *J. Exp. Biol.* 203.Pt 24, pp. 3745–3754. DOI: 10.1016/S0169-5347(01)02393-X.

Tyrrell, L. P., Goller, B., Moore, B. A., Altshuler, D. L., and Fernández-Juricic, E. (2018). "The orientation of visual space from the perspective of hummingbirds". In: *Front. Neurosci.* 12.16, pp. 1–9. DOI: 10.3389/fnins.2018.00016.

Wallace, D. J., Greenberg, D. S., Sawinski, J., Rulla, S., Notaro, G., and Kerr, J. N. (2013). "Rats maintain an overhead binocular field at the expense of constant fusion". In: *Nature* 498.7452, pp. 65–69. DOI: 10.1038/nature12153.

Wallman, J. and Pettigrew, J. D. (1985). "Conjugate and disjunctive saccades in two avian species with contrasting oculomotor strategies". In: *J. Neurosci.* 5.6, pp. 1418–1428. DOI: 10.1523/jneurosci.05-06-01418.1985.

Wilkie, R. M., Wann, J. P., and Allison, R. S. (2008). "Active gaze, visual look-ahead, and locomotor control". In: *J. Exp. Psychol. Hum. Percept. Perform.* 34.5, pp. 1150–1164. DOI: 10.1037/0096-1523.34.5.1150.

Yorzinski, J. L., Patricelli, G. L., Babcock, J. S., Pearson, J. M., and Platt, M. L. (2013). "Through their eyes: selective attention in peahens during courtship". In: *J. Exp. Biol.* 216.16, pp. 3035–3046. DOI: 10.1242/jeb.087338.

Yorzinski, J. L. and Platt, M. L. (2014). "Selective attention in peacocks during predator detection". In: *Anim. Cogn.* 17.3, pp. 767–777. DOI: 10.1007/s10071-013-0708-x.

Yorzinski, J. L., Patricelli, G. L., Platt, M. L., and Land, M. F. (2015). "Eye and head movements shape gaze shifts in Indian peafowl". In: *J. Exp. Biol.* 218.23, pp. 3771–3776. DOI: 10.1242/jeb.129544.

Zeil, J., Narendra, A., and Stürzl, W. (2014). "Looking and homing: how displaced ants decide where to go". In: *Philos. Trans. R. Soc. B Biol. Sci.* 369.1636. DOI: 10.1098/rstb.2013.0034.

Zeiler, M. D., Taylor, G. W., Troje, N. F., and Hinton, G. E. (2009). "Modeling pigeon behaviour using a conditional restricted Boltzmann machine". In: *ESANN 2009 Proceedings, 17th Eur. Symp. Artif. Neural Networks - Adv. Comput. Intell. Learn.*, pp. 391–396.

2

Optimal marker pack design for a time-independent labelling algorithm

Contents

2.1 Abstract **21**
2.2 Introduction **22**
2.3 Methods **27**
2.3.1 Derivation of an optimal marker pack 28
2.3.2 Construction of a bird's headpack 38
2.3.3 Validation experiments 38
2.4 Results **47**
2.4.1 Choice of $l_{min,res}$ 47
2.4.2 Choice of $l_{max,design}$ 48
2.4.3 Comparison between EE and temporal labelling 56
2.5 Discussion **62**
References **67**

2.1 Abstract

The way an animal moves provides valuable information about the cognitive processes taking place as the animal interacts with its environment. However, accurately extracting the three-dimensional posture of an animal over time is not an easy task. Marker-based motion capture is the 'gold standard' for acquiring precise

measurements of three-dimensional motion, but commercially available systems are largely optimised for human movements, and typically perform poorly when applied to animals. Motion capture experiments with animals often require large volumes to capture relevant manoeuvres. They usually involve movements that are very different from human ones, such as flight, making human-optimised tracking algorithms unusable. Dense arrangements of small markers are also frequently used, due to the limited area available for markers' attachment in many animals. All these factors push motion capture systems to their limits, and ultimately require additional postprocessing work to identify the individual markers reliably across the recorded frames. Here, we present a method that largely reduces the labelling effort with minimal postprocessing work, with the design of an optimal arrangement of markers (a marker pack), that maximises the performance of a custom labelling algorithm. The algorithm relies on the representation of each of the markers in the pack in a high-dimensional space, the 'EE' space. Our optimal arrangement of markers is such that the markers' representations in EE space are as distinct as possible, while fulfilling the required constraints. This allows us to apply an off-the-shelf clustering algorithm in EE space to efficiently label the markers across many frames. In this chapter, we derive an optimal marker pack design solving the constrained optimisation problem, apply it to the case of tracking head movements of a bird in flight, and carry out experiments to validate our design choices. We also compare the EE labelling approach to a commercial labelling method that relies on differences across frames, and show that our method is more robust to lower sampling frequencies and faster rotational speeds, owing to the markers' representations in EE space being independent of the frame at which they are computed. Our approach is readily applicable to experiments with other animals and opens up new avenues for studies based on the collection of large amounts of data.

2.2 Introduction

The motion of an animal, from large-scale movements like migrations (Flack et al., 2018) to small-scale movements like eye saccades (Kjærsgaard et al., 2008; Land et al.,

1994; Michaiel et al., 2020), contains valuable information about its behavioural state. It often reflects the cognitive processes that take place as the animal interacts with its environment, such as decision-making (Strandburg-Peshkin et al., 2013) or navigation (Kano et al., 2018). Precisely tracking an animal's pose can also be an integral part of sophisticated biomechanics or behavioural experiments, such as closed-loop approaches or virtual-reality setups (Naik et al., 2020). However, accurately extracting this three-dimensional (3D) pose is a challenging problem (Naik, 2021; Jensen et al., 2020).

Studies carried out in the field generally use animal-borne IMUs to characterise an animal's 3D motion (e.g., Kano et al., 2018). In the lab, higher accuracy is achievable and other methods are preferred. An approach widely used in the animal behaviour literature consists of attaching markers to the animal and recording its behaviour with video cameras, often high-speed ones. The invididual markers' positions are then tracked by applying semi-automatic tools to the video data (Hedrick, 2008; Ros et al., 2017b; Ros et al., 2017a; Lin et al., 2014; Chin et al., 2019; Roderick et al., 2019; Ingersoll et al., 2018). In most cases, multiple jointly-calibrated cameras are used, so that the markers' 3D coordinates can be computed and the animal's pose extracted (e.g., Eckmeier et al., 2008). The markers used range from infrared LEDs (Ros et al., 2017a; Lin et al., 2014) and polystyrene spheres (Ros et al., 2017b), to coloured balls (Brighton et al., 2019) and ink marks (Kress et al., 2015). Alternatively, conspicuous features of the animal can be used, such as the bill tip of a bird (Ros et al., 2017b; Eckmeier et al., 2008) or the wingtips of a flying lizard (Khandelwal et al., 2022).

Although this approach is well-established in the field, the postprocessing work becomes infeasible for large amounts of data, as it still requires a large fraction of manually labelled frames (Kress et al., 2015; Ravi et al., 2022; Khandelwal et al., 2022). Recent advances in computer vision and deep learning have benefited video-based methods hugely, and have facilitated high-throughput analyses by further automating the postprocessing pipeline (Mathis et al., 2020b). A large number of video-based 2D tracking solutions have been developed in the past few

years (Mathis et al., 2018; Pereira et al., 2020; Graving et al., 2019; Walter et al., 2021), that do not require the animals to carry markers or to have particularly conspicuous features. These perform very well even in changing visual environments, but are generally sensitive to occlusions if no prior model of the animal's motion is assumed (Nath et al., 2019).

Some of these video-based 2D tracking solutions can also be used with multiple cameras to extract 3D pose (Nath et al., 2019; Joska et al., 2021). However, performance usually degrades if features are not seen by all cameras (Dunn et al., 2021), and they often require postprocessing methods beyond standard triangulation to achieve sufficient accuracy (Dunn et al., 2021; Joska et al., 2021; Zimmermann et al., 2020). Algorithms specifically developed for 3D pose extraction in animals are emerging, many focusing on specific species (Arac et al., 2019; Zimmermann et al., 2020; Günel et al., 2019). Solutions that are species-agnostic also exist (Karashchuk et al., 2021; Dunn et al., 2021), but have mostly been applied for small animals and/or reduced volumes. Generally speaking, the error in deep learning video-based animal pose estimation methods is typically in the order of centimetres (Joska et al., 2021; Li et al., 2021; Zuffi et al., 2017; Biggs et al., 2019), which contrasts unfavourably with the sub-millimetre accuracy that is possible with marker-based motion capture (Marshall et al., 2021; Naik, 2021). Motion capture systems are still considered the 'gold standard' for recording human and animal movement (Dunn et al., 2021; Naik, 2021), and as such, they are often used as ground truth for video-based methods (e.g., Dunn et al., 2021).

Optical motion capture systems usually have a set of synchronised, overlapping cameras, that emit and sense strobes of infrared light. The retroreflective markers reflect this light, which if captured by two or more cameras, enables the determination of the markers' positions in 3D space. By placing retroreflective markers on the animal, we are able to reconstruct its 3D movements. Usually, accompanying software is used to label the markers' trajectories and extract the 3D pose of the tracked subject. However, these tools are largely optimised for human motion, and are pushed to their limits when applied to non-human animals (Naik, 2021).

Animals have diverse shapes and sizes, and move in very different ways to humans, from flying to making fast saccadic head movements. As a result, tracking solutions that rely on prior models of human motion are likely to fail when directly applied to animals. To capture manoeuvres of interest in animals of medium to large size, we often require a large volume to be covered. Additionally, the areas in which markers can be attached are often small: for birds, typically under 50×50 mm^2 (Naik, 2021; Minano et al., 2021). This affects the performance of the motion capture system, since markers in a dense arrangement are more likely to occlude each other in several camera views, leading to the computation of merged or ghost markers (see Naik, 2021, appendix B). We can minimise this risk by using small markers, but these are also harder to detect. Rigid arrangements of markers (or 'marker packs') should not be too prominent or heavy either, to avoid interfering with the animal's natural behaviour or welfare. An animal's fast motion may also lead to markers' displacements per frame that are very similar to the distances between markers in the pack, which further challenges common tracking algorithms. All of these factors result in commercially available tracking software considerably underperforming when tracking the motion of animals (Naik, 2021), which strongly limits the possibilities of high-throughput analyses.

Due to these challenges, studies using motion capture on animals usually require some trade-offs. For example, with animals of relatively larger size, such as dogs (Zhang et al., 2018) and horses (Byström et al., 2021), larger markers and greater distances between markers can be used, both of which improve the performance of the motion capture system. Other studies focus on manoeuvres that take place in relatively small volumes: for example, in Singh et al. (2022), bats fly in a $5 \times 2 \times 3$ m volume; and in Graham et al. (2021), flying snakes cross gaps of approximately 1 m length. Similarly, studies using motion capture in birds either record birds flying in a reduced volume (Gutierrez et al., 2017; Quinn et al., 2019), limit the number of markers on the bird (Gutierrez et al., 2017; Minano et al., 2021) or analyse movements that are much slower than flying, such as walking (Naik, 2021; Zeiler et al., 2009), pecking (Theunissen et al., 2017b; Theunissen et al.,

2017a) or head-stabilisation while standing (Theunissen et al., 2017b). Additionally, most studies using marker-based motion capture on animals report using custom methods to deal with mislabellings and marker dropout (see for example Marshall et al., 2021; Genzel et al., 2021).

Here, we present the design of a rigid arrangement of markers that allows us to automatically label large amounts of data in a reliable and efficient way, and overcome many of the issues described. We focus on off-line labelling (i.e., not real-time) and on animal tracking applications. Our method deals with many of the current limitations on motion capture experiments with animals, particularly with birds, without adding significant postprocessing steps. Instead, it deals with many of the issues already at data acquisition time. Based on a per-frame labelling method, our approach does not require the displacement of a marker across frames to be much smaller than the distance between markers. This makes it more robust than commercially available tracking approaches that rely on temporal differences under many scenarios. The design is also parametrised as a function of the animal's size, making it easily translatable to other species. Our method also provides a systematic way to design an optimal maker pack. The importance of designing a 'good' marker pattern when using motion capture systems is often highlighted among practitioners (Naik, 2021, Chapter 6 and Appendix B). However, this requirement is usually stated in an ambiguous way, and it is not explicitly linked to the effect on labelling performance, ultimately leading to designs based on trial-and-error.

The proposed method offers other important advantages: it is not a 'black box', and thus allows for further inspection and improvement, and it can easily be made part of a more sophisticated labelling pipeline. It is also non-intrusive, minimally affecting the animal's welfare or behaviour, and finally, it facilitates the collection of large amounts of data across many individuals by further automating the data postprocessing. This opens up new possibilities for large-scale analyses and for the collection of large ground truth datasets to benchmark and improve the accuracy of video-based 3D pose estimation methods (which are currently lacking; Biggs et al., 2019). Although our main focus is tracking motion capture markers, the

same method could be applied to track the 3D positions of unlabelled markers extracted from video data from multiple calibrated cameras, as long as the distances between markers stay fairly constant (note that we assume that the set of markers defines a rigid body). This is the most popular approach in the literature, and our method would reduce the manual labelling effort involved. Additionally, our design process leads to several unique marker packs, that could also be used to identify and track individuals in a collective.

In the following sections, we describe the rationale behind an optimal marker pack design, apply it to the case of tracking head movements in flying hawks, and validate our hypotheses on the design and the proposed labelling algorithm with three sets of experiments.

2.3 Methods

We describe the design of an optimal marker pack, that maximises the performance of a time-independent labelling algorithm. This algorithm relies on each of the markers in the pack being represented as a single point in a high-dimensional space. We define this high-dimensional space based on the geometry of the pack: the coordinates of each marker in this space are a function of its distance to the rest of the markers in the pack.

In practice, we will have noise in the measured distances between markers when these are recorded across many frames. As a result, each marker will be represented by a cluster of points in this high-dimensional space, rather than a single point. However, if these clusters are sufficiently far apart, we can label the markers across all recorded frames by applying a classic clustering algorithm, such as k-means, to their coordinates in this high-dimensional space. This is the main idea behind the design and the labelling algorithm we present.

This results in an efficient way of labelling large amounts of data. Additionally, because the coordinates of the markers in the high-dimensional space are independent of the frame at which they are computed, this approach is potentially more robust

than tracking approaches based on temporal differences. Temporal approaches often rely on the motion of the pack between consecutive frames producing a marker displacement much smaller than the distance between the markers in the pack. This is usually the case in motion capture experiments with humans, but not for many non-human animals, which typically have very small areas in which we can attach marker packs.

Our aim is thus to determine an arrangement of markers such that the distances between the markers' clusters in the high-dimensional space is maximised. In this section, we analytically derive the geometry of an optimal marker pack, discuss its construction as a 'headpack' for tracking head movements in flying hawks, and describe a set of experiments to validate our design choices and assess the performance of the proposed labelling algorithm.

2.3.1 Derivation of an optimal marker pack

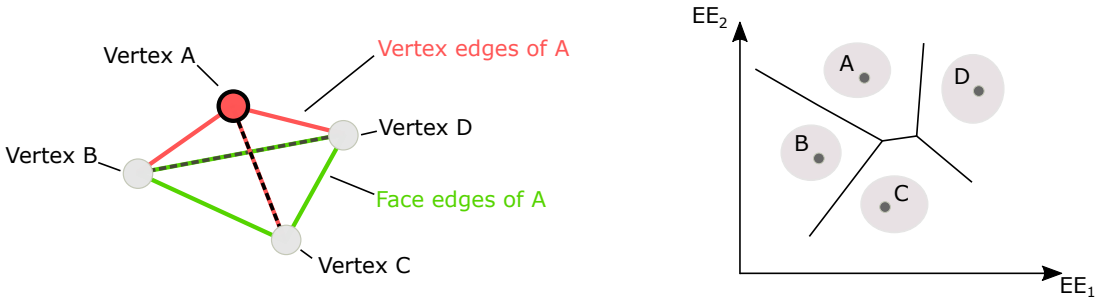
In this section we analyse the design problem and derive the optimal solutions for the marker pack. To do so, we first describe the marker pack as an irregular tetrahedron, then introduce its representation in the high-dimensional space, and finally discuss the optimisation problem.

2.3.1.1 Marker pack as a tetrahedron

We analyse the case of a marker pack of four motion capture markers. Although only three markers are required to extract the pose of a rigid body, additional markers allow us to derive the pack's pose even if at some frames some markers are not reconstructed correctly. However, too many markers constrained to a small volume (such as the top of the bird's head) are more likely to occlude each other in several camera views, which may lead to an incorrect determination of the markers' 3D position. We select a total of four markers as a good trade-off between these two constraints.

A pack of four markers is geometrically an irregular tetrahedron whose four vertices correspond to the markers' centres and whose six edges correspond to the

segments connecting them. For any given vertex v_i, there are three edges that converge on it; we refer to their lengths as the set of *vertex edges* of v_i. For this vertex v_i there are also three edges that do not connect to it; we refer to their lengths as the set of *face edges* of v_i (see Figure 2.1A). Note that the set of face edges defines the face opposite to the vertex v_i. We will also work with the concept of an *opposite edge*: for any edge l in the tetrahedron, there is only one edge that doesn't share any vertices with it. This is the opposite edge to l, and we denote it by $\tilde{l}$ (dashed segments in Figure 2.1A). Note that the face edges of v_i can also be defined as the opposite edges of the vertex edges of v_i. In defining a tetrahedron from the lengths of its edges, we follow the approach and nomenclature from Wirth et al. (2009) with only minor modifications (e.g., they refer to 'face-triples', 'vertex-triples' and 'opposed edges', instead of 'face edges', 'vertex edges' and 'opposite edges').



(a) Geometry of tetrahedron (b) Analogy of EE space in 2D

Figure 2.1: A four-marker pack as a tetrahedron. A pack of four markers can be interpreted geometrically as an irregular tetrahedron. (a) Each vertex in the tetrahedron has a set of vertex edges and face edges, marked in red and green respectively for vertex A in the figure. The face edges of A define the face opposite to the vertex A. Two edges are called opposite (dashed edges) if they don't have any vertex in common. (b) Two-dimensional analogy of the representation of the marker pack in EE space. The axes labelled EE_1, EE_2 represent the coordinates of the markers in a hypothetical 2-dimensional EE space. In this space, each vertex of the tetrahedron (or equivalently, each marker of the marker pack) produces a cluster of points when recorded across many frames, due to measurement error. The solid boundary lines between the clusters define the regions in which all points are closest to a certain cluster, as in a Voronoi diagram.

2.3.1.2 Representation of vertices in EE space

We define the coordinates of each vertex of the tetrahedron v_i in the custom high-dimensional space as a vector of six components: the first three correspond to the lengths of the vertex edges of v_i sorted in descending order; the next three represent

the lengths of the face edges of v_i, also in descending order. Mathematically, we can use set-builder notation to express this as:

$$\begin{aligned}\vec{V_i} = [&sort(\{l_{i,k} | i \neq k\}); \\ &sort(\{l_{j,k} | j \neq k, j \neq i, k \neq i\})], \quad i,j,k \in \{1,2,3,4\},\end{aligned} \tag{2.1}$$

where $sort()$ denotes a function that sorts in descending order, $\vec{V_i}$ is the feature vector for the vertex v_i, and $l_{j,k}$ denotes the length of the edge connecting the vertices j and k.

We refer to this high-dimensional space as *EE space*, which stands for *extended edges space*. As we have mentioned, in practice the distances between the markers in the pack (i.e., the lengths of the tetrahedron's edges) will be measured with some degree of noise, and therefore the representation of a marker in EE space across many frames will produce a cluster of points. An analogy in 2D space is shown in Figure 2.1B.

If each marker is represented by a cluster in EE space, we can apply an unsupervised clustering algorithm to these coordinates to identify the individual markers (that is, to label them). We use MATLAB's implementation of the k-means algorithm, and call the approach *EE labelling*. The more separate these clusters are from each other, the better the labelling performance of the algorithm will be. Note however that there are constraints involved (for example, due to limitations on the size of the marker pack) and that beyond a certain point, the labelling performance won't improve significantly by increasing the distance between the clusters. We will take these aspects into account during the design process and explore them with the validation experiments.

With all this in mind, we can define more precisely the goal of our constrained design optimisation: to find sets of six edge lengths that maximise the distance between the markers' clusters (more precisely, their centroids) in EE space, while fulfilling the required constraints. In the optimisation, we assume that the vertices' clusters in EE space correspond to normally-distributed noise around the vertices' theoretical locations.

2.3.1.3 Analysis of possible solutions

Before formalising the optimisation problem, it is useful to review the requirements to define a tetrahedron from the lengths of its edges, and the number of possible solutions. To uniquely define a non-degenerate tetrahedron (that is, a tetrahedron with non-zero volume), we need to specify the following items (Wirth et al., 2009):

1. The lengths of its six edges. These need to fulfil two conditions, which we discuss below.

2. A pairing between opposite edges. There are in total 15 possible ways of arranging six edges' lengths in pairs.

3. The geometric interpretation of a triple of non-opposite edges. That is, we need to define whether a triple of non-opposite edges define a vertex or a face in the tetrahedron. Note that by defining just one set of non-opposite edges' lengths as a set of vertex or face edges, the geometric interpretation for the rest of the edges' lengths sets is fully defined (provided the pairing between opposite edges is known). This choice is also related to the geometric concept of dual polyhedra: two tetrahedra A and B are *dual* to each other if the four sets of vertex edges of A correspond to the four sets of face edges of B, and vice-versa. So by determining whether a triple of non-opposite edges define a vertex or a face, we are choosing between defining a certain tetrahedron or its dual.

We have mentioned in item 1 that a set of six lengths defines the edges of a non-degenerate tetrahedron if they fulfil two constraints. The first one is that the triangle inequality condition must be satisfied for each of the sets of face edges. This is a necessary but not sufficient condition (see Appendix A.1 for a simple example).

The second one is that the Cayley-Menger determinant D must be positive. The Cayley-Menger determinant is a three-dimensional analogue of Heron's formula for the area of a triangle. For a tetrahedron, it is a sixth-degree polynomial in six variables (the edges' lengths), and it is related to the tetrahedron's volume V

by $D = 288V^2$ (Wirth et al., 2009). Its expression is, in principle, different for a tetrahedron and its dual (see Appendix A.2). With the requirement that the Cayley-Menger determinant is positive for the edges' lengths, the set of conditions is sufficient (Wirth et al., 2009, Theorem 3.1).

For the same six edges' lengths, different pairings of opposite edges will in principle produce different representations of the tetrahedron's vertices in EE space. A tetrahedron and its dual will also have different representations of their vertices in EE space. Therefore, in our optimisation problem, we need to explore all 15 possible pairings of opposite edges, and for each one of them, study the cases of both dual tetrahedra.

2.3.1.4 Optimisation procedure

To solve the optimisation problem, we parametrised the lengths of the six edges of the tetrahedron using their maximum and minimum values, l_{min} and l_{max}. This makes two edges' lengths, which we call l_e and l_f, equal to these values: $l_e = l_{max}$ and $l_f = l_{min}$. The four remaining edges' lengths l_a, l_b, l_c, l_d are the independent variables to optimise, and we assume $l_a \leqslant l_b \leqslant l_c \leqslant l_d$. For convenience we use the following adimensional forms for the edges' lengths:

$$\begin{aligned} a &= \frac{l_a - l_{min}}{l_{max} - l_{min}}, \\ b &= \frac{l_b - l_{min}}{l_{max} - l_{min}}, \\ c &= \frac{l_c - l_{min}}{l_{max} - l_{min}}, \\ d &= \frac{l_d - l_{min}}{l_{max} - l_{min}}. \end{aligned} \tag{2.2}$$

We performed a grid search over the space defined by a, b, c, d to find the values that maximised h, the minimum square-distance between the vertices' representations in EE space. We performed searches considering all possible pairings of opposite edges, for both a tetrahedron and its dual. In all cases h_{max}, the maximum value of h at the optimal solutions, has the following form:

$$h_{max} = \alpha(l_{max} - l_{min})^2 \tag{2.3}$$

where α varies across pairings of opposite edges. The solutions found for each possible pairing between opposite edges are shown in Table 2.1 and the adimensional edges' lengths for each solution are presented in Table 2.2. Further details on the optimisation procedure and the grid search are included in Appendix A.3.

We find that a subset of solutions presents an interesting relationship with the golden ratio ϕ ($\phi = \frac{1+\sqrt{5}}{2}$): the edges' lengths that are not equal to l_{min} or l_{max} are simple functions of ϕ (see Tables 2.1 and 2.2). We call these solutions *golden solutions*. The golden solutions present the largest α values, and therefore the largest minimum distance between vertices' in EE space for any given l_{min} and l_{max} (see Equation 2.3). However, the size of the region of feasible designs each golden solution defines in the l_{min}, l_{max} plane varies. This is because there are constraints to the values of l_{min} and l_{max} we can select for our design, and some of these constraints depend on the specific solution selected. We explain this further in the following section. Further details on the golden solutions and on the regions of feasible designs are included in Appendices A.4 and A.5 respectively. Note that our golden solutions do not correspond to the golden sextuples described in Wirth et al. (2009).

2.3.1.5 Selected optimal solution

From the set of golden solutions (marked in bold in Table 2.2), we selected one for which the region of feasible designs in the plane of the design parameters l_{min} and l_{max} was largest. The maximum value h_{max} of the minimum square-distance between vertices in EE space for this solution has the form $h_{max} = \alpha(l_{max} - l_{min})^2$, with $\alpha = 0.58$ (see Table 2.1). The edges' lengths for this solution are:

$$\begin{aligned} {l_a}^* &= \frac{1}{\phi^2} l_{max} + \frac{1}{\phi} l_{min}, \\ {l_b}^* &= \frac{1}{\phi} l_{max} + \frac{1}{\phi^2} l_{min}, \\ {l_c}^* &= l_{max}, \\ {l_d}^* &= l_{max}, \\ {l_e}^* &= l_{max}, \\ {l_f}^* &= l_{min}, \end{aligned} \tag{2.4}$$

Table 2.1: Results for all pairings between opposite edges. For each possible pairing between opposite edges we computed the optimal sets of edges' lengths that maximised h, the minimum distance between vertices' in EE space. The pairs between opposite edges are represented in brackets. Note we parametrised all edges' lengths using their maximum and minimum values, and therefore for two of the edges we have $l_e = l_{max}$ and $l_f = l_{min}$. The number of solutions found for each pairing is shown. The maximum value of h for any pairing has the form $h_{max} = \alpha(l_{max} - l_{min})^2$; we report the value of α for each pairing. The solutions with largest α, or *golden solutions*, are highlighted in bold.

Pairing ID	Pairs of opposite edges	Number of solutions	$\alpha = \frac{h_{max}}{(l_{max}-l_{min})^2}$
1	$(l_f, l_a), (l_b, l_c), (l_d, l_e)$	3	0.50
2	$(l_f, l_a), (l_b, l_d), (l_c, l_e)$	**1**	**0.58**
3	$(l_f, l_a), (l_c, l_d), (l_b, l_e)$	4	0.50
4	$(l_f, l_b), (l_a, l_c), (l_d, l_e)$	**1**	**0.58**
5	$(l_f, l_b), (l_a, l_d), (l_c, l_e)$	**3**	**0.58**
6	$(l_f, l_b), (l_c, l_d), (l_a, l_e)$	**2**	**0.58**
7	$(l_a, l_b), (l_f, l_c), (l_d, l_e)$	4	0.50
8	$(l_a, l_b), (l_f, l_d), (l_c, l_e)$	**2**	**0.58**
9	$(l_a, l_b), (l_c, l_d), (l_f, l_e)$	**1**	**0.58**
10	$(l_f, l_c), (l_a, l_d), (l_b, l_e)$	4	0.50
11	$(l_f, l_c), (l_b, l_d), (l_a, l_e)$	2	0.51
12	$(l_a, l_c), (l_f, l_d), (l_b, l_e)$	2	0.51
13	$(l_a, l_c), (l_b, l_d), (l_f, l_e)$	2	0.51
14	$(l_b, l_c), (l_f, l_d), (l_a, l_e)$	2	0.51
15	$(l_b, l_c), (l_a, l_d), (l_f, l_e)$	4	0.50

with ϕ being the golden ratio.

The pairs of opposite edges are $({l_f}^*, {l_b}^*), ({l_a}^*, {l_e}^*)$ and $({l_c}^*, {l_d}^*)$, and the triple of non-opposite edges $\{{l_f}^*, {l_a}^*, {l_c}^*\}$ is defined as making up a vertex. This results in the following sets of vertex edges:

$$\begin{aligned} \{v_1\} &= \{{l_f}^*, {l_a}^*, {l_c}^*\}, \\ \{v_2\} &= \{{l_a}^*, {l_b}^*, {l_d}^*\}, \\ \{v_3\} &= \{{l_e}^*, {l_c}^*, {l_b}^*\}, \\ \{v_4\} &= \{{l_e}^*, {l_d}^*, {l_f}^*\}, \end{aligned} \tag{2.5}$$

Table 2.2: Solutions per pairing of opposite edges. The table shows for each pairing of opposite edges, the optimal adimensional edges' lengths found, a^*, b^*, c^*, d^*. Note an adimensional edge length equal to 1 corresponds to an edge length of l_{max}, and an adimensional edge length equal to 0 corresponds to an edge length of l_{min}, as per Equations 2.2. For all solutions, the two remaining adimensional edges' lengths of the tetrahedron are $e^* = 1$ and $f^* = 0$. For the golden solutions, the adimensional edges' lengths are a function of the golden ratio ϕ. The selected golden solution is highlighted in bold. Every solution can be built as a tetrahedron or its dual.

Pairing ID	Solution ID	a^*	b^*	c^*	d^*
1	1	0.0	0.0	0.5	0.5
1	2	0.5	0.5	0.5	0.5
1	3	0.5	0.5	1.0	1.0
2	1	0.0	0.0	$1/\phi^2$	$1/\phi$
3	1	0.0	0.0	0.0	0.5
3	2	0.0	0.0	0.5	1.0
3	3	0.5	0.5	0.5	0.5
3	4	0.5	0.5	1.0	1.0
4	1	$1/\phi^2$	$1/\phi$	1.0	1.0
5	1	0.0	0.0	$1/\phi^2$	$1/\phi$
5	2	0.0	$1/\phi^2$	$1/\phi$	1.0
5	3	$1/\phi^2$	$1/\phi$	1.0	1.0
6	1	0.0	$1/\phi^2$	$1/\phi$	1.0
6	**2**	$\mathbf{1/\phi^2}$	$\mathbf{1/\phi}$	**1.0**	**1.0**
7	1	0.0	0.0	0.5	0.5
7	2	0.0	0.5	1.0	1.0
7	3	0.5	0.5	0.5	0.5
7	4	0.5	1.0	1.0	1.0
8	1	0.0	0.0	$1/\phi^2$	$1/\phi$
8	2	0.0	$1/\phi^2$	$1/\phi$	1.0
9	1	0.0	$1/\phi^2$	$1/\phi$	1.0
10	1	0.0	0.0	0.0	0.5
10	2	0.0	0.5	0.5	1.0
10	3	0.5	0.5	0.5	0.5
10	4	0.5	1.0	1.0	1.0
11	1	0.0	0.41	0.41	0.71
11	2	0.29	0.59	0.59	0.59
12	1	0.29	0.59	0.59	1.0
12	2	0.41	0.41	0.41	0.71
13	1	0.0	0.41	0.41	0.71
13	2	0.29	0.59	0.59	1.0
14	1	0.29	0.59	0.59	0.59
14	2	0.41	0.41	0.41	0.71
15	1	0.0	0.0	0.0	0.5
15	2	0.0	0.5	0.5	0.5
15	3	0.5	0.5	0.5	1.0
15	4	0.5	1.0	1.0	1.0

where $\{v_i\}$ denotes the set of vertex edges for the vertex v_i. The vertices in Equation 2.5 are numbered in no particular order.

Given this solution, we would typically finalise our design by selecting a pair values for the l_{min} and l_{max} parameters that make h_{max} as large as possible (see Equation 2.3). However, there are limitations to the values of l_{min} and l_{max} we can select. Some limitations result from the conditions required for a set of edges' lengths to form a non-degenerate tetrahedron (see Section 2.3.1.3). Some other limitations are due to the practical implementation of the tetrahedron as a pack of motion capture markers. To explore the region of possible tetrahedra we can build with the selected solution, we represent these constraints in the l_{min}, l_{max} plane, along with the square-root of the h_{max} function; this is shown in Figure 2.2. We represent the square-root of h_{max}, rather than h_{max} itself, because its units (mm) are more easily interpretable.

The boundaries that limit the values of the design parameters l_{min} and l_{max} in Figure 2.2 result from the following constraints:

- The triangle inequality condition applied to each face of the tetrahedron. In the l_{min}, l_{max} plane this results in at most three linear boundaries, each defining a semiplane of feasible designs. For the selected solution they reduce to one linear boundary, shown in green in Figure 2.2.

- The Cayley-Menger determinant condition, which guarantees that the tetrahedron has a positive volume. It defines a region delimited by a sixth-degree bivariate polynomial in l_{min} and l_{max}, shown in yellow in Figure 2.2.

- The minimum edge length due to the motion capture system's resolution, $l_{min,res}$. It refers to the minimum distance between two markers such that these are still identified as two separate markers by the motion capture system. Its value will be affected by many factors, such as marker size, the cameras' parameters for postprocessing the grayscale image, and the number of cameras for which the markers are visible. Here, we are not concerned about estimating a precise value for it, but rather an estimate with enough safety margin. We

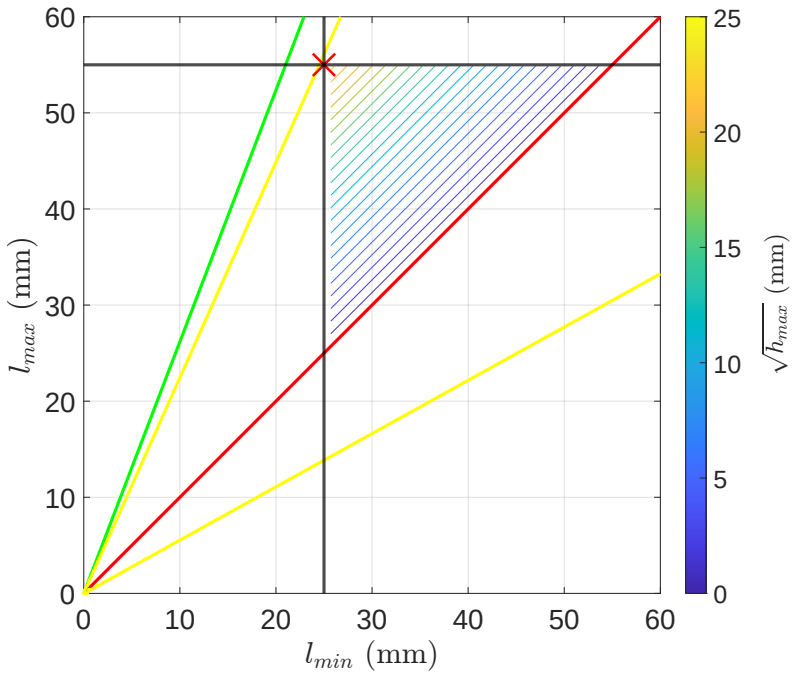


Figure 2.2: Design plane l_{min}, l_{max} for the selected solution. The points in the striped region represent combinations of l_{min}, l_{max} parameters that produce non-degenerate tetrahedra for the selected solution. The following boundaries are represented: triangular faces condition (green); positive Cayley-Menger determinant (yellow); $l_{min,res}$ minimum distance between markers determined by the motion capture system (l_{min} constant grey line); $l_{max,size}$ maximum admissible distance between markers for a bird headpack (l_{max} constant grey line); and $l_{min} < l_{max}$ (red). The Cayley-Menger determinant condition (yellow) is always more restrictive than the triangular conditions per face (green). Lines of constant $\sqrt{h_{max}}$ (mm) are shown in the feasible region. Note that the larger the difference between l_{min} and l_{max}, the larger the value for h_{max} (see Equation 2.3). The design point (red cross) corresponds to the feasible design with largest h_{max} ($\sqrt{h_{max}} = 22.9$ mm).

estimate a conservative value of $l_{min,res} = 25$ mm, which we validate later with experiments. This constraint is shown as a constant l_{min} line in grey in Figure 2.2.

- The maximum edge length due to size constraints on the marker pack, $l_{max,size}$. It relates to the maximum size the marker pack can have for its practical use. In our case, we are designing a marker pack to attach to the head of the bird, so we are limited by the size of the bird's head and by what the animal would be comfortable flying with. We estimate a value of $l_{max,size} = 55$ mm from measurements of the birds' heads. It is represented as a constant l_{max} grey line in Figure 2.2.

- By definition we have $l_{min} < l_{max}$, which restricts the design space to the

semiplane above the $l_{min} = l_{max}$ diagonal. This boundary is shown in red in Figure 2.2.

These constraints define the region of feasible designs in the l_{min}, l_{max} plane for a certain solution.

2.3.2 Construction of a bird's headpack

We used the selected optimal solution to design a 'headpack', a pack of motion capture markers to track head movements in hawks. The general workflow for the design of an optimal marker pack is shown in Figure 2.3.

For our application, we set $l_{min,res} = 25$ mm and $l_{max,size} = 55$ mm. Since we would like our design to have the largest h_{max} value possible, our chosen design point in the l_{min}, l_{max} plane is:

$$\begin{aligned} l_{max,design} &= l_{max,size} = 55\,\text{mm}, \\ l_{min,design} &= l_{min,res} = 25\,\text{mm}. \end{aligned} \tag{2.6}$$

The design point is marked with a red cross in Figure 2.2. With these values we can apply Equation 2.4 to obtain the remaining edges lengths, and the expressions in 2.5 to obtain the sets of vertex edges. Note that we can easily derive the distance $l_{i,j}$ between two vertices i and j as the element in common (intersection) between their sets of vertex edges, $l_{i,j} = \{v_i\} \cap \{v_j\}$. With the distances between vertices $l_{i,j}$ we can compute the coordinates of the vertices in an auxiliary reference frame (see Appendix A.6 for a possible way of doing this). The 3D coordinates of the vertices can then be imported to a CAD software package to design the marker pack for 3D printing. Figure 2.4 shows the process of designing the 3D-printed headpack for the birds. Additional design considerations for the birds' headpacks are discussed in Appendix A.7.

2.3.3 Validation experiments

We carried out experiments in the motion capture lab, a room of 12 m x 5.3 m x 3.3 m equipped with 22 motion capture cameras (Vantage V16; Vicon Motion Systems Ltd., Oxford, UK). The experiments had three objectives:

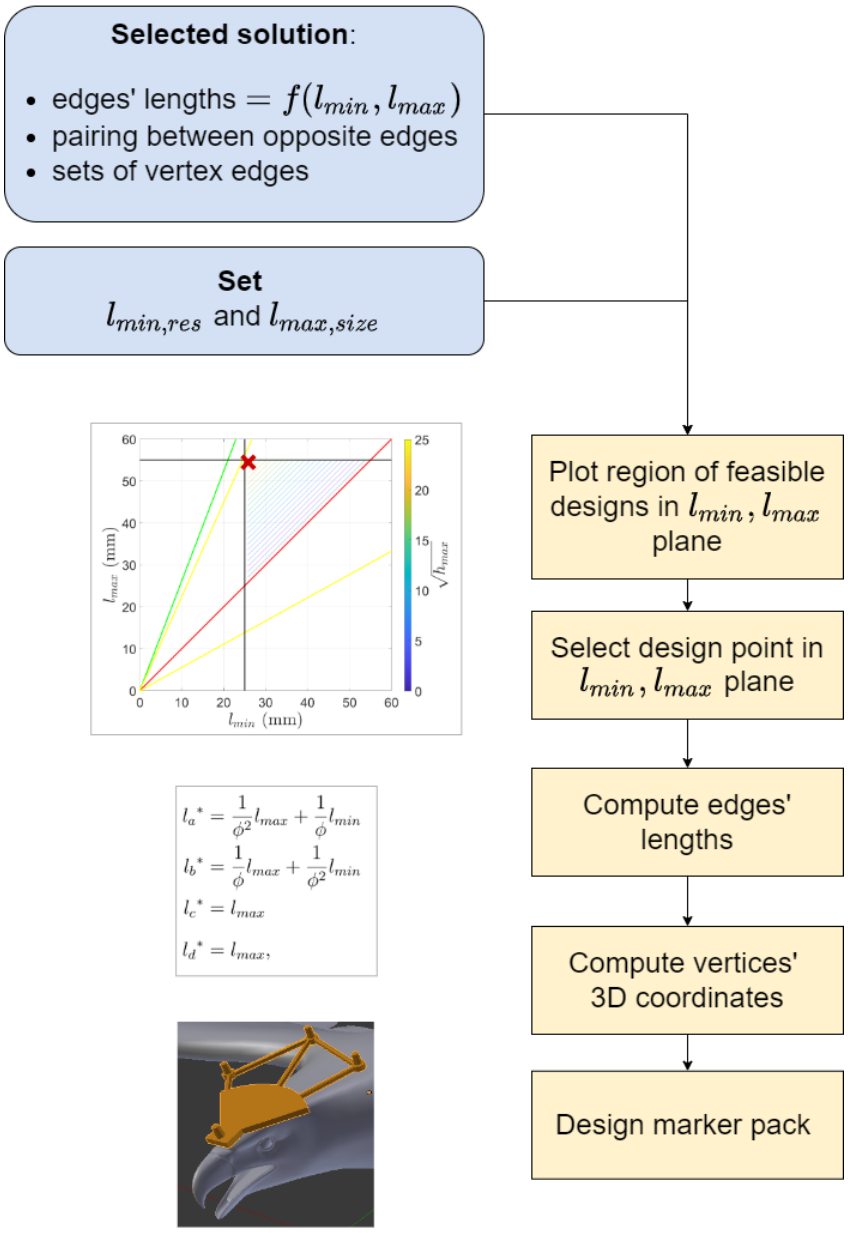


Figure 2.3: General workflow for the design of an optimal marker pack. Inputs are represented by blue rounded boxes and actions by orange square boxes. The process starts off with a selected optimal solution and the values for $l_{min,res}$ annd $l_{max,size}$. The selected solution implies knowledge of the edges' lengths as a function of l_{min}, l_{max}, the pairs of opposite edges and the sets of vertex edges that define a unique tetrahedron. With this data we can plot in the l_{min}, l_{max} plane the region of feasible designs, and select the desired values of l_{min} and l_{max}. We can then compute the edges' lengths and the vertices' 3D coordinates, which we can import in a CAD software to design the marker pack.

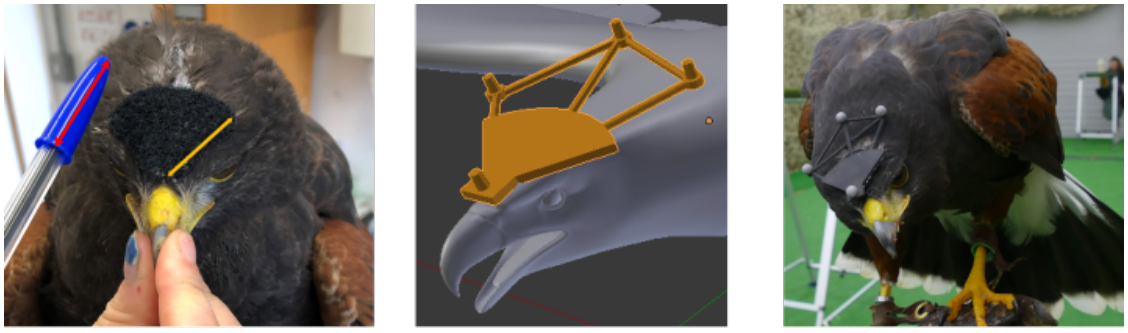


Figure 2.4: Construction of bird headpack. (a) Attachment point on the bird's head. We designed our bird headpack with a base that mimicked the Velcro pad on the bird's head (radius of Velcro's circular sector is approximately 30 mm; the pen cap is shown for reference and is 35 mm long). (b) We used a CAD of a red-tailed hawk, *Buteo jamaicensis*, to test different placements of the headpack on the bird's head. The bird CAD is a 3D model from Turbosquid[1]; it was developed for animation purposes and it is not anatomically correct, but it is sufficient for our purpose. (c) Final design for $l_{max} = 55$ mm on one of the birds. We added a lip to the headpack plate to facilitate its detachment.

[1] https://www.turbosquid.com/3d-models/rigged-realistic-3d-model-1615068

1. To validate our choice of $l_{min,res}$, the minimum distance between markers for them to be distinguishable by the motion capture system, using data across the motion capture volume.

2. To validate our design choice of $l_{max,design}$ for an optimal headpack for the birds. The selected $l_{max,design}$ matches the maximum admissible size on the bird's head, but ideally we would like to use headpacks that are as compact as possible, to minimise the interference with the bird. Therefore, we would like to determine the smallest l_{max} value for an optimal headpack that still guarantees a good labelling performance of the EE algorithm.

3. To evaluate the performance of the EE-labelling approach. To do this, we compare the EE-labelling to computed ground truth labels, and to the labelling from the commercial software Nexus v2.8.0 (Vicon Motion Systems Ltd., Oxford, UK). Nexus is the software associated with the motion capture system; we refer to it as the 'temporal algorithm' because its algorithm relies on temporal differences across frames.

For the first objective, we collected data with two markers placed on rigid supports of different lengths. For the second objective, we recorded headpacks of different sizes, under two different types of motion: spinning on a drill, and moving while attached to the birds in flight. For the third objective, we used the data of the headpacks moving on the drill and on the birds, and compared the labelling performance of the algorithms considered. The experiments are described in the following sections.

2.3.3.1 Experiments with two markers on supports

With this set of experiments, we wanted to estimate $l_{min,res}$, the minimum distance between two markers at which they can still be distinguished by the motion capture system. We were interested in good performance across the volume and over a reasonably long trial (45 s; a typical perching trial of a bird takes around 2 s).

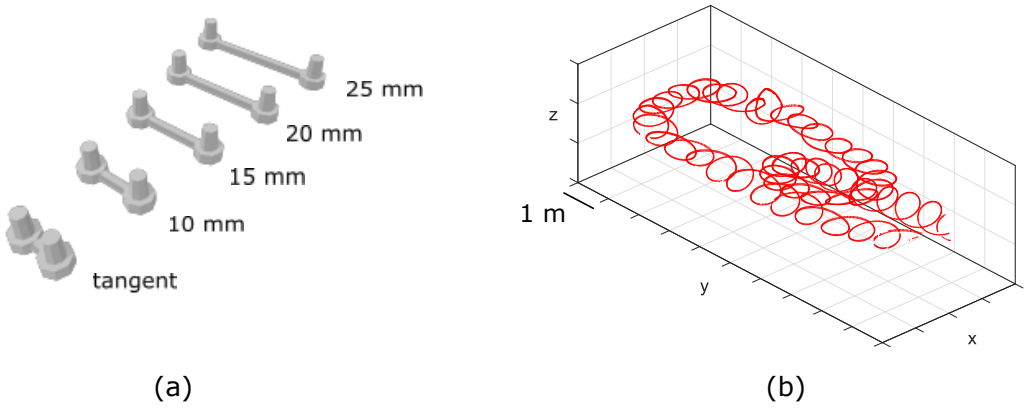


Figure 2.5: Experiments with two markers on supports. (a) CAD models of the supports used. (b) Example of a lasso trajectory, collected with $l_{min} = 25$ mm spinning at approximately 1.8 m height. The grid is represented at 1 m spacing.

To answer this question, we 3D-printed supports for two markers of 6.4 mm diameter. These fixed the markers' centres at 5 different lengths: 6.4 mm (i.e., the markers are tangent), 10 mm, 15 mm, 20 mm and 25 mm (see Figure 2.5a). We used them in two types of trial: static and lasso trials.

In the static trials, we recorded the markers attached to each of the supports, for 5 s while stationary on a table with good visibility from multiple cameras. We did two repetitions, 10 minutes apart. In the lasso trials, we attached a 30 cm thread

Table 2.3: Headpacks used in drill and bird experiments. The values for l_{min} and l_{max} for each design are shown, along with whether the design is optimal or not, and the type of trials it was used in (drill or bird trials). In the last column, $\sqrt{h_{max}}$ represents the minimum distance between any two vertices in the marker pack in EE space, for each design; note the low value for the suboptimal headpack. The two largest headpacks were not used in the bird trials.

l_{max} (mm)	l_{min} (mm)	optimal design?	used in drill trials?	used in bird trials?	$\sqrt{h_{max}}$ (mm)
45	22.6	✗	✓	✓	0.85
30	25	✓	✓	✓	3.8
45	25	✓	✓	✓	15.3
55	25	✓	✓	✓	22.9
75	33.4	✓	✓	✗	31.7
100	44.6	✓	✓	✗	42.2

to the centre of each support and used it to wave the markers around the capture volume in circles parallel to the floor, while walking along a predefined path (see Figure 2.5b). We did two repetitions, approximately 60 minutes apart.

All trials were recorded at 200 Hz and with a common calibration, which we don't expect to drift significantly between trials. Additional details on the design of the two-marker supports are included in Appendix A.8.

2.3.3.2 Experiments with headpacks

To inspect our choice of $l_{max,design}$, and evaluate the performance of the EE-labelling algorithm, we designed five optimal headpacks of increasing l_{max} values. For comparison, we additionally considered a suboptimal headpack, whose design was simply guided by trial and error with the motion capture system.

The six headpack designs we tested are summarised in Table 2.3. Note how the minimum distance between vertices in EE space $\sqrt{h_{max}}$, for each optimal design increases with increasing l_{max} (as per Equation 2.3), and notice its low value for the suboptimal headpack.

To design the optimal headpacks, we set $l_{min,res} = 25$ mm. For each value of l_{max} considered, we selected the point in the l_{min}, l_{max} plane within the region of feasible designs that had the largest possible h_{max} value (see Figure 2.6). For the three smallest optimal designs, the points of largest h_{max} fell on the $l_{min,res}$

boundary. For the two largest optimal designs, they fell on the Cayley-Menger determinant boundary. To avoid having a planar headpack, which may have more visibility issues (Naik, 2021), for the two largest designs we selected a design point slightly offset from the Cayley-Menger boundary in the l_{min} positive direction (by 0.01 mm). The selected design points for the optimal headpacks are marked with red crosses in Figure 2.6.

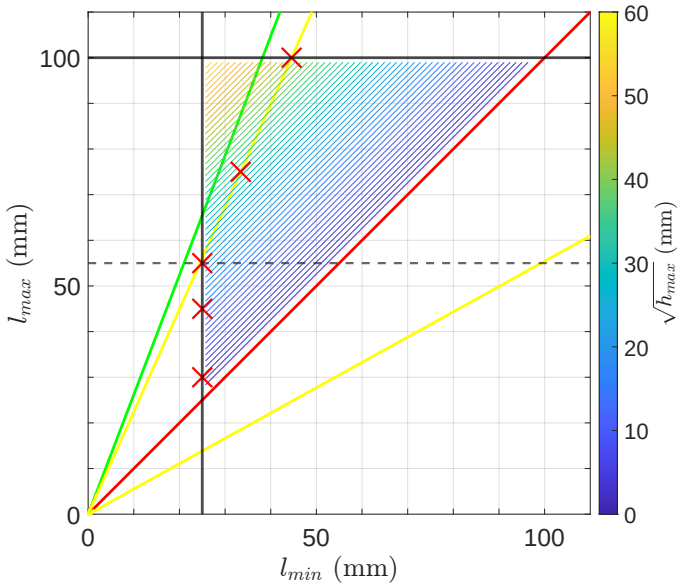


Figure 2.6: Design points for optimal headpacks. Each optimal headpack corresponds to a design point in the l_{min}, l_{max} plane (red crosses). The minimum distance between vertices in EE space, $\sqrt{h_{max}}$ is represented as a series of isolines. The boundaries represented correspond to the triangle inequality condition (green), the Cayley-Menger determinant condition (yellow), the minimum size due to the motion capture system's resolution (grey vertical line), and the maximum admissible size in the bird's head (grey horizontal line). The red line represents the $l_{min} = l_{max}$ boundary. The region of feasible designs is that delimited by the red, grey and yellow boundaries; note that in this case the striped area extends beyond the region of feasible designs. The dashed line marks the value for $l_{max,size}$ we estimated from measurements of the birds' heads.

We recorded the headpacks in two types of trials: drill trials and bird flights. In the drill trials, we attached each headpack to the contour of a drill accessory consisting on a Velcro pad of 15 cm diameter (see inset in Figure 2.7). We fixed the drill with the pad to a table which was approximately centred in the motion capture volume, using a 3-axis vice. The vice allowed us to orient the pad parallel to each of the planes of the motion capture coordinate system. We recorded the headpacks rotating at 4 different speeds, two sampling frequencies, and three orientations of the angular velocity vector, with three repetitions per condition (see Table 2.4). To

Table 2.4: Drill trials conditions. We recorded the headpacks rotating on a drill across a range of speeds, rotational axes and sampling frequencies; these are reported in the table. The drill speeds are the average values for each of the speed settings considered. Using a three-axis vice, we aligned the drill with one of the axes of the motion capture coordinate system $(\vec{x}, \vec{y}, \vec{z})$. We explored two sampling frequencies for the motion capture data, 100 Hz and 200 Hz. Note that the drill speeds are below the Nyquist frequencies for both sampling frequencies considered ($f_{Nyquist} = f_{sampling}/2$). We recorded three repetitions for each condition.

$\lvert\vec{\omega}_{drill}\rvert$	18,	25,	35,	40	(Hz)
$\vec{\omega}_{drill}$ axis	$\vec{x}$,	$\vec{y}$,	$\vec{z}$		
$f_{sampling}$	100,	200			(Hz)

measure the drill speed independently of the headpack markers, we attached a larger marker of 10 mm diameter to the contour of the pad, which we call the 'speed marker'. We covered the speed marker for one of the repetitions per condition, to later test if this affected the labelling performance of the temporal algorithm. We measured the drill's rotational speed at each of the settings with an optical tachometer as external validation prior to the experiments; the results are reported in Appendix A.9.

In the bird trials, we tested the three smallest optimal headpacks (up to $l_{max} = 55$ mm) and the suboptimal one on four different birds. The birds were trained to fly between two perches, 9 m apart. The lateral position of the perches was randomised between three options, 1 m apart (see schematic in Figure 2.8). We recorded perching trials, in which the bird flew back and forth between the perches, and obstacle avoidance trials, in which we added a set of obstacles 1.5 m ahead of the end perch. The set of obstacles was made up of four styrofoam pillars, of 2 m height and 0.3 m diameter each. For the analysis, we selected the trials that had two valid take-offs and two valid landings per recording; the final number of trials analysed are shown in Table 2.5. To ensure good quality data, we re-calibrated the system before recording all drill trials with the same headpack, and all bird trials with the same bird-headpack pair.

2.3.3.3 Labelling performance metrics

To assess the performance of the different labelling algorithms, we computed ground truth labels for the headpack data. We used an approach based on iteratively solving the procrustes problem (Schönemann, 1966; Horn et al., 1988), which is

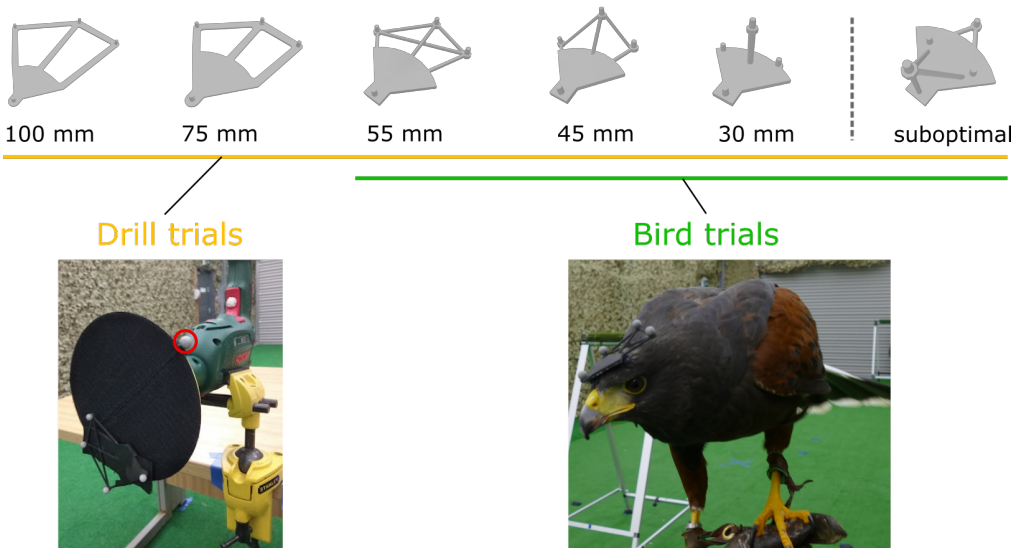


Figure 2.7: Experiments with headpacks. We 3D-printed five optimal headpacks with different l_{max} values, and one suboptimal one; their CAD models are shown at the top. We tested all headpacks on a set of drill trials, across different speed settings, rotation axes, and sampling frequencies. The left image shows the setup for a drill trial in which the rotation axis was aligned with the y-axis of the motion capture coordinate system. The larger marker used for independently estimating the drill speed ('speed marker') is highlighted with a red circle. We tested the three smallest optimal headpacks and the suboptimal one on a set of bird trials, in which the birds wore the headpacks while executing perching and obstacle avoidance manoeuvres.

Table 2.5: Number of bird trials analysed for each type of headpack. The table shows the number of trials analysed per bird and headpack type. The initials D, R, C, T correspond to the individual birds (based on their names, Drogon, Ruby, Charmander and Toothless). The column under 'Total' shows the total number of trials for the same headpack across all birds. All trials were recorded at 200 Hz.

	Perching trials					Obstacle avoidance trials				
Headpack	D	R	C	T	**Total**	D	R	C	T	**Total**
suboptimal	7	6	7	6	26	8	5	7	7	27
$l_{max} = 30$ mm	6	7	7	4	24	7	7	8	6	28
$l_{max} = 45$ mm	7	7	7	6	27	5	8	5	8	26
$l_{max} = 55$ mm	8	6	6	7	27	7	8	7	7	29

reliable but computationally intensive. We adapted the approach for the drill and the bird trials to make the most of the information available about the motion in each case (see summary diagrams in Appendix A.10). As a quality check, we computed across all trials the maximum distance between a ground truth labelled marker and its theoretical position, in a coordinate system linked to the headpack. In the drill trials, the maximum distance observed was 4 mm. In the bird trials,

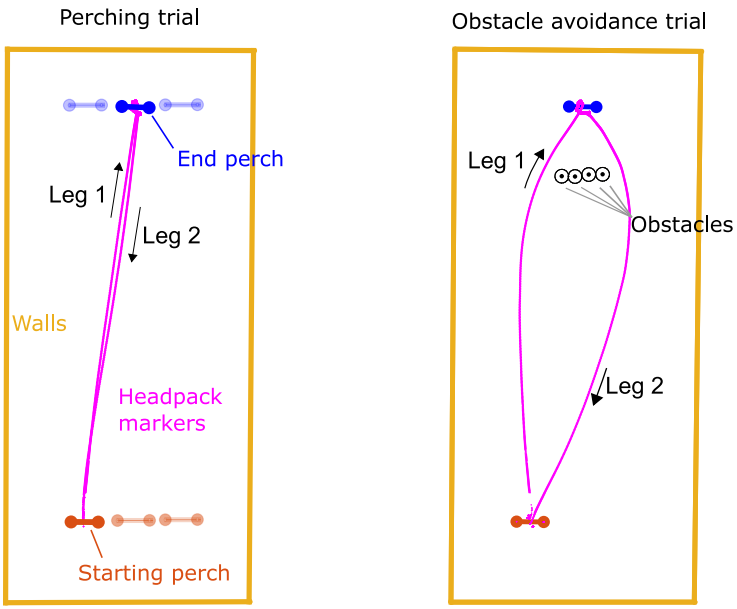


Figure 2.8: Types of bird trials. We recorded the headpacks on the birds while they executed two types of trials: perching and obstacle avoidance. The top-down view of two sample trajectories of each type is shown. In the perching trials (left) the bird flew from the starting perch (red) to the end perch (blue) and back. The perches were 9 m apart. Their lateral position was randomised across trials between three stations each, distributed with 1 m spacing (transparent). In the obstacle avoidance trial, we placed a set of four obstacles 1.5 m ahead of the end perch (black cylinders, centres highlighted with black markers). The reconstructed headpack markers are shown in magenta as well as the walls' positions (dark yellow).

the maximum distance for the frames with four headpack markers reconstructed was 4.5 mm; we focus on the frames with four markers because that is the scope of the EE-labelling algorithm. Note that if there were incorrect ground truth labels in any of the trials, we would expect this value to be around l_{min} or above. Since the values obtained are well below the lowest l_{min} of all headpacks, this is a good validation that our ground truth labels are reliable.

We computed the labels from the temporal algorithm (the 'temporal labels') by running a postprocessing pipeline in Nexus 2.8. The pipeline consisted of two steps: reconstruction and labelling. In the reconstruction step, the 3D position of the markers is computed and in the labelling step, their identity is determined across frames. For the labelling step, we defined the required templates in Nexus: six templates for the drill trials (one per headpack), and 16 templates for the bird trials (one per bird-headpack pair). Further details on the templates' definition

are included in Appendix A.11.

We found the temporal approach often switched labels, or started new labels along the trial (especially after marker occlusions). To make an unbiased comparison to the EE algorithm, we decided to focus on a temporal labelling that was consistent across the trial. To do this, we defined a mapping between the temporal and the ground truth labels using as reference the first frame with four markers fully labelled, and with at least one temporal label in the top 4 most frequent labels of the recording.

To assess the labelling performance of the EE and the temporal algorithms, we computed for each trial the percentage of frames with all four headpack markers correctly labelled. We focused on four-marker frames since that is the scope of the EE algorithm; otherwise the six EE coordinates cannot be computed. Both the EE algorithm and the ground truth labelling algorithm aim to label headpack markers only, so in all trials we pre-processed the data to remove any other markers present.

2.4 Results

We present the results of the experiments carried out to validate our choice of $l_{min,res}$, to explore our choice of $l_{max,design}$ for an optimal headpack for the birds, and to assess the performance of the EE labelling algorithm.

2.4.1 Choice of $l_{min,res}$

We used the data from the static trials to compute the mean distance between the two markers for each support. The results from aggregating the two repetitions per support are shown in Table 2.6. All mean distances are within 0.1 mm of their intended value, except for the tangent support, for which the system did not distinguish the markers as two separate ones. We used the measured mean distances as reference distances for each support. For the tangent support, we used the theoretical distance between the markers (6.4 mm).

With the data from the lasso trials, we computed the percentage of frames per trial in which the two markers were detected at a distance within 1 mm of its

Table 2.6: Distance measured for each two-markers support in the static trials. The table shows the mean distance d and the standard deviation σ for each support. The frames across the two repetitions per support are aggregated. Two markers were seen in all frames, except when using the tangent support, for which only one marker was reconstructed.

two-marker support	d (mm)	σ (mm)	$n_{samples}$
tangent	-	-	2590
10 mm	10.08	0.02	2462
15 mm	15.00	0.02	2805
20 mm	19.91	0.02	2265
25 mm	25.02	0.03	2258

reference value. The results per trial are shown in Figure 2.9, and the aggregated results are shown in Table 2.7. Only the support with the markers at 25 mm distance reaches over 95% of correct frames considering both trials aggregated. This confirms that our design choice for $l_{min,res}$ was reasonable.

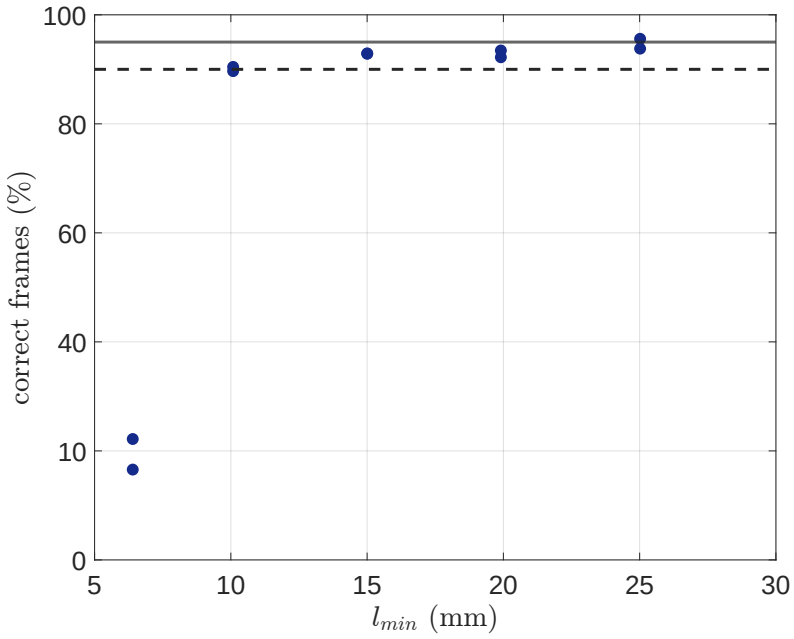


Figure 2.9: Percentage of correct frames for each two-markers support. The percentage of correct frames for each trial is represented. The x-axis l_{min} represents the design length of the two-marker support. Reference lines at 90% (dashed) and 95% (continuous) are shown.

2.4.2 Choice of $l_{max,design}$

With the data from the headpack trials, we want to determine what is the most compact headpack (i.e., the lowest l_{max} value) that still guarantees a good labelling

Table 2.7: Percentage of correct frames in lasso trials, and mean distance measured. The fraction of correct frames per support is reported, along with the mean d and standard deviation σ of the distance measured in the correct frames. All are computed by aggregating the data from the two repetitions per support.

two-marker support	correct frames (%)	d (mm)	σ (mm)	$n_{samples}$
tangent (6.4 mm)	19	6.34	0.42	8042
10 mm	90	9.95	0.22	8463
15 mm	93	15.07	0.20	8804
20 mm	93	20.07	0.21	8363
25 mm	95	25.20	0.20	8702

performance of the EE algorithm. First, we verify our expectation that optimal headpacks with larger l_{max} lead to more distant representations of their markers in EE space. Then, we examine how the EE labelling algorithm performs with the different headpacks.

Larger l_{max} leads to more distinct clusters in EE space

We measure the separation between clusters in EE space relative to their spread with the following metric ρ:

$$\rho = min\left(\frac{d_{min,i}}{\sigma_i}\right), \tag{2.7}$$

where $d_{min,i}$ represents the distance from the centroid of cluster i to its closest neighbouring centroid. The denominator σ_i represents the spread of cluster i, defined as the square-root of the trace of the covariance matrix of the EE coordinates:

$$\sigma_i = \sqrt{\sum_{k=1}^{6} {\sigma_{i,k}}^2}, \tag{2.8}$$

where $\sigma_{i,k}$ represents the variance of the kth coordinate in EE space for cluster i. Note that the EE coordinates are the edges' lengths. The $min()$ function returns the minimum over the four marker clusters. We computed the ρ metric for each trial using the ground truth labelled data.

Figure 2.10 shows violin plots for ρ, grouped by headpack type, for the drill and the bird trials. The mean and median values for ρ over all trials with the same

headpack type are shown in Table 2.8. In both drill and bird trials, the mean ρ value increases with optimal headpacks of increasing size, and is lowest for the suboptimal headpack. This validates our design objective and confirms that larger optimal headpacks have more distant clusters relative to the clusters' spread in EE space. We also find larger variability in the ρ values per trial when larger headpacks are used, both in drill and bird trials. This could be explained by a decrease in the variance of the edges' lengths for larger headpacks.

Table 2.8: Mean, median and standard deviation ρ for each headpack type. The mean ($\overline{\rho}$), median ($med(\rho)$) and standard deviation ($std(\rho)$) for the ρ metric are reported for the drill and the bird trials, aggregating all trials recorded with the same headpack.

	drill trials			bird trials		
headpack	$\overline{\rho}$	$med(\rho)$	$std(\rho)$	$\overline{\rho}$	$med(\rho)$	$std(\rho)$
suboptimal	2.0	1.5	1.1	1.9	2.2	0.9
$l_{max} = 30$ mm	10.7	7.3	6.4	7.6	7.4	1.9
$l_{max} = 45$ mm	11.0	10.1	3.7	31.1	30.7	7.4
$l_{max} = 55$ mm	72.7	53.2	43.3	69.7	66.5	22.3
$l_{max} = 75$ mm	80.8	51.4	60.0	-	-	-
$l_{max} = 100$ mm	144.4	82.7	118.3	-	-	-

In the violin plots, trials of different conditions are aggregated. We check whether the headpack type is the main effect behind the ρ increase by inspecting the rest of the trial conditions; these are represented on Figures 2.11 and 2.12. In both the drill and the bird trials, the headpack type appears to drive the increasing ρ values. A Kruskal-Wallis test confirms significant differences between the distributions of ρ when grouped by headpack type, in both drill ($\chi^2(5, N = 432) = 368.4$, $p < 0.01$) and bird trials ($\chi^2(3, N = 214) = 195.6$, $p < 0.01$). In particular, the difference between the mean ranks of the distributions is tested.

No significant differences are found in the distribution of the ρ values per trial when grouped by any other trial condition (speed setting, sampling frequency, bird or type of manoeuvre), except when grouped by the rotational axis in the drill trials ($\chi^2(2, N = 432) = 30.4$, $p < 0.01$). This is likely due to the large ρ values observed for most of the drill trials recorded rotating around the z-axis (see Figure 2.11b, yellow markers). This could be explained by an increased visibility of the

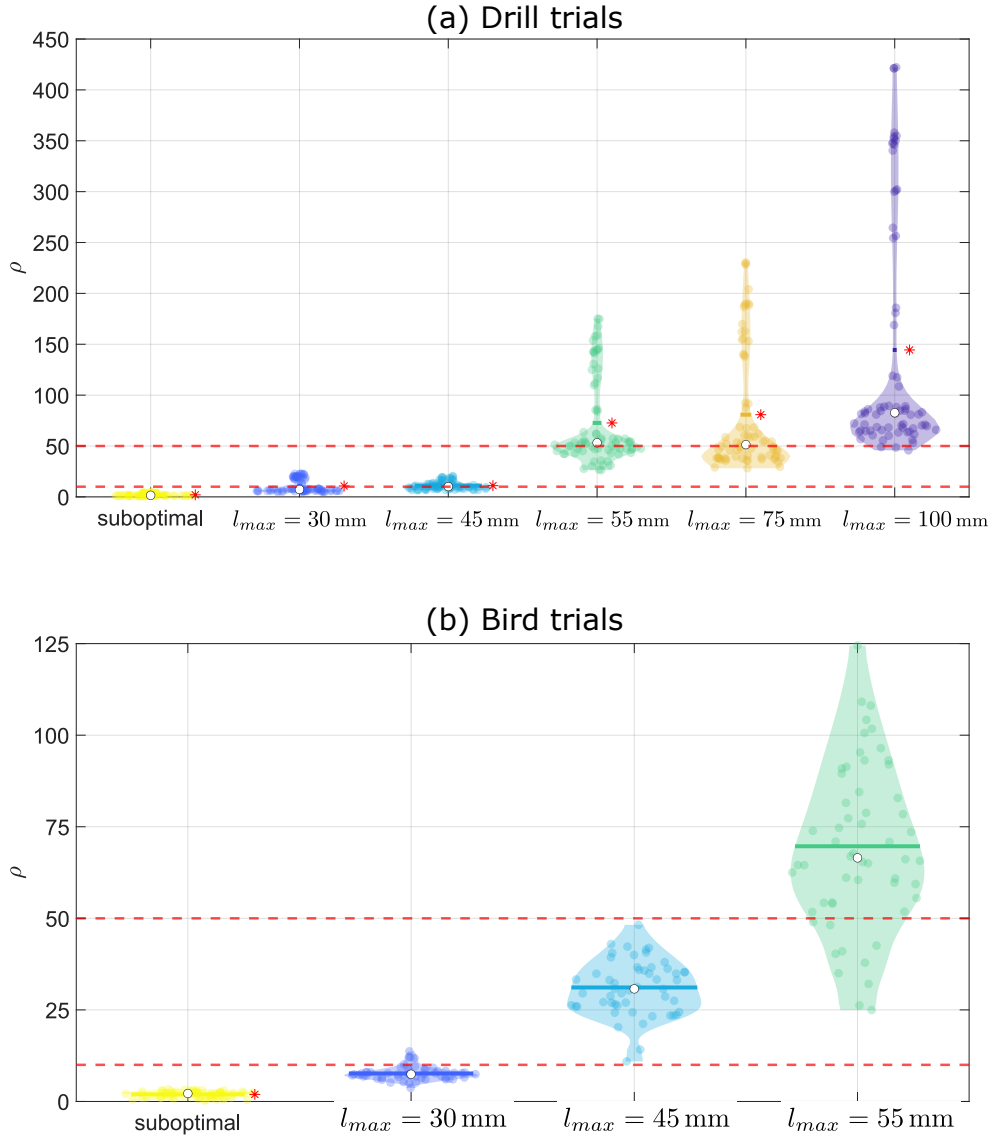


Figure 2.10: Violin plots for ρ per headpack type. Results for drill trials (a) and bird trials (b). The data points are coloured by headpack type. The mean value is marked in the distribution with a coloured horizontal bar; when too narrow to be visible its location is marked by a red star. The median is represented by a white marker. Each optimal headpack is identified by its l_{max} value in the x-axis. Reference lines (red dashed) are plotted at $\rho = 10$ and 50.

markers by the motion capture cameras when the headpacks rotate around the z-axis. With more cameras contributing to the estimate of the markers' positions, the variance of the edges' lengths $\sigma^2_{i,k}$ decreases and ρ increases. We find larger ρ values for the z-axis trials within all headpacks, except in those recorded with $l_{max} = 45$ mm. Upon further inspection, we find the trials with $l_{max} = 45$ mm have larger residuals in the reconstruction of the markers than the rest of the drill trials (see Figure A.5); this could explain the absence of the z-axis effect. Further details on the residuals and their potential effect on ρ are included in Appendix A.12.

EE labelling improves with larger optimal headpacks

We have confirmed that optimal headpacks of larger l_{max} produce more distinct clusters in EE space (i.e., more distant clusters relative to their spread). But how compact can we make the headpacks while still retaining a good labelling performance of the EE algorithm? To inspect this, we computed the labelling performance of the EE algorithm as the fraction of frames correctly labelled for each trial. We inspect the frames with four markers only, since that is the scope of the EE algorithm.

Figures 2.13a and b show the fraction of candidate frames correctly labelled by the EE algorithm, for the drill and the bird trials, grouped by headpack type. The mean and median values across all trials with the same headpack are shown in Table 2.9.

When used with an optimal headpack, the EE labelling algorithm labels correctly 100% of the candidate frames on average across headpack sizes, in both drill and bird trials. In fact in the drill trials, all candidate frames are labelled correctly when any optimal headpack is used. In the bird trials, which cover a larger capture volume and are more representative of our desired application, we find differences between the smaller sizes, with $l_{max} = 45$ mm outperforming $l_{max} = 30$ mm. In contrast, when using the suboptimal headpack the EE algorithm underperforms, labelling on average 95.0% of the candidate frames correctly in the drill trials, and only 76.6% in the bird trials (see Table 2.9).

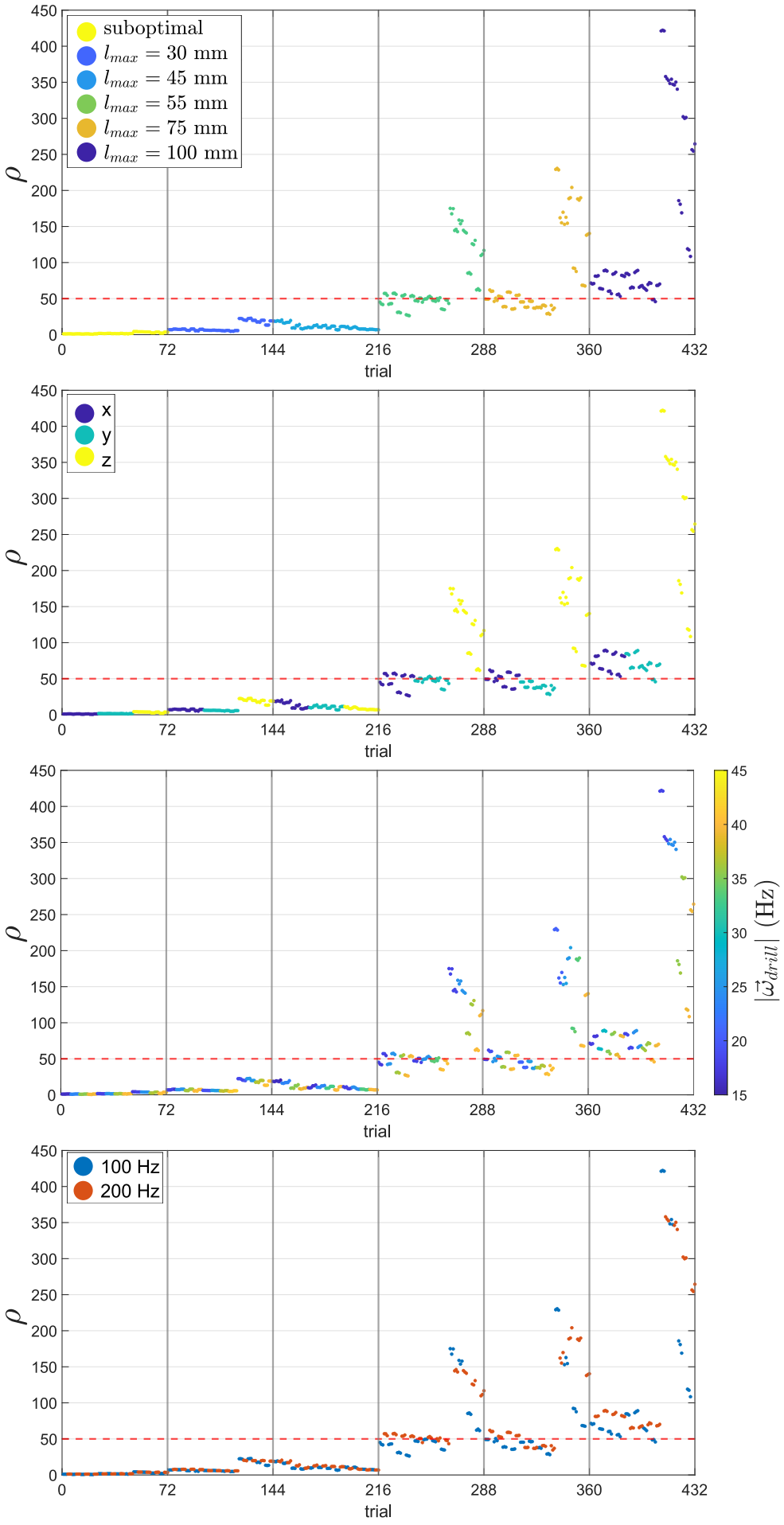


Figure 2.11: ρ in drill trials. The ρ metric for each drill trial is represented. Each datapoint is coloured by headpack type in (a), drill rotational axis in (b), drill speed in (c) and motion capture sampling frequency in (d). The vertical lines indicate the last trial within a set using the same headpack.

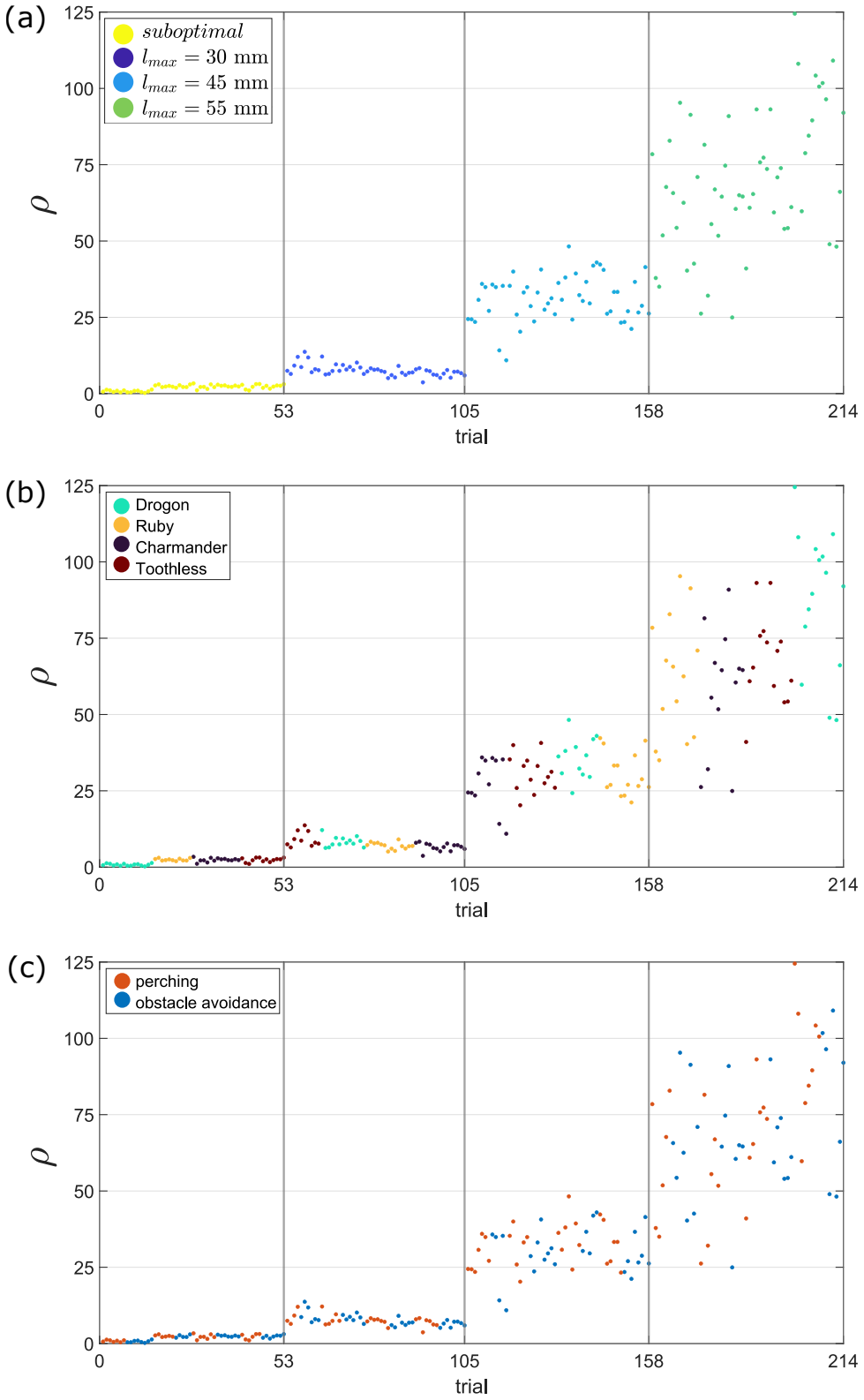


Figure 2.12: ρ in bird trials. The ρ metric for each bird trial is represented. Each datapoint is coloured by headpack type in (a), bird in (b) and type of manoeuvre in (c). The vertical lines indicate the last trial within a set using the same headpack; note they are not equally spaced for the birds dataset.

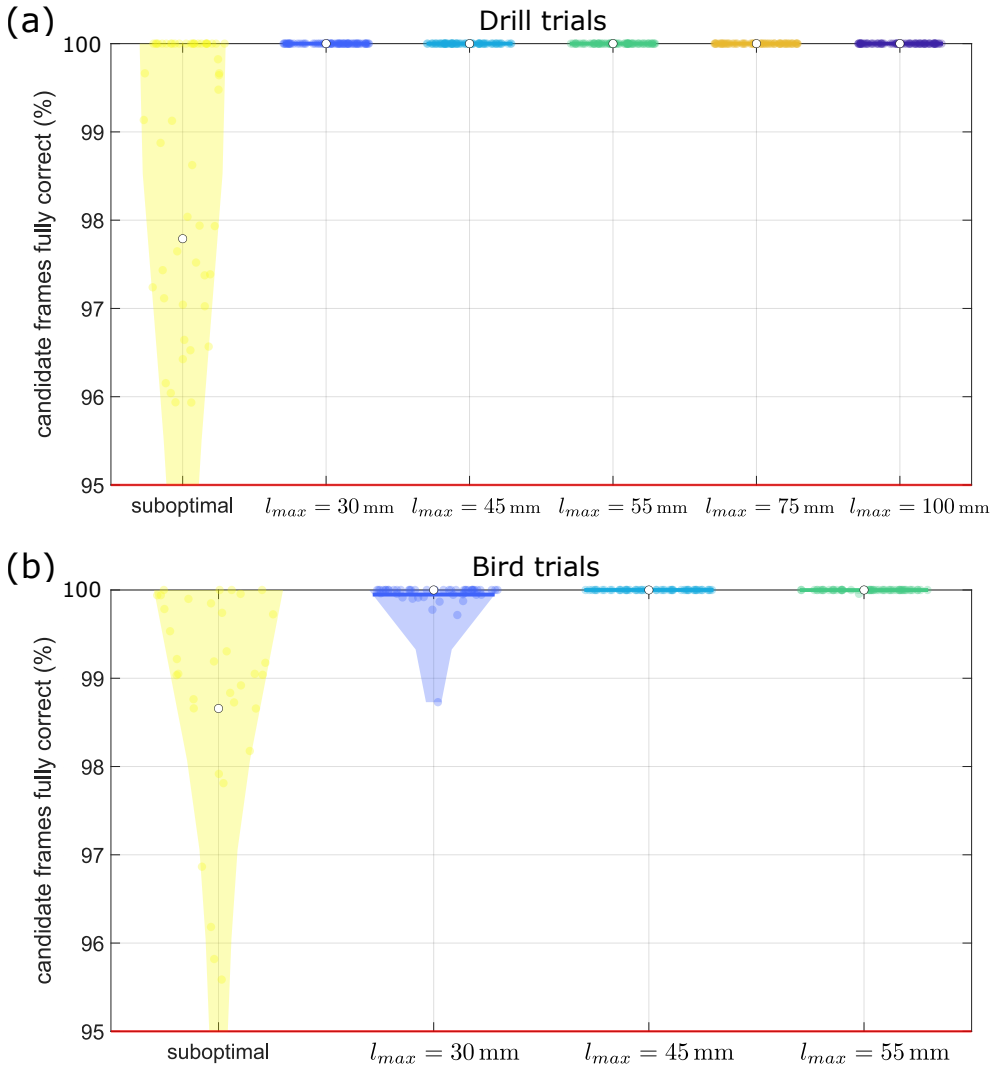


Figure 2.13: Labelling performance of EE algorithm across headpack types. Figure (a) shows the results for the drill trials, and figure (b) shows the results for the bird trials. Data points and probability density function are coloured by headpack type. Note that the y-axis origin is at 95 %. The mean value is marked in the distribution with a coloured horizontal bar, and the median with a white marker. The mean value for the suboptimal headpack is out of the figure bounds (see Table 2.9).

Inspecting the rest of the trial conditions, the labelling performance of the EE algorithm appears to be driven by headpack type and size in Figures 2.14 to 2.17 (panels a and b). A Kruskal-Wallis test confirms significant differences in the distribution of the labelling performance when trials are grouped by headpack type in both drill ($\chi^2(5, N = 432) = 268.1$, $p < 0.01$) and bird datasets ($\chi^2(3, N = 214) = 154.3$, $p < 0.01$). Additionally in the drill trials, significant differences are found in the distribution of the labelling performance when trials are grouped

Table 2.9: Mean and median labelling performance of the EE algorithm for each headpack type The mean and median values for the percentage of correctly labelled frames with the EE labelling algorithm are reported for the drill and the bird trials, aggregating all trials recorded with the same headpack. The two largest optimal headpacks were not tested in the bird trials due to size constraints.

	drill trials		bird trials	
headpack	mean (%)	median (%)	mean (%)	median (%)
suboptimal	94.97	97.79	76.62	98.66
$l_{max} = 30$ mm	100.00	100.00	99.95	100.00
$l_{max} = 45$ mm	100.00	100.00	100.00	100.00
$l_{max} = 55$ mm	100.00	100.00	100.00	100.00
$l_{max} = 75$ mm	100.00	100.00	-	-
$l_{max} = 100$ mm	100.00	100.00	-	-

by rotational axis ($\chi^2(2, N = 432) = 27.0$, $p < 0.01$). This matches the results obtained for the ρ metric and reflects that the design and the labelling performance are linked. It also points to an interesting result for the suboptimal headpack: when it rotates around the z-axis, the EE algorithm performs particularly well, presumably due to the increased visibility in this position reducing the uncertainty in the edges' lengths, and thus enabling the distinction of the closest clusters in EE space. No significant differences between the distributions of the labelling performance are found when grouping the drill trials by speed setting, sampling frequency, or number of repetition; or when grouping the bird trials by bird or type of manoeuvre. This suggest the headpack used is a main factor driving the labelling performance of the EE algorithm, as we intended.

The headpack with $l_{max} = 45$ is the most compact headpack with a 100% of correctly labelled frames across all trials, so we determine it would be a good candidate for our application. Additionally, the minimum distance between its clusters seems to be robust to markers' reconstructions with large residuals (see Appendix A.12). However, this should be tested directly.

2.4.3 Comparison between EE and temporal labelling

We compare the labelling performance of the EE and the temporal algorithms in Figures 2.14 to 2.17.

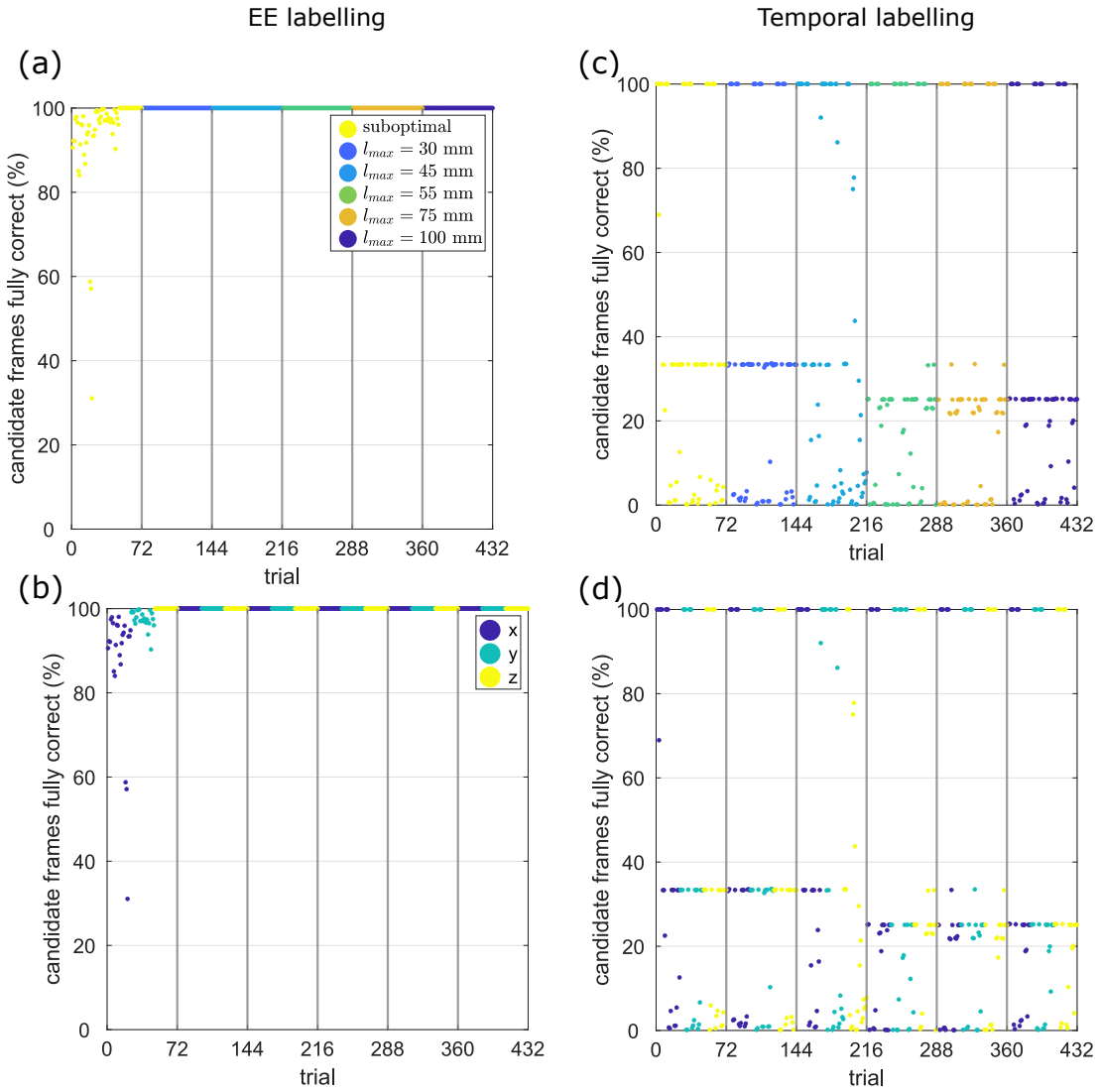


Figure 2.14: Labelling performance of EE and temporal algorithm in drill trials - headpack type and rotation axis effect. The labelling performance per trial is represented, with colour showing headpack type in (a) and (c), and rotational axis in (b) and (d). Left panels (a and b) show the results for the EE labelling algorithm, and right panels show the results for the temporal algorithm (c and d). Vertical lines indicate the last trial within a set using the same headpack.

In the drill trials, the EE labelling algorithm is more robust than the temporal algorithm to the different sampling frequencies and drill speeds considered (Figure 2.15, compare panel a to c, and b to d). As discussed in the previous section, the EE algorithm labelled all candidate frames correctly when used with any of the optimal headpacks; this is the case regardless of the drill speed or sampling frequency of the trial (see Figures 2.14a, 2.15a and 2.15b). In contrast, the temporal algorithm only reaches 100% of labelling performance in trials that are recorded at either the highest sampling frequency (200 Hz), the lowest drill speeds (18 or 25 Hz), or both. A Kruskal-Wallis test confirms that for the temporal algorithm, the distribution of the labelling performance per trial is significantly different when grouped by speed setting ($\chi^2(3, N = 432) = 109.3$, $p < 0.01$) or by sampling frequency ($\chi^2(1, N = 432) = 156.8$, $p < 0.01$). No significant differences are found when the labelling performance of the temporal algorithm is grouped by headpack type (see Figure 2.14c) or rotational axis (see Figure 2.14d). No significant differences were found when grouping the labelling performance of the temporal algorithm by the number of the repetition either, showing that the occlusion of the speed marker did not significantly affect the performance.

These results confirm our expectations on the performance of a time-independent labelling algorithm: because the EE algorithm relies on features that are independent of the frame they are computed at, it is more robust to high speeds and low sampling frequencies than an algorithm based on temporal differences. Note that the fastest drill speed measured over all trials (42 Hz) is below the Nyquist frequency for both sampling frequencies considered ($f_{Nyquist} = f_{sampling}/2$), so the motion should be resolvable without aliasing distortion.

We inspect why the temporal algorithm presents such low labelling performance in many of the drill trials. The labelling performance of the temporal algorithm appears to group around certain values (see panels c and d in Figures 2.14 and 2.15). We selected four trials whose labelling performance values were within these groups, and compared their temporal labels to the EE and ground truth labels in Figure 2.16.

The selected trial for which the temporal algorithm labels correctly less than 5% of the candidate frames shows most markers either unlabelled or with a mix of labels assigned to them (Figure 2.16e, left). In the trial with 25% of frames correctly labelled, the labels seem randomly assigned to the four available markers (Figure 2.16d, left). The trial with 33% labelling performance shows one marker correctly labelled across the trial, but the rest of them consistently mislabelled (Figure 2.16c, left). The trial with 70% of the candidate frames correctly labelled shows markers with noisy reconstructions (Figure 2.16b, left), which likely affects the performance of the temporal algorithm. The EE and ground truth labels are also shown for comparison (Figures 2.16b to e, centre and right columns). Note how the EE algorithm outperforms the temporal one in all cases, and only mislabels a few markers when using the suboptimal headpack (Figure 2.16b, centre).

In the bird trials, the EE labelling algorithm also outperforms the temporal approach when the optimal headpacks are used, with a minimum of 98.7% of candidate frames correctly labelled across these trials, versus 30.4% for the temporal method (see Figure 2.17a and b). The suboptimal headpack performs better with the temporal labelling than with the EE approach in many trials; this makes sense considering we designed it by trial-and-error in the motion capture lab. We did not find significant differences in the distribution of the labelling performance of the temporal algorithm when grouping the bird trials by headpack type or type of manoeuvre recorded. We found significant differences when grouped by birds, if a significance level of $\alpha = 0.05$ is considered ($\chi^2(3, N = 214) = 8.8, p = 0.03$). This could be due to the Drogon trials with the suboptimal headpack, which show the most extreme outliers for the labelling performance of the temporal algorithm, possibly due to these trials presenting larger residuals than most bird trials (see Figure A.5c).

We found that in many bird trials in which the EE labelling outperforms the temporal one, there is often a problem of consistency in the temporal labelling. At some point of the trial (often at the start) the temporal labels change, either by switching between existing labels or by starting new labels. Examples of this are shown in Figure 2.18. From visual inspection this appeared to happen more

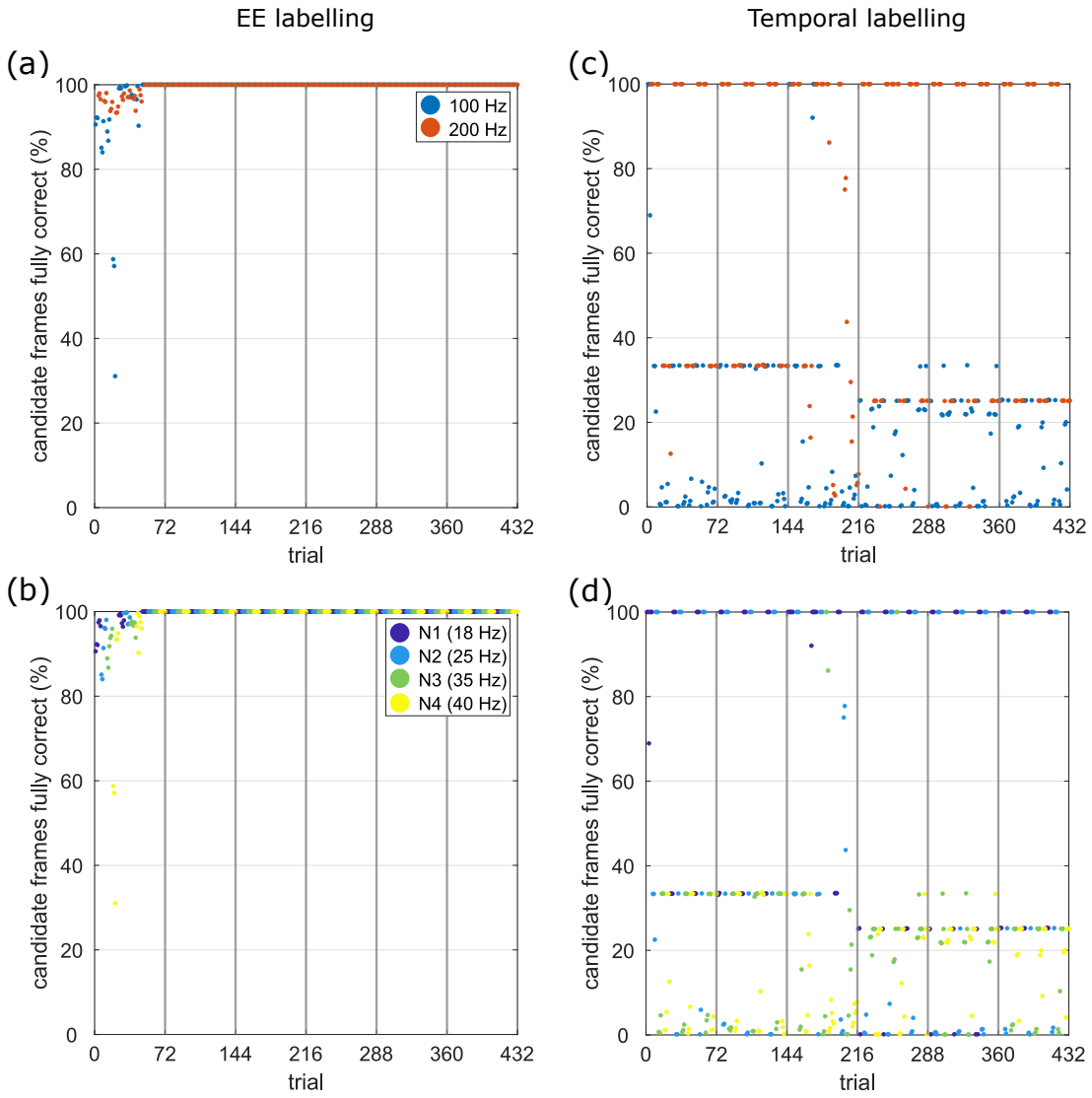


Figure 2.15: Labelling performance of EE and temporal algorithm in drill trials - sampling frequency and drill speed effect. The labelling performance per trial is represented, with colour showing sampling frequency in (a) and (c), and speed setting in (b) and (d). Left panels (a and b) show the results for the EE labelling algorithm, and the right panels show the results for temporal algorithm (c and d). Vertical lines indicate the last trial within a set using the same headpack.

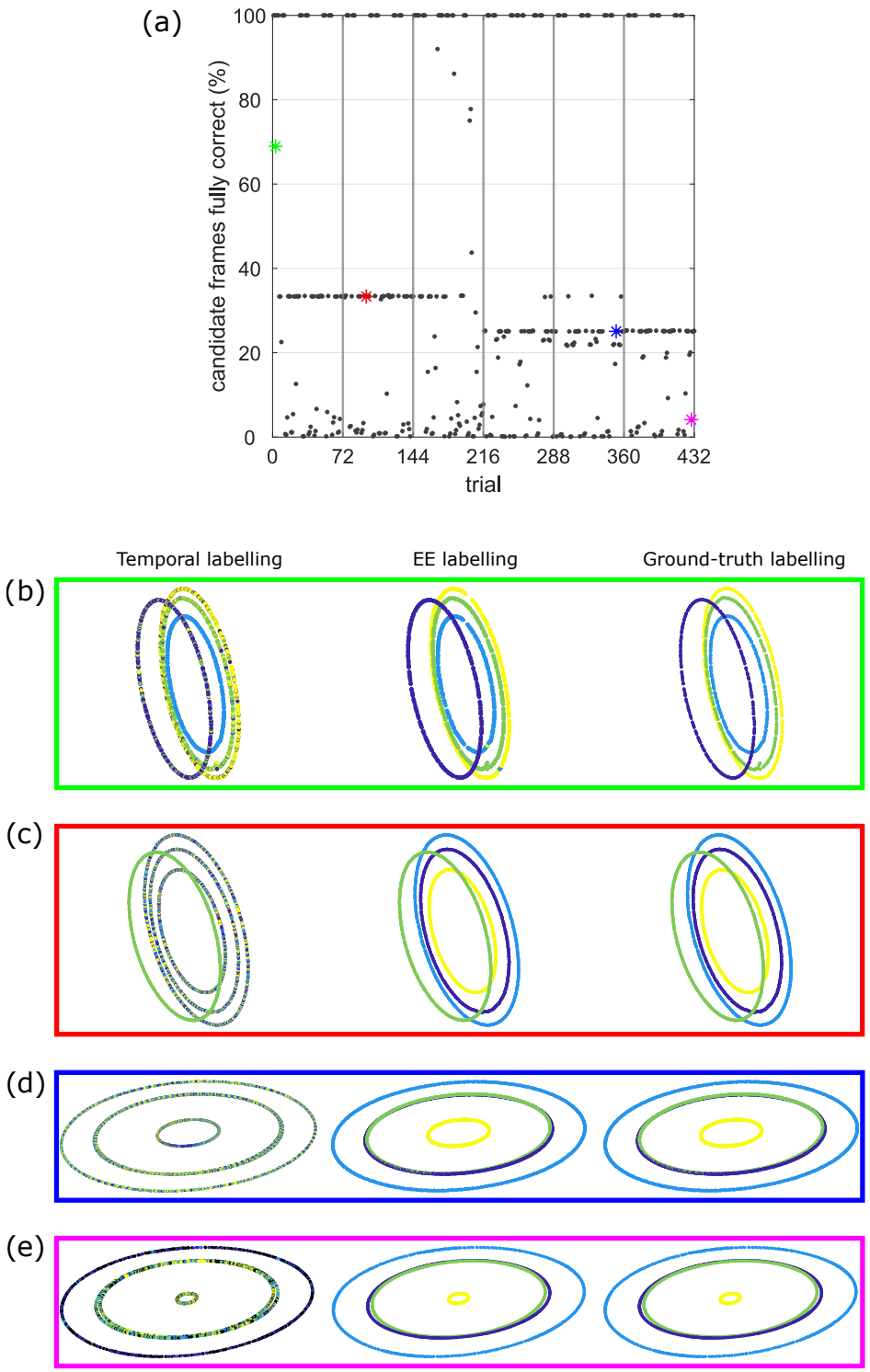


Figure 2.16: Labellings comparison in selected drill trials. Figure (a) shows the labelling performance of the temporal algorithm. Figures (b) to (e) show for the selected trials the labelling output from the temporal algorithm (left), the EE algorithm (centre) and the ground truth algorithm (right). The labelling performance of the selected trials is marked with stars in (a); their colour matches that of the insets (b) to (e). In (b) to (e) markers are coloured by their assigned label, and those left unlabelled by the temporal algorithm are shown in black. The conditions of each of the selected drill trials are: (b) suboptimal headpack rotating around the x-axis at 18 Hz, and a sampling frequency of 100 Hz; (c) optimal headpack with $l_{max} = 30$ mm, rotating around the x-axis at 40 Hz, and a sampling frequency of 200 Hz; (d) optimal headpack with $l_{max} = 75$ mm, rotating around the z-axis at 35 Hz, and a sampling frequency of 200 Hz; (e) optimal headpack with $l_{max} = 100$ mm, rotating around the z-axis at 40 Hz, and a sampling frequency of 100 Hz. Ground truth and EE labels could be computed for all markers in all trials shown, except for trial (b) in which only 95.75% of the markers have an EE label (only markers with an EE label are shown). The two EE labels switched for some of the frames in (b) correspond to the two closest clusters defined by the suboptimal headpack in EE space.

frequently after markers were occluded (for example if the headpack is occluded by the wings, when the bird flaps with large amplitude at take-off).

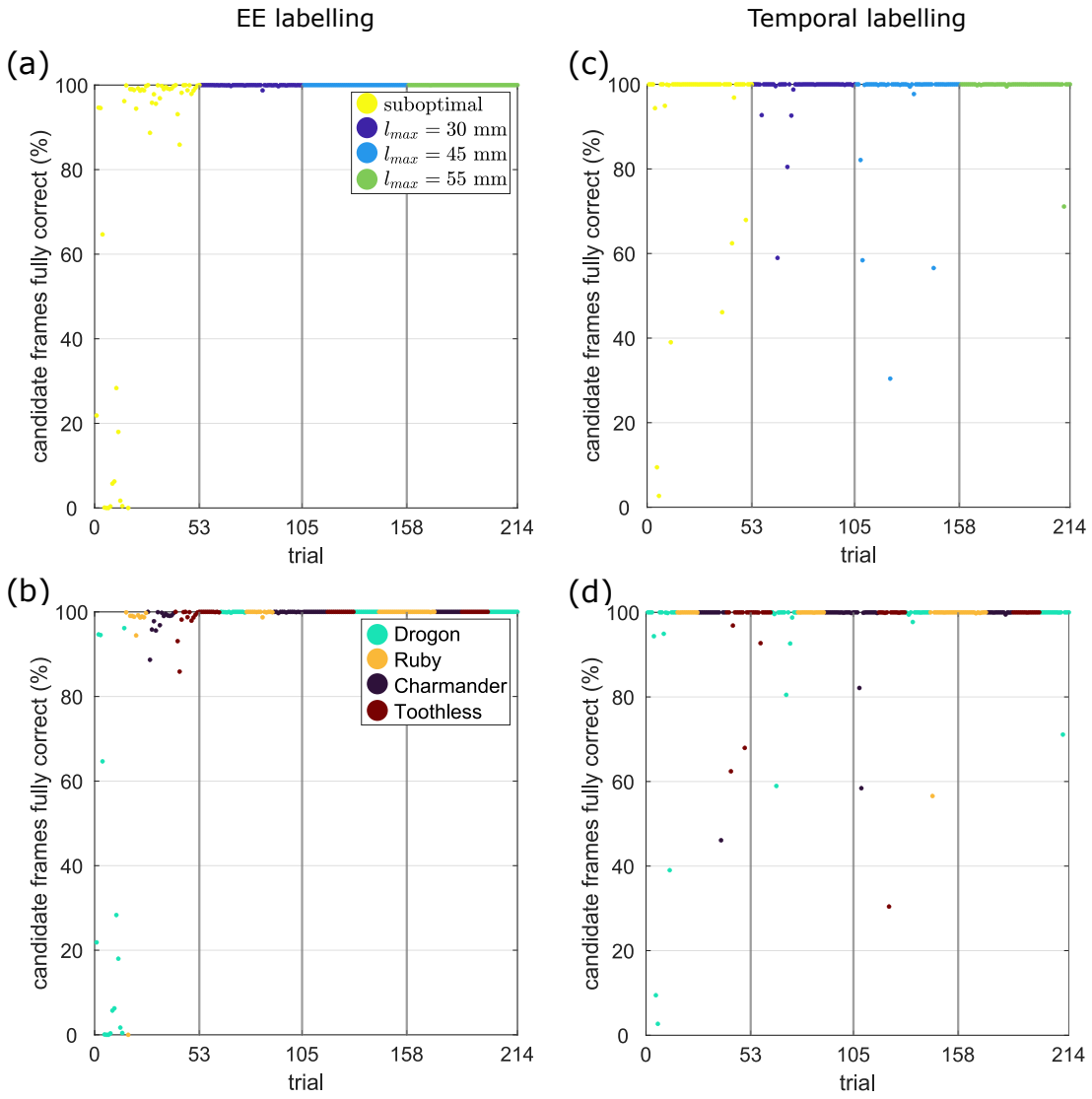


Figure 2.17: Labelling performance of EE and temporal algorithm in bird trials. The labelling performance per trial is represented, with colour showing headpack type in (a) and (c), and bird executing the flight in (b) and (d). Left panels (a and b) show results for EE labelling algorithm, and right panels show results for temporal algorithm (c and d). Vertical lines indicate the last trial within a set using the same headpack; note they are not equally spaced for the bird dataset.

2.5 Discussion

Marker-based motion capture systems are currently the gold standard for acquiring precise measurements of motion. However, these systems are pushed to their limits

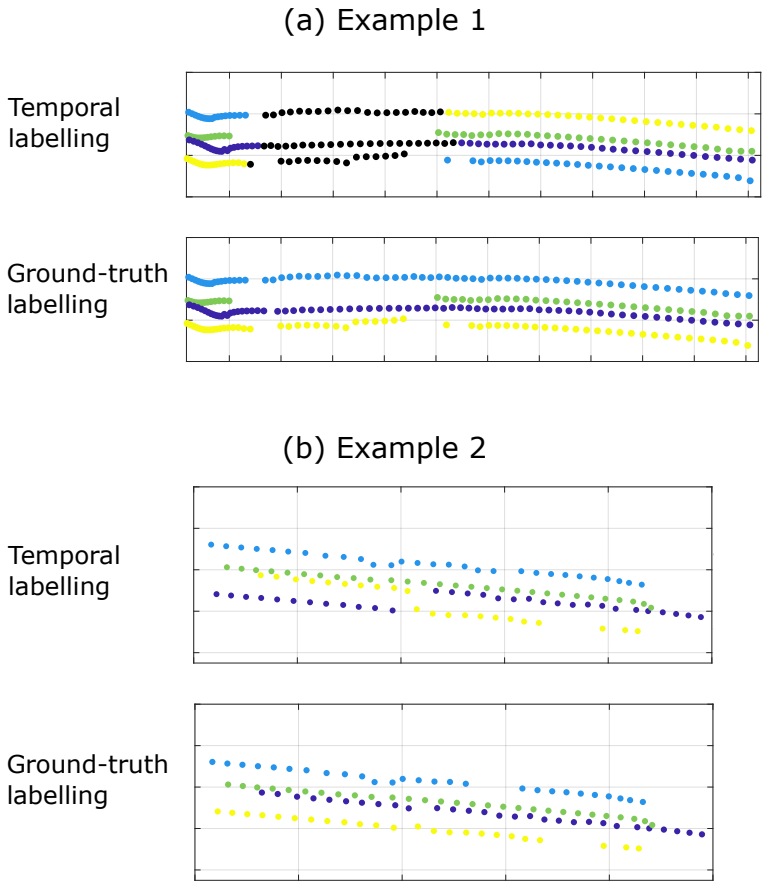


Figure 2.18: Examples of label switching in the temporal algorithm. Two examples of labels switched by the temporal algorithm in two sections of two bird trials are shown. For each example, the temporal and ground truth labels are represented. The markers' positions progress in time from left to right. Markers left unlabelled, or labelled with a different labelling than the selected one for the temporal algorithm, are show in black. In (a), the temporal algorithm switches the yellow and light blue labels after an unlabelled section (shown in black). In (b) the temporal algorithm switches the yellow and dark blue labels.

in animal studies. Animal motion capture experiments usually require dense marker packs, often with small markers, to capture fast motions in large volumes. All of these aspects challenge most commercially available systems. As a result, most animal experiments that use marker-based motion capture end up requiring manual labelling for a large fraction of frames or otherwise intense postprocessing work. For example, in the recent work by Marshall et al. (2021), the authors relied on a model based on pairwise distances between markers to label the data. They also developed a temporal convolutional neural network trained on part of the data, to manage mislabellings and marker dropout. Similarly, the most widespread

video-based tracking solution used in animal behaviour and biomechanics (Hedrick, 2008) still requires a large fraction of manually labelled frames (Ravi et al., 2022; Ingersoll et al., 2018; Khandelwal et al., 2022; Quinn et al., 2019; Kress et al., 2015). This bottleneck in the postprocessing of motion data is currently limiting the possibilities of high-throughput studies.

In this chapter, we have presented a method that automates the labelling of markers without adding significant steps in the postprocessing pipeline. Instead, our method deals with the labelling problem already when acquiring the data. We described the design of an optimal marker pack that maximises the performance of a per-frame labelling algorithm we call 'EE labelling'. This algorithm clusters the markers in the pack based on their representation in a high-dimensional space ('EE space'). The coordinates of the markers in this space are a function of geometric features of the pack that can be defined at each frame. Our goal was to design an arrangement of markers such that the representations of the markers in EE space are as distinct as possible within the design constraints.

We focus on the particular case of tracking head movements of birds in flight, over a large number of trials. For this application, we would like a marker pack (or 'headpack') as compact as possible, since it needs to be attached to the birds' heads. However, a smaller headpack is also expected to reduce the labelling performance of the EE labelling approach. With experiments, we confirm that optimal headpacks of larger size (i.e., larger l_{max}) produce more distinct representations of the markers in EE space, and that this translates to an improved labelling performance. For example, a suboptimal headpack which had two clusters very close to each other in EE space showed a much lower performance on average than an optimal headpack of similar size, in terms of candidate frames correctly labelled by the EE algorithm. Although we found no difference between optimal headpacks of different sizes in the drill trials (all of them labelled correctly 100% of the candidate frames across all trials), we did found differences between the two smallest optimal headpacks in the bird trials. Note that the bird trials covered a much larger volume and are more representative of our desired application. We selected the optimal headpack of $l_{max} =$

45 mm as a good design for our application, since it was the most compact headpack that showed a good labelling performance across all bird and drill trials recorded.

The experiments with different headpacks rotating on a drill showed that compared to the temporal approach, the EE labelling algorithm is more robust to reduced sampling frequencies and higher rotational speeds. This was one of our objectives when defining a labelling algorithm that only relied on features that can be computed per frame. In the bird trials, the EE labelling algorithm combined with an optimal headpack also outperformed the temporal algorithm, with all trials having over 98.7% of the candidate frames correctly labelled (versus a minimum of 30.4% of frames correctly labelled with the temporal algorithm). We find that in the bird trials, the lower performance of the temporal algorithm is likely due to its sensitivity to occlusions, as labels seem to be frequently switched after a few frames of some markers not being visible. Since our approach does per-frame labelling, it does not present this issue and mislabellings are not carried on to subsequent frames.

The method we present here is readily applicable to motion capture studies in other animals and has large potential to reduce the postprocessing workload involved. However, it also has some opportunities for improvement. For example, other definitions for the high-dimensional EE space could be considered. In fact, a feature vector in which the first three coordinates are the sorted lengths of the vertex edges, and the next three coordinates are the corresponding opposite edges' lengths could have some benefits over the one we present here. This is because applying the sorting function only once (instead of twice as we did in our case) would lead to a larger space in which the markers' clusters are represented. This can be more easily seen with a 2D analogy: in an x, y plane, if we restrict the first coordinate x to be larger than the second coordinate y, we are limiting the space to the lower semiplane defined by $x > y$. The more restrictions of this type we apply, the more limited the space. In any case the procedure followed would be easily translatable if a different high-dimensional space is considered. The EE labelling approach could also be further improved, for example by defining a threshold distance beyond which markers aren't assigned to a cluster in EE space.

Since we didn't have a good estimate of what this distance could be *a priori*, this is not explicitly done in the work presented here.

Some of the results presented are specific to our system and setup, such as the estimated minimum distance between markers. However, the same procedure could be easily applied to obtain similar metrics in other setups. Although the proposed EE labelling method focuses on labelling frames in which all markers are visible, it would still be very useful as part of a more sophisticated labelling pipeline, for example as a first labelling pass, or to provide a reliable template of the marker pack based on the measured data. The EE labelling algorithm could also be adapted for real-time tracking. We didn't focus on this application because many animal behaviour studies with motion capture don't seem to require it: real-time is the feature most voted as 'least important' in a short survey run on researchers studying animal collective behaviour presented in Naik (2021). However, it would be a functionality that would extend the use of our method further.

The 74 different marker pack designs we determined during the optimisation procedure could be very useful in collective behaviour studies, since they allow us to distinguish an animal's identity from the marker pack worn. This is likely to be useful even if a commercial tracking software is used. As detailed in Naik (2021), determining over 20 unique marker patterns manually can be quite challenging, and there is a lack of clear guidelines in this respect. Each of the designs listed in Tables 2.1 and 2.2 is optimal for a given pairing of opposite edges (note that each solution comprises a tetrahedron and its dual). However, those that are not golden solutions will have a lower minimum distance between clusters in EE space than the solution selected here. As we have seen in our experiments, this may not be an issue depending on the experimental setup; for example in the drill trials, we found no difference in performance across optimal headpacks of different sizes. We recommend exploring the feasible region for each design in the l_{min}, l_{max} plane to make an informed decision.

Our method helps collect large and reliable motion capture datasets in animals, by enabling further automation in areas that are currently postprocessing bottlenecks.

Such datasets could improve video-based methods for 3D pose estimation in animals, as ground truth motion capture data is currently lacking for many common species (Biggs et al., 2019). Large motion capture datasets could also improve the performance of algorithms that extract dense surface models of animals from video or image data (Zuffi et al., 2017; Zuffi et al., 2019; Zuffi et al., 2018; Sanakoyeu et al., 2020; Li et al., 2021; Wang et al., 2021; Badger et al., 2020), which in turn would enable their use to study biological movement (Mathis et al., 2020a). Currently most of them rely on priors that are not necessarily anatomically correct (Zuffi et al., 2017; Zuffi et al., 2019; Zuffi et al., 2018). Additionally, rich motion capture datasets could also allow us to explore novel data-driven models of animal motor control, following previous works in humans (Merel et al., 2017) or quadrupeds (Zhang et al., 2018).

References

Arac, A., Zhao, P., Dobkin, B. H., Carmichael, S. T., and Golshani, P. (2019). "DeepBehavior: A deep learning toolbox for automated analysis of animal and human behavior imaging data". In: *Front. Syst. Neurosci.* 13.May, pp. 1–12. DOI: 10.3389/fnsys.2019.00020.

Badger, M., Wang, Y., Modh, A., Perkes, A., Kolotouros, N., Pfrommer, B. G., Schmidt, M. F., and Daniilidis, K. (2020). "3D Bird Reconstruction: A Dataset, Model, and Shape Recovery from a Single View". In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. Vol. 12363 LNCS, pp. 1–17. DOI: 10.1007/978-3-030-58523-5_1.

Biggs, B., Roddick, T., Fitzgibbon, A., and Cipolla, R. (2019). "Creatures Great and SMAL: Recovering the Shape and Motion of Animals from Video". In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 11365 LNCS, pp. 3–19. DOI: 10.1007/978-3-030-20873-8_1.

Brighton, C. H. and Taylor, G. K. (2019). "Hawks steer attacks using a guidance system tuned for close pursuit of erratically manoeuvring targets". In: *Nat. Commun.* 10.1, pp. 1–28. DOI: 10.1038/s41467-019-10454-z.

Byström, A., Hardeman, A. M., Serra Bragança, F. M., Roepstorff, L., Swagemakers, J. H., van Weeren, P. R., and Egenvall, A. (2021). "Differences in equine spinal kinematics between straight line and circle in trot". In: *Sci. Rep.* 11.1, pp. 1–13. DOI: 10.1038/s41598-021-92272-2.

Chin, D. D. and Lentink, D. (2019). "Birds repurpose the role of drag and lift to take off and land". In: *Nat. Commun.* 10.1, pp. 1–10. DOI: 10.1038/s41467-019-13347-3.

Dunn, T. W., Marshall, J. D., Severson, K. S., Aldarondo, D. E., Hildebrand, D. G., Chettih, S. N., Wang, W. L., Gellis, A. J., Carlson, D. E., Aronov, D., Freiwald, W. A., Wang, F., and Ölveczky, B. P. (2021). "Geometric deep learning enables 3D kinematic profiling across species and environments". In: *Nat. Methods* 18.5, pp. 564–573. DOI: 10.1038/s41592-021-01106-6.

Eckmeier, D., Geurten, B. R., Kress, D., Mertes, M., Kern, R., Egelhaaf, M., and Bischof, H. J. (2008). "Gaze strategy in the free flying zebra finch (Taeniopygia guttata)". In: *PLoS One* 3.12. DOI: 10.1371/journal.pone.0003956.

Flack, A., Nagy, M., Fiedler, W., Couzin, I. D., and Wikelski, M. (2018). "From local collective behavior to global migratory patterns in white storks". In: *Science (80-.).* 360.6391, pp. 911–914. DOI: 10.1126/science.aap7781.

Genzel, D. and Yartsev, M. M. (2021). "The fully automated bat (FAB) flight room: A human-free environment for studying navigation in flying bats and its initial application to the retrosplenial cortex". In: *J. Neurosci. Methods* 348.July 2020, p. 108970. DOI: 10.1016/j.jneumeth.2020.108970.

Graham, M. and Socha, J. J. (2021). "Dynamic movements facilitate extreme gap crossing in flying snakes". In: *J. Exp. Biol.* 224.20. DOI: 10.1242/jeb.242923.

Graving, J. M., Chae, D., Naik, H., Li, L., Koger, B., Costelloe, B. R., and Couzin, I. D. (2019). "DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning". In: *Elife* 8, pp. 1–42. DOI: 10.7554/eLife.47994.

Günel, S., Rhodin, H., Morales, D., Campagnolo, J., Ramdya, P., and Fua, P. (2019). "DeepFly3D, a deep learning-based approach for 3D limb and appendage tracking in tethered, adult Drosophila". In: *Elife* 8, pp. 1–23. DOI: 10.7554/eLife.48571.

Gutierrez, E., Quinn, D. B., Chin, D. D., and Lentink, D. (2017). "Lift calculations based on accepted wake models for animal flight are inconsistent and sensitive to vortex dynamics". In: *Bioinspiration and Biomimetics* 12.1, pp. 1–15. DOI: 10.1088/1748-3190/12/1/016004.

Hedrick, T. L. (2008). "Software techniques for two- and three-dimensional kinematic measurements of biological and biomimetic systems". In: *Bioinspiration and Biomimetics* 3.3. DOI: 10.1088/1748-3182/3/3/034001.

Horn, B. K. P., Hilden, H. M., and Negahdaripour, S. (1988). "Closed-form solution of absolute orientation using orthonormal matrices". In: *J. Opt. Soc. Am. A* 5.7, p. 1127. DOI: 10.1364/josaa.5.001127.

Ingersoll, R., Haizmann, L., and Lentink, D. (2018). "Biomechanics of hover performance in Neotropical hummingbirds versus bats". In: *Sci. Adv.* 4.9. DOI: 10.1126/sciadv.aat2980.

Jensen, G. W., van der Smagt, P., Heiss, E., Straka, H., and Kohl, T. (2020). "SnakeStrike: a low-cost open-source high-speed multi-camera motion capture system". In: *Front. Behav. Neurosci.* 14.August, pp. 1–16. DOI: 10.3389/fnbeh.2020.00116.

Joska, D., Clark, L., Muramatsu, N., Jericevich, R., Nicolls, F., Mathis, A., Mathis, M. W., and Patel, A. (2021). "AcinoSet: a 3D pose estimation dataset and baseline models for cheetahs in the wild". In: Icra, pp. 13901–13908. DOI: 10.1109/icra48506.2021.9561338.

Kano, F., Walker, J., Sasaki, T., and Biro, D. (2018). "Head-mounted sensors reveal visual attention of free-flying homing pigeons". In: *J. Exp. Biol.* 221.17, pp. 1–13. DOI: 10.1242/jeb.183475.

Karashchuk, P., Rupp, K. L., Dickinson, E. S., Walling-Bell, S., Sanders, E., Azim, E., Brunton, B. W., and Tuthill, J. C. (2021). "Anipose: a toolkit for robust markerless 3D pose estimation". In: *Cell Rep.* 36.13, p. 109730. DOI: 10.1016/j.celrep.2021.109730.

Khandelwal, P. C. and Hedrick, T. L. (2022). "Combined effects of body posture and three-dimensional wing shape enable efficient gliding in flying lizards". In: *Sci. Rep.* 12.1, pp. 1–11. DOI: 10.1038/s41598-022-05739-1.

Kjærsgaard, A., Pertoldi, C., Loeschcke, V., and Hansen, D. W. (2008). "Tracking the gaze of birds". In: *J. Avian Biol.* 39.4, pp. 466–469. DOI: 10.1111/j.0908-8857.2008.04288.x.

Kress, D., Van Bokhorst, E., and Lentink, D. (2015). "How lovebirds maneuver rapidly using super-fast head saccades and image feature stabilization". In: *PLoS One* 10.6, pp. 1–24. DOI: 10.1371/journal.pone.0129287.

Land, M. F. and Lee, D. N. (1994). "Where we look when we steer". In: *Nature* 369.6483, pp. 742–744. DOI: 10.1038/369742a0.

Li, C., Ghorbani, N., Broomé, S., Rashid, M., Black, M. J., Hernlund, E., Kjellström, H., and Zuffi, S. (2021). "hSMAL: detailed horse shape and pose reconstruction for motion pattern recognition". In: pp. 2–7. arXiv: 2106.10102.

Lin, H. T., Ros, I. G., and Biewener, A. A. (2014). "Through the eyes of a bird: modelling visually guided obstacle flight". In: *J. R. Soc. Interface* 11.96, pp. 1–12. DOI: 10.1098/rsif.2014.0239.

Marshall, J. D., Aldarondo, D. E., Dunn, T. W., Wang, W. L., Berman, G. J., and Ölveczky, B. P. (2021). "Continuous whole-body 3D kinematic recordings across the rodent behavioral repertoire". In: *Neuron* 109.3, 420–437.e8. DOI: 10.1016/j.neuron.2020.11.016.

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., and Bethge, M. (2018). "DeepLabCut: markerless pose estimation of user-defined body parts with deep learning". In: *Nat. Neurosci.* 21.9, pp. 1281–1289. DOI: 10.1038/s41593-018-0209-y.

Mathis, A., Schneider, S., Lauer, J., and Mathis, M. W. (2020a). "A primer on motion capture with deep learning: principles, pitfalls, and perspectives". In: *Neuron* 108.1, pp. 44–65. DOI: 10.1016/j.neuron.2020.09.017.

Mathis, M. W. and Mathis, A. (2020b). "Deep learning tools for the measurement of animal behavior in neuroscience". In: *Curr. Opin. Neurobiol.* 60, pp. 1–11. DOI: 10.1016/j.conb.2019.10.008.

Merel, J., Tassa, Y., TB, D., Srinivasan, S., Lemmon, J., Wang, Z., Wayne, G., and Heess, N. (2017). "Learning human behaviors from motion capture by adversarial imitation". In: arXiv: 1707.02201.

Michaiel, A. M., Abe, E. T., and Niell, C. M. (2020). "Dynamics of gaze control during prey capture in freely moving mice". In: *Elife* 9, pp. 1–27. DOI: 10.7554/eLife.57458.

Minano, S. and Taylor, G. K. (2021). "Through hawks' eyes: reconstructing a bird's visual field in flight to study gaze strategy and attention during perching and obstacle avoidance". In: *bioRxiv*. DOI: 10.1101/2021.06.16.446415.

Naik, H. (2021). "XR for all: closed-loop visual stimulation techniques for human and non-human animals". PhD thesis. Munich: Technische Universität München.

Naik, H., Bastien, R., Navab, N., and Couzin, I. D. (2020). "Animals in virtual environments". In: *IEEE Trans. Vis. Comput. Graph.* 26.5, pp. 2073–2083. DOI: 10.1109/TVCG.2020.2973063.

Nath, T., Mathis, A., Chen, A. C., Patel, A., Bethge, M., and Mathis, M. W. (2019). "Using DeepLabCut for 3D markerless pose estimation across species and behaviors". In: *Nat. Protoc.* 14.7, pp. 2152–2176. DOI: 10.1038/s41596-019-0176-0.

Pereira, T. D., Tabris, N., Li, J., Ravindranath, S., Papadoyannis, E. S., Wang, Z. Y., Turner, D. M., McKenzie-Smith, G., Kocher, S. D., Falkner, A. L., Shaevitz, J. W.,

and Murthy, M. (2020). “SLEAP: multi-animal pose tracking”. In: *bioRxiv*. DOI: 10.1101/2020.08.31.276246.

Quinn, D., Kress, D., Chang, E., Stein, A., Wegrzynski, M., and Lentink, D. (2019). “How lovebirds maneuver through lateral gusts with minimal visual information”. In: *Proc. Natl. Acad. Sci. U. S. A.* 116.30, pp. 15033–15041. DOI: 10.1073/pnas.1903422116.

Ravi, S., Siesenop, T., Bertrand, O. J., Li, L., Doussot, C., Fisher, A., Warren, W. H., and Egelhaaf, M. (2022). “Bumblebees display characteristics of active vision during robust obstacle avoidance flight”. In: *J. Exp. Biol.* 225.4. DOI: 10.1242/jeb.243021.

Roderick, W. R., Chin, D. D., Cutkosky, M. R., and Lentink, D. (2019). “Birds land reliably on complex surfaces by adapting their foot-surface interactions upon contact”. In: *Elife* 8, pp. 1–23. DOI: 10.7554/eLife.46415.

Ros, I. G., Bhagavatula, P. S., Lin, H. T., and Biewener, A. A. (2017a). “Rules to fly by: pigeons navigating horizontal obstacles limit steering by selecting gaps most aligned to their flight direction”. In: *Interface Focus* 7.1, p. 20160093. DOI: 10.1098/rsfs.2016.0093.

Ros, I. G. and Biewener, A. A. (2017b). “Pigeons (C. livia) follow their head during turning flight: Head stabilization underlies the visual control of flight”. In: *Front. Neurosci.* 11.DEC, pp. 1–12. DOI: 10.3389/fnins.2017.00655.

Sanakoyeu, A., Khalidov, V., McCarthy, M. S., Vedaldi, A., and Neverova, N. (2020). “Transferring dense pose to proximal animal classes”. In: *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 5232–5241. DOI: 10.1109/CVPR42600.2020.00528.

Schönemann, P. H. (1966). “A generalized solution of the orthogonal procrustes problem”. In: *Psychometrika* 31.1, pp. 1–10.

Singh, S. K., Zhang, L.-B., and Zhao, J.-S. (2022). “Reconstruction of flight parameters of a bat for flapping robots”. In: *J. Biomech. Eng.* 144.5. DOI: 10.1115/1.4052889.

Strandburg-Peshkin, A., Twomey, C. R., Bode, N. W., Kao, A. B., Katz, Y., Ioannou, C. C., Rosenthal, S. B., Torney, C. J., Wu, H. S., Levin, S. A., and Couzin, I. D. (2013). “Visual sensory networks and effective information transfer in animal groups”. In: *Curr. Biol.* 23.17, R709–R711. DOI: 10.1016/j.cub.2013.07.059.

Theunissen, L. M., Reid, T., and Troje, N. F. (2017a). “Pigeons use distinct stop phases to control pecking”. In: *J. Exp. Biol.* 220.3, pp. 437–444. DOI: 10.1242/jeb.147850.

Theunissen, L. M. and Troje, N. F. (2017b). “Head stabilization in the pigeon: Role of vision to correct for translational and rotational disturbances”. In: *Front. Neurosci.* 11.OCT, pp. 1–12. DOI: 10.3389/fnins.2017.00551.

Walter, T. and Couzin, I. D. (2021). “Trex, a fast multi-animal tracking system with markerless identi cation, and 2d estimation of posture and visual elds”. In: *Elife* 10, pp. 1–73. DOI: 10.7554/eLife.64000.

Wang, Y., Kolotouros, N., Daniilidis, K., and Badger, M. (2021). “Birds of a feather: capturing avian shape models from images”. In: *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 14734–14744. DOI: 10.1109/CVPR46437.2021.01450.

Wirth, K. and Dreiding, A. (2009). “Edge lengths determining tetrahedrons”. In: *Elem. der Math.* 64, pp. 160–170. DOI: 10.4171/em/129.

Zeiler, M. D., Taylor, G. W., Troje, N. F., and Hinton, G. E. (2009). “Modeling pigeon behaviour using a conditional restricted Boltzmann machine”. In: *ESANN 2009*

Proceedings, 17th Eur. Symp. Artif. Neural Networks - Adv. Comput. Intell. Learn., pp. 391–396.

Zhang, H., Starke, S., Komura, T., and Saito, J. (2018). "Mode-adaptive neural networks for quadruped motion control". In: *ACM Trans. Graph.* 37.4, pp. 1–11. DOI: 10.1145/3197517.3201366.

Zimmermann, C., Schneider, A., Alyahyay, M., Brox, T., and Diester, I. (2020). "FreiPose: a deep learning framework for precise animal motion capture in 3D spaces". In: *bioRxiv*. DOI: 10.1101/2020.02.27.967620.

Zuffi, S., Kanazawa, A., Berger-Wolf, T., and Black, M. (2019). "Three-D safari: learning to estimate zebra pose, shape, and texture from images 'in the wild'". In: *Proc. IEEE Int. Conf. Comput. Vis.* Vol. 2019-Octob, pp. 5358–5367. DOI: 10.1109/ICCV.2019.00546.

Zuffi, S., Kanazawa, A., and Black, M. J. (2018). "Lions and tigers and bears: capturing non-rigid, 3D, articulated shape from images". In: *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* Pp. 3955–3963. DOI: 10.1109/CVPR.2018.00416.

Zuffi, S., Kanazawa, A., Jacobs, D., and Black, M. J. (2017). "3D menagerie: modeling the 3D shape and pose of animals". In: *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*. Vol. 2017-Janua, pp. 5524–5532. DOI: 10.1109/CVPR.2017.586.

3

Through hawks' eyes: synthetically reconstructing the visual field of a bird in flight

Contents

3.1 Abstract 73
3.2 Introduction 73
3.2.1 Related work 75
3.2.2 Contribution 81
3.3 Methods 83
3.3.1 Bird flight experiments 83
3.3.2 Motion capture data postprocessing 84
3.3.3 Modelling the bird's view 85
3.3.4 Hybrid model of the lab environment 93
3.4 Results 98
3.4.1 Pursuit flight 98
3.4.2 Obstacle avoidance flights 104
3.5 Discussion 106
3.5.1 Key features of rendering method 106
3.5.2 Behavioural analyses: pursuit flight 109
3.5.3 Behavioural analyses: obstacle avoidance 110
3.5.4 Limitations and future work 111
3.6 Supplementary videos 113
References 114

3.1 Abstract

Diurnal birds of prey strongly rely on vision to execute flight manoeuvres that are key to their survival, such as intercepting fast-moving targets or navigating through clutter. A better understanding of the precise role played by vision during these manoeuvres is not only relevant within the field of animal behaviour, but also could have applications in visually-guided autonomous drones. Here, we present a method that uses computer vision tools to tackle the problem of visuomotor control in bird flight, and demonstrate its use to answer behavioural questions. Combining motion capture data from live Harris' hawks with a hybrid 3D model of the environment, we produce RGB, semantic, depth and optic flow outputs that characterise the visual experience of the bird in flight. Contrary to previous approaches, our method allows us to consider different camera models and alternative gaze strategies for the purposes of hypothesis testing, further allows us to consider visual input over the complete visual field of the bird, and is not limited by the technical specifications and performance of a head-mounted camera light enough to attach to a bird's head in flight. We present pilot data from three sample flights: a pursuit flight, in which a hawk intercepts a moving target, and two obstacle avoidance flights. With this approach, we aim to provide a reproducible method that facilitates the collection of large volumes of data across many individuals, opening up new avenues for data-driven models of animal behaviour.

3.2 Introduction

From intercepting moving targets to manoeuvring through clutter, birds use vision to coordinate flight manoeuvres with an agility and flexibility beyond the reach of current autonomous systems. Nevertheless, the links between their vision, guidance, and control are complex and poorly understood. Much more is known about the role of vision in insect flight (Taylor et al., 2008), because the size and sentience

of birds complicates experimental identification of the input-output relationships characterising their visuomotor control (Altshuler et al., 2018). To control the visual experience of a bird, previous works (Bhagavatula et al., 2011; Schiffner et al., 2015; Dakin et al., 2016; Ros et al., 2016) have investigated avian flight behaviour in abstract visual environments, such as corridors with vertically or horizontally striped walls (see Figure 3.1b). This has proven a useful first step in exploring how birds use visual self-motion cues, and in isolating its effects on flight control. For example, budgerigars flying through narrow corridors regulate flight speed in response to optic flow from sliding gratings projected onto the walls (Schiffner et al., 2015). These methods are inherited from similar studies of insects (Baird et al., 2021; Tammero et al., 2002a; Tammero et al., 2002b; Altshuler et al., 2018), so may oversimplify the rich visual input available to birds with their high visual acuity and complex neural organization (Altshuler et al., 2018). Such abstract visual inputs are far from those experienced naturally in flight, so caution is warranted if using them to draw conclusions about the birds' strategies in the wild.

Here, we present a method for reconstructing the scene contained within the visual field of a bird flying through a structured environment, as a first step towards understanding how birds use visual information to guide and control their flight. Our synthetic reconstruction method is designed to simulate the physical stimulus to a bird's optomotor system, and avoids the main practical difficulties associated with making invasive physiological measurements of its response. In developing our method, we aim to support the analysis of large quantities of data across multiple individuals, using environments that may vary experimentally across trials. To achieve this, we focus on developing a non-invasive method for tracking the animal's self-motion, and on methods to reconstruct the geometry of the environment that allow us to capture different levels of complexity. Specifically, we combine high-speed motion capture data with a three-dimensional (3D) reconstruction of the laboratory environment, which we use to generate synthetic visual inputs modelling those that may be available to the bird in flight. These visual inputs characterise the bird's

visual experience of its own self-motion in detail, opening up several new avenues of research in bio-inspired computer vision and behavioural modelling.

In summary, we present a computer vision approach to understanding the visuomotor guidance and control of bird flight, and demonstrate the possibilities of using the data that we generate to answer mechanistic questions in behaviour. We present pilot data from three sample flights: one pursuit flight, in which a hawk intercepts a moving target pulled across the ground, and two obstacle avoidance flights, in which a hawk flies between two perches around a set of obstacles placed at varying distances from the take-off perch.

3.2.1 Related work

Here we review the main approaches that have been used previously to investigate the role of active vision in avian flight, in order to highlight the contribution of the novel method that we present.

3.2.1.1 Active vision in birds

Bird vision is an active process, so understanding how birds interrogate their visual environment is key to unravelling the cognitive processes coordinating their flight manoeuvres. Birds mainly use head movements controlled by their neck motor system to look around the environment. This is because their eyes have a limited range of motion within the orbit, such that the largest eye movements driven by the oculomotor system are small compared to those made by the head (Yorzinski et al., 2015; Mitkus et al., 2018). The head movements of birds in flight display a characteristic saccade-and-fixate strategy recalling primate eye movements (Eckmeier et al., 2008; Ozawa, 2010; Kress et al., 2015; Ros et al., 2017). This strategy supports the use of optic flow in guidance and control, because the rotational component of the optic flow field is eliminated during fixation, leaving only the translational component containing depth information (Eckmeier et al., 2008). A small body of previous work has assessed how head movements are affected by the visual environment experienced in flight. For example, the frequency and

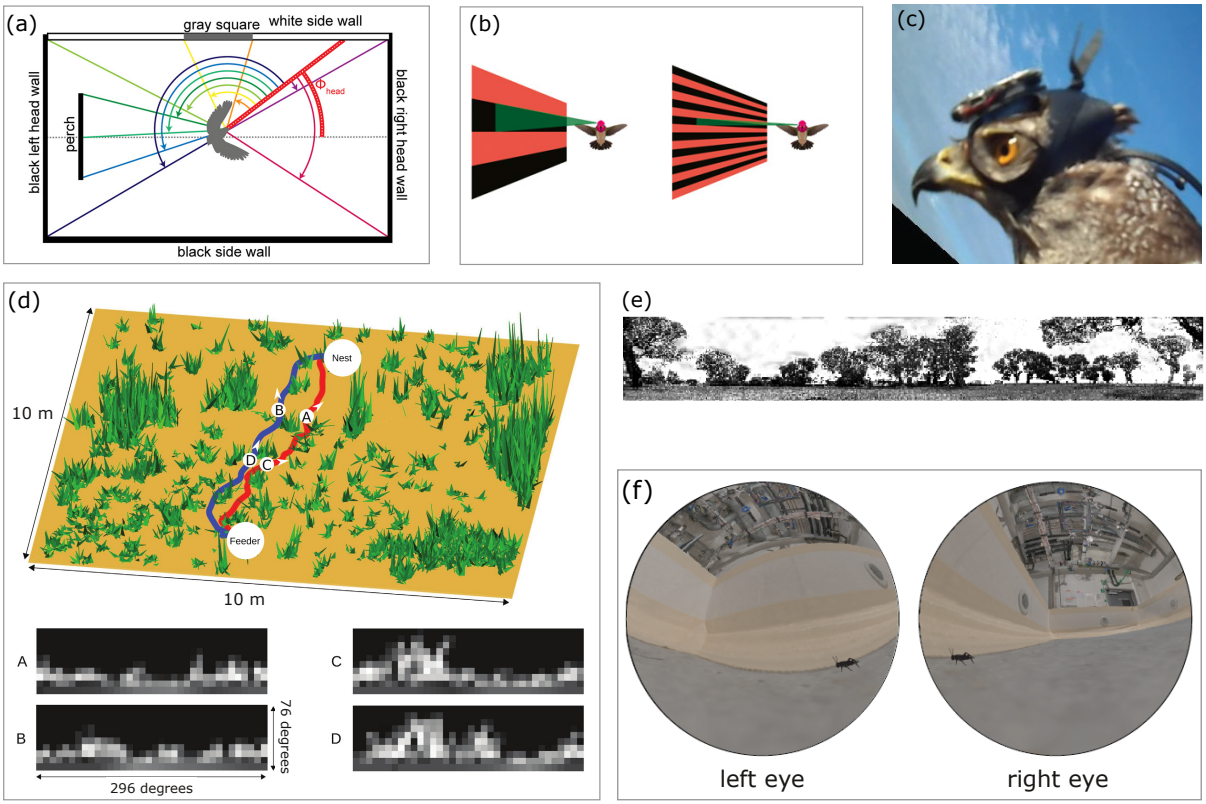


Figure 3.1: Related work. (a) Overhead view of a lovebird turning in flight; the alignment of the forward direction of its head is compared against several features of the lab environment (Kress et al., 2015). (b) Hummingbirds flying along striped corridors (Dakin et al., 2016). (c) Northern goshawk with head-mounted camera (Kane et al., 2015). (d) Reconstruction of the environment around a nest of homing ants, and reconstructed views (A,B,C,D) at different instances of their recorded paths (Ardin et al., 2016). (e) Reconstruction of an ant's field of view, combining laser-scanned data and camera images (Stürzl et al., 2015). (f) Reconstructed view from each eye of a mouse hunting a cricket, using laser-scanned data and texture from high-resolution images (Holmgren et al., 2021).

amplitude of head saccades has been shown to vary in homing pigeons with the structure of the landscape they are flying through (Kano et al., 2018), the structure of the clutter they are negotiating (Ros et al., 2016), and the presence of another individual when they are flying in pairs (Taylor et al., 2019).

Although head pose is the primary determinant of a bird's gaze direction, it can be measured most directly using eye-tracking cameras. These have so far been used to identify where birds look when assessing mates (Yorzinski et al., 2013), watching predators (Yorzinski et al., 2014; Yorzinski, 2021), or inspecting the environment (Yorzinski et al., 2015). Weight limitations have so far restricted this technology

to terrestrial use cases, but whilst miniaturisation may be possible, eye-tracking cameras will always be challenging to use in flight, owing to the difficulty of keeping the camera steady without occluding the frontal field of view. Most studies therefore take head orientation as a proxy for gaze direction, which is reasonable provided that the eyes move little relative to the head in flight, and is often sufficient to identify the features to which a bird is attending. For example, work on lovebirds flying to a perch (Kress et al., 2015) investigated the alignment of the bird's head with the edges of the perch and flight arena (see Figure 3.1a), whilst work on pigeons negotiating a forest of vertical poles (Lin et al., 2014) investigated the alignment of the bird's head with the obstacles it had to traverse. Both studies analysed the problem in two dimensions, however, focusing only on changes in head azimuth, and reducing these extended features of the environment to single points in the visual field. A complete understanding of the problem requires fuller knowledge of what a bird sees in flight, which is what motivates the visual reconstruction method of the present study.

3.2.1.2 Head-mounted cameras

Head-mounted video cameras can be used to sample the view that a bird experiences as a result of its self-motion through the environment (Figure 3.1c), and have been used so to analyse aerial attack behaviours in hawks and falcons (Kane et al., 2015; Kane et al., 2014; Ochs et al., 2016). This approach can be especially informative regarding a bird's behaviour in its natural habitat, subject to the extreme limitations of pixel count, dynamic range, and field of view of any camera small enough to mount on the head. Payload is conventionally limited to $\leq 5\%$ of a bird's body mass on welfare grounds (Fair et al., 2010), but much more stringent limits may be required to ensure natural behaviour if the load is carried on the head (Kane et al., 2014). The 20 g cameras that have been used previously (Kane et al., 2015; Kane et al., 2014) are twice the weight of many small birds, so are only suitable for studying very large species such as raptors. Even so, it is simply not possible for a small camera to cover a bird's full field of view at an appropriate optical or sampling resolution. For example, the vertical field of view (31°) of the

head-mounted cameras used to study Harris' hawks (Kane et al., 2015) compares unfavourably with the vertical extent of the binocular overlap (100°) in this species (Potier et al., 2016), and their 30 Hz frame rate is well below these birds' 81 Hz flicker fusion frequency (Potier et al., 2020). The possibilities for analysing head-mounted video data are further impacted by their low frame rates, and the motion blur associated with low shutter speeds and rolling shutters (Kane et al., 2014). Finally, head-mounted cameras are prone to wobble unless surgically attached to the head (Lev-Ari et al., 2018; Hazan et al., 2015), which is an undesirable intervention. Head-mounted video cameras therefore have less utility for studying visually-guided flight in birds than might first be imagined.

3.2.1.3 Synthetic reconstruction

An alternative to using head-mounted cameras to record an animal's view of the world is to recreate its visual input synthetically. Some works have explored the possibility of mimicking an animal's visual input directly using specially designed video cameras (Stuerzl et al., 2010) or event-based cameras (Zhu et al., 2021), but rendering avoids most of the shortcomings of camera systems, offering complete control over the visual detail and visual field (Holmgren et al., 2021). A rendering approach is also well suited to scientific inference, because of the possibility of defining alternative scenarios or viewpoints for hypothesis testing.

Even so, rendering requires detailed knowledge of the animal's head pose and/or eye movements (see above), and typically a 3D model of its visual environment. State-of-the-art research on image-based rendering points to the possibility of reconstructing novel views given only a set of camera images with known pose (Tancik et al., 2022; Mildenhall et al., 2020), but these are not yet established techniques among non-experts. Standard modelling approaches such as simultaneous localization and mapping (SLAM) suffer from accumulating noise and drift when covering large areas (Schonberger et al., 2016), and automating the postprocessing of the resulting meshes can be challenging (Risse et al., 2018; Stürzl et al., 2015).

Nevertheless, the quality of the dense maps that can now be captured with consumer-level hand-held devices has improved greatly in the past few years, achieving results comparable to more expensive laser-scanning methods even in complex natural environments (Tatsumi et al., 2022; Gollob et al., 2021).

Most examples of synthetic reconstruction of an animal's visual field to date have focused on insects. Several studies have investigated terrestrial navigation in ants (Zeil et al., 2014; Ardin et al., 2015; Ardin et al., 2016) using fully synthetic models of their environment (Figure 3.1d), and a basic geometric reconstruction of the laboratory environment has been used to investigate the role of optic flow in bee flight (Ravi et al., 2022; Ravi et al., 2019). More detailed 3D models of the natural environment obtained using laser scanners, structure-from-motion (SfM), and photographic reconstruction techniques (Figure 3.1e) have been used to study homing flight behaviour in bees and wasps (Stürzl et al., 2015; Stuerzl et al., 2016; Schulte et al., 2019).

This focus on insects reflects the fact that they are simpler to model than vertebrates, owing to their lower sampling resolution, and the fact that their eyes are rigidly fixed with respect to their head, but the same general approaches have more recently been extended to vertebrates. In particular, by tracking the animal's head and eye movements together, and by combining laser-scanning and photographic texture to produce a high-resolution 3D reconstruction of its lab environment, one recent study was able to analyse the optic flow experienced by mice when pursuing prey within a 1 m^2 arena (Holmgren et al., 2021). Additionally, an earlier work in zebra finches (Eckmeier et al., 2013) followed an approach similar to what we present here: they reconstructed a bird's perspective during a single manoeuvring flight, considering a simple geometric model of the flight arena, to use as a naturalistic stimuli for neural recordings. However, their work was at a much smaller scale than we present here (their flight arena was under 1 m^3 and they only reconstructed one flight), and they estimated the bird's perspective based on a single direction. The growing interest in reconstructing an animal's experience of its environment at the much larger spatial scales relevant to ecology and conservation (Tuia et al.,

2022) has also led to demonstrations of mapping sensors worn by a dog (McClune, 2018), and a method for embedding an animal's track within an aerial view of its environment obtained by a mobile camera used for tracking (Haalck et al., 2020).

A key advantage of synthetic reconstruction techniques is that they can be used to study the appearance of the visual environment in the context of an accurate model of the animal's visual system. This is important in capturing detail appropriate to the optical resolution and sampling resolution of the eyes (Holmgren et al., 2021; Neumann, 2002), and also in matching the response properties of the motion vision system (Windsor et al., 2017; Bian et al., 2018). Even so, rendering an RGB image that looks realistic to a human eye is not the same as reconstructing the stimulus to another animal's visual system, owing to variation in the number of photoreceptor types and their spectral response curves (e.g. whereas humans have trichromatic vision, birds are tetrachromatic). This may not be necessary in the context of motion vision, however, because despite the possible importance of chromatic motion vision cues in humans (Ruppertsberg et al., 2003), changes in luminance are always sufficient for motion vision. For example, synthetic reconstruction has been used to study the effectiveness of movement displays of model lizards in naturalistic scenes under different lighting scenarios and wind conditions, assuming a colour space similar to that of humans (Bian et al., 2018; Bian et al., 2019; Bian et al., 2021). These studies of territorial displays in lizards are perhaps the only examples to date in which the environment has been rendered at a level of precision comparable to that experienced by the well-developed visual system of a bird in flight, albeit that they are rendered from the perspective of a stationary observer.

Synthetic renderings also lend themselves to being presented to an animal moving in a virtual reality (VR) environment. This approach has been used in tethered and free-flight studies of insect flight for some time (Kern et al., 2005; Taylor et al., 2008; Windsor et al., 2017), but impressive VR environments have more recently been developed for freely-moving mice, flies, and zebra fish (Stowers et al., 2017), which enable stimulus presentation under open or closed loop experimental conditions (Naik et al., 2020). Such environments are currently limited to volumes

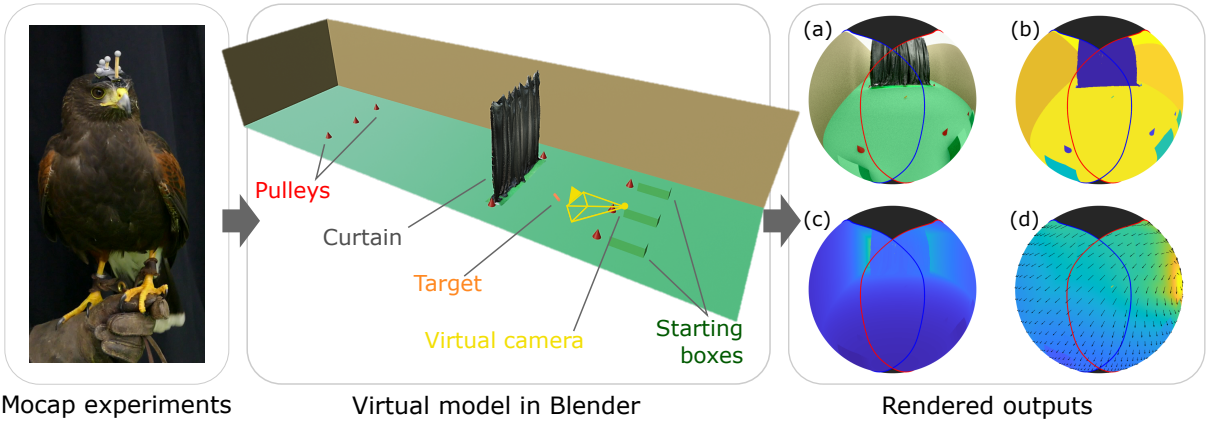


Figure 3.2: Summary of method for reconstructing the visual information contained within the visual field of a bird in flight. The head movements of birds executing flight manoeuvres are recorded in a large motion capture lab (left panel). The lab environment is modelled in Blender, using geometric primitives and a dense 3D map to model objects with a more complex geometry (centre panel). The measured head pose is then used to define a virtual camera that is representative of the bird's visual field (right panel). With this approach we can generate detailed information describing the visual scene that the bird experiences in flight, including: (a) basic RGB renderings; (b) semantic maps; (c) depth maps; or (d) optic flow. In the virtual model of the lab shown (centre panel), the pulleys and target are displayed at twice their actual size for clarity. The spheres representing the bird's visual field in the right panel show the animal's view as it flies through the lab. The retinal margins of the left eye (blue line) and right eye (red line) are shown for reference, as well as the blind area caudal to the area sampled by the retina (black fill). Note that because the spheres are represented using an orthographic projection, not all the visual field of the bird is visible.

of approximately 1 m^3, however, so have yet to find use for larger animals making larger-scale movements.

3.2.2 Contribution

Here, we combine high-speed motion capture with 3D modelling of the laboratory environment, to render an approximation of the visual stimuli experienced by a bird in flight. By defining a virtual camera that replicates the bird's head pose within a 3D model of its environment, we reconstruct RGB renderings, semantic maps, depth maps, and optic flow over the complete visual field of the bird (see Figure 3.2). We make no attempt to model the spectral response of the bird's visual system, although it would in principle be possible to use this to analyse the information contained in the RGB renderings given knowledge of the ambient lighting conditions.

The method that we present allows for the possibility of defining a hybrid model

of the lab environment, combining basic geometric primitives with dense 3D models. A representation based on geometric primitives alone may be sufficient to investigate how the bird's gaze strategy relates to the main elements in the scene (Minano et al., 2021). However, if elements of complex shape or texture are present, then modelling them using geometric primitives may be out of the question, which limits the scope for high-throughput analysis. We address this by using a SLAM application on a consumer-level handheld mapping device to capture the more complex elements of the scene, which we combine with a more basic geometric representation of the environment. In this way, we limit the issues of drift and noise that are typical of the large, dense 3D maps produced by SLAM or SfM (Schonberger et al., 2016), whilst also minimising the modelling effort.

We further facilitate the integration of the captured meshes with the motion capture data by transforming the mesh at acquisition time to the motion capture coordinate system using an ArUco fiducial marker (Romero-Ramirez et al., 2018; Garrido-Jurado et al., 2014). The novelty of our method therefore lies in its fusion of several different computer vision techniques, and in their application to the experimental and computational study of animal behaviour. Our aim is to provide a reproducible method that facilitates the collection and creation of large volumes of data across many individuals, with a view to future work using behavioural cloning and inverse reinforcement learning.

Our method has several advantages over previous approaches used to characterise the visual experience of a bird in flight. First, it allows us to consider visual input over the complete visual field of the bird, as opposed to reducing the analysis to the field of view of a head-mounted camera (Ochs et al., 2016; Kane et al., 2015; Kane et al., 2014) or to a unique gaze direction (Land, 1999; Eckmeier et al., 2008; Yorzinski et al., 2013; Yorzinski et al., 2014; Yorzinski et al., 2015; Kress et al., 2015; Yorzinski, 2021). Second, it enables us to characterise the visual input of the bird without having to contain the animal within an abstract visual environment (Bhagavatula et al., 2011; Schiffner et al., 2015; Dakin et al., 2016; Ros et al., 2016). Third, we are not limited by the technical specifications and

performance of a head-mounted camera light enough to attach to a bird's head in flight (Ochs et al., 2016; Kane et al., 2015; Kane et al., 2014). Fourth, our method allows us to consider different camera models and alternative gaze strategies for the purposes of hypothesis testing.

3.3 Methods

Here we describe the key details of our experimental method and reconstruction method.

3.3.1 Bird flight experiments

We recorded Harris' hawks (*Parabuteo unicinctus*) flying in a large motion capture lab, using 22 motion capture cameras sampling at 200 Hz (Vantage V16; Vicon Motion Systems Ltd., Oxford, UK). Here we report the results of three flights from two different birds: one involving pursuit of a dummy target pulled along an erratic course; the other two involving obstacle avoidance and perching manoeuvres.

To track the bird's head, we used a custom 'headpack', comprising a rigid arrangement of four or five retroreflective 6.4 mm diameter spherical markers that we fixed to the bird on a Velcro patch glued to its head. We attached additional 6.4 mm or 14 mm diameter markers to the main static elements of the scene, to allow us to locate these within the motion capture coordinate system. In the pursuit flight, we tracked the target using three 6.4 mm diameter markers.

In the pursuit flight, the bird was trained to chase a 2.54 cm wide and 15 cm long artificial target with a food reward attached, within a flight volume of 19.2 m$\times$ 5.3 m$\times$ 3.3 m. To further challenge the bird's manoeuvring, we hung a black curtain across the room from floor to ceiling, acting as an obstacle and creating a gap of approximately one wingspan ($1.0 - 1.1$ m) on either side. A linear motor pulled the target on a line passing around a series of pulleys, from one of three starting positions to one of three end positions, via one of the two gaps (Figure 3.3a). The average speed of the target was 5.6 m/s. We used a randomised experimental

design to prevent the bird from learning the target's path, although we only present the results of a single flight here. The bird began flying freely from its falconer's fist as soon as the target appeared from inside the starting box in which it was hidden. Before recording the pursuit, we recorded a 'gaze calibration trial' to estimate the bird's gaze direction within the headpack coordinate system. In it the falconer held the bird on their fist whilst we displayed a piece of food with a marker attached at a range of 0.5-1.5 m.

In the obstacle avoidance flights, the bird flew between two perches set 9 m apart within a flight volume of 12.0 m$\times$ 5.3 m $\times$3.3 m (Figure 3.3b). To reach the opposite perch, the bird had to avoid a set of four 0.3 m diameter styrofoam pillars (height: 2 m) placed 1.5 m ahead of the end perch. The bird began flying freely from the starting perch when one of the two falconers whistled from the end perch. It received a food reward upon landing, and returned promptly to the starting perch when signalled by the other falconer, for a further food reward. The two obstacle avoidance flights that we present here to illustrate our method are part of a larger set of perching flights recorded with and without obstacles present. We used the complete set of flights from the same bird on the same day to calibrate the bird's gaze direction within the headpack coordinate system (see Section 3.3.3). Additional details on the motion capture experiments with the birds are included in Appendix B.1.

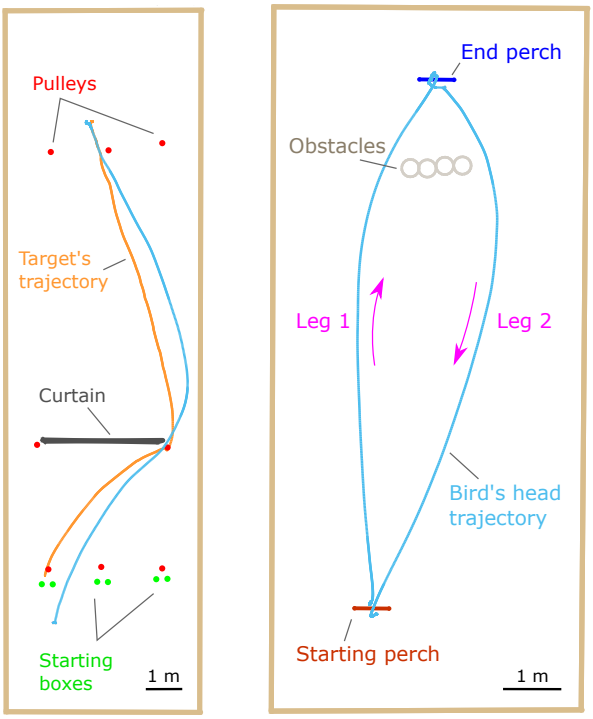
3.3.2 Motion capture data postprocessing

We used Nexus v2.8.0 software (Vicon Motion Systems Ltd., Oxford, UK) to estimate the 3D positions of all of the unlabelled retroreflective markers. In the pursuit flight, we labelled the headpack markers manually within Nexus, using its semi-automatic labelling functionality. In the obstacle avoidance flights, we labelled the headpack markers using the ground-truth labelling approach described in Chapter 2 (Section 2.3.3.3 and Appendix A.10). In both trials, we used custom scripts written in MATLAB R2020b (The Mathworks Inc., Natick, MA) to label the stationary markers, to label the target markers, to compute and interpolate

the position and orientation of the headpack and the target, and to handle missing marker data (see Appendix B.2). To assess the quality of our labelling, we computed, in a coordinate system fixed to the headpack, the maximum distance between a labelled marker and the corresponding marker in an ideal headpack template. The maximum distance deviation was 4.6 mm in the pursuit flight, and 2.8 mm in the obstacle avoidance flights (see Figure B.3 in Appendix B.2). Hence, although the ideal distance deviation without measurement error would of course be zero, the actual distance deviation is in all cases smaller than the diameter of the headpack markers, which implies that the markers are correctly labelled, albeit that their precise positioning is subject to measurement error.

3.3.3 Modelling the bird's view

We modelled the bird's view of the scene in Blender 2.93.5 (Blender Online Community, 2021) by defining a virtual camera whose movements are matched to those of the bird's head. However, as the headpack is arbitrarily placed on the bird's head, it is necessary to relate the headpack coordinate system to the bird's visual coordinate system. To do so, we make use of three assumptions (Minano et al., 2021). First, we assume that the bird's eye movements are negligibly small relative to those of its head. This is a common assumption in similar studies, especially those involving flight (Kano et al., 2018; Brighton et al., 2017; Ros et al., 2017; Kress et al., 2015; Kane et al., 2014; Eckmeier et al., 2008). Second, we assume that the bird fixates its gaze on points in the environment that can be specified from first principles. For the pursuit case, we assume that the bird fixates its gaze on the food presented to it during calibration. For the obstacle avoidance case, we assume that the bird fixates its gaze on the perch centre as it lands (Potier et al., 2016; Kress et al., 2015). Third, we assume that the bird holds its eyes level when fixating the food presented to it, and also when in flight. Head levelling behaviour has been reported multiple times in the bird flight literature (Brighton et al., 2019; Ros et al., 2017; Warrick et al., 2002), and is confirmed for the gaze calibration by the reference video.



(a) Pursuit trial (b) Obstacle avoidance trial

Figure 3.3: Pursuit and obstacle avoidance flights. An overhead view of the motion capture trajectories collected on the pursuit and obstacle avoidance trials is shown. Note the different scale bar for each panel (bottom right). (a) The pursuit trial consists of one flight. The bird's head trajectory (light blue) is shown as it flies to intercept a moving target (orange). The bird takes off spontaneously from the fist of a falconer standing 1 m behind the starting boxes (green markers). The target is attached to a line that passes around a series of pulleys (red). It begins hidden within the leftmost starting box in the flight shown, and is pulled by a linear motor located at the top of the image (not shown). The curtain (black) is represented here as the best-fitting plane to the reference markers attached to its edges. (b) The obstacle avoidance trial consists of two flights, each corresponding to one leg of a back-and-forth flight (magenta arrows) between two perches, with a bank of tall cylindrical obstacles located 1.5 m ahead of the end perch. The walls (brown) are estimated as planes from the motion capture cameras' positions and orientations. The pursuit flight takes 2.5 s and the obstacle flights around 2 s each. Further details on the positions of the reference markers for the different objects can be found in Appendix B.4.

We identified specific periods in the collected data in which we expect these assumptions to hold most reliably. We used the data during these periods to estimate $\vec{v}_{gaze}$, the bird's gaze direction, and $\vec{n}_{sagittal}$, the normal to the bird's head symmetry plane (i.e., the sagittal plane). These vectors define the basis of the visual coordinate system: the x-axis is defined parallel to $\vec{n}_{sagittal}$ (pointing to the left side of the head) and the y-axis parallel to $-\vec{v}_{gaze}$. We defined the x and y axes in this way so that the visual coordinate system is close to parallel to the motion capture coordinate system at the start of the trial; this facilitates

the interpretation of the computed rotations. The following sections describe the estimation of $\vec{v}_{gaze}$ and $\vec{n}_{sagittal}$ for the pursuit and the obstacle avoidance flights; additional details are included in Appendix B.3.

3.3.3.1 Estimate for the pursuit flight

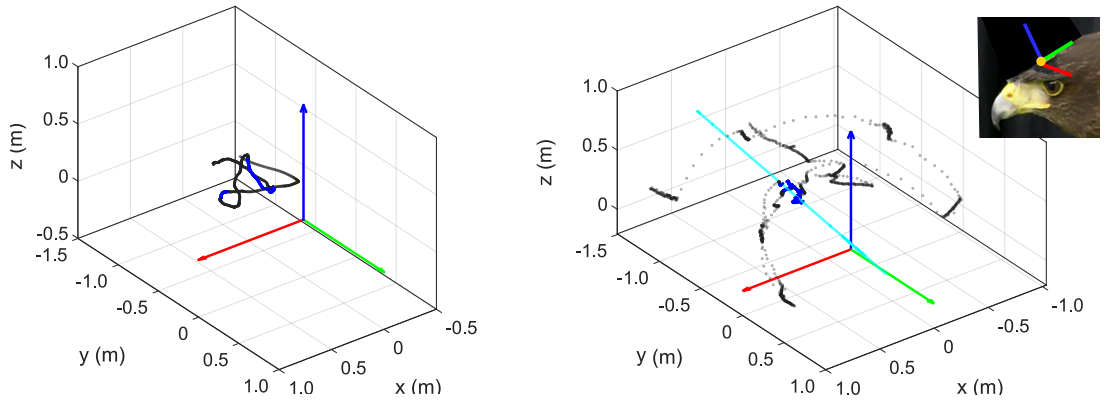
For the pursuit flight, we estimated the visual coordinate system using data from the gaze calibration trial. In the corresponding reference video, we identified a 'fixation phase': a range of frames in which the bird is likely to be fixating on the food reward while holding its head level. We identified these frames by a characteristic behaviour of the bird, in which it lowers its head and prepares its wings for a downstroke (see Appendix B.3.1, Figure B.5).

We estimated the bird's forward gaze direction $\vec{v}_{gaze}$ by fitting a 3D line to the food marker's trajectory during the fixation phase, in an auxiliary coordinate system linked to the headpack (RMSE = 47.5 mm for $N = 462$ samples, see Figure 3.4).

For the same range of frames, we estimated the orientation of the imaginary line connecting the bird's eyes, assuming the animal keeps its head approximately level during that period. We identify this line with $\vec{n}_{sagittal}$, the normal to the bird's head symmetry plane (i.e. the sagittal plane), which we define as positive when pointing to the left side of the bird's head. We compute $\vec{n}_{sagittal}$ as the unit vector perpendicular to $\vec{v}_{gaze}$ that best approximates the normal to the sagittal plane in the headpack coordinate system. We do this by solving the following least-squares problem:

$$\begin{aligned} \theta^* &= \arg\min_{\theta} \sum_{i=1}^{N} (\vec{z}_{world,i} \cdot \vec{n}_{sagittal}(\theta))^2, \\ \vec{n}_{sagittal}(\theta) &= (cos\theta)\vec{a} + (sin\theta)\vec{b}, \\ \vec{n}_{sagittal} &= \vec{n}_{sagittal}(\theta^*), \end{aligned} \tag{3.1}$$

where $\vec{z}_{world,i}$ is the world's z-axis in the headpack coordinate system (positive opposite to gravity), $\vec{a}$ and $\vec{b}$ are an arbitrary orthonormal basis of the plane perpendicular to $\vec{v}_{gaze}$, and θ is the angle between $\vec{n}_{sagittal}$ and $\vec{a}$. The fit over the $N = 462$ fixation frames yields a root-mean square residual of 0.04 (see Figure 3.5).



(a) Motion capture coordinate system

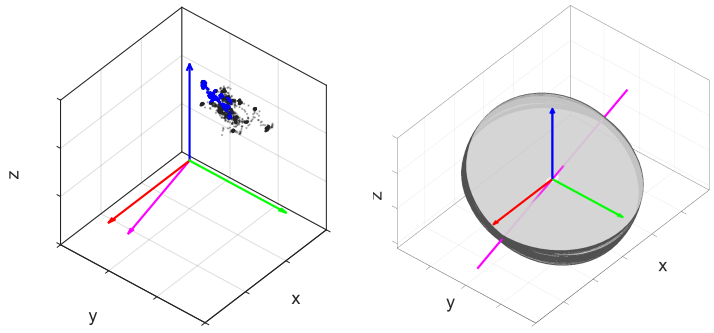
(b) Headpack coordinate system

Figure 3.4: Estimation of gaze direction for the pursuit flight. The trajectory of the marker attached to the food reward in the gaze calibration trial is represented in two coordinate systems: one that translates with the headpack and is parallel to the motion capture coordinate system (a), and one fixed to the headpack (b). The estimated gaze direction $\vec{v}_{gaze}$ is the orthogonal regression line (cyan) to the samples in the fixation phase (blue) in the headpack coordinate system (b). The samples outside this phase are shown for reference (black, semi-transparent). The food reward is presented to the bird at a distance between 0.5 and 1.5 m. Note how the position of the food marker in the fixation phase in (b) falls within a small region, in contrast to the curved trajectory it follows in (a), suggesting the bird is actively positioning the food reward on a small area of its visual field during that phase. Inset in (b) shows the approximate orientation of the headpack coordinate system relative to the bird's head.

The origin of the visual coordinate system in the pursuit flight is defined as the midpoint between the two lateral markers of the headpack (see appendix Figure B.6a). We estimated this point to be close to the midpoint between the eyes, the ideal origin of the visual coordinate system (Potier et al., 2016). From reference images of the headpack on the bird's head, we estimate that the origin is < 10 mm from the midpoint between the bird's eyes in the direction perpendicular to the sagittal plane, and < 20.1 mm in the direction perpendicular to the headpack's baseplate (see Appendix B.3.2).

3.3.3.2 Estimate for the obstacle avoidance flights

For the obstacle avoidance flights, we estimated the visual coordinate system using data from 15 trials with the same bird and headpack placement. Note that each trial consists of two flights, which correspond to the two legs of the trial (see Figure 3.3). From those 15 trials, 8 were recorded with obstacles in place, and 7 without them.



(a) Estimation of sagittal plane (b) Local horizon

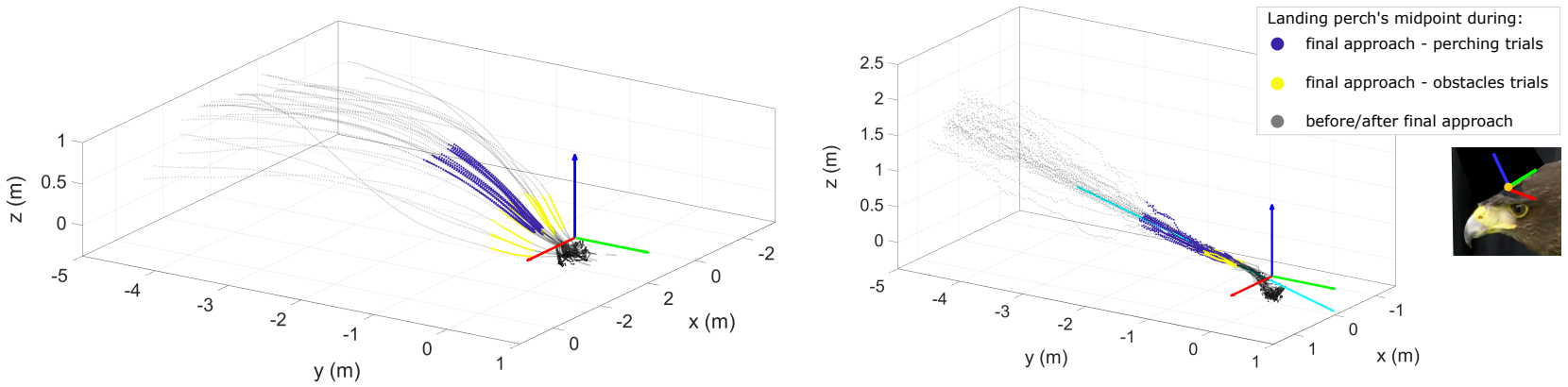
Figure 3.5: Estimation of the normal to the sagittal plane for the pursuit flight. We estimated the normal to the sagittal plane $\vec{n}_{sagittal}$ (magenta), assuming the bird keeps its eyes level during the fixation phase of the gaze calibration trial. In (a), dots represent the direction of the world's z-axis in the headpack coordinate system, during the fixation phase (blue markers) and outside the fixation phase (black, semi-transparent). Samples outside the fixation phase show larger variation in roll (i.e. rotation around the y-axis). In (b), the local horizon (i.e. the plane perpendicular to the world's z-axis in the headpack coordinate system) is represented for every frame in the fixation phase (gray circular planes). The direction of the $\vec{n}_{sagittal}$ vector (magenta) is the best-fit line that is perpendicular to the previously estimated $\vec{v}_{gaze}$ direction and is contained in all these planes, in least-squares sense.

We identified two phases in all the flights within this set of trials: a final approach phase, in which we expect the bird to fixate on the centre of the perch; and a mid-flight phase, in which we expect the bird to keep its head level. We used these two phases to estimate $\vec{v}_{gaze}$ and $\vec{n}_{sagittal}$, respectively.

We estimated the bird's forward gaze direction $\vec{v}_{gaze}$ using data from the final approach phase. This phase was defined based on the bird's distance to the landing perch (0.5 to 1.0 m for the obstacle trials, 0.5 to 2.0 m for the no obstacle trials; note that the obstacles were placed 1.5 m ahead of the end perch, as per Figure 3.3b). We fitted a 3D line to the trajectory of the midpoint of the landing perch during this phase, in a headpack coordinate system (see Figure 3.6). The root-mean square error of the fit was RMSE = 65.6 mm, for N = 1274 samples. Results were similar to those obtained by fitting each trial individually (see Appendix B.3.3, Table B.3). We computed the normal to the sagittal plane $\vec{n}_{sagittal}$ following the same approach as in the pursuit flight, using the data of the mid-flight phase. We defined this phase to be when the bird was > 2 m away from either perch and flying at a speed > 2.5 m/s; for all flights these conditions defined a continuous range of frames.

We solved for $\vec{n}_{sagittal}$ using the set of equations in 3.1, and obtained a root-mean square residual of 0.08 for $N = 5056$ samples (see Figure 3.7).

We defined the origin of the visual coordinate system as the centroid of the headpack markers projected onto the headpack's baseplate plane. This is different to the definition used in the pursuit flight because we used a different headpack design. From reference images, we estimate this point is within 12 mm of the midpoint between the bird's eyes in the direction perpendicular to the sagittal plane, and within 17 mm in the direction perpendicular to the headpack's baseplate (see Appendix B.3.2).



(a) Motion capture coordinate system

(b) Headpack coordinate system

Figure 3.6: Estimation of the gaze direction for the obstacle avoidance flights. The gaze direction $\vec{v}_{gaze}$ is estimated over 8 trials with obstacles in place (obstacle trials) and 7 trials without them (perching trials). Note that each trial consists of two flights, which correspond to the two legs of the trial (see Figure 3.3). The landing perch's midpoint (grey and coloured dots) is represented from 5 m away until landing, in two coordinate systems: one that translates with the headpack and is parallel to the motion capture coordinate system (a), and one fixed to the headpack (b). The estimated gaze direction $\vec{v}_{gaze}$ is the orthogonal regression line (cyan) to the samples in the final approach phase (plotted blue for perching trials, and yellow for obstacle trials) in the headpack coordinate system (b). Note the straightness of the perch's midpoint trajectories in the headpack coordinate system in (b), and compare this to the curvature of the lines in (a), which confirms unequivocally that the head pose was being stabilised in relation to the perch. Inset in (b) shows the approximate orientation of the headpack coordinate system relative to the bird's head.

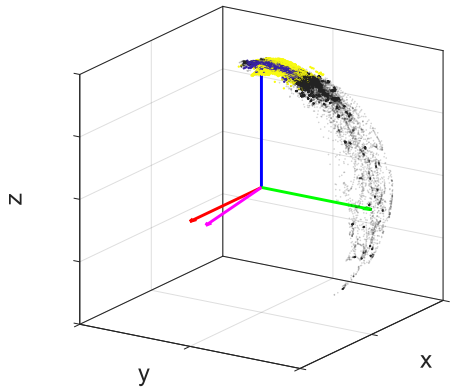


Figure 3.7: Estimation of the normal to the sagittal plane for the obstacle avoidance flights. We estimated the normal to the sagittal plane $\vec{n}_{sagittal}$ (magenta) assuming the bird keeps its eyes level during the mid-flight phase. The dots represent the direction of the world's z-axis in the headpack coordinate system during the mid-flight phase, for perching (blue) and obstacle trials (yellow). Note that each trial consists of two flights, which correspond to the two legs of the trial (see Figure 3.3). Samples outside of the mid-flight phase (black, transparent), show qualitatively larger variation in pitch (i.e., rotation around the x-axis) and roll (i.e., rotation around the y-axis).

3.3.3.3 Retinal margins

The retinal margins define the set of directions in the visual field that project on the right or left retina; we represented the retinal margins of Harris' hawks' eyes in the estimated visual coordinate system, using data from Potier et al., 2016 (Figure 3.8). We identified the estimated gaze direction $\vec{v}_{gaze}$ with the direction of maximum binocular overlap, based on the description of Fig. S1 in Potier et al., 2016. Further details on the computation of the retinal margins are presented in Appendix B.3.5.

3.3.3.4 Virtual camera

We modelled the scene viewed by the bird in Blender, using a 360° virtual camera whose translation and rotation per frame were matched to those of the estimated visual coordinate system fixed to the bird's head. As any vergence movements of the eyes are unknown, we modelled the binocular visual field as a monocular camera. We selected a resolution of 5 pixels per degree latitude and longitude, which results in pixels within the bird's visible visual field having a length and width approximately 10 times larger than the minimum resolution angle of the bird (Potier et al., 2016). For each flight, we rendered the view from this virtual

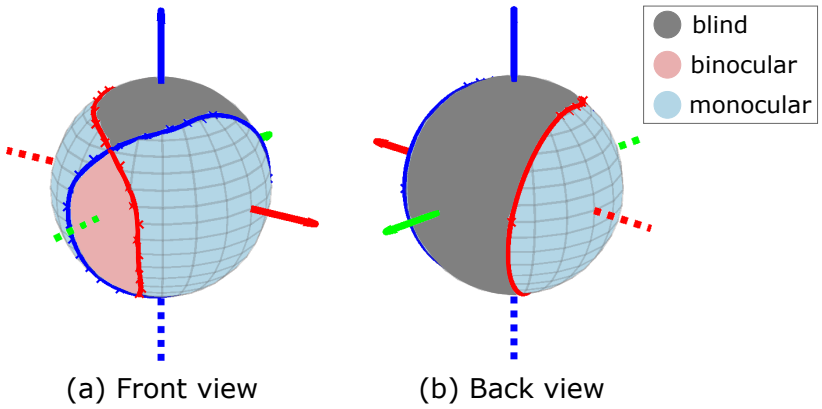


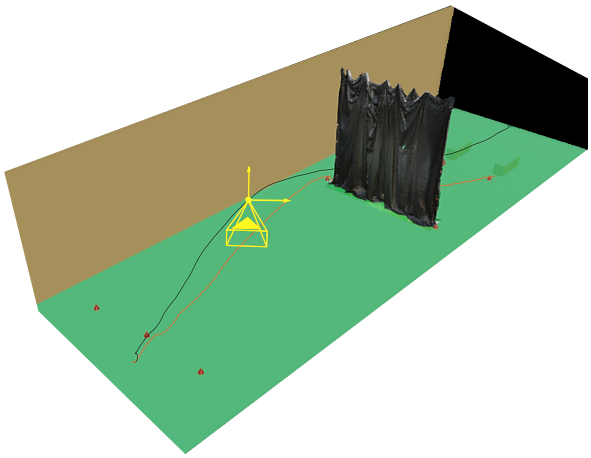
Figure 3.8: Retinal margins of Harris' hawks eyes. The retinal margins for the left (blue) and right (red) eyes are shown on the unit sphere in the visual coordinate system, from the front (a) and the back (b). The data points for the retinal margins from Potier et al., 2016 are represented by crosses. The blind (grey), binocular (pink) and monocular (blue) areas are shown. The parallels on the sphere are plotted every 9° in latitude and the meridians every 18° in longitude.

camera in the bird's visual coordinate system using the Cycles rendering engine, and produced RGB, depth, semantic and optic flow data per frame.

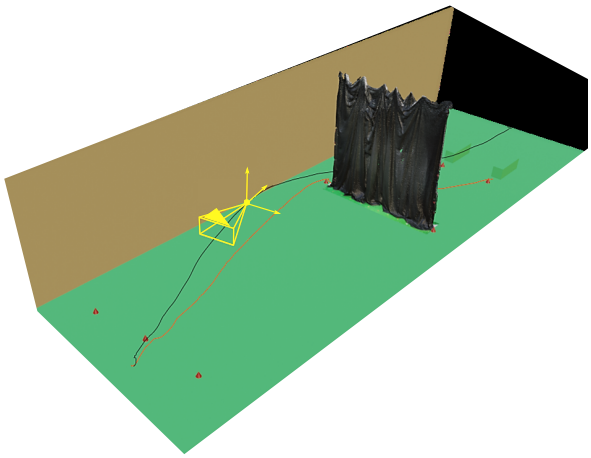
To test the effect of the bird's head movements, we also defined a trajectory coordinate system representing the pose of a horizon-levelled virtual camera whose optical axis was directed tangent to the bird's head trajectory (see Figure 3.9). Specifically, we set the y-axis of the trajectory coordinate system parallel to the birds' head velocity vector and its x-axis parallel to the floor plane, whilst fixing its origin to that of the visual coordinate system (see Appendix B.3.6). We again rendered the view from a virtual camera following the motion of this trajectory coordinate system for all three flights.

3.3.4 Hybrid model of the lab environment

We model the lab environment using a hybrid approach, which uses a combination of geometric primitives for the simple geometries in the scene, and dense 3D meshes for the most complicated ones. The dense 3D meshes are captured with a mobile device, and expressed at acquisition time in the same coordinate system as the motion capture trajectories. In this way we aim to minimise the modelling and postprocessing effort, while producing realistic representations of the environment.



(a) Visual coordinate system



(b) Trajectory coordinate system

Figure 3.9: The two coordinate systems used for rendering. For each flight, we produce one set of renderings in which the virtual camera pose matches that of the visual coordinate system, and one set in which it matches the pose of the trajectory coordinate system (see text for definitions). The figure shows the pose of the virtual camera (yellow), the bird's head trajectory (black line) and the target's trajectory (orange line), for a selected frame in the final stages of the interception manoeuvre during the pursuit flight (ceiling and walls removed for clarity). In (a), the camera tracks the estimated visual coordinate system, which we expect to be representative of the pose of the bird's visual field. In (b), the camera tracks the trajectory coordinate system, which represents a horizon-level camera whose optical axis is tangent to the head trajectory. Although the camera is represented schematically as a pyramid, a 360° virtual camera was considered.

We demonstrate the use of a hybrid model of the lab for the pursuit flight, modelling the curtain using a dense 3D mesh, and the rest of the objects in the scene as geometric primitives (see middle panel in Figure 3.2). In the obstacle avoidance flights, we represented all the main elements of the scene using geometric primitives (see Figure 3.10). The synthetic scene can easily be made more photo-realistic as

required: as an example, we provide an RGB rendering in which we added texture to the walls replicating the camouflage netting photographically (see Supplementary videos `ESM_1.avi` and `ESM_2.avi`). Additional details on the 3D model of the lab are included in Appendix B.4.

3.3.4.1 Geometric primitives

We modelled the floor, ceiling and walls of the lab as planes. We determined the positions of the walls and ceiling from estimated points in the motion capture system's scaffolding, which we derived from the cameras' positions and orientations (see Appendix B.2.4). For the RGB renderings, we added texture to the walls representing the camouflage netting that was hung from the walls of the lab to provide a natural-looking texture and prevent the birds from perching. In the pursuit flight, we modelled the pulleys as cones and the boxes covering the target's initial position as cuboids (see middle panel in Figure 3.2). We modelled the target as a cylinder of 15 cm length and 2.54 cm diameter. We placed additional reference markers on the curtain's edges and on the wall netting at the curtain gap, but only used them to measure deviation from our modelled geometry (see Appendices B.1.2 and B.4.3). In the obstacle avoidance flights, we modelled the obstacles as vertical cylinders, and the perches as horizontal cylinders, thereby reducing each A-frame perch to the top rung on which the bird landed (Figure 3.10). The position, orientation and size of these geometric elements was determined from the retroreflective markers attached to the corresponding objects, and from measurements of the dimensions of the real objects. Further details on the definition of the geometric primitives are included in Appendix B.4.

3.3.4.2 Dense 3D map

To capture a dense 3D map of the curtain, we used the open-source *SemanticPaint* framework (Golodetz et al., 2015; Golodetz et al., 2018), which is built on top of InfiniTAM v3 (Prisacariu et al., 2017). We used the ASUS ZenFone AR smartphone as a mobile mapping sensor, and to perform visual-inertial odometry (ZenFone

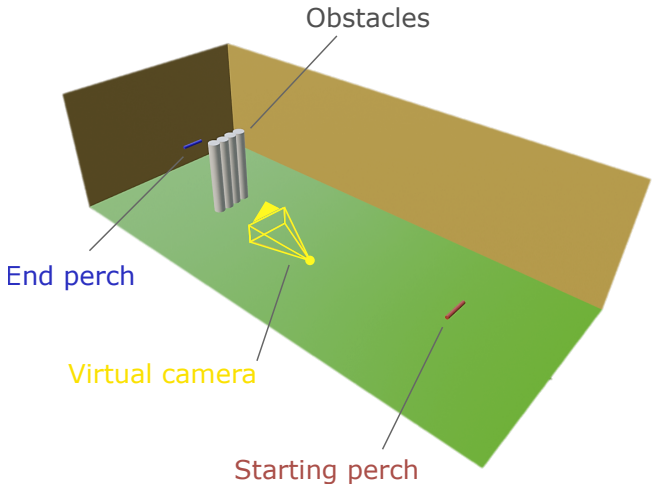


Figure 3.10: Model of the lab environment for the obstacle avoidance flights. In the obstacle avoidance flights, we modelled the lab environment using basic geometric primitives to characterise the main elements in the environment. The walls, floor and ceiling were computed from the positions and orientations of the motion capture cameras. The obstacles were modelled as vertical cylinders and the perches as horizontal ones, reducing them to their top rung. The positions of the perches and obstacles were determined from the reference retroreflective markers attached to them. Their dimensions were derived from the motion capture markers, or from direct measurements of the objects.

ZS571KL, ASUS, Taipei, Taiwan). To facilitate the integration of the mesh in the virtual model of the lab, we captured it in the motion capture coordinate system, using ArUco fiducial markers (Romero-Ramirez et al., 2018; Garrido-Jurado et al., 2014). The voxel size was set to 10 mm and the truncation distance to 40 mm (4× the voxel size).

The coordinate transformations applied to express the mesh in the motion capture coordinate system are summarised in Figure 3.11. By default, the 3D mesh is captured in the SLAM world coordinate system, which is defined by the first camera pose. To compute the required transform from the SLAM world coordinate system to the motion capture coordinate system, we used an ArUco calibration plate. This consisted of an ArUco fiducial marker of size 28.8 × 28.8 cm fixed to an acrylic plastic sheet with three retroreflective markers (10 mm diameter) on three of its corners. When brought into camera view, the coordinates of the ArUco marker's corners are computed in the SLAM world coordinate system. Since we also placed retroreflective markers on these corners, their coordinates in the motion capture coordinate system are also known. By defining an auxiliary coordinate

system with these three points, we can compute the transform from the SLAM world coordinate system to the motion capture one.

We used the open-source software MeshLab (Cignoni et al., 2008) to crop the mesh, remove duplicate vertices, and remove isolated pieces. We found that the floor plane of the mesh was slightly deviated from the motion capture system's xy-plane (2.4°, see Appendix B.4.3), likely due to drift. We used MATLAB's Point Cloud Processing functions to fit a plane to the floor of the mesh (mean error = 3.8 cm), and to compute the required transform to match the fitted plane with the motion capture's XY plane. The transformed mesh deviated on average by 0.093 m from the reference markers placed on the curtain's edges, as they were registered during the pursuit trial, and by 0.089 m from their position recorded just before capturing the mesh. Further details on the postprocessing of the mesh and the deviation metrics are included in Appendix B.4.3.

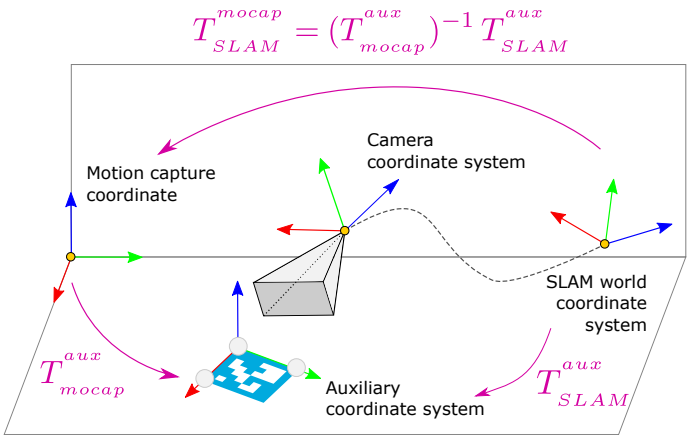


Figure 3.11: Capturing the 3D dense map in the motion capture coordinate system. By default, the 3D mesh is captured in the SLAM world coordinate system, defined by the first camera pose. The camera's trajectory (dashed line) and pose in this coordinate system are computed via visual-inertial odometry. When the ArUco marker (shown in cyan) is brought into camera view, the coordinates of three of its corners are computed in the camera coordinate system, and thus via visual-inertial odometry they are also known in the SLAM world coordinate system. Since we placed three retroreflective markers on these corners (grey circles), their coordinates are also known in the motion capture coordinate system. The pose of an auxiliary coordinate system defined by these three points is thus known in the motion capture coordinate system and the SLAM world coordinate system, which allows us to compute the required transform. The transforms between coordinate systems are shown in magenta, where T_A^B represents the transform from A to B.

3.4 Results

For all three flights, we rendered the view from: (i) a virtual camera in the visual coordinate system, representing the translation and rotation of the bird's visual field inclusive of all head movements; and (ii) a virtual camera following the trajectory coordinate system, representing the translation and rotation of the bird exclusive of rotational head movements. Videos containing each of the rendered outputs per frame are included as Supplementary information (see Supplementary videos `ESM_1.avi` to `ESM_15.avi`); a brief description of them is included in the section 3.6, and further details on how they were computed is provided in Appendix B.5. As a demonstration of how these synthetic data can offer new insights into the bird's behaviour, we used the RGB and semantic data per frame to inspect the birds' gaze strategy, following a similar approach to the preliminary analysis presented by Minano et al. (2021).

3.4.1 Pursuit flight

The RGB renderings in the visual coordinate system for the pursuit trial show that the target remains within the bird's area of binocular overlap for almost the entire duration of the pursuit (see snapshots in Figure B.12, and Supplementary videos `ESM_1.avi` and `ESM_2.avi`). In contrast, the target is not held steady or centered within the trajectory coordinate system, for which the RGB renderings display pronounced pitch oscillations. This comparison confirms that the bird uses its rotational head movements to stabilize its gaze, and to keep the moving target reasonably well centered within its visual field. To analyse these properties quantitatively, we make use of the semantic data.

Figure 3.13 shows the frequency with which the target appears at each point in the visual field during the flight. For each pixel, the figure displays the metric:

$$h_i = \frac{n_i}{NA_i}, \tag{3.2}$$

where n_i denotes the number of frames over which the ith pixel saw the target, where N denotes the number of frames analysed, and where A_i denotes the solid

angle covered by the ith pixel, normalised by the maximum solid angle that any pixel covers. Note that as the output is an equirectangular projection of a sphere, different pixels subtend different solid angles in the figure. We only consider those pixels falling within the visible areas of the bird's visual field, and exclude from the analysis any frames in which the head transform was interpolated, and any frames after interception (identified as the point at which the distance to the target first reaches a local minimum in the final approach phase). The results show that the target is held $\pm 10°$ longitude and from $-10°$ to $4°$ latitude in the visual coordinate system for most of the flight (Figure 3.13a). This contrasts with the results in the trajectory coordinate system, in which the target is not confined to this central area at all (Figure 3.13b). However, a limitation of these results is that they are affected by the apparent size of the target, and refer to data aggregated across the whole of the flight. How does the target's position in the visual field vary quantitatively through the flight?

Figure 3.14 plots the evolution of the target's edge contour in the visual coordinate system, where the visual field is shown in equirectangular projection and cropped close to the binocular area. The sections of flight before and after the curtain are plotted separately for clarity (Figure 3.14a and Figure 3.14b respectively). In the first section of the flight, the target begins drifting across the visual field, but then seems to be stabilised at approximately $10°$ longitude (Figure 3.14a). Target tracking seems to be lost temporarily as the bird turns around the curtain (green contours in Figure 3.14a), but is quickly recovered with the target now stabilised at $-10°$ longitude (Figure 3.14b). Towards the end of the flight, the target gradually becomes centred in the visual field, looming until interception. The same evolution can be seen by inspecting the longitudinal position of the boundaries and midpoint of the target through time (Figure 3.14c). The target's boundaries also remain between $-10°$ and $4°$ latitude for most of the flight, except at takeoff, when turning around the curtain, and at final interception (Figure 3.14d).

For comparison, we computed the equivalent path of the target as seen from the trajectory coordinate system (Figure 3.15). In the first part of the flight, the

target shows considerably more oscillations in the vertical direction, likely reflecting wingbeat oscillations (Figure 3.15a). However just before turning around the curtain, the target appears to be aligned longitudinally with the bird's head velocity vector. In the second part of the flight (Figure 3.15b), the target is clearly not aligned with the head's velocity vector, drifting out of the central area of the trajectory coordinate system (Figure 3.15b). The comparison of Figures 3.15b and Figure 3.14b reflects how the estimated gaze direction diverges from the bird's velocity vector in the final phase of interception. In summary, although the moving target was not perfectly stabilised in the bird's visual field and sometimes appeared to be held at $\pm 10^{\circ}$ longitude, rather than being centered, the effect of the bird's head movements was to keep the target contained within the area of binocular overlap (Figure 3.14a).

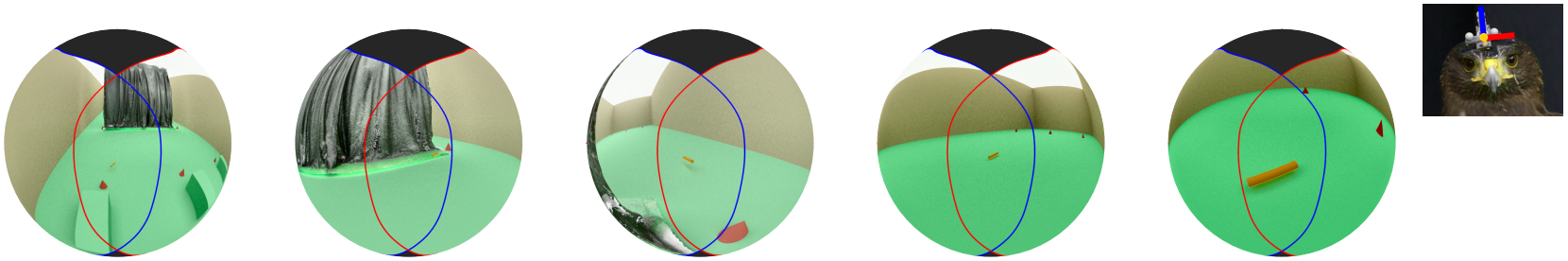
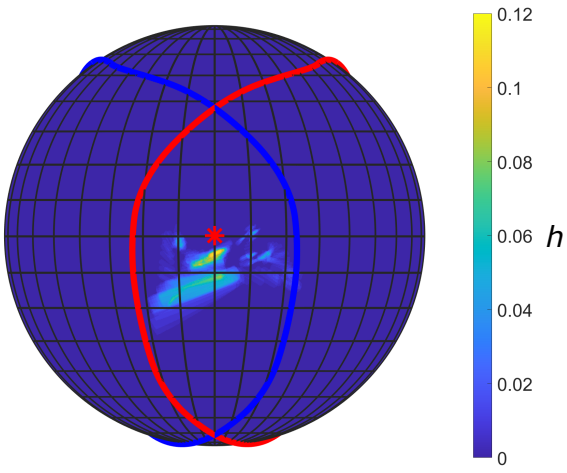
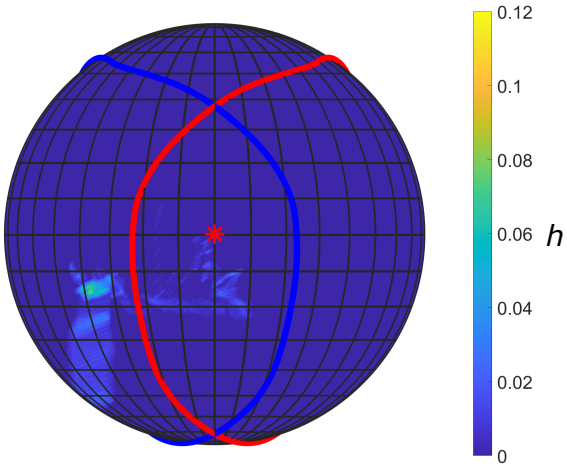


Figure 3.12: Snapshots of the rendered pursuit flight. The spheres representing the bird's visual field show the animal's point of view as it is flying through the lab; the retinal margins for the left (blue line) and right (red line) eyes are shown for reference, as well as the blind area in the visual field (black). Note that because the spheres are represented using an orthographic projection (right panel), not all the visual field of the bird is visible. The inset shows the approximate orientation of the estimated visual coordinate system on the bird's head. The full flight takes 2.5 s.



(a) Visual coordinate system



(b) Trajectory coordinate system

Figure 3.13: Semantic heatmap for the target. The heatmaps show the value of h as the frequency with which the target appears at each pixel in the visual field throughout a flight, normalised by the pixels' subtended solid angle. Results are shown for a camera following the visual coordinate system (a), and the trajectory coordinate system (b). The virtual camer axis is marked by a red star: it corresponds to the estimated gaze direction $\vec{v}_{gaze}$ in the visual coordinate system, and the velocity vector direction in the trajectory coordinate system. In the virtual coordinate system (a), the target appears centred in the binocular area throughout the flight. In the trajectory coordinate system (b), it doesn't appear constrained to this region. We considered $n = 337$ frames, from the takeoff frame to the interception frame, and excluded the frames in which the head translation or rotation was interpolated. Note that the visible area in the bird's visual field extends beyond what is shown in this projection (orthographic). The colourbar maximum value was set to match the maximum value observed in the visual coordinate system data. The visual field is represented as seen from the bird's point of view

.

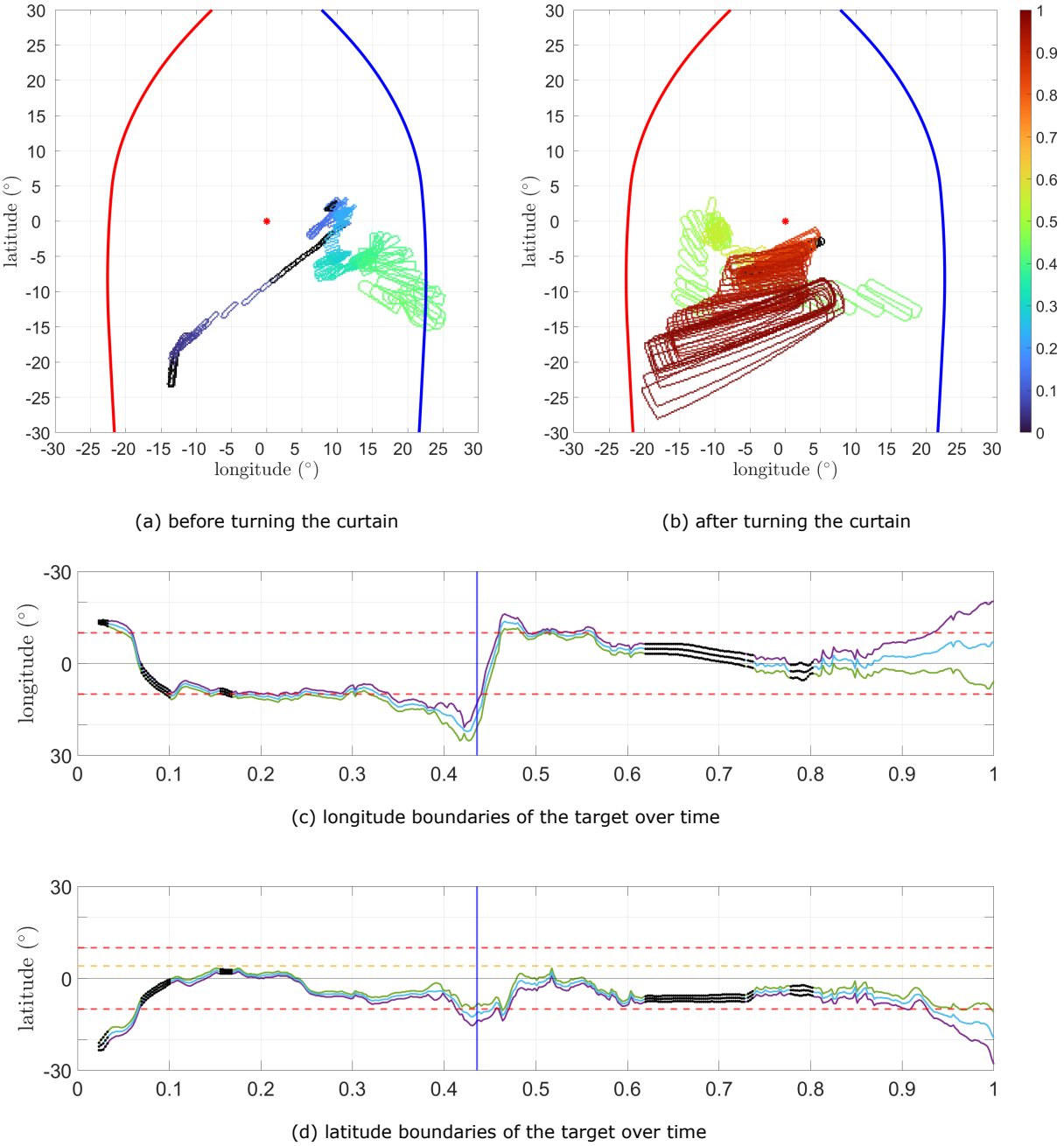
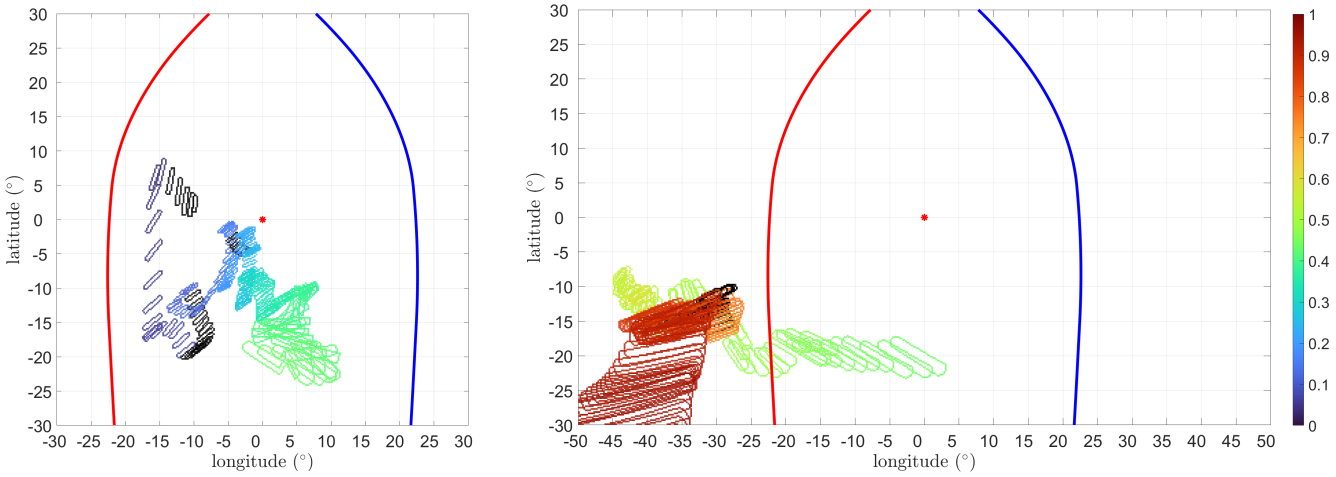


Figure 3.14: Trajectory of the target in the bird's visual field. The contour of the target is represented per frame in an equirectangular projection of the area around the estimated gaze direction $\vec{v}_{gaze}$ (red star), for the frames before (a) and after (b) turning around the curtain. The colormap indicates time through the flight. The extension of the target in longitude and latitude over time is shown in (c) and (d) respectively, with the maximum (purple) and minimum (green) values represented per frame, and their mean (blue). For the frames in which the bird's head transform was interpolated, these values are shown in black. The blue vertical line in (c) and (d) indicates the frame that defines the data split before and after turning the curtain, into (a) and (b). Reference lines are shown in (c) and (d) at ±10° (red dashed lines) and at 4° latitude in (d) (yellow dashed line). The normalised time is 0 at the takeoff frame, identified when the bird dips its head just before the takeoff jump; note that it takes some frames for target to be visible but the linear motor pulling it was already triggered. The normalised time is 1 at interception, when the bird and the target reach a local minimum in distance in the final terminal phase. The data shown covers 2.145 s.



(a) before turning the curtain

(b) after turning the curtain

Figure 3.15: Trajectory of the target in the trajectory coordinate system. The contour of the target is represented per frame in an equirectangular projection of the area around the camera axis (red star) in the trajectory coordinate system. The virtual camera's axis corresponds to the head velocity vector. The retinal margins of the bird's right (red) and left (blue) eyes, relative to the virtual camera's axis, are show for reference. The data is split following the same criteria as in Figure 3.14, separating the frames before (a) and after (b) turning around the curtain. The colormap indicates time through the trial. Note that in the first part of the flight the target shows oscillations in the vertical direction, likely reflecting the unstabilised wingbeat induced motion of the head in the trajectory coordinate system. In the second part of the flight, the target is not constrained to the equivalent of the binocular area, showing that the velocity vector of the head trajectory is not aligned with the target.

3.4.2 Obstacle avoidance flights

The RGB renderings of the visual coordinate system for the obstacle avoidance trials show the obstacles located towards the centre of the visual field (see Supplementary videos `ESM_8.avi` and `ESM_9.avi` for leg 1, and Supplementary videos `ESM_12.avi` and `ESM_13.avi` for leg 2). This is not the case for the renderings of the same flights in the trajectory coordinate system, in which the obstacles are not centred, and where wingbeat oscillations are clearly visible (see Supplementary videos `ESM_10.avi` and `ESM_11.avi` for the flight corresponding to leg 1 of the trial, and Supplementary videos `ESM_14.avi` and `ESM_15.avi` for the flight corresponding to leg 2). This confirms that the bird actively stabilises its visual field against the pitch oscillations associated with its wingbeat, and also directs its gaze so as to keep the obstacles broadly centered. In this case, however, close inspection of the semantic data reveals a more subtle interpretation of how the bird is directing its gaze.

As the obstacles occupy a much larger portion of the field of view than does the target shown in the cropped equirectangular projections of Figure 3.14, we display the evolution of the obstacles' edge contour using an orthographic projection to reduce distortion in Figure 3.16. The obstacles' edge contour is represented from the point of takeoff until the point at which the landing perch is visible without occlusion. The data show that the nearside edge of the obstacles seen by the bird remains aligned longitudinally with the centre of the visual coordinate system in both flights (Figure 3.16, top row). This alignment does not appear for the semantic data rendered within the trajectory coordinate system (Figure 3.16, bottom row), which reinforces the role of the bird's head movements in fixating the obstacles' nearside edge.

Figure 3.17 plots how the longitudinal positions of the nearside edge, the farside edge, and the midpoint of the obstacles and landing perch evolve through time. For the obstacle avoidance flight comprising the first leg of the trial, the nearside edge of the obstacles appears to remain approximately aligned with the centre of the visual field until the point at which the landing perch first becomes fully visible (Figure 3.17a). Beyond this point, the midpoint of the landing perch becomes the object most closely aligned with the centre of the visual field. For the obstacle avoidance flight comprising the second leg of the trial, the nearside edge of the obstacles is again aligned closely with the bird's estimated gaze direction until the point at which the landing perch becomes visible. The bird then appears to make a head saccade, such that its new gaze direction aligns with the nearside edge of the landing perch, which remains the case for approximately the next 0.8 s (Figure 3.17b). The bird then seems to make another head saccade, such that its estimated gaze direction becomes most aligned with the midpoint of the perch. Both head saccades can be seen in the corresponding RGB rendered video (Supplementary videos `ESM_12.avi` and `ESM_13.avi`).

In summary, when avoiding the obstacles, the bird stabilized and centered the nearside edge of the obstacles in its visual coordinate system. The bird switched to fixating the perch once this was fully visible, looking at the nearside edge of

the perch first in one of the flights, and then at its midpoint before landing. It is important to note that the latter statement is inevitable for the final set of frames that we used to calibrate the bird's gaze direction under the assumption that the bird would look at the midpoint of the perch when landing. However, the coincidence of the edge of the obstacles with the centre of the visual coordinate system provides strong internal support for the reliability of its calibration, and the midpoint of the perch was found to be centered within the visual coordinate system from well before the start of the set of frames that were used in the calibration (Figure 3.17).

3.5 Discussion

Here we have presented a method generating synthetic visual inputs characterising the visual experience of a bird whose head pose is recorded in flight using a motion capture system or other measurement system providing similar data (e.g. a head mounted inertial measurement unit). Although we have limited our analysis of these data to presenting simple metrics for three flights as an illustration of the method, these are already sufficient to show the potential of using synthetic data generated using our method to investigate how birds use vision to guide and control their flight. Comparable analyses would be exceedingly challenging using video data from head-mounted cameras, given the attendant limitations on payload, resolution, field of view, and motion blur, and would be much more limited if relying on point estimates of the bird's gaze direction in relation to prominent visual features in the environment. Our method allows us to define alternative gaze strategies, by comparing our renderings in the visual coordinate system to renderings in the trajectory coordinate system, in order to inspect counterfactual scenarios. Again, this would be difficult or impossible to do using any of the other reviewed approaches.

3.5.1 Key features of rendering method

Our rendering method uses a hybrid approach to model the lab environment, which combines basic geometric primitives defined using the motion capture data

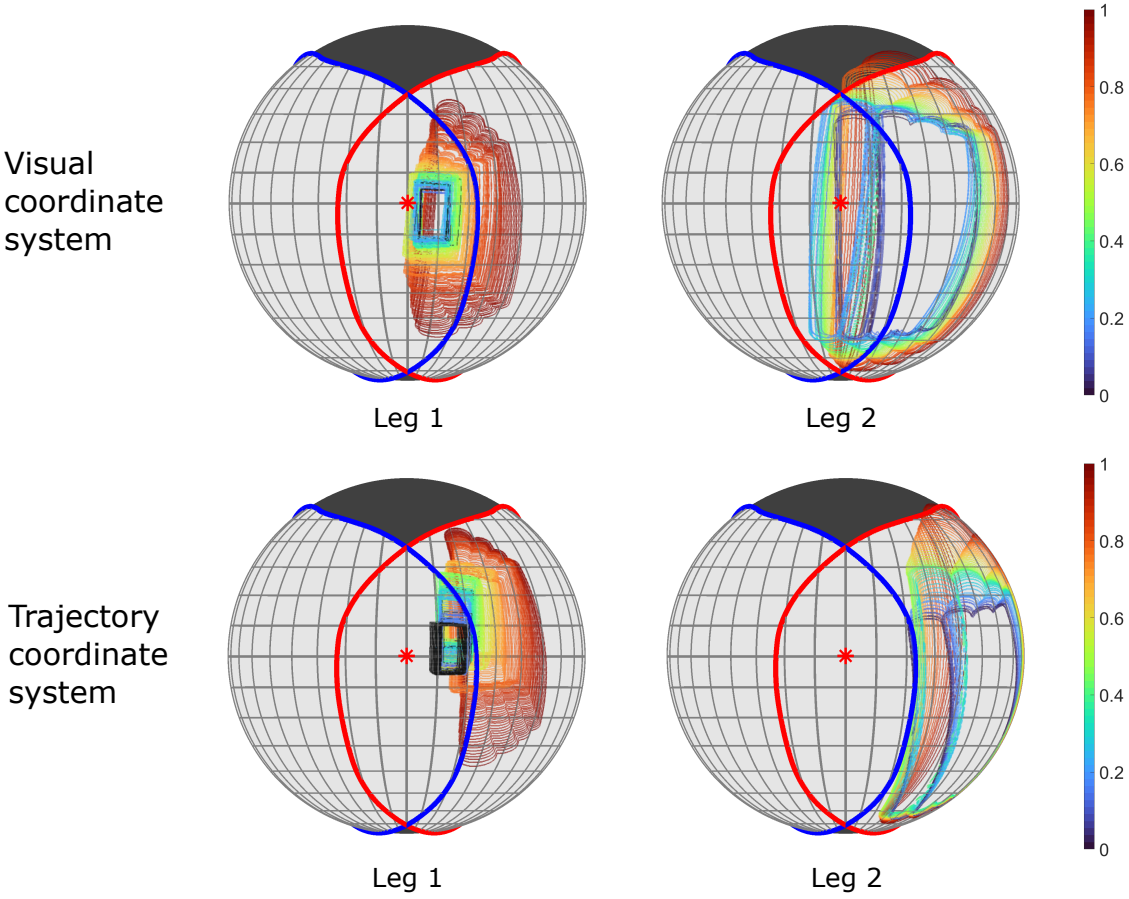


Figure 3.16: Trajectory of the obstacles in the visual field sphere. The contour of the obstacles is represented in the virtual camera's visual field per frame (coloured semi-transparent), for both obstacle avoidance flights, from takeoff until the perch is fully visible (i.e. until the second blue dash vertical line in Figure 3.17, included). The colormap indicates time through each flight. For the frames in which the camera's transform was interpolated (because not enough markers were reconstructed), the contour of the obstacles is shown in black. The retinal margins of the bird are shown for the left eye (blue line) and right eye (red line), and the virtual camera's axis is shown for reference (red star). Note that the virtual camera's axis corresponds to the estimated gaze direction $\vec{v}_{gaze}$ in the visual coordinate system (top row), and to the head velocity vector in the trajectory coordinate system (bottom row). In the visual coordinate system, the leftmost edge of the obstacles stays largely aligned with the estimated sagittal plane in both flights. In contrast, in the trajectory coordinate system, the obstacles are not stabilised in the vertical direction and they are not aligned with the head velocity vector. The visible area in the bird's visual field extends beyond what is shown in this orthographic projection.

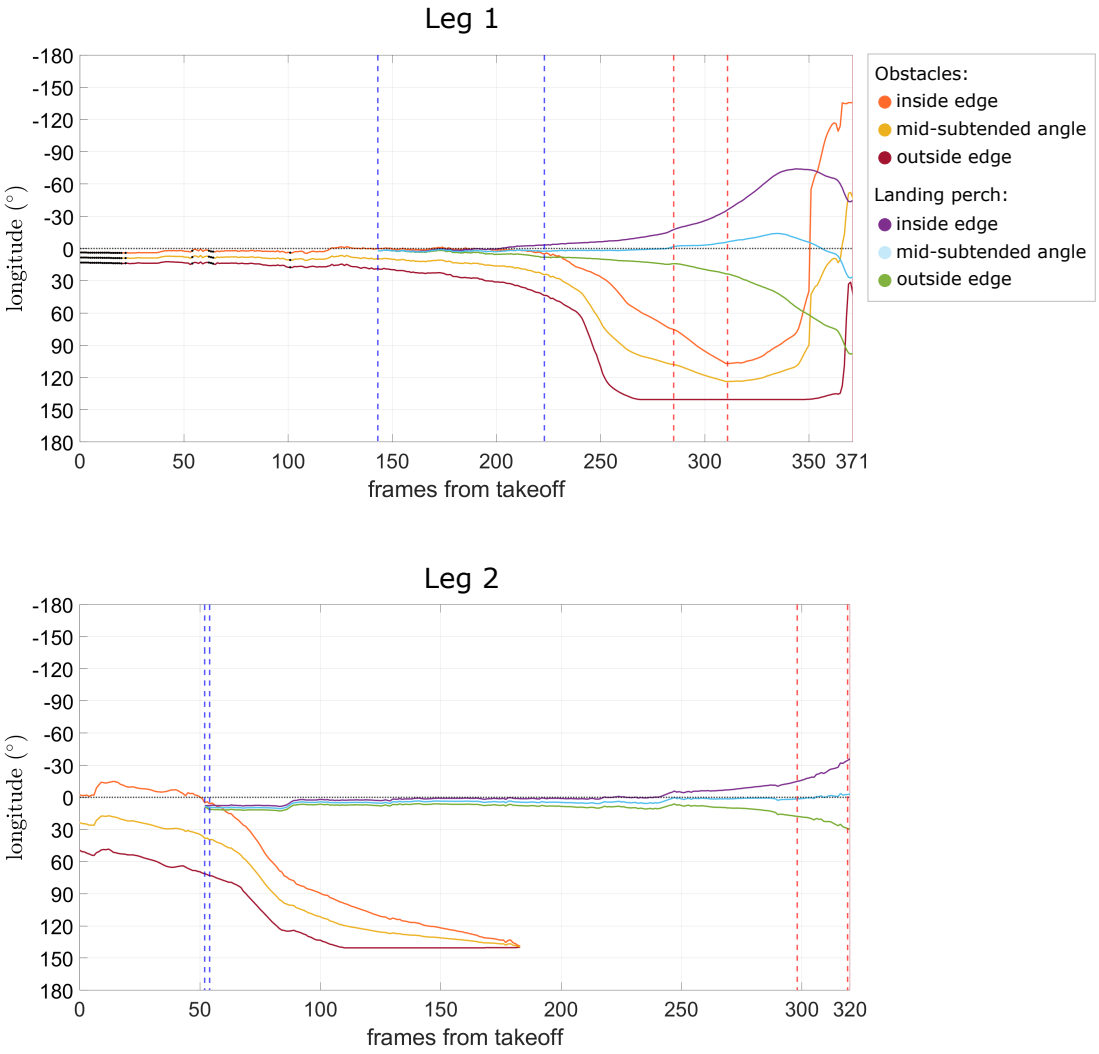


Figure 3.17: Extension in longitude of the obstacles set and the landing perch per frame, in the obstacle avoidance flights. The evolution in time of the obstacles' and the landing perch's extension in longitude is represented for the two obstacle avoidance flights, corresponding to leg 1 (top panel) and leg 2 (bottom panel) of the trial. The obstacles' inside and outside edges relative to the bird's turn are shown in orange and red, and the landing perch's in purple and green. For reference the midpoint of the angle subtended by the obstacles is shown in yellow, and in blue for the landing perch. Black markers denote frames in which the bird's head transform was interpolated. The range of frames when the landing perch per leg is partially occluded by the obstacles is marked between two vertical dashed lines. Note that because objects curve when they are close to the poles of the spherical virtual camera, it may be that the landing perch is fully visible but its longitudinal extension overlaps with the obstacles' one (e.g. around 225 frames from takeoff; see Supplementary videos `ESM_8.avi` and `ESM_9.avi`). The first and last frames used to estimate the bird's gaze direction $\vec{v}_{gaze}$ are marked with vertical red dashed lines. In leg 1, the obstacles appear within the central part of the visible field after 350 frames, but in reality the would be occluded by the bird's body (see Supplementary video `ESM_8.avi`). In leg 2, we can visually identify two potential head saccades: one 87 frames after takeofff, seems to align the estimated gaze direction with the left edge of the landing perch, as seen by the bird; the other one 242 frames after takeoff, seems to align the estimated gaze direction with the perch's centre (see Supplementary videos `ESM_12.avi` and `ESM_13.avi`). The sampling rate is 200 Hz.

with dense 3D maps of features with more complex geometry. In this way, we avoid the accumulated drift and noise typical of large 3D maps, whilst reducing the modelling effort for the most intricate shapes, and making the most of the accurate motion capture data. To facilitate the integration of the dense meshes within the geometric model of the environment, they are expressed in the same coordinate system as the motion capture data at the point of acquisition. We have demonstrated the applicability of this hybrid approach for a simple example of reconstructing a curtain that was hung to act as an obstacle to the bird. However, the method would be especially relevant for modelling natural-looking environments with more complicated features, such as small trees. As an illustration of the viability of this approach, Figure 3.18 shows the 3D mesh of a set of trees that we placed around the lab and captured using *SemanticPaint*, in this case using a Kinect v1 as the mapping sensor.

3.5.2 Behavioural analyses: pursuit flight

Our analysis of the synthetic reconstructions demonstrate the potential of using these to provide further insight into the bird's behaviour. During the pursuit flight, we found that the target is held within the area of the bird's binocular overlap for most of the flight. Modelling a counterfactual gaze strategy, in which the virtual camera's principal axis is aligned with the bird's head velocity, corroborates our conclusion that the bird actively directs its gaze to keep the target in this region of the visual field. Inspecting the evolution of the target's position in the visual field through the flight, we find that the bird fixates the target at $\pm 10^\circ$ longitude from the estimated gaze direction, and only centres it towards the terminal interception phase. In common with most other raptors, Harris' hawks have two points of high-resolution vision per retina: one projecting frontally and the other laterally (Mitkus et al., 2018; Potier et al., 2016). Whilst it would be premature to draw any firm conclusions from data for a single flight of a single bird, we hypothesise that these locations at $\pm 10^\circ$ longitude in the visual field may correspond to the projections of the frontal foveae of the bird's left and right eyes. These are estimated

to be located at between 9° and 16° longitude from the forward direction of the head in most raptors (Frost et al., 1990; Kane et al., 2014; Tucker, 2000), and so unlike in humans, they may not project to a single point in the visual field—even when the eyes converge (Tyrrell et al., 2018). This being so, our results could indicate that the bird tracks the target with one or other of its frontal foveae throughout the flight, before centering it in the visual field prior to interception and producing a symmetric looming stimulus. It is also interesting that when the bird loses track of its target as it makes a sharp turn at the curtain, it appears to reacquire it with the other fovea.

3.5.3 Behavioural analyses: obstacle avoidance

In both obstacle avoidance flights that we analysed, we found that the inside edge of the obstacles was aligned with the longitudinal centre of the visual field for substantial portions of the flights in which the obstacles were visible. Again this was not observed in the counterfactual gaze strategy that we considered, in which the virtual camera's principal axis was aligned with the bird's head velocity. This is in accordance with previous findings from other birds (Kress et al., 2015) and humans (Raudies et al., 2012; Rothkopf et al., 2009), which both seem to fixate on the edges of objects than can be perceived as obstacles, and on the centre of objects perceived as goals. Moreover, in the flight corresponding to the second leg of the trial, the bird seemed to align its visual field first with the edge of the landing perch and then with its centre (see Supplementary videos `ESM_12.avi` and `ESM_13.avi`, around frames 2029 and 2184 as numbered in the video). This is also in line with previous reports in lovebirds (Kress et al., 2015) and may reflect a strategy based on aiming at intermediate goals. The agreement with previous works indicates that our approach is reasonable. However, as with the pursuit flight, these hypotheses can only be confirmed or rejected by analysing the complete set of flights across different individuals.

3.5.4 Limitations and future work

We aimed to develop a reconstruction method that would enable the collection of large amounts of data from many individuals. Some key steps that we have taken towards this goal include our non-invasive tracking of head pose using marker-based motion capture, our programmatic definition of geometric primitives based on motion capture markers placed on objects in the environment, and our automated integration of the dense 3D maps on the motion capture coordinate system. The key limitation in our current pipeline is the need to calibrate the visual coordinate system against the coordinate system of the headpack, which we achieve using two alternative methods sharing common assumptions. A second potential limitation of the reconstruction method is its representation of a binocular system using a monocular visual field.

One way of overcoming both limitations would be to combine data from calibrated stereo cameras with the motion capture data. This could allow the pose of the head to be calibrated directly within the headpack coordinate system, also allowing measurement of the stereo baseline of the eyes. The potential for exploiting this approach is demonstrated by recent work by Naik (2021), who developed an annotation tool to obtain the 3D coordinates of identifiable features on a bird, such as its bill and eyes, with respect to its head. This will be a key focus of our future work, as it will allow us to streamline the process and obtain more accurate estimations of the visual coordinate system. Other potential improvements could involve automating the correction of the dense 3D maps, potentially using the reference motion capture markers in the mapped objects as fiducials. An improved ArUco plate would also be an easy way to improve floor alignment. Exploring the possibilities for tracking eye movements would also be of interest in providing better estimates of the range of the motions, or defining a more accurate visual coordinate system (Holmgren et al., 2021).

There is still much to learn about how animals interact with their environment, and innovative methods such as the one we propose open a wide range of possibilities for behavioural analysis. Understanding the role of active vision in bird flight may

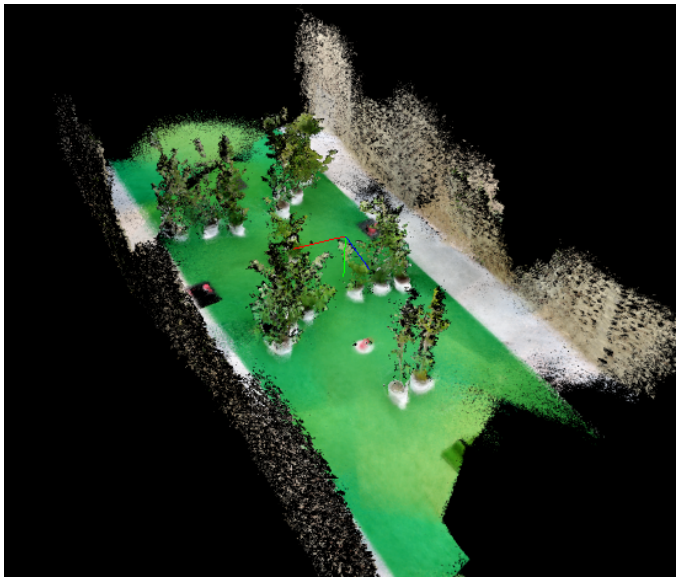


Figure 3.18: 3D model of a forest environment in the lab. We used a mix of 20 small trees (laurel and bay trees) of under 2 m height to recreate a forest environment in the motion capture lab. The 3D model was captured using *SemanticPaint* and a Kinect v1 as a mapping sensor. Note that in this case the world coordinate system of the captured map is not aligned with the motion capture coordinate system.

reveal efficient processing strategies that could be translated to autonomous systems. Human-inspired active vision has already been explored for robotic applications (Seara et al., 2004; Seara et al., 2002; Seara et al., 2001), as active observers have been shown to solve basic vision problems more efficiently than passive ones, with many ill-posed problems becoming well-posed (Aloimonos et al., 1988). Visual-guidance algorithms could also be tested directly using the data generated with our approach, similar to the method followed when studying navigation in homing ants by Ardin et al. (2015). The generated synthetic data could also be useful as naturalistic visual stimuli in studies on motion sensitive neurons in birds, similarly to the work by (Eckmeier et al., 2013). Finally, large datasets collected with this or similar methods could be combined with inverse reinforcement learning frameworks, leading to data-driven models offering new insights into the bird's gaze strategy in flight; similar approaches already exist that make use of human motion capture data to generate active-sensing behaviours in synthetic humanoids (Merel et al., 2020). In conclusion, we see many exciting opportunities for mutual collaboration between the animal behaviour and computer vision communities in all of these areas.

Table 3.1: Description of supplementary videos. We provide supplementary videos for the rendered output per trial. The description field in the table specifies the output data represented in the video, the coordinate system the virtual camera is tracking in the video and the projection used.

Trial		Filename	Description (output, coordinate system tracked, projection)
Pursuit		ESM_1.avi	RGB, visual coordinate system, orthographic
		ESM_2.avi	RGB, visual coordinate system, equirectangular
		ESM_3.avi	Depth, visual coordinate system, orthographic
		ESM_4.avi	Semantic, visual coordinate system, orthographic
		ESM_5.avi	Optic flow, visual coordinate system, orthographic
		ESM_6.avi	RGB, trajectory coordinate system, orthographic
		ESM_7.avi	RGB, trajectory coordinate system, equirectangular
Obstacle avoidance	Leg 1	ESM_8.avi	RGB, visual coordinate system, orthographic
		ESM_9.avi	RGB, visual coordinate system, equirectangular
		ESM_10.avi	RGB, trajectory coordinate system, orthographic
		ESM_11.avi	RGB, trajectory coordinate system, equirectangular
	Leg 2	ESM_12.avi	RGB, visual coordinate system, orthographic
		ESM_13.avi	RGB, visual coordinate system, equirectangular
		ESM_14.avi	RGB, trajectory coordinate system, orthographic
		ESM_15.avi	RGB, trajectory coordinate system, equirectangular

3.6 Supplementary videos

We provide the rendered outputs as video supplementary material. For the pursuit flight, we include videos for the RGB, semantic, depth and optic flow synthetic data generated. For the obstacle avoidance flights, we include RGB videos. All videos are reproduced at 20Hz, (1/10 of the real speed), except the optic flow video which is reproduced at 5 Hz (1/40 of the real speed). The frame numbering shown in the videos follows the motion capture data system's numbering. A description of each of the video files included is presented in Table 3.1.

The RGB outputs are presented using two projections: equirectangular, in which the geometry appears distorted, but which shows the complete field of view of the bird; and orthographic, in which the distortion is reduced but doesn't include the most peripheral regions of the bird's visual field. The rest of the rendered outputs are only represented in orthographic projection. Note that for the orthographic case, the point of view is as if looking frontally to the bird's visual field sphere

(see Figure B.12). In all videos, a red contour around the figure indicates the head transform for that frame was interpolated. The retinal margins and the blind areas of the bird's visual field are overlaid on the rendered output.

In the optic flow video, the colormap represents for each frame the instantaneous angular speed per pixel, in degrees per second. The colourbar is in logarithmic scale and capped at $10°/s$ in the lower bound and $1000°/s$ in the upper bound. The vector field results from the transformation of the output data from pixel space to the surface of the unit sphere. Further details on the computation of the videos are included in Appendix B.5.

References

Aloimonos, J., Weiss, I., and Bandyopadhyay, A. (1988). "Active vision". In: *Int. J. Comput. Vis.* 1.4, pp. 333–356. DOI: 10.1007/BF00133571.

Altshuler, D. L. and Srinivasan, M. V. (2018). "Comparison of visually guided flight in insects and birds". In: *Front. Neurosci.* 12, p. 157. DOI: 10.3389/fnins.2018.00157.

Ardin, P., Mangan, M., Wystrach, A., and Webb, B. (2015). "How variation in head pitch could affect image matching algorithms for ant navigation". In: *J. Comp. Physiol. A* 201.6, pp. 585–597. DOI: 10.1007/s00359-015-1005-8.

Ardin, P., Peng, F., Mangan, M., Lagogiannis, K., and Webb, B. (2016). "Using an insect mushroom body circuit to encode route memory in complex natural environments". In: *PLoS Comput. Biol.* 12.2. DOI: 10.1371/journal.pcbi.1004683.

Baird, E., Boeddeker, N., and Srinivasan, M. V. (2021). "The effect of optic flow cues on honeybee flight control in wind". In: *Proc. R. Soc. B Biol. Sci.* 288.1943. DOI: 10.1098/rspb.2020.3051.

Bhagavatula, P. S., Claudianos, C., Ibbotson, M. R., and Srinivasan, M. V. (2011). "Optic flow cues guide flight in birds". In: *Curr. Biol.* 21.21, pp. 1794–1799. DOI: 10.1016/j.cub.2011.09.009.

Bian, X., Chandler, T., Laird, W., Pinilla, A., and Peters, R. (2018). "Integrating evolutionary biology with digital arts to quantify ecological constraints on vision-based behaviour". In: *Methods Ecol. Evol.* 9.3, pp. 544–559. DOI: 10.1111/2041-210X.12912.

Bian, X., Chandler, T., Pinilla, A., and Peters, R. A. (2019). "Now you see me, now you don't: Environmental conditions, signaler behavior, and receiver response thresholds interact to determine the efficacy of a movement-based animal signal". In: *Front. Ecol. Evol.* 7.APR, pp. 1–16. DOI: 10.3389/fevo.2019.00130.

Bian, X., Pinilla, A., Chandler, T., and Peters, R. (2021). "Simulations with Australian dragon lizards suggest movement-based signal effectiveness is dependent on display structure and environmental conditions". In: *Sci. Rep.* 11.1, pp. 1–11. DOI: 10.1038/s41598-021-85793-3.

Blender Online Community (2021). *Blender - a 3D modelling and rendering package*. Stichting Blender Foundation, Amsterdam.

Brighton, C. H. and Taylor, G. K. (2019). "Hawks steer attacks using a guidance system tuned for close pursuit of erratically manoeuvring targets". In: *Nat. Commun.* 10.1, pp. 1–28. DOI: 10.1038/s41467-019-10454-z.

Brighton, C. H., Thomas, A. L., and Taylor, G. K. (2017). "Terminal attack trajectories of peregrine falcons are described by the proportional navigation guidance law of missiles". In: *Proc. Natl. Acad. Sci.* 114.51, p. 201714532. DOI: 10.1073/pnas.1714532114.

Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., and Ranzuglia, G. (2008). "MeshLab: an Open-Source Mesh Processing Tool". In: *Eurographics Ital. Chapter Conf.* Ed. by V. Scarano, R. D. Chiara, and U. Erra. The Eurographics Association. DOI: 10.2312/LocalChapterEvents/ItalChap/ItalianChapConf2008/129-136.

Dakin, R., Fellows, T. K., and Altshuler, D. L. (2016). "Visual guidance of forward flight in hummingbirds reveals control based on image features instead of pattern velocity". In: *Proc. Natl. Acad. Sci.* 113.31, pp. 8849–8854. DOI: 10.1073/pnas.1603221113.

Eckmeier, D., Geurten, B. R., Kress, D., Mertes, M., Kern, R., Egelhaaf, M., and Bischof, H. J. (2008). "Gaze strategy in the free flying zebra finch (Taeniopygia guttata)". In: *PLoS One* 3.12. DOI: 10.1371/journal.pone.0003956.

Eckmeier, D., Kern, R., Egelhaaf, M., and Bischof, H. J. (2013). "Encoding of naturalistic optic flow by motion sensitive neurons of nucleus rotundus in the zebra finch (Taeniopygia guttata)". In: *Front. Integr. Neurosci.* 7.SEP, pp. 1–17. DOI: 10.3389/fnint.2013.00068.

Fair, J., Paul, E., and Jones, J. (2010). *Guidelines to the use of wild birds in research.* Tech. rep. August, pp. 48–92.

Frost, B. J., Wise, L. Z., Morgan, B., and Bird, D. (1990). "Retinotopic representation of the bifoveate eye of the kestrel (Falco sparverius) on the optic tectum". In: *Vis. Neurosci.* 5.3, pp. 231–239. DOI: 10.1017/S0952523800000304.

Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. J., and Marín-Jiménez, M. J. (2014). "Automatic generation and detection of highly reliable fiducial markers under occlusion". In: *Pattern Recognit.* 47.6, pp. 2280–2292. DOI: 10.1016/j.patcog.2014.01.005.

Gollob, C., Ritter, T., Kraßnitzer, R., Tockner, A., and Nothdurft, A. (2021). "Measurement of forest inventory parameters with apple ipad pro and integrated lidar technology". In: *Remote Sens.* 13.16, pp. 1–35. DOI: 10.3390/rs13163129.

Golodetz, S., Cavallari, T., Lord, N. A., Prisacariu, V. A., Murray, D. W., and Torr, P. H. S. (2018). "Collaborative Large-Scale Dense 3D Reconstruction with Online Inter-Agent Pose Optimisation". In: *IEEE Trans. Vis. Comput. Graph.* 24.11, pp. 2895–2905. DOI: 10.1109/TVCG.2018.2868533.

Golodetz, S., Sapienza, M., Valentin, J. P. C., Vineet, V., Cheng, M.-M., Arnab, A., Prisacariu, V. A., Kähler, O., Ren, C. Y., Murray, D. W., Izadi, S., and Torr, P. H. S. (2015). "SemanticPaint: a framework for the interactive segmentation of 3D scenes". In: *arXiv Prepr.*, pp. 1–33. DOI: 10.1145/2751556.

Haalck, L., Mangan, M., Webb, B., and Risse, B. (2020). "Towards image-based animal tracking in natural environments using a freely moving camera". In: *J. Neurosci. Methods* 330.108455. DOI: 10.1016/j.jneumeth.2019.108455.

Hazan, Y., Kra, Y., Yarin, I., Wagner, H., and Gutfreund, Y. (2015). "Visual-auditory integration for visual search: a behavioral study in barn owls". In: *Front. Integr. Neurosci.* 9, pp. 1–12. DOI: 10.3389/fnint.2015.00011.

Holmgren, C. D., Stahr, P., Wallace, D. J., Voit, K.-M., Matheson, E. J., Sawinski, J., Bassetto, G., and Kerr, J. N. D. (2021). "Visual pursuit behavior in mice maintains the pursued prey on the retinal region with least optic flow". In: *Elife* 10, pp. 1–34. DOI: 10.7554/eLife.70838.

Kane, S. A., Fulton, A. H., and Rosenthal, L. J. (2015). "When hawks attack: animal-borne video studies of goshawk pursuit and prey-evasion strategies". In: *J. Exp. Biol.* 218.2, pp. 212–222. DOI: 10.1242/jeb.108597.

Kane, S. A. and Zamani, M. (2014). "Falcons pursue prey using visual motion cues: new perspectives from animal-borne cameras". In: *J. Exp. Biol.* 217.2, pp. 225–234. DOI: 10.1242/jeb.092403.

Kano, F., Walker, J., Sasaki, T., and Biro, D. (2018). "Head-mounted sensors reveal visual attention of free-flying homing pigeons". In: *J. Exp. Biol.* 221.17, pp. 1–13. DOI: 10.1242/jeb.183475.

Kern, R., Van Hateren, J. H., Michaelis, C., Lindemann, J. P., and Egelhaaf, M. (2005). "Function of a fly motion-sensitive neuron matches eye movements during free flight". In: *PLoS Biol.* 3.6, pp. 1130–1138. DOI: 10.1371/journal.pbio.0030171.

Kress, D., Van Bokhorst, E., and Lentink, D. (2015). "How lovebirds maneuver rapidly using super-fast head saccades and image feature stabilization". In: *PLoS One* 10.6, pp. 1–24. DOI: 10.1371/journal.pone.0129287.

Land, M. F. (1999). "Motion and vision: why animals move their eyes". In: *J. Comp. Physiol. - A Sensory, Neural, Behav. Physiol.* 185.4, pp. 341–352. DOI: 10.1007/s003590050393.

Lev-Ari, T. and Gutfreund, Y. (2018). "Interactions between top-down and bottom-up attention in barn owls (Tyto alba)". In: *Anim. Cogn.* 21.2, pp. 197–205. DOI: 10.1007/s10071-017-1150-2.

Lin, H. T., Ros, I. G., and Biewener, A. A. (2014). "Through the eyes of a bird: modelling visually guided obstacle flight". In: *J. R. Soc. Interface* 11.96, pp. 1–12. DOI: 10.1098/rsif.2014.0239.

McClune, D. W. (2018). "Joining the dots: reconstructing 3D environments and movement paths using animal-borne devices". In: *Anim. Biotelemetry* 6, p. 5. DOI: 10.1186/s40317-018-0150-6.

Merel, J., Tunyasuvunakool, S., Ahuja, A., Tassa, Y., Hasenclever, L., Pham, V. U., Erez, T., Wayne, G., and Heess, N. (2020). "Catch & carry: reusable neural controllers for vision-guided whole-body tasks". In: *ACM Trans. Graph.* 39.4, pp. 1–14. DOI: 10.1145/3386569.3392474.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., and Ng, R. (2020). "NeRF: representing scenes as neural radiance fields for view synthesis". In: *ECCV*. Vol. 12346 LNCS, pp. 405–421. DOI: 10.1007/978-3-030-58452-8_24.

Minano, S. and Taylor, G. K. (2021). "Through hawks' eyes: reconstructing a bird's visual field in flight to study gaze strategy and attention during perching and obstacle avoidance". In: *bioRxiv*. DOI: 10.1101/2021.06.16.446415.

Mitkus, M., Potier, S., Martin, G. R., Duriez, O., and Kelber, A. (2018). "Raptor vision". In: *Oxford Res. Encycl. Neurosci.* March. Oxford University Press, pp. 1–38. DOI: 10.1093/acrefore/9780190264086.013.232.

Naik, H. (2021). "XR for all: closed-loop visual stimulation techniques for human and non-human animals". PhD thesis. Munich: Technische Universität München.

Naik, H., Bastien, R., Navab, N., and Couzin, I. D. (2020). "Animals in virtual environments". In: *IEEE Trans. Vis. Comput. Graph.* 26.5, pp. 2073–2083. DOI: 10.1109/TVCG.2020.2973063.

Neumann, T. R. (2002). "Modeling insect compound eyes: space-variant spherical vision". In: *Proc. 2nd Int. Work. Biol. Motiv. Comput. Vis. (BMCV 2002)* LNCS 2525, pp. 360–367. DOI: 10.1007/3-540-36181-2_36.

Ochs, M. F., Zamani, M., Gomes, G. M. R., de Oliveira Neto, R. C., and Kane, S. A. (2016). "Sneak peek: Raptors search for prey using stochastic head turns". In: *Auk* 134.1, pp. 104–115. DOI: 10.1642/auk-15-230.1.

Ozawa, Y. (2010). "Vision and movement in birds". PhD thesis. Oxford: University of Oxford.

Potier, S., Bonadonna, F., Kelber, A., Martin, G. R., Isard, P.-F., Dulaurent, T., and Duriez, O. (2016). "Visual abilities in two raptors with different ecology". In: *J. Exp. Biol.* 219.17, pp. 2639–2649. DOI: 10.1242/jeb.142083.

Potier, S., Lieuvin, M., Pfaff, M., and Kelber, A. (2020). "How fast can raptors see?" In: *J. Exp. Biol.* 223.1. DOI: 10.1242/jeb.209031.

Prisacariu, V. A., Kähler, O., Golodetz, S., Sapienza, M., Cavallari, T., Torr, P. H. S., and Murray, D. W. (2017). "InfiniTAM v3: a framework for large-scale 3D reconstruction with loop closure". In: *arXiv Prepr.* arXiv: 1708.00783.

Raudies, F., Mingolla, E., and Neumann, H. (2012). "Active gaze control improves optic flow-based segmentation and steering". In: *PLoS One* 7.6. DOI: 10.1371/journal.pone.0038446.

Ravi, S., Bertrand, O., Siesenop, T., Manz, L. S., Doussot, C., Fisher, A., and Egelhaaf, M. (2019). "Gap perception in bumblebees". In: *J. Exp. Biol.* 222.2, pp. 1–10. DOI: 10.1242/jeb.184135.

Ravi, S., Siesenop, T., Bertrand, O. J., Li, L., Doussot, C., Fisher, A., Warren, W. H., and Egelhaaf, M. (2022). "Bumblebees display characteristics of active vision during robust obstacle avoidance flight". In: *J. Exp. Biol.* 225.4. DOI: 10.1242/jeb.243021.

Risse, B., Mangan, M., Stürzl, W., and Webb, B. (2018). "Software to convert terrestrial LiDAR scans of natural environments into photorealistic meshes". In: *Environ. Model. Softw.* 99, pp. 88–100. DOI: 10.1016/j.envsoft.2017.09.018.

Romero-Ramirez, F. J., Muñoz-Salinas, R., and Medina-Carnicer, R. (2018). "Speeded up detection of squared fiducial markers". In: *Image Vis. Comput.* 76, pp. 38–47. DOI: 10.1016/j.imavis.2018.05.004.

Ros, I. G. and Biewener, A. A. (2016). "Optic flow stabilizes flight in ruby-throated hummingbirds". In: *J. Exp. Biol.* 219.16, pp. 2443–2448. DOI: 10.1242/jeb.128488.

Ros, I. G. and Biewener, A. A. (2017). "Pigeons (C. livia) follow their head during turning flight: Head stabilization underlies the visual control of flight". In: *Front. Neurosci.* 11.DEC, pp. 1–12. DOI: 10.3389/fnins.2017.00655.

Rothkopf, C. A. and Ballard, D. H. (2009). "Image statistics at the point of gaze during human navigation". In: *Vis. Neurosci.* 26.1, pp. 81–92. DOI: 10.1017/S0952523808080978.

Ruppertsberg, A. I., Wuerger, S. M., and Bertamini, M. (2003). "The chromatic input to global motion perception". In: *Vis. Neurosci.* 20.4, pp. 421–428. DOI: 10.1017/S0952523803204077.

Schiffner, I. and Srinivasan, M. V. (2015). "Direct evidence for vision-based control of flight speed in budgerigars". In: *Sci. Rep.* 5, pp. 1–7. DOI: 10.1038/srep10992.

Schonberger, J. L. and Frahm, J. M. (2016). "Structure-from-Motion revisited". In: *2016 IEEE Conf. Comput. Vis. Pattern Recognit.* Pp. 4104–4113. DOI: 10.1109/CVPR.2016.445.

Schulte, P., Zeil, J., and Stürzl, W. (2019). "An insect-inspired model for acquiring views for homing". In: *Biol. Cybern.* 113.4, pp. 439–451. DOI: 10.1007/s00422-019-00800-1.

Seara, J. F., Lorch, O., and Schmidt., G. (2001). "Gaze Control for Goal-Oriented Humanoid Walking". In: *Proc. IEEE/RAS Int. Conf. Humanoid Robot. (Humanoids), S. 187–195, Tokio, Japan.*

Seara, J. F. and Schmidt, G. (2004). "Intelligent gaze control for vision-guided humanoid walking: Methodological aspects". In: *Rob. Auton. Syst.* 48.4 SPEC. ISS. Pp. 231–248. DOI: 10.1016/j.robot.2004.07.003.

Seara, J. F., Strobl, K. H., and Schmidt, G. (2002). "Information management for gaze control in vision guided biped walking". In: *IEEE Int. Conf. Intell. Robot. Syst.* 1.October, pp. 31–36. DOI: 10.1109/IRDS.2002.1041357.

Stowers, J. R., Hofbauer, M., Bastien, R., Griessner, J., Higgins, P., Farooqui, S., Fischer, R. M., Nowikovsky, K., Haubensak, W., Couzin, I. D., Tessmar-Raible, K., and Straw, A. D. (2017). "Virtual reality for freely moving animals". In: *Nat. Methods* 14.10, pp. 995–1002. DOI: 10.1038/nmeth.4399.

Stuerzl, W., Boeddeker, N., Dittmar, L., and Egelhaaf, M. (2010). "Mimicking honeybee eyes with a 280 deg field of view catadioptric imaging system". In: *Bioinspiration and Biomimetics* 5.3. DOI: 10.1088/1748-3182/5/3/036002.

Stuerzl, W., Zeil, J., Boeddeker, N., and Hemmi, J. M. (2016). "How wasps acquire and use views for homing". In: *Curr. Biol.* 26, pp. 470–482. DOI: 10.1016/j.cub.2015.12.052.

Stürzl, W., Grixa, I., Mair, E., Narendra, A., and Zeil, J. (2015). "Three-dimensional models of natural environments and the mapping of navigational information". In: *J. Comp. Physiol. A Neuroethol. Sensory, Neural, Behav. Physiol.* 201.6, pp. 563–584. DOI: 10.1007/s00359-015-1002-y.

Tammero, L. F. and Dickinson, M. H. (2002a). "Collision-avoidance and landing responses are mediated by separate pathways in the fruit fly, Drosophila melanogaster". In: *J. Exp. Biol.* 205.18, pp. 2785–2798. DOI: 10.1242/jeb.205.18.2785.

Tammero, L. F. and Dickinson, M. H. (2002b). "The influence of visual landscape on the free flight behavior of the fruit fly Drosophila melanogaster". In: *J. Exp. Biol.* 205.3, pp. 327–343. DOI: 10.1007/bf00355543.

Tancik, M., Casser, V., Yan, X., Pradhan, S., Mildenhall, B., Srinivasan, P., Barron, J. T., and Kretzschmar, H. (2022). "Block-NeRF: Scalable Large Scene Neural View Synthesis". In: *arXiv*. arXiv: 2202.05263.

Tatsumi, S., Yamaguchi, K., and Furuya, N. (2022). "ForestScanner: A mobile application for measuring and mapping trees with LiDAR-equipped iPhone and iPad". In: *bioRxiv*, pp. 95–100. DOI: 10.1101/2021.12.11.472207.

Taylor, G. K., Bacic, M., Bomphrey, R. J., Carruthers, A. C., Gillies, J., Walker, S. M., and Thomas, A. L. R. (2008). "New experimental approaches to the biology of flight control systems". In: *J. Exp. Biol.* 211.2, pp. 258–266. DOI: 10.1242/jeb.012625.

Taylor, L. A., Taylor, G. K., Lambert, B., Walker, J. A., Biro, D., and Portugal, S. J. (2019). "Birds invest wingbeats to keep a steady head and reap the ultimate benefits

of flying together". In: *PLoS Biol.* 17.6, pp. 1–20. DOI: 10.1371/journal.pbio.3000299.

Tucker, V. A. (2000). "The deep fovea, sideways vision and spiral flight paths in raptors". In: *J. Exp. Biol.* 203.Pt 24, pp. 3745–3754. DOI: 10.1016/S0169-5347(01)02393-X.

Tuia, D., Kellenberger, B., Beery, S., Costelloe, B. R., Zuffi, S., Risse, B., Mathis, A., Mathis, M. W., van Langevelde, F., Burghardt, T., Kays, R., Klinck, H., Wikelski, M., Couzin, I. D., van Horn, G., Crofoot, M. C., Stewart, C. V., Berger-Wolf, T., et al. (2022). "Perspectives in machine learning for wildlife conservation". In: *Nat. Commun.* 13.1, pp. 1–15. DOI: 10.1038/s41467-022-27980-y.

Tyrrell, L. P., Goller, B., Moore, B. A., Altshuler, D. L., and Fernández-Juricic, E. (2018). "The orientation of visual space from the perspective of hummingbirds". In: *Front. Neurosci.* 12.16, pp. 1–9. DOI: 10.3389/fnins.2018.00016.

Warrick, D. R., Bundle, M. W., and Dial, K. P. (2002). "Bird maneuvering flight: blurred bodies, clear heads". In: *Integr. Comp. Biol.* 42.1, pp. 141–148. DOI: 10.1093/icb/42.1.141.

Windsor, S. P. and Taylor, G. K. (2017). "Head movements quadruple the range of speeds encoded by the insect motion vision system in hawkmoths". In: *Proc. R. Soc. B Biol. Sci.* 284.1864. DOI: 10.1098/rspb.2017.1622.

Yorzinski, J. L., Patricelli, G. L., Babcock, J. S., Pearson, J. M., and Platt, M. L. (2013). "Through their eyes: selective attention in peahens during courtship". In: *J. Exp. Biol.* 216.16, pp. 3035–3046. DOI: 10.1242/jeb.087338.

Yorzinski, J. L. and Platt, M. L. (2014). "Selective attention in peacocks during predator detection". In: *Anim. Cogn.* 17.3, pp. 767–777. DOI: 10.1007/s10071-013-0708-x.

Yorzinski, J. L. (2021). "Great-tailed grackles can independently direct their eyes toward different targets". In: *Exp. Brain Res.* 239.7, pp. 2119–2126. DOI: 10.1007/s00221-021-06122-8.

Yorzinski, J. L., Patricelli, G. L., Platt, M. L., and Land, M. F. (2015). "Eye and head movements shape gaze shifts in Indian peafowl". In: *J. Exp. Biol.* 218.23, pp. 3771–3776. DOI: 10.1242/jeb.129544.

Zeil, J., Narendra, A., and Stürzl, W. (2014). "Looking and homing: how displaced ants decide where to go". In: *Philos. Trans. R. Soc. B Biol. Sci.* 369.1636. DOI: 10.1098/rstb.2013.0034.

Zhu, L., Mangan, M., and Webb, B. (2021). "Spatio-temporal memory for navigation in a mushroom body model". In: *Conf. Biomim. Biohybrid Syst.* Ed. by V. Vouloutsi, A. Mura, F. Tauber, T. Speck, T. J. Prescott, and P. F. M. J. Verschure. Vol. 12413 LNAI. 201808060165. Berlin, Heidelberg: Springer-Verlag, pp. 415–426. DOI: 10.1007/978-3-030-64313-3_39.

4

Analysis of Harris' hawks' gaze strategies in perching and obstacle avoidance

Contents

4.1 Abstract 120
4.2 Introduction 122
4.3 Methods 124
4.3.1 Motion capture experiments 125
4.3.2 Visual coordinate system 128
4.3.3 Computational model in Blender 136
4.4 Results 136
4.4.1 Stabilisation of the landing perch and the obstacles in the visual field 137
4.4.2 Longitude angle subtended by the landing perch and the obstacles along the trajectory 141
4.4.3 Gaze strategy in perching 148
4.4.4 Gaze strategy around obstacles 157
4.5 Discussion 164
References 171

4.1 Abstract

In animals with well-developed visual systems, vision is an active process: in goal-directed behaviours, gaze is actively directed towards the information that will enable

a successful execution of the task. Consequently, gaze motion patterns provide information about the cognitive processes that take place during the execution of these tasks, and can reveal important aspects about the information required, as well as how and when it is acquired. Here, we investigate the gaze strategy used by Harris' hawks when executing two flight manoeuvres key to their survival: perching and obstacle avoidance. In birds, eye movements relative to the head are limited, and head orientation is frequently used as a proxy for gaze direction. We tracked the head movements of four hawks in flight with a high-speed motion capture system. We used the data to model the birds' visual fields in flight in a computational environment, as virtual cameras moving in a 3D model of the lab. Using a rendering engine, we generated semantic descriptions of the birds' visual input at every timestep along each flight. We used a subset of the generated synthetic data to characterise the birds' gaze strategy around the landing perch and the obstacles. We found that Harris' hawks use distinct gaze strategies for these two elements: during perching manoeuvres, the landing perch is centred in the visual field for a substantial period; in an obstacle avoidance flight, the obstacle's edge is aligned frontally with the bird's sagittal plane. This suggests that the birds adapt their gaze strategy around goals (or elements to intercept), and obstacles (or elements with which to avoid colliding), and is in line with previous reports of similar behaviour in humans and other birds. However in relation to the landing perch, we found that a centring gaze strategy was not always sustained throughout the flight. Our results suggest that the birds spend more time aligned with the perch's edge, and less time aligned with the perch's centre, when the trajectory involves more turning. This behaviour has previously been reported in lovebirds while turning, and may reflect a steering strategy based on fixating intermediate targets. In relation to the obstacles, we estimated that the birds stopped fixating the edge of the obstacles frontally in their visual fields when they were 1.5 m ahead of them on average, and 0.3 s before passing around them. This study allowed us to explore different methods of analysis, and define a set of hypotheses to verify with the full dataset in the future. Understanding how birds actively direct their visual input in

flight may not only reveal important aspects about their sensory-motor decisions in flight, but also inspire new vision-based strategies for autonomous flying systems.

4.2 Introduction

In animals with well-developed visual systems, vision is in large part an active process: through a variety of eye, head and body movements, animals are able to manipulate the position and orientation of their gaze relative to their environment (Land et al., 2012). In visually-guided tasks, top-down (task driven) mechanisms of gaze control typically dominate over bottom-up (saliency driven) ones (Land, 2006; Mole et al., 2021). As a result, gaze motion patterns can be very informative about the cognitive processing the animals are engaged in (Rothkopf et al., 2007), and provide valuable insight into what information is required to successfully execute a task, how it is acquired or when (Hayhoe, 2017).

For many animals, gaze stabilisation and control are vital in guiding their locomotion (Land et al., 2012). In humans, active gaze control in locomotion has been widely studied, from walking (Matthis et al., 2022; Matthis et al., 2018; Pelz et al., 2007), to steering while driving (Land et al., 1994; Mole et al., 2021; Pekkanen et al., 2017; Lappi, 2022) or cycling (Wilkie et al., 2008). Humans have been reported to direct their gaze with a constant lookahead of 1.5 to 2 s when walking over terrains of varying complexity (Matthis et al., 2018), and while steering cars (Land et al., 1994; Mole et al., 2021) and bicycles (Wilkie et al., 2008). When steering while driving, they have been shown to fixate their gaze on the inside edge of the road, particularly on the stationary point of the curve (Land et al., 1994). When walking, they have been reported to more frequently target the centre of objects they aim to approach, and the edge of objects they mean to avoid (Rothkopf et al., 2007; Rothkopf et al., 2009; Raudies et al., 2012). However, animals have a wide range of visual systems (Martin, 2007) and modes of locomotion, and therefore gaze strategies could be dramatically different across species (Yorzinski, 2019).

Very few studies have investigated how non-human animals use gaze to guide their locomotion (Yorzinski, 2019; Billington et al., 2020), presumably due to challenges such as experimentally reproducing realistic visual environments, or adapting eye-tracking equipment to a variety of species (Billington et al., 2020). Regarding terrestrial locomotion, a set of studies investigated the acquisition of visual information in cats relative to their step cycle (Volgushev et al., 2022; Zubair et al., 2019; Rivers et al., 2014). An eye-tracking setup was used with peahens in Yorzinski (2019) to study visual attention while jumping on and off a perch. In aerial locomotion, the role of active vision in flying insects has been the most studied: flies, bees and hawkmoths have all been reported to shape their visual input with specific head or body movements (Schilstra et al., 1998; Cellini et al., 2020; van Breugel et al., 2012; Windsor et al., 2017; Egelhaaf et al., 2012; Boeddeker et al., 2010). A recent work by Ravi et al. (2022) explored the gaze strategy of bumblebees around obstacles, and showed that the insects' head stabilisation enabled them to detect obstacles based on motion contrast against the background. For the case of birds, the role of gaze in flight has mostly been studied by inspecting head stabilisation in relation to the environment. For example, saccade-fixation patterns, in which periods of gaze stabilisation are separated by fast gaze shifts, have been reported in head movements of pigeons (Ros et al., 2017b), zebra finches (Eckmeier et al., 2008) and lovebirds (Kress et al., 2015) during turning manoeuvres. Head motion patterns have also been investigated in pigeons flying through a forest of horizontal poles (Ros et al., 2017a), and homing over a natural landscape (Kano et al., 2018). However, only a few studies have specifically related objects in the environment to the birds' gaze direction in flight. In the work by (Kress et al., 2015), birds making tight turns were found to fixate high-contrast features frontally in their visual fields. In Lin et al. (2014), the head orientation of pigeons in flight was analysed in relation to a set of vertical obstacles they traversed, but to develop a gap-aiming steering model of their flight paths, rather than to inspect the features likely fixated by the birds. Both of these works analysed the gaze strategy of the birds, focusing on head movements in yaw in a two-dimensional approach.

Here, we present a study investigating the gaze strategies of Harris' hawks executing perching and obstacle avoidance manoeuvres. We carried out motion capture experiments with four birds, and used the data to recreate the position and orientation of the birds' visual fields in flight in a computational environment. With a rendering engine, we generated detailed information about the visual scene the birds experience while flying to the perch and avoiding obstacles. Contrary to previous approaches, our method allows us to analyse the full visual field of the birds and the full 3D movements of their heads. We found that the birds use distinct gaze strategies around the landing perch and the obstacles: while the perch is stabilised and centred in the birds' visual field through most of the perching manoeuvre, the inside edge of the obstacles relative to the turn is fixated frontally when the birds fly around them. In relation to the perch, our results suggest that the birds are more likely to deviate from a perch-centring strategy, and instead to fixate the edge of the perch frontally, when the trajectory involves more turning. This study allowed us to explore different methods of analysis and to define clear hypotheses that can be verified with the full dataset in the future.

4.3 Methods

We tracked the head movements of $N = 4$ Harris' hawks executing perching and obstacle avoidance manoeuvres in a motion capture lab, and reconstructed their visual experience synthetically over $n = 97$ flights. Specifically, we used the motion capture data collected across all flights per bird to estimate a coordinate system representative of each bird's visual field. For each flight, we modelled the bird's visual field in a 3D computer graphics package (Blender, version 2.93.5, Blender Online Community, 2021), using a virtual camera moving through a model of the lab environment. Using Blender's rendering engine (Cycles), we produced semantic maps per frame for all flights: these specify the object to which each pixel rendered by the virtual camera back-projects. We combined the semantic maps generated with data available in the literature defining the different regions in the visual fields of Harris' hawks (Potier et al., 2016), to analyse the hawks' gaze strategies in

flight. The methodology applied here follows the one presented in Chapter 3 for the obstacle avoidance trial; the obstacle avoidance trial presented in Chapter 3 for demonstration is one of the Drogon trials analysed here.

4.3.1 Motion capture experiments

The birds flew in a volume of size 12.0 m × 5.3 m ×3.3 m, instrumented with 22 motion capture cameras (Vicon Vantage V16; 200 Hz sampling rate). To track the birds' heads, we designed 3D-printed 'headpacks' with four spherical retroreflective markers each (6.4 mm diameter), which we fixed to the birds with a Velcro patch glued to their heads (see Appendix C.1). For each bird, we attached the headpack just before recording its set of trials, and removed it immediately after.

We recorded two types of trial: perching and obstacle avoidance. In the perching trials, the birds flew from a starting perch to an end perch (first leg of the trial), and back (second leg). The perches were placed 9 m apart, along the longitudinal axis of the room. In each trial, we randomly selected the perches' lateral positions from one of three options (set 1 m apart, see Figure 4.1, left panel). In the obstacle avoidance trials, we followed the same procedure, but added four cylindrical styrofoam pillars of 0.3 m diameter and 2 m height, 1.5 m ahead of the end perch (Figure 4.1, right panel). Note that each recorded trial consists of two flights, which correspond to the two legs of the trial. To locate the main elements in the environment, we fixed large markers (14 mm diameter) to the edges of the perches and the tops of the obstacles.

In this chapter, we include results for $n = 97$ flights across $N = 4$ birds; the number of flights per bird is shown in Table 4.1, and the recorded trajectories are represented in Figures 4.2 and 4.3. Note that the dataset is not balanced within obstacle avoidance flights. This is because two of the birds flew over the obstacles, rather than around them, in some of the flights executed in leg 2 of the trials (see trajectories in black in Figure 4.3). This happened for 4 flights executed by Charmander and 1 flight executed by Ruby; we deal with these flights separately in the following sections. The flights analysed in this chapter are part of a larger experimental campaign in which we recorded $n = 279$ trials in total. Here, we use a

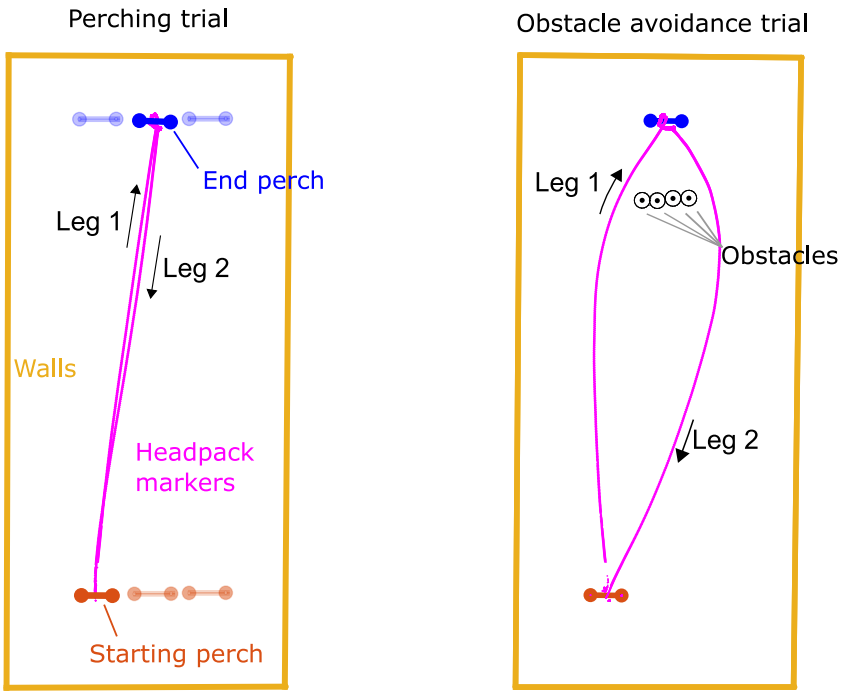


Figure 4.1: Experimental setup for the perching and obstacle avoidance trials. A top-down view of the experimental setup for the perching and obstacle avoidance trials is shown. In the perching trials (left), the bird flew from the starting perch (red) to the end perch (blue) in leg 1 of the trials, and back in leg 2 of the trials. The perches were 9 m apart and we randomly selected their lateral positions from one of three options, set 1 m apart (shown in semi-transparent red and blue). In the obstacle avoidance trials, we additionally placed a set of four obstacles 1.5 m ahead of the end perch (black white-filled circles, centres highlighted with black markers). The trajectories of the markers on the bird's headpack are shown (magenta), as well as the estimated position of the walls (yellow).

subset of the data to consolidate our analytical approach, and plan to verify our hypotheses with the full dataset as part of the future work. Although the resulting dataset is not balanced across birds, we consider that it is still useful for our aim of consolidating the method of analysis. Further details on the motion capture experiments and the full dataset are included in Appendix C.1.

We used the motion capture system's software Vicon Nexus v2.8.0 (Vicon Motion Systems Ltd., Oxford, UK) to extract the unlabelled 3D coordinates of the retroreflective markers, and used custom scripts in MATLAB (version 2020b, The Mathworks Inc., 2020) to label them, handle missing data, extract the headpacks' positions and orientations, and compute the simplified geometry of the environment. The headpack markers' labels were computed following the ground-truth labelling

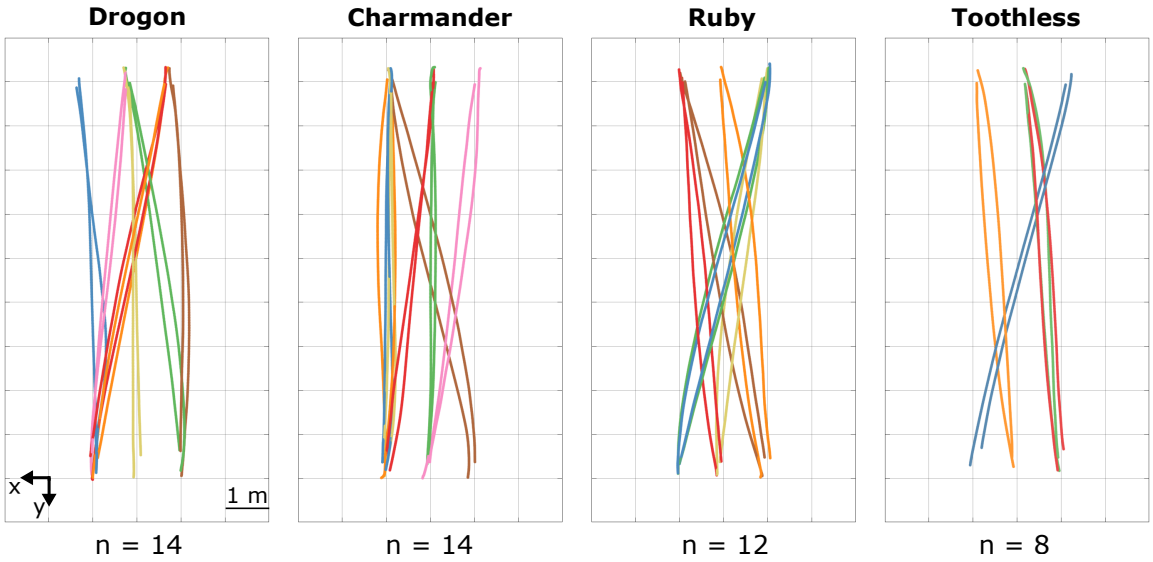


Figure 4.2: Perching flights per bird. The figure shows a top-down view of the trajectories of the birds' heads in the perching trials, grouped by bird. Each trial consists of two flights, which correspond to the two legs of the trial. Flights of the same trial are represented in the same colour. The total number of flights per bird n is shown. The trajectories are represented in a grid of 1 m spacing, in the motion capture coordinate system; its orientation is shown in the bottom-left corner. The takeoff position for leg 1 of the trials would be located at the bottom of the image.

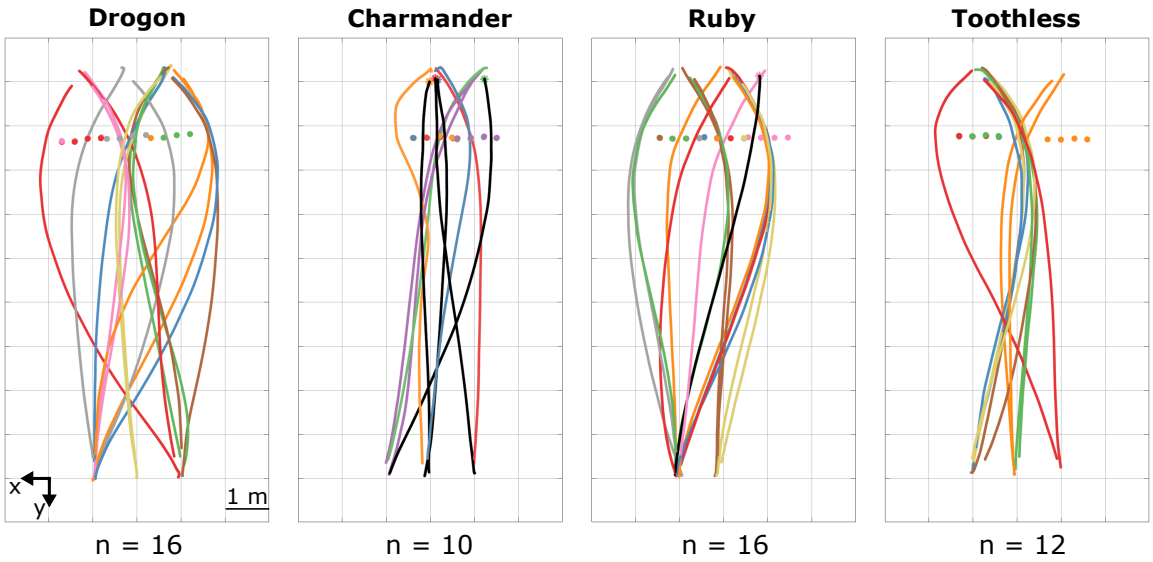


Figure 4.3: Obstacle avoidance flights per bird. The figure shows a top-down view of the trajectories of the birds' heads in the obstacle avoidance trials, grouped by bird. Each trial consists of two flights, which correspond to the two legs of the trial. Flights of the same trial are represented in the same colour. The total number of flights per bird n is shown. For Charmander and Ruby, n includes the flights in which they flew over the obstacles rather than around them; these trajectories are shown in black. The obstacles' centres per trial are shown (colour markers, same colour as the corresponding trial). The trajectories are represented in a grid of 1 m spacing, in the motion capture coordinate system; its orientation is shown in the bottom-left corner. The takeoff position for leg 1 of the trials would be located at the bottom of the image.

Table 4.1: Number of flights analysed per bird, trial type and trial leg. For two birds, the number of obstacle avoidance flights executed in leg 1 and leg 2 of the trials is not the same due to trial failure: the birds flew over the obstacles in the second leg of the trial, rather than around them. These cases are marked with a star in the table (*); we excluded these flights from the analyses in the main text.

	Perching		Obstacle avoidance	
Bird	Leg 1	Leg 2	Leg 1	Leg 2
Drogon	7	7	8	8
Charmander	7	7	5	1*
Ruby	6	6	8	7*
Toothless	4	4	6	6

method from Chapter 2 (Section 2.3.3.3 and Appendix A.10). The labelled markers deviated less than 3.7 mm from their corresponding positions in a reference template across all flights (see Appendix C.2). To estimate the walls and the ceiling locations, we used the positions and orientations of the motion capture cameras themselves, which were mounted on scaffolding around the walls. This postprocessing approach follows the one described for the obstacle avoidance flights in Chapter 3.

4.3.2 Visual coordinate system

To model the bird's view of the scene in Blender, we defined for each flight a virtual camera that replicated the bird's head movements. This is a reasonable approach, since in birds, the large size of their eyes relative to their heads limits the eyes' motion, so gaze movements across the environment are largely executed via head movements (Mitkus et al., 2018; Yorzinski et al., 2015). However, the virtual camera axes should be aligned with a coordinate system that is representative of the bird's visual field.

A headpack is arbitrarily placed on the bird's head, so a coordinate system defined relative to the headpack markers is not necessarily aligned with the bird's visual field. One option would be to attach the headpack to the bird's skull in a precise and repeatable manner, but this would require us to restrain the animal every time the headpack is attached, and would likely involve a surgical intervention (Lev-Ari et al., 2018; Hazan et al., 2015). This type of procedure limits the number

of trials that can be recorded due to animal welfare considerations, which conflicts with our aim of collecting large amounts of data across many individuals. Instead, we estimate a coordinate system that is representative of each bird's visual field using the collected data. We refer to it as the 'visual coordinate system', and we define it using data recorded for the same bird and headpack placement. In this way, we can compare our results across birds and days, and to the previous literature.

4.3.2.1 Assumptions

To define the visual coordinate system, we make use of three assumptions: first, we assume that the bird's eye movements relative to the head are small in flight, such that the bird's head orientation is a good proxy for gaze orientation. Similar studies frequently make this assumption (Kano et al., 2018; Brighton et al., 2017; Ros et al., 2017b; Kress et al., 2015; Kane et al., 2014; Eckmeier et al., 2008). Although we found no reports in the literature for the range of eye motions in Harris' hawks, we found estimates for birds in the same Family (Accipitridae): $\leq 6°$ is estimated for goshawks (*Accipiter gentilis*) in Kane et al. (2015); $\sim 5°$ is estimated for red-tailed hawks (*Buteo jamaicensis*), in O'Rourke et al. (2010); and $\sim 8°$ is estimated for Cooper's hawks (*Accipiter cooperi*), in O'Rourke et al. (2010). Second, we assume that the bird fixates its gaze on the perch's centre upon landing; we observed this during the experiments and in previous ones (see appendix Figure C.3), and similar reports exist in the literature (Potier et al., 2016; Kress et al., 2015). Third, we assume that the bird keeps its eyes level in flight. This has also been reported previously in Harris' hawks and other birds (Brighton et al., 2019; Ros et al., 2017b; Warrick et al., 2002).

We use these assumptions to estimate for each bird its gaze direction $\vec{v}_{gaze}$ and the normal to the symmetry plane of its head (i.e., the sagittal plane), $\vec{n}_{sagittal}$. The axes of the visual coordinate system are defined as follows: the x-axis is defined parallel to $\vec{n}_{sagittal}$ (pointing to the left side of the head), the y-axis is defined parallel to $-\vec{v}_{gaze}$, and the z-axis follows the definition of a right-handed coordinate system. In the following sections we describe the procedure to estimate $\vec{v}_{gaze}$

and $\vec{n}_{sagittal}$ per bird. This method follows the approach described in Chapter 3, Section 3.3.3.2 for the obstacle avoidance flights.

Table 4.2: Estimated gaze direction vectors $\vec{v}_{gaze}$ per bird. The estimated gaze direction $\vec{v}_{gaze}$ per bird is expressed in an auxiliary coordinate system fixed to the bird's headpack. Therefore, the vectors are not expressed in the same coordinate system across birds, since the headpack designs and their placement on the birds' heads vary (see further details in Appendix C.3). RMSE stands for root mean square error, where the error is the distance between the sample points and the fitted line. The number of samples n corresponds to the number of frames in the final approach phase, across all trials per bird. The number of perching flights (N_P) and obstacle avoidance flights (N_{OA}) considered in each fit per bird is also included.

Bird	$\vec{v}_{gaze}$	RMSE (mm)	n	N_P	N_{OA}
Drogon	$[0.13, -0.94, 0.31]$	65.6	1274	14	16
Charmander	$[0.00, -0.97, 0.23]$	64.7	1328	14	10
Ruby	$[0.16, -0.96, 0.22]$	70.6	1353	12	16
Toothless	$[-0.07, -0.95, 0.31]$	46.7	847	8	12

4.3.2.2 Gaze direction estimate

If, as we have assumed, the bird is fixating the midpoint of the landing perch in its visual field during the final approach to the perch, the trajectory of this point in a coordinate system fixed to the headpack would be close to a straight line. Figure 4.4 confirms this is a good approximation for the collected data. To estimate the gaze vector $\vec{v}_{gaze}$ for each bird, we fitted a 3D line to the trajectory of the landing perch midpoint in an auxiliary coordinate system fixed to the headpack. For the fit, we considered the frames in the 'final approach' phase across all flights per bird. We define the final approach phase based on the distance to the landing perch (0.5 m to 2.0 m for perching trials, 0.5 m to 1.0 m for obstacle trials; see Appendix C.4).

The results for each bird are shown in Figure 4.4 and Table 4.2. Note that the $\vec{v}_{gaze}$ vectors across birds are not expressed in the same auxiliary coordinate system, since this is defined relative to the headpack. Further details on the headpack coordinate system for each headpack design are included in Appendix C.3. For one of the birds (Drogon), we additionally fitted a gaze direction vector $\vec{v}_{gaze}$ per trial, and verified that the results were similar to the gaze direction vector computed

by aggregating all trials (see Chapter 3, Appendix B.3.3). Additional details on the $\vec{v}_{gaze}$ fitting procedure are included in Appendix C.4.

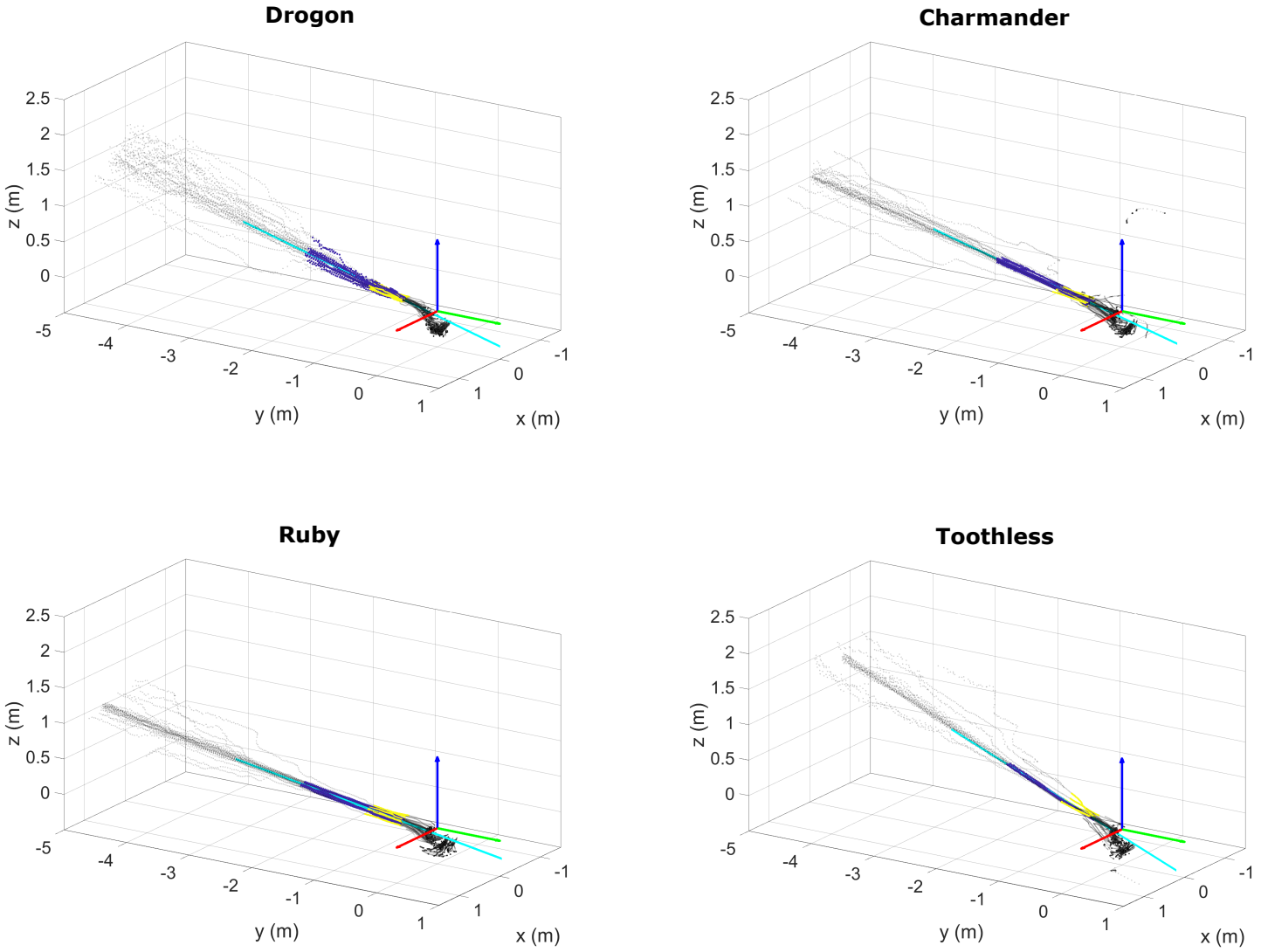


Figure 4.4: Gaze direction estimate per bird. The trajectory of the landing perch's midpoint is represented in the headpack coordinate system for all flights per bird, from the point the bird is 5 m away from the landing perch in y-direction until landing. The direction of $\vec{v}_{gaze}$ is estimated as the orthogonal regression line (cyan) to the samples in the 'final approach' phase (blue for the perching flights, yellow for the obstacle avoidance flights). The samples outside the final approach phase are shown for reference (grey, semi-transparent markers). The number of samples and flights per bird is included in Table 4.2.

4.3.2.3 Normal to the sagittal plane estimate

To estimate the normal to the sagittal plane $\vec{n}_{sagittal}$ for each bird, we computed the unit vector perpendicular to $\vec{v}_{gaze}$ that best approximates the local horizon in the headpack coordinate system, during the 'mid-flight' phase across all flights. We defined the mid-flight phase as the period in which the bird was > 2 m away from either perch and flying at a speed > 2.5 m/s. We identified $\vec{n}_{sagittal}$ as the solution to the least-squares problem

$$\begin{aligned} \theta^* &= \arg\min_{\theta} \sum_{i=1}^{N} (\vec{z}_{world,i} \cdot \vec{n}_{sagittal}(\theta))^2, \\ \vec{n}_{sagittal}(\theta) &= (cos\theta)\vec{a} + (sin\theta)\vec{b}, \\ \vec{n}_{sagittal} &= \vec{n}_{sagittal}(\theta^*), \end{aligned} \tag{4.1}$$

where $\vec{z}_{world,i}$ is the world's z-axis in the headpack coordinate system at frame i (positive opposite to gravity), $\vec{a}$ and $\vec{b}$ are an arbitrary orthonormal basis of the plane perpendicular to $\vec{v}_{gaze}$, and θ is the angle between $\vec{n}_{sagittal}$ and $\vec{a}$. The $\vec{n}_{sagittal}$ vectors per bird are shown in Figure 4.5 and Table 4.3. In Table 4.3, RMSE stands for root mean square error, where the error is the residual for each sample i; that is, the projection of the fitted vector $\vec{n}_{sagittal}(\theta^*)$ on the world's z-axis at frame i: $\vec{z}_{world,i} \cdot \vec{n}_{sagittal}(\theta^*)$. Note that the world's z-axis is parallel to the z-axis of the motion capture coordinate system.

4.3.2.4 Trajectory coordinate system

For each trial, we defined an additional coordinate system for the virtual camera to track, which we call the trajectory coordinate system. Its x-axis is always parallel to the floor plane of the lab, and its y-axis is always parallel to the bird's head velocity. Thus, the optical axis of a virtual camera tracking the trajectory coordinate system points tangentially along the bird's head trajectory, while keeping the horizon level. We use the trajectory coordinate system as a 'null model' for comparison with the visual coordinate system, following the approach in Chapter 3. The mathematical definition of the trajectory coordinate system is included in Appendix B.3.6.

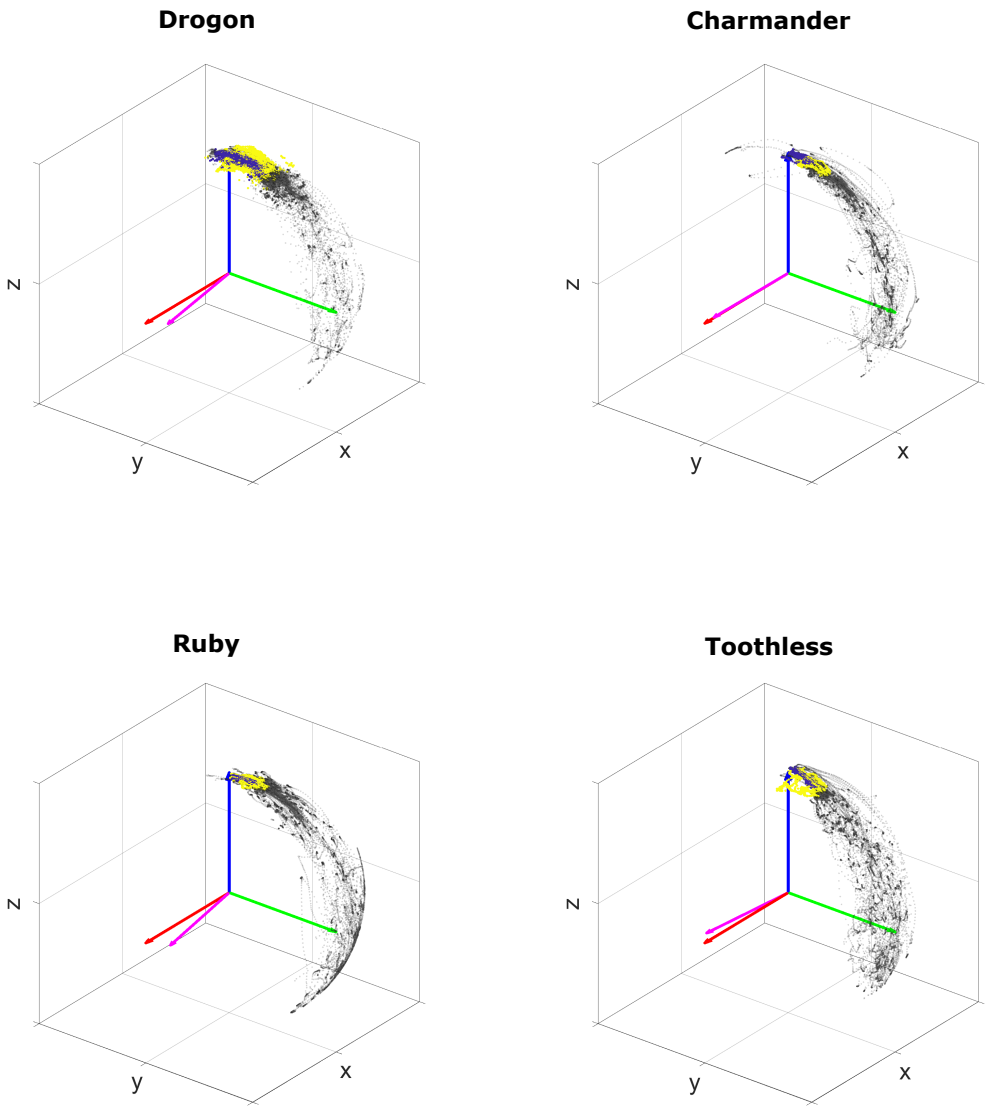


Figure 4.5: Normal to the sagittal plane estimate per bird. The trajectory of the world's z-axis unit vector $\vec{z}_{world,i}$ is represented in the headpack coordinate system, for all flights per bird. The normal to the sagittal plane $\vec{n}_{sagittal}$ (magenta) is estimated as the vector perpendicular to $\vec{v}_{gaze}$ (not shown) that minimises its projection on the world's z-axis $\vec{z}_{world,i}$, across all frames in the 'mid-flight' phase. The samples in the mid-flight phase are shown in colour (blue for the perching flights, yellow for the obstacle avoidance flights), while the samples outside the mid-flight phase are shown in grey (semi-transparent markers). The samples outside the mid-flight phase show qualitatively larger variation in pitch (i.e., rotation around the x-axis) and roll (i.e., rotation around the y-axis). The number of samples and flights per bird is included in Table 4.3.

Table 4.3: Estimated normals to the sagittal plane $\vec{n}_{sagittal}$ per bird. The estimated normal to the sagittal plane $\vec{n}_{sagittal}$ per bird is expressed in an auxiliary coordinate system fixed to the bird's headpack. Therefore, the vectors are not expressed in the same coordinate system across birds, since the headpack designs and their placement on the birds' heads vary (see further details in Appendix C.3). RMSE stands for root mean square error, where the error is the residual of the least-squares problem solved in Equation 4.1; note it is dimensionless. The number of samples n is the number of frames in the mid-flight phase across all trials per bird. The number of perching flights (N_P) and obstacle avoidance flights (N_{OA}) considered in each fit per bird is also included.

Bird	$\vec{n}_{sagittal}$	RMSE	n	N_P	N_{OA}
Drogon	$\begin{bmatrix}0.99, & 0.14, & -0.01\end{bmatrix}$	0.08	5056	14	16
Toothless	$\begin{bmatrix}1.00, & -0.07, & 0.02\end{bmatrix}$	0.06	3543	8	12
Charmander	$\begin{bmatrix}1.00, & 0.01, & 0.01\end{bmatrix}$	0.03	4641	14	10
Ruby	$\begin{bmatrix}0.99, & 0.16, & -0.02\end{bmatrix}$	0.03	5350	12	16

4.3.2.5 Origin of the visual and the trajectory coordinate system

The origin of both the visual and the trajectory coordinate systems is defined as the projection of the headpack markers' centroid onto the plane defined by the headpack's baseplate (i.e., the surface of the headpack that attaches to the bird's head). However, following the methodology to represent a bird's visual field in a sphere described in Martin (2007), a better approximation for the origin of the visual coordinate system would be the midpoint between the tips of the eyes' corneas. We estimated the offset between our selected origin and the midpoint between the eyes for one of the birds (Drogon), using reference pictures. We estimated the selected origin to be within 12 mm of the midpoint between the bird's eyes in the transverse direction, and within 17 mm in the dorso-ventral direction (see Appendix B.3.2). Since all headpacks used are of comparable size, and in all cases we defined the origin to be at the headpack baseplate, we expect similar displacements for the rest of the birds.

4.3.2.6 Regions in Harris' hawks visual field

We incorporated in our analysis the boundaries of the different regions in the visual field of Harris' hawks (monocular, binocular and blind areas), from data

available in the literature (Potier et al., 2016). The method is described in Chapter 3 (Section 3.3.3.3 and Appendix B.3.5) and the identified regions are shown in Figure 3.8.

4.3.3 Computational model in Blender

We used the Python API in Blender to model, for each trial, the simplified geometry of the lab, and to define a virtual camera that tracked the visual coordinate system or the trajectory coordinate system. We modelled the lab environment using geometric primitives, following the method described in Chapter 3 (Section 3.3.4 and Appendix B.4.2). We found this model to be sufficient for our aim of investigating the bird's gaze strategy in relation to the main elements in the scene, but note that the method allows for the creation of a more detailed model (see Chapter 3, Section 3.3.4).

We used a monocular 360° virtual camera to model the birds' binocular visual system, given that any vergence movements of the eyes were unknown. We used an equirectangular projection to output the data from the 360° camera view as 2D matrices. We selected a resolution of 5 pixels per degree latitude and longitude; this makes the length and width of all pixels within the visible regions of the birds' visual fields larger than the minimum resolution angle of Harris' hawks (Potier et al., 2016). For each flight, we rendered semantic maps per frame for a virtual camera tracking the visual coordinate system and the trajectory coordinate system.

4.4 Results

We use the per-frame semantic data, from $n = 97$ flights across $N = 4$ birds to investigate the birds' gaze strategy during perching and obstacle avoidance manoeuvres. We first investigate whether the landing perch or the obstacles are stabilised in specific areas of a bird's visual field (such as the binocular area), next inspect if this stabilisation is sustained along the flight, and finally analyse in further detail the gaze strategy in relation to the two objects.

4.4.1 Stabilisation of the landing perch and the obstacles in the visual field

We investigate whether the landing perch or the obstacles are stabilised in preferential areas of a bird's visual field for a substantial proportion of the flight, and consistently across trials. For this we compute spherical histograms that represent the frequency with which the contour of the landing perch or the obstacles appears at each point in a bird's visual field. More specifically, the spherical histograms show the following metric for each pixel in the visual field sphere:

$$h_i = \frac{n_i}{NA_i}, \tag{4.2}$$

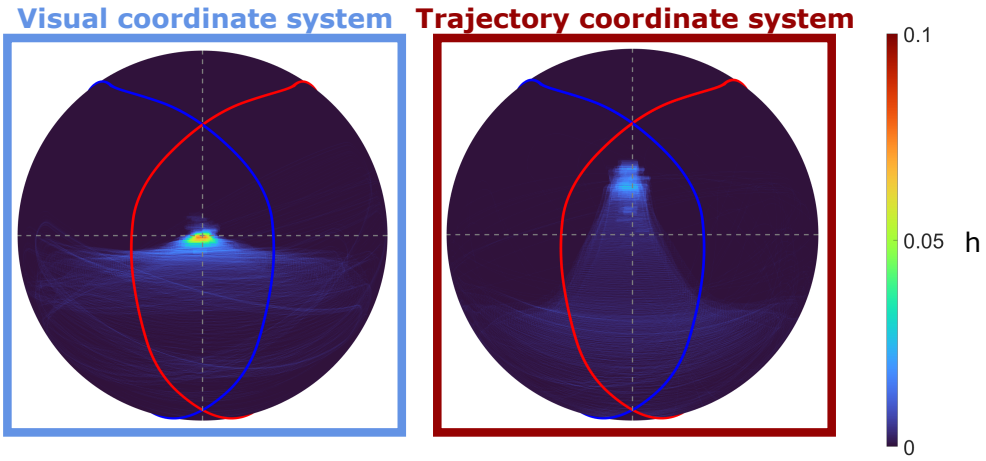
where n_i denotes the number of frames over which the ith pixel saw the contour of the object of interest, N denotes the total number of frames analysed per bird, and A_i denotes the solid angle covered by the ith pixel, normalised by the maximum solid angle of a pixel in the visual field. We apply this normalisation to a pixel's solid angle so that A_i takes values in the range $[0, 1]$, which in turn keeps the interpretation of the metric h_i as approximately expressing a fraction of frames. A schematic describing the way the spherical histograms were computed is shown in the appendix Figure C.5. This approach is similar to that presented in Chapter 3, Section 3.4.1.

We compute the spherical histograms using data from virtual cameras tracking the visual coordinate system and the trajectory coordinate system. For each histogram, we group the data by bird, type of trial (perching or obstacle avoidance), and leg of the trial (1 or 2). We only consider the semantic data within the visible region of a bird's visual field, and exclude the frames in which the head pose was interpolated. When representing the spherical histograms of the obstacles' contour in the main text, we horizontally mirror the semantic data from the required flights for clarity, so that the obstacles are always on the same side of the virtual camera when passed. This transformation produces different h_i values per pixel compared to the case in which no transformation is applied, but we find that the differences are not substantial due to the birds' gaze strategy around the obstacles. The spherical histograms for the obstacles without data mirroring are included in Appendix C.5.

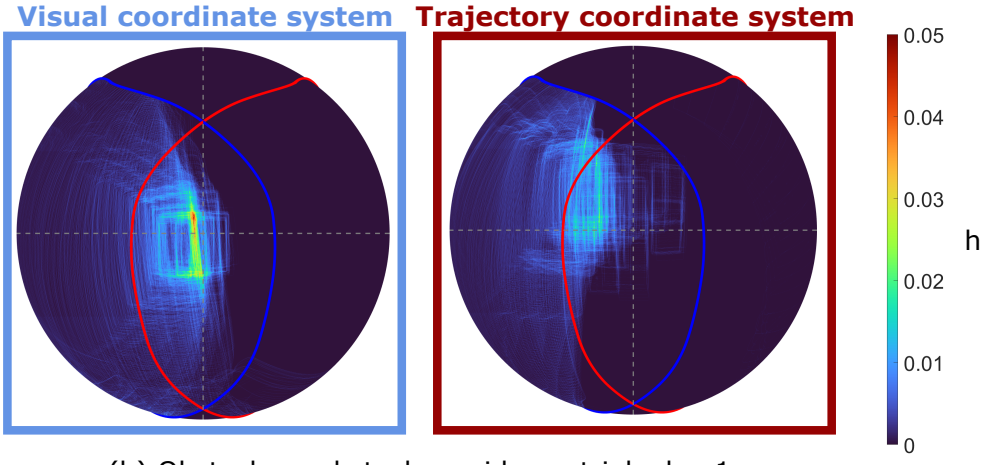
We summarise the main results from the spherical histograms in Figures 4.6 to 4.8; we include the spherical histograms aggregating all flights per bird, type of trial, and leg of the trial in Appendix C.5. In Figure 4.6, we first highlight that, across all birds and types of manouevres, both the landing perch and the obstacles are more stabilised and centred in the visual field when the camera tracks the visual coordinate system than when it tracks the trajectory coordinate system. This suggests that the birds are stabilising their head pose against oscillations likely induced by the wingbeat motion, and is in agreement with the results obtained for the two flights analysed in Chapter 3 (see Figure 3.16, and the supplementary videos `ESM_8.avi` and `ESM_10.avi`). Additionally, notice how the spherical histograms in the trajectory coordinate system reflect the shape of the trajectory relative to the landing perch or the obstacles (see Figure C.6 for additional examples, and the panels with red contours in Figures C.7 to C.12).

The spherical histograms in the visual coordinate system also suggest that certain features of the objects of interest are placed in specific areas of the visual field. The landing perch is most frequently aligned with the plane at 0° latitude, and centred relative to the estimated sagittal plane (i.e., the 0° longitude plane). This is the case in both perching (Figure 4.7a and b) and obstacle avoidance flights (Figures 4.7c and d). The spherical histogram for the obstacles in Figure 4.8 shows that the edge closest to the turn in the trajectory is most frequently aligned with the estimated sagittal plane of the bird. This occurs for the flights executed in both legs of the obstacle avoidance trial, although more clearly for those in leg 1 (see Figures C.11 and C.12 for further examples). None of these features are stabilised in the visual field when the virtual camera tracks the trajectory coordinate system (Figure 4.6). This suggests that all birds are actively stabilising their head poses to maintain these features in specific areas of their visual fields. We inspect whether this is sustained or varies along the flight in the next section.

Note that in the perching flights, the period in which the landing perch is visible is very similar for both legs of the trial. In contrast, in the obstacle avoidance flights, the sections in which the landing perch or the obstacles are visible are very



(a) Landing perch - perching trials, leg 1
Drogon, n = 7 flights, 2356 frames



(b) Obstacles - obstacle avoidance trials, leg 1
Drogon, n = 8 flights, 2860 frames

Figure 4.6: Comparison of spherical histograms in the visual coordinate system and the trajectory coordinate system. The figure shows the spherical histograms for the landing perch (a) and the obstacles (b), for Drogon's flights executed in leg 1 of the trials. The results are represented in the visual coordinate system (blue box) and the trajectory coordinate system (red box). The visual field spheres are represented using an orthographic projection on the xz plane of the visual or the trajectory coordinate system; note that this omits the most peripheral parts of the bird's visual field (see Figure 3.8). The retinal margins from the right (red) and left (blue) eyes are shown for reference, as well as the planes at 0° latitude and 0° longitude (grey dotted lines). The colourbars' maximum values correspond to the rounded maximum h_i values observed for the histograms in the visual coordinate system.

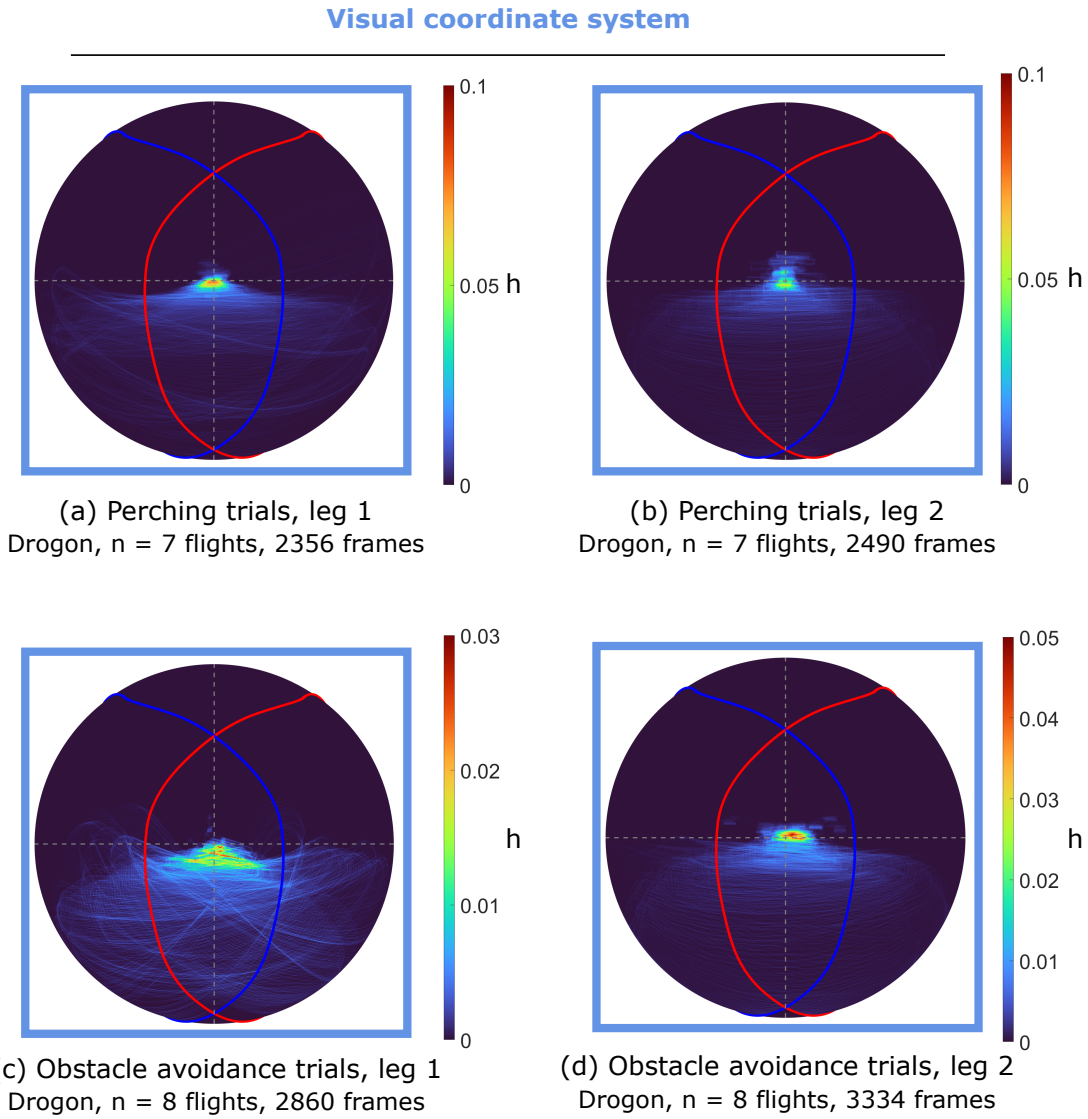


Figure 4.7: Spherical histograms showing the stabilisation of the landing perch in the visual field. The figure shows the spherical histograms for the landing perch in the visual coordinate system (blue box) for Drogon's perching flights (a and b), and obstacle avoidance flights (c and d), grouped by leg of the trial. The visual field spheres are represented using an orthographic projection on the xz plane of the visual or the trajectory coordinate system; note that this omits the most peripheral parts of the bird's visual field (see Figure 3.8). The retinal margins from the right (red) and left (blue) eyes are shown for reference, as well as the planes at 0° latitude and 0° longitude (grey dotted lines). The colourbars' maximum values correspond to the rounded maximum h_i values observed. Note that figure (a) is the same as Figure 4.6a.

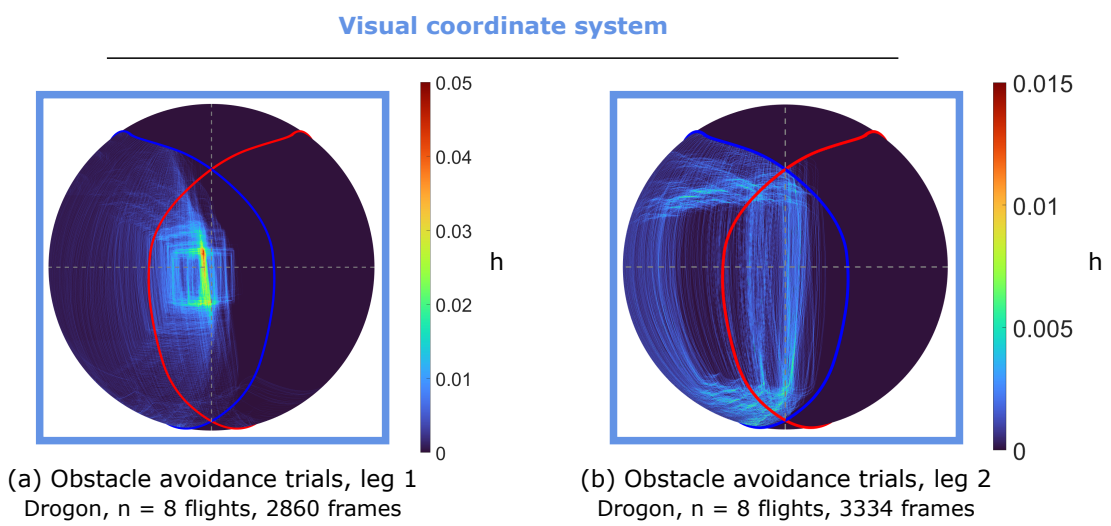


Figure 4.8: Spherical histograms showing the stabilisation of the obstacles in the visual field. The figure shows the spherical histograms for the obstacles in the visual coordinate system (blue box), for Drogon's flights executed in leg 1 (a) and 2 (b) of the obstacle avoidance trials. The visual field spheres are represented using an orthographic projection on the xz plane of the visual or the trajectory coordinate system; note that this omits the most peripheral parts of the bird's visual field (see Figure 3.8). The retinal margins from the right (red) and left (blue) eyes are shown for reference, as well as the planes at 0° latitude and 0° longitude (grey dotted lines). The colourbars' maximum values correspond to the rounded maximum h_i values observed.

different for each leg (see Figure 4.9). This is reflected in the maximum values observed for h_i in the spherical histograms per leg (compare the colourbars in Figures 4.7c and d, and Figures 4.8a and b).

The spherical histograms for the five flights in which two birds flew over the obstacles are included for reference in Appendix C.5 (Figure C.13). In them, the top edge of the obstacles appears centred and stabilised in the visual coordinate system. This would be in agreement with a strategy of fixating the edge of the obstacles closest to the inside of the turn, but it would need to be confirmed with more flights. We exclude these flights from subsequent analyses.

4.4.2 Longitude angle subtended by the landing perch and the obstacles along the trajectory

Since the landing perch and the obstacles are stabilised latitudinally, we focus on how their subtended longitude angle varies along the trajectory. This is represented in Figures 4.10 to 4.12, which show the maximum and minimum

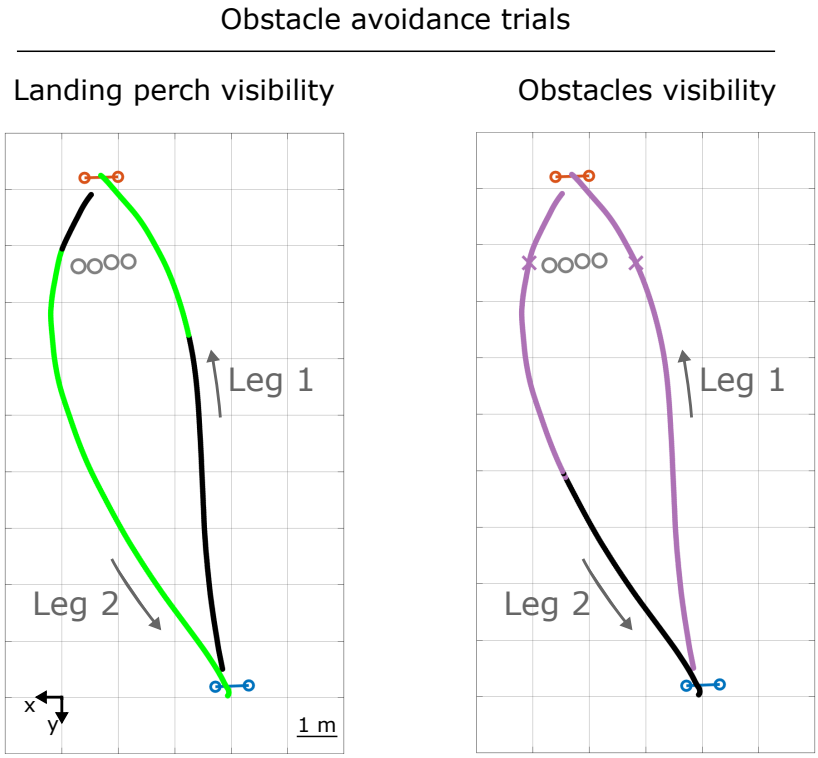


Figure 4.9: Visibility of the landing perch and the obstacles in an obstacle avoidance trial. The figure shows the periods in which the landing perch (left panel) and the obstacles (right panel) are visible, for the two flights in an obstacle avoidance trial. We consider an element to be visible when at least part of it falls within the visible region of the bird's visual field. For the landing perch, we additionally require that it is not partially occluded by the obstacles. The period in which the landing perch is visible (marked in green) is longer in leg 2 of the trial than in leg 1, whereas the opposite occurs for the obstacles (visibility period marked in purple). The point in which the bird's y-coordinate crosses the obstacles' mean y-coordinate is highlighted (purple cross in the right panel). The obstacles appear as visible to the bird after being passed, but in some of the frames just prior to landing they are likely occluded by the bird's body (see for example the section from frame 1078 in the Supplementary video `ESM_8.avi`; note that the orthographic projection used in the video does not capture the bird's full visual field).

longitude angles registered for the landing perch (Figures 4.10 and 4.11) and the obstacles (Figure 4.12) in each bird's visual field. The results are grouped by bird, type of trial and leg of the trial, and plotted as a function of the normalised distance to the landing perch in y-direction, $\hat{y}_{perch}$:

$$\hat{y}_{perch}(t) = \frac{y(t) - y_{L\,perch}}{y_{TO\,perch} - y_{L\,perch}}, \tag{4.3}$$

where $y(t)$ denotes the bird's y-coordinate at frame t, $y_{TO\,perch}$ denotes the y-coordinate of the takeoff perch, and $y_{L\,perch}$ denotes the y-coordinate of the landing perch, all expressed in the motion capture coordinate system. With this representa-

tion, the location of the landing perch and the obstacles is at a constant value of $\hat{y}_{perch}$, which facilitates the visualisation of the data; we show later that the curves are qualitatively similar when represented against time.

Within the obstacle avoidance flights, we focus the analysis on the legs of the trial in which the landing perch or the obstacles are visible for longer. That is, for the landing perch, we focus on the obstacle avoidance flights executed in leg 2, and for the obstacles, we focus on the flights executed in leg 1 (see Figure 4.9). This way, the number of frames analysed is comparable between the obstacle avoidance and the perching flights. We follow the same approach in subsequent sections. For reference, we include the curves for the remaining flights in Appendix C.6. Additionally, in Figure 4.12 we mirror the data of the required flights so that the obstacles are always passed on the same side relative to the bird; this is for clarity only. The data without mirroring is included in Appendix C.6 (Figures C.19 and C.20). The curves representing the evolution of the latitude angle subtended by the landing perch along the trajectory are also included in Appendix C.6.

The curves shown in Figures 4.10 to 4.12 confirm that the centring behaviour around the perch, and the edge-fixating behaviour around the obstacles are largely sustained along the flight. They also show the different stimuli these distinct gaze strategies induce on the birds' visual field. When approaching the landing perch, the birds experience a bilateral looming stimulus, in both perching and obstacle avoidance trials (see Figures 4.10 and 4.11). In contrast, when approaching the obstacles, the birds experience a unilateral looming stimulus. Additionally, the fact that this edge-fixating behaviour when navigating around obstacles has been reported previously in humans (Raudies et al., 2012; Rothkopf et al., 2009; Rothkopf et al., 2007) and in smaller birds (Kress et al., 2015) suggests that our estimated visual coordinate system is reasonable.

In the following sections, we inspect further the birds' gaze strategy around the landing perch and the obstacles, focusing on their extension in longitude angle along the flight. For the landing perch, we inspect the ways in which the birds deviate from an ideal symmetric expansion of the perch; from visual inspection in Figures 4.10 and

4.11 we can see it is not perfectly symmetric in many flights. Similarly, for the gaze strategy around the obstacles, we inspect how much the birds deviate from a perfect alignment with the obstacles' inside edge, and estimate for each flight the point in the trajectory at which the obstacles' edge is no longer fixated in the visual field.

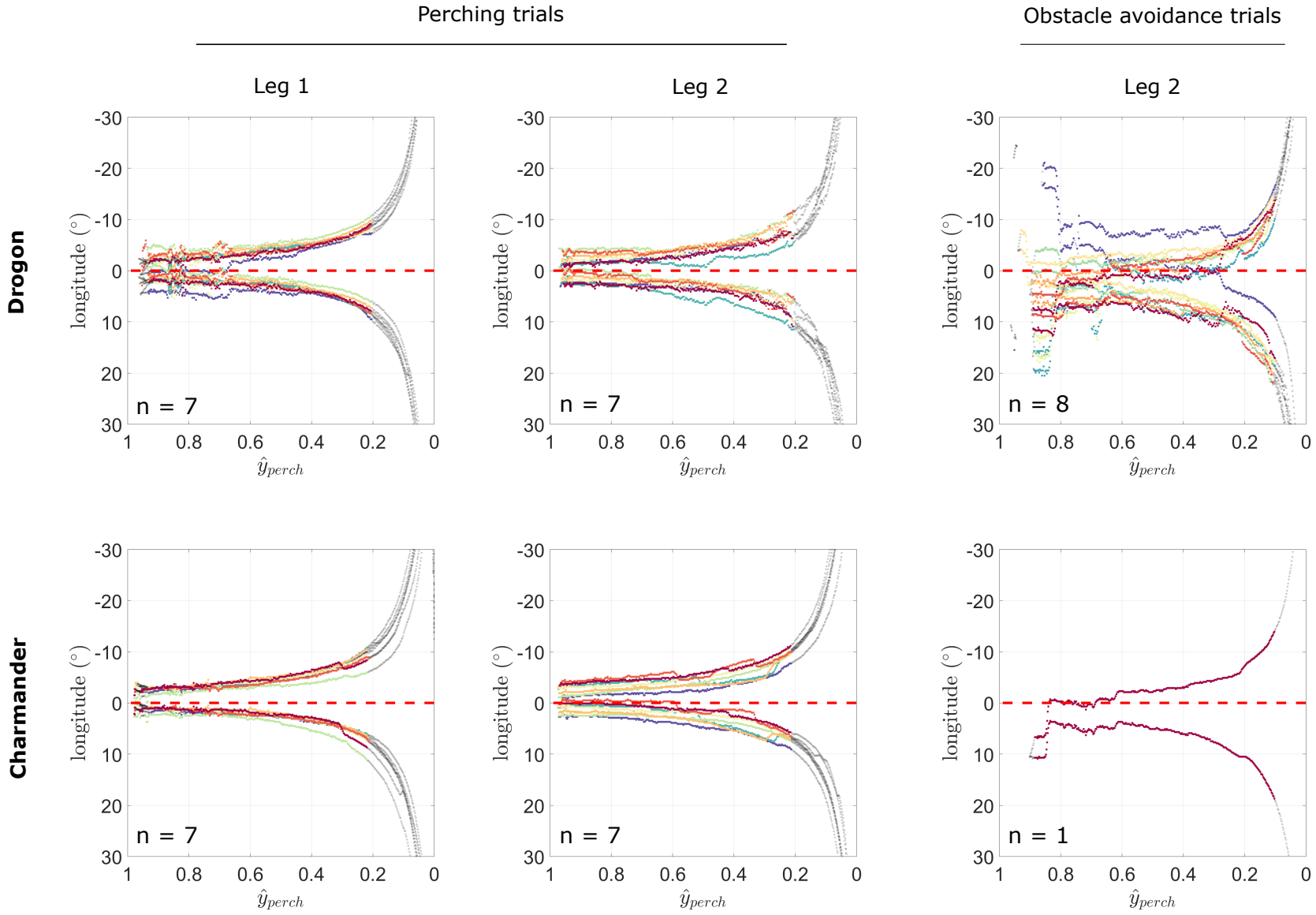


Figure 4.10: Longitude angle subtended by the landing perch on a bird's visual field along the trajectory - Drogon and Charmander flights. The figure shows the minimum and maximum longitude angles registered for the landing perch in the bird's visual field against $\hat{y}_{perch}$, the normalised distance to the perch in the y direction. Each subfigure shows the results across all flights per bird, type of trial and leg of the trial. Data from the same flight are shown in the same colour, and data outside the range of frames of interest are shown in grey (see Section 4.4.3). The longitude angle subtended by the perch is the difference between the curves for the same flight. The number of trials n per group is indicated in the bottom-left corner of each subfigure.

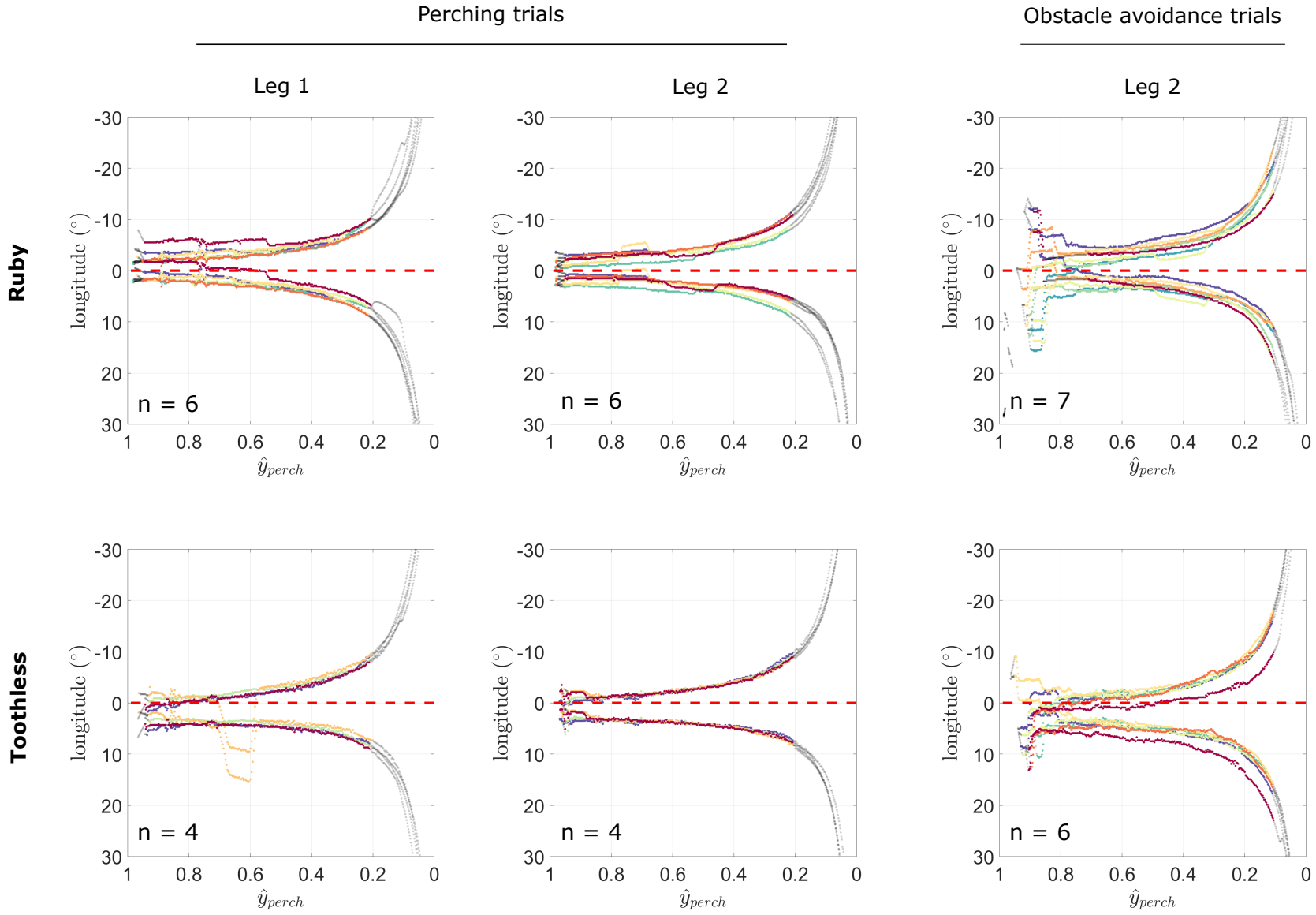


Figure 4.11: Longitude angle subtended by the landing perch on a bird's visual field along the trajectory - Ruby and Toothless flights. The figure shows the minimum and maximum longitude angles registered for the landing perch in the bird's visual field against $\hat{y}_{perch}$, the normalised distance to the perch in y direction. Each subfigure shows the results across all flights per bird, type of trial and leg of the trial. Data from the same flight are shown in the same colour, and data outside the range of frames of interest are shown in grey (see Section 4.4.3). The longitude angle subtended by the perch is the difference between the curves for the same flight. The number of trials n per group is indicated in the bottom-left corner of each subfigure.

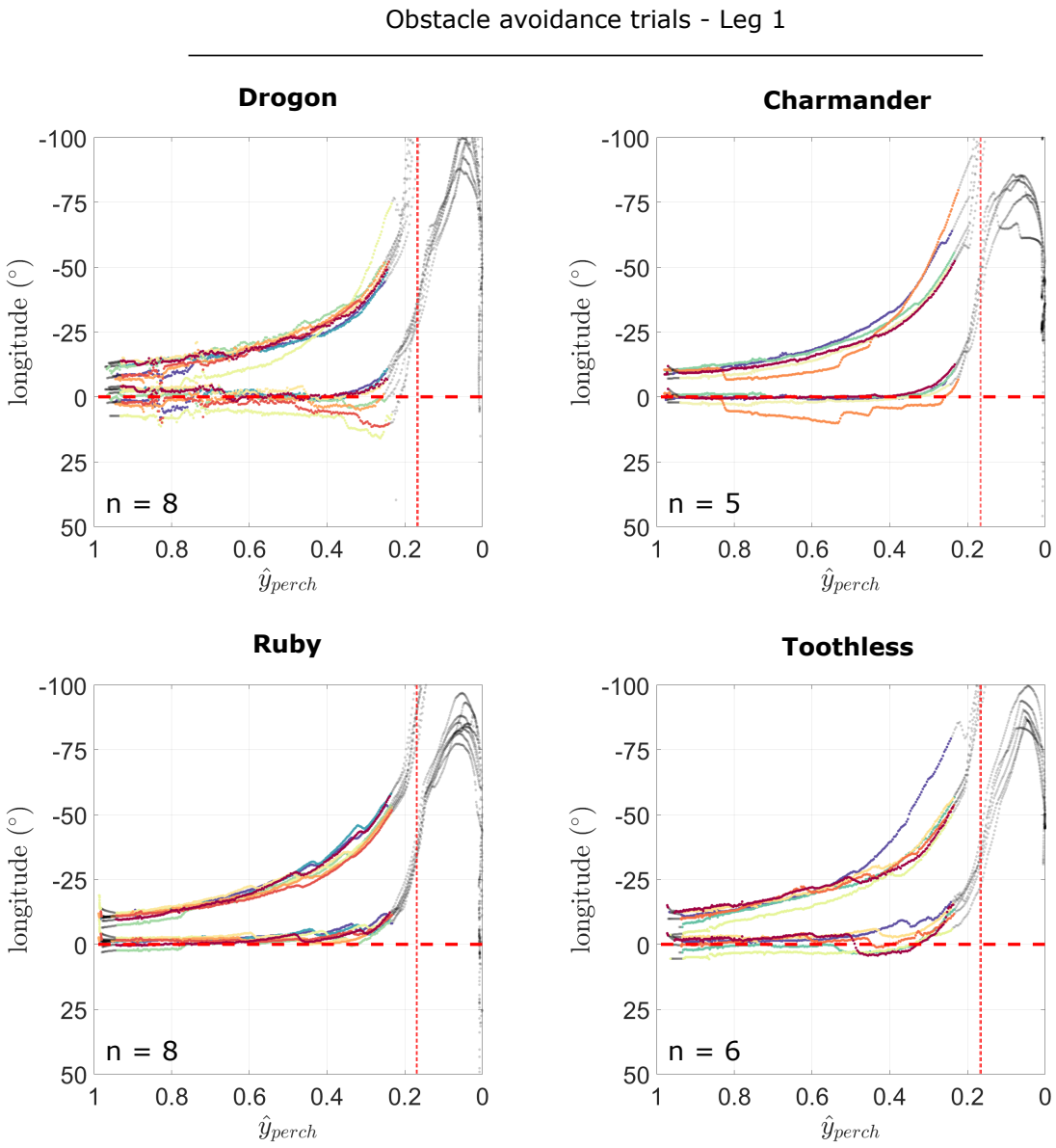


Figure 4.12: Longitude angle subtended by the obstacles on a bird's visual field along the trajectory. The figure shows the minimum and maximum longitude angles registered for the obstacles in the bird's visual field against $\hat{y}_{perch}$, the normalised distance to the perch in the y direction. Each subfigure shows the results across all flights per bird. Data from the same flight are shown in the same colour, and data outside the range of frames of interest are shown in grey (see Section 4.4.4). The longitude angle subtended by the obstacles is the difference between the curves for the same flight. For each flight, the point at which the bird passes the mean y-coordinate of the obstacles is marked with a $\hat{y} = constant$ line (red dashed lines). The number of trials n per group is indicated in the bottom-left corner of each subfigure.

4.4.3 Gaze strategy in perching

We first quantify how much the birds deviate from an ideal symmetric expansion of the landing perch, and then inspect possible factors affecting larger deviations.

4.4.3.1 Deviation from an ideal symmetric expansion of the landing perch

For each flight, we simulate an ideal symmetric expansion of the landing perch using the motion capture data, and assuming that at every frame during the flight: (1) the top edge of the landing perch is aligned with the 0° latitude plane of the bird's visual field, and (2) the estimated sagittal plane of the bird (i.e., the 0° longitude plane) is aligned with the bisector of the angle defined by the perch's edges and the bird's head. Note that this is not exactly the same as assuming the bird's sagittal plane is aligned with the perch's centre, but the difference is only relevant at very close distances to the perch (see Appendix C.7).

To assess the deviation from this ideal case, we compute for each flight the root-mean-square error (RMSE) between the longitude angles of the perch's edges, assuming the perch followed an ideal symmetric expansion, and the perch's minimum and maximum longitude angles per frame, derived from the semantic data. We focus the analysis on a range of frames of interest per flight, which cover the period before the final landing phase. This is because from Figures 4.10 and 4.11, and from previous reports in the literature (Potier et al., 2016; Kress et al., 2015) we expect the landing perch to be very close to centred relative to the sagittal plane towards the end of the landing manoeuvre; indeed, we assume this in estimating the visual coordinate system for each bird (see Section 4.3.2).

Specifically, we focus on the range of frames from takeoff until the bird is at a normalised distance to the perch of $\hat{y}_{perch} = 0.1$ (approximately 0.9 m). We exclude from this subset the frames in which the perch is partially occluded by the obstacles, the frames in which the bird's head transform was interpolated, and the frames used to estimate the bird's gaze direction $\vec{v}_{gaze}$ (see Section 4.3.2). The range of frames of interest is highlighted with coloured markers in Figures 4.10 and 4.11.

Figure 4.13 shows the results for the perching flights, and Figure 4.14 for the obstacle avoidance flights. In both figures, the flights with RMSE values closer to the 10th, 50th and 95th percentile are shown, along with a histogram of the RMSE values over all flights. Note how the representation of the longitude angle subtended by the landing perch against time (Figures 4.13 and 4.14) is qualitatively similar to the representation against the normalised distance to the landing perch $\hat{y}_{perch}$ in Figures 4.10 and 4.11.

Figure 4.14 shows that the distribution of RMSE for the obstacle flights is shifted to higher values, compared to the perching flights (Figure 4.13); the mean RMSE over the obstacle flights is 4.9°, which contrasts with a mean value of 1.6° for the perching flights. When grouping the flights by bird, type of trial and leg of the trial, the results show the same pattern: all birds deviate more from an ideal perch-centring strategy in the obstacle avoidance flights than in the perching ones, in the range of frames considered (see Table 4.4). The larger mean RMSE over Toothless' perching flights executed in leg 1 of the trials is due to a large head saccade in one of the flights, that temporarily places the landing perch far from the bird's sagittal plane (see yellow curve in the bottom-left panel of Figure 4.11).

The sampled flights in Figures 4.13 and 4.14 suggest that larger deviations from the ideal case often involve periods in which the bird's sagittal plane is aligned with one of the perch's edges. For example, the curves for the 95th percentile in Figure 4.13 suggest the bird fixated frontally the right edge of the perch (shown in red) for around 100 frames (500 ms) after takeoff. In Figure 4.14, the curves corresponding to the 10th and 50th percentiles show the left edge of the perch (shown in blue) aligned with the bird's sagittal plane for approximately 100 frames (500 ms), shortly after the perch is fully visible. The 95th percentile curves in Figure 4.14 show first a period of frames in which neither the centre nor the edges of the perch appear aligned with the bird's sagittal plane; then a short period in which the right edge is frontally fixated; and finally a period in which the perch appears centred. We inspect further the periods of alignment with the perch's centre or edges in the next section.

Although the flights executed in leg 1 of the obstacle avoidance trials are not the focus of the main text, we include a figure equivalent to Figure 4.14 for that data in Appendix C.8. In these flights, the edge-aligning behaviour is less evident, presumably due to the obstacles being visible from the perch well before takeoff, and only for a short period during flight (see Figure 4.9). We additionally computed the mean position of the landing perch centre in the bird's visual field in latitude and longitude, for the range of frames of interest; these show a similar pattern between perching and obstacle avoidance flights as the RMSE distribution. The results are included for reference in Appendix C.9.

Table 4.4: Mean root-mean-square error (RMSE) of the landing perch's longitudinal extension relative to an ideal symmetric case, over all flights per bird. The table shows the mean RMSE values over all flights with the same bird, type of trial and leg of the trial. The standard deviation σ and the total number of flights per group n is also reported. Note that the sample size for Charmander's obstacle avoidance flights is reduced to $n = 1$, due to excluding the flights in which the bird flies over the obstacles.

	Perching trials		Obstacle avoidance trials
Bird	Leg 1	Leg 2	Leg 2
Drogon	$1.5°$ ($\sigma = 0.8°$, $n = 7$)	$1.5°$ ($\sigma = 0.8°$, $n = 7$)	$6.5°$ ($\sigma = 2.8°$, $n = 8$)
Charmander	$1.5°$ ($\sigma = 0.3°$, $n = 7$)	$1.8°$ ($\sigma = 1.0°$, $n = 7$)	$4.2°$ ($\sigma = 0.0°$, $n = 1$)
Ruby	$1.4°$ ($\sigma = 1.3°$, $n = 6$)	$1.5°$ ($\sigma = 0.6°$, $n = 6$)	$4.3°$ ($\sigma = 1.1°$, $n = 7$)
Toothless	$2.9°$ ($\sigma = 2.0°$, $n = 4$)	$1.3°$ ($\sigma = 0.3°$, $n = 4$)	$3.5°$ ($\sigma = 1.3°$, $n = 6$)

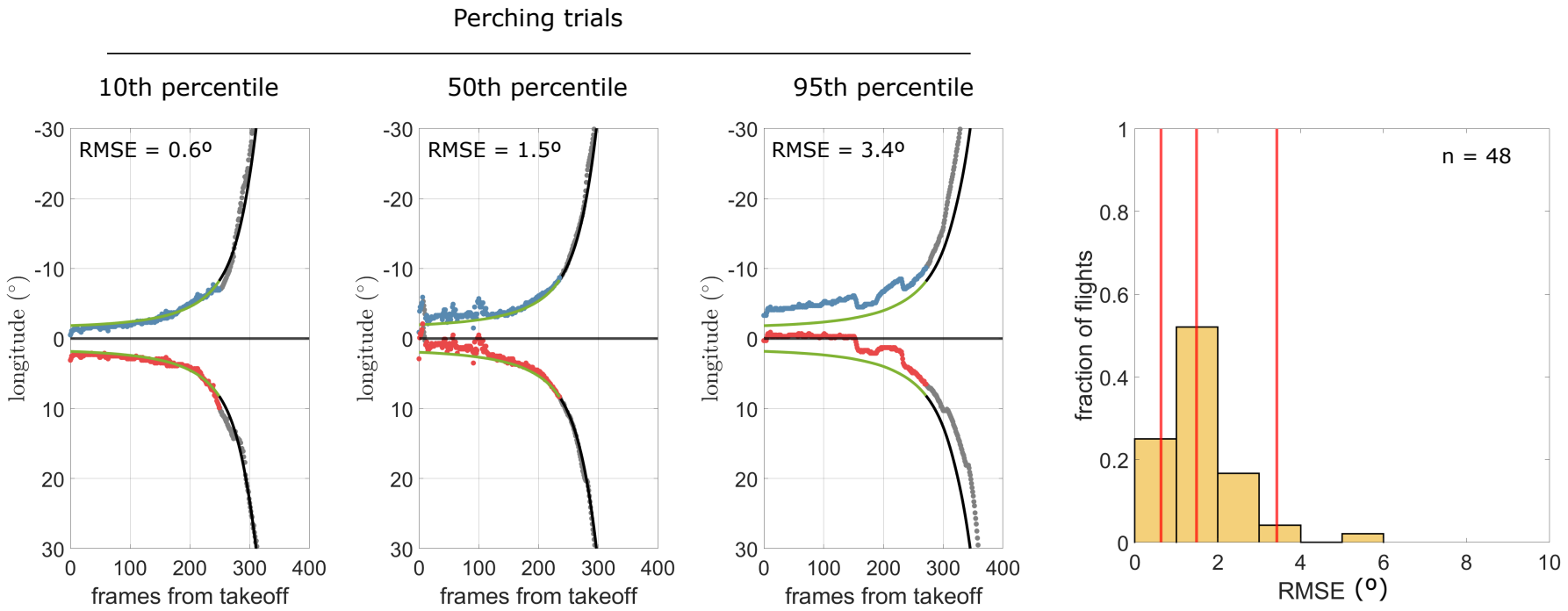


Figure 4.13: Deviation from an ideal symmetric expansion of the landing perch - Perching flights. We inspected how much each perching flight deviated from an ideal symmetric expansion of the landing perch in longitude (i.e., azimuth) direction. We characterised the deviation with the root-mean-square error (RMSE); the distribution of its value across the $n = 48$ flights analysed is shown (right panel). The three panels from the left show the minimum (blue) and maximum (red) longitude angles registered for the perch per frame, for flights with a RMSE value corresponding to the 10th, 50th and 95th percentile. The edges of the perch in an ideal symmetric expansion case are also shown (green). Only the samples within the subset of frames of interest are in colour; the samples outside this range are shown in grey for the data derived from the semantic maps, and black for the ideal case. The location of the selected percentiles on the histogram is marked with red vertical lines.The sampling frequency is 200 Hz.

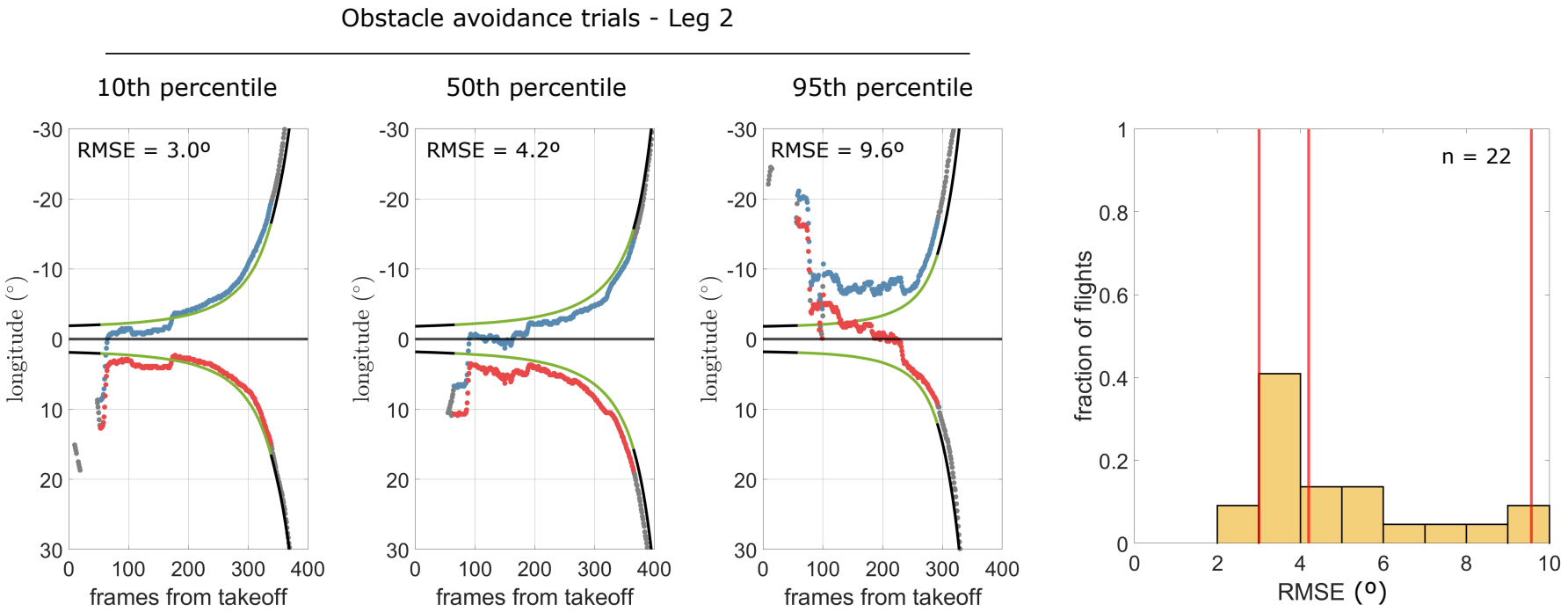


Figure 4.14: Deviation from an ideal symmetric expansion of the landing perch - Obstacle avoidance flights. We inspected how much each obstacle avoidance flight deviated from an ideal symmetric expansion of the landing perch in longitude (i.e., azimuth) direction. We characterised the deviation with the root-mean-square error (RMSE); the distribution of its value across the $n = 22$ flights analysed is shown (right panel). The three panels from the left show the minimum (blue) and maximum (red) longitude angles registered for the perch per frame, for flights with a RMSE value corresponding to the 10th, 50th and 95th percentile. The edges of the perch in an ideal symmetric expansion case are shown (green). Only the samples within the subset of frames of interest are in colour; the samples outside this range are shown in grey for the data derived from the semantic maps, and black for the ideal case. The location of the selected percentiles on the histogram is marked with red vertical lines. The sampling frequency is 200 Hz.

4.4.3.2 Alignment of gaze with the landing perch's centre and edges

We compute for each trial the fraction of frames in which the bird's sagittal plane is aligned with the perch's centre, or with a perch edge, in the range of frames of interest (see definition in Section 4.4.3.1, and colour markers in Figures 4.10 and 4.11). To do so, we registered for each frame the perch feature (its centre or its edges) that was closest in longitude angle to the sagittal plane, and less than 3° from it; we selected this maximum deviation as a rough estimate of the range of eye motion in Harris' hawks (O'Rourke et al., 2010; Kane et al., 2015). For each flight, we focused the analysis on the edge that was fixated for longer; from visual inspection we did not find examples in which both edges appeared aligned with the bird's sagittal plane within the same flight.

We first inspect whether a larger deviation from the ideal symmetric expansion of the perch is correlated with longer periods aligned with one of the perch's edges; this is shown in Figure 4.15. The correlation coefficients (ρ) suggest there is a moderate positive correlation between the two magnitudes, in both perching and obstacle avoidance flights (perching flights: $\rho = 0.66$, $p < 0.01$, 95%CI $= [0.46, 0.80]$; obstacle avoidance flights: $\rho = 0.51$, $p = 0.01$, 95%CI $= [0.12, 0.77]$).

We then inspect the distribution of the duration of the edge-alignment periods and centre-alignment periods, across all flights. This is shown in Figure 4.16. In the majority of perching flights (over 56% of them), the bird's sagittal plane is aligned with the perch's centre for over 80% of the frames (top-left panel), but with the perch's edge less than 20% of the time (bottom-left panel). In contrast, in the obstacle avoidance flights, both distributions become more uniform: the fraction of frames in which the bird's sagittal plane is aligned with the perch's centre reduces to 49.7% on average (bottom-left panel), and the mean period aligned with the landing perch's edge increases to 34.1% (bottom right panel). Since the distributions are not close to normal, we report their medians and interquartile ranges to better characterise their shape in Table 4.5.

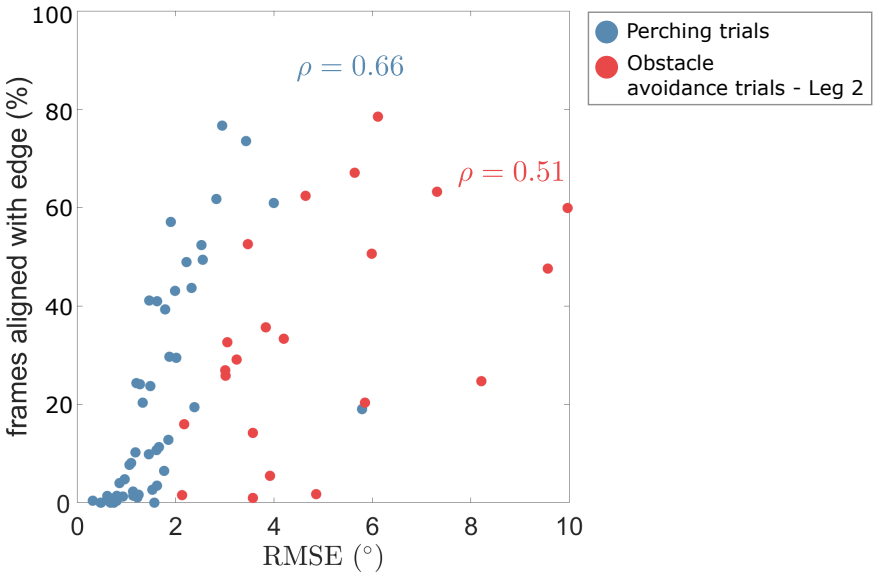


Figure 4.15: Fraction of frames aligned with one of the perch's edges per flight, against the RMSE relative to the ideal symmetric expansion of the perch. The figure represents for each flight analysed, the fraction of frames in which the bird's sagittal plane is aligned with one landing perch's edges against the root-mean-square error measuring the deviation from an ideal symmetric expansion of the landing perch. Perching flights (blue, $n = 48$ flights) and obstacle avoidance flights (red, $n = 22$ flights) are shown, as well as the correlation coefficient for each case.

Table 4.5: Median and interquartile range for the fraction of frames (%) aligned with the landing perch's centre, and the fraction of frames (%) aligned with a landing perch's edge, over all flights. The table shows the median and interquartile range (IQR) for the distribution of fraction of frames aligned with the landing perch's centre, and a landing perch's edge per flight. Flights across all birds are aggregated and the total number of flights per group n is reported. The obstacle avoidance flights are all executed in leg 2 of the trials.

	Perching flights	Obstacle avoidance flights
frames aligned with centre	89 % (IQR = 40 %, $n = 48$)	60 % (IQR = 48 %, $n = 22$)
frames aligned with edge	10 % (IQR = 39 %, $n = 48$)	31 % (IQR = 37 %, $n = 22$)

The median values across flights grouped by bird, type of trial and leg of the trial are show in Tables 4.6 and 4.7. Since the sample size per group is small, we report the minimum and maximum values observed rather than the interquartile range (see values in brackets). Overall, the median values per group show a similar pattern to the median values obtained when aggregating all flights per trial type. For example, the median number of frames aligned with the perch's centre is over 85% in 6 out of the 8 groups of perching flights, and over 59% in 3 out of 4 groups

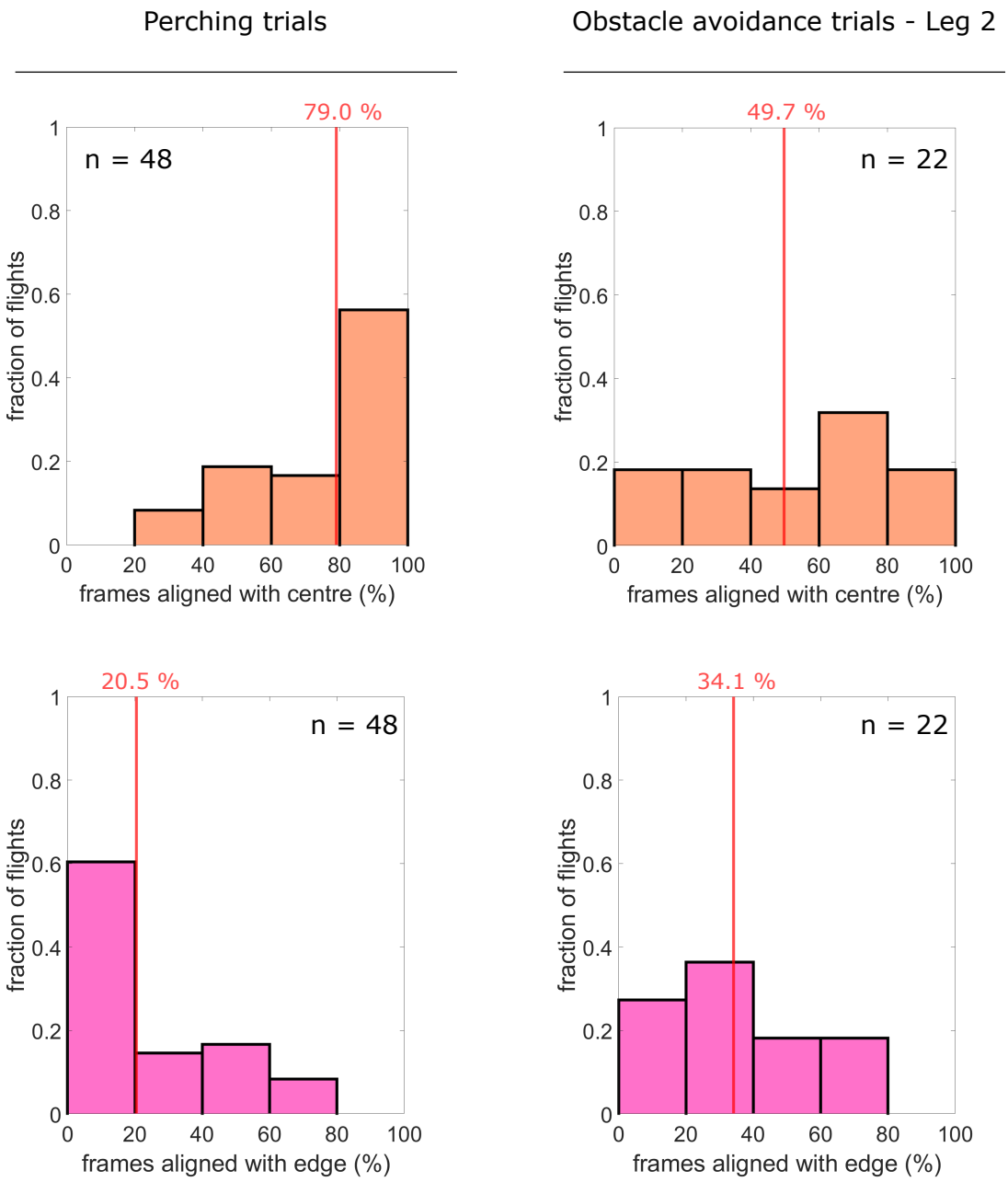


Figure 4.16: Histograms for the fraction of frames aligned with the landing perch's centre and the landing perch's edge per flight. The figure shows histograms for the fraction of frames aligned with the landing perch's centre (top row panels) and the landing perch's edge (bottom row panels) per flight. The data is grouped by perching flights (left panels) and obstacle avoidance flights executed in leg 2 of the trials (right panels). For each histogram, the mean value is highlighted with a red vertical line and the total number of flights n considered is indicated at the top. The median and interquartile range for each distribution are reported in Table 4.5.

Table 4.6: Median fraction of frames (%) aligned with the landing perch's centre, across all flights per bird. We computed for each flight the fraction of frames for which the landing perch's centre is closest and within ±3° from the estimated sagittal plane of the bird, in longitudinal direction; the table shows the median value over all flights per bird, trial type, and trial leg. We also report the minimum and maximum values observed per group (in brackets) and the number of flights n. Note that the sample size for the obstacle avoidance flights by Charmander is reduced to $n = 1$, due to excluding the flights in which the bird flies over the obstacles.

	Perching trials		Obstacle avoidance trials
Bird	Leg 1	Leg 2	Leg 2
Drogon	87 % [39, 99] ($n = 7$)	96 % [50, 100] ($n = 7$)	22 % [5, 94] ($n = 8$)
Charmander	97 % [71, 99] ($n = 7$)	70 % [23, 100] ($n = 7$)	59 % [59, 59] ($n = 1$)
Ruby	94 % [39, 100] ($n = 6$)	89 % [57, 100] ($n = 6$)	70 % [8, 92] ($n = 7$)
Toothless	59 % [48, 76] ($n = 4$)	86 % [59, 95] ($n = 4$)	64 % [0, 82] ($n = 6$)

Table 4.7: Median fraction of frames (%) aligned with the landing perch's edge, across all flights per bird. We computed for each flight the fraction frames for which the landing perch's edge is closest and within ±3° from the estimated sagittal plane of the bird, in longitudinal direction; the table shows the median value over all flights per bird, trial type, and trial leg. We also report the minimum and maximum values observed per group (in brackets) and the number of flights n. Note that the sample size for the obstacle avoidance flights by Charmander is reduced to $n = 1$, due to excluding the flights in which the bird flies over the obstacles.

	Perching trials		Obstacle avoidance trials
Bird	Leg 1	Leg 2	Leg 2
Drogon	11 % [1, 62] ($n = 7$)	4 % [0, 49] ($n = 7$)	49 % [2, 63] ($n = 8$)
Charmander	3 % [1, 30] ($n = 7$)	30 % [0, 77] ($n = 7$)	33 % [33, 33] ($n = 1$)
Ruby	6 % [0, 61] ($n = 6$)	10 % [0, 43] ($n = 6$)	14 % [1, 67] ($n = 7$)
Toothless	37 % [20, 52] ($n = 4$)	14 % [5, 41] ($n = 4$)	31 % [16, 78] ($n = 6$)

of obstacle avoidance flights. The median number of frames aligned with the perch's edge is under 15% for 6 out of the 8 groups of perching flights, and over 30% in 3 out of 4 groups of obstacle avoidance flights. These are all comparable to the median values obtained when grouping flights by trial type (see Table 4.5). However, we also find variation across groups, possibly indicative of differences between birds (Tables 4.6 and 4.7). A Wilcoxon signed-rank paired test found no significant differences in the median fraction of frames aligned with the perch's edge or the perch's centre between perching flights executed at leg 1 or leg 2 of the trials (alignment with perch edge: $p = 1$, $T = 150$; alignment with perch centre:

$p = 0.94$, $T = 147$). However, a larger sample size would be required to further inspect if the differences across birds are significant.

Finally, we inspected which edge was most frequently fixated across trials. We found that in 83.3% of the perching flights, the edge fixated for longer was also the closest to the sagittal plane on average over the first 50 frames of visibility of the perch. For the obstacle avoidance flights, this occurred in 86.4% of the cases.

4.4.4 Gaze strategy around obstacles

We now inspect how much the birds deviate from an ideal alignment with the obstacles' inside edge per flight, and estimate the instant in the trajectory at which the bird no longer fixates the edge of the obstacles frontally. We explore across birds whether these instants correspond to similar distances to the obstacles, and similar time periods before passing around them.

4.4.4.1 Deviation from an ideal alignment with the obstacles' inside edge

For the ideal gaze strategy around the obstacles, we simply assume the bird's sagittal plane is aligned with the inside edge of the obstacles for the whole flight. Note that this is not what the birds do: they always stop aligning with the obstacles' edge frontally at some point before passing them (see Figure 4.12). Since the period in which the bird is no longer fixating the obstacles' edge will only add error to the RMSE metric, we limit the range of frames in our analysis to the set from takeoff up to 25 frames (0.125 s) before the obstacles are passed, excluding the frames in which the head pose of the bird was interpolated. We select 25 frames prior to the point of passing the obstacles as an estimate of the time it takes the bird to fly a distance of one body length (around 60 cm), at an approximate speed of 5 m/s. We highlighted the range of frames of interest with colour markers in Figure 4.12.

Figure 4.14 shows the flights with the RMSE closest to the 10th, 50th, 95th and 100th percentile, and a histogram of the values computed for the $n = 27$ flights considered. Note how the representation of the longitude angle subtended by the obstacles against time in Figure 4.17 is qualitatively similar to the representation

against the normalised distance to the landing perch $\hat{y}_{perch}$ in Figure 4.12. As we did in Figure 4.12, we mirrored the required data relative to the y-axis of the plot in Figure 4.17, so that the edge of the obstacles that is closest to the turn is always the edge of maximum longitude angle.

The flight corresponding to the 50th percentile suggests there is a reasonable agreement with the ideal case for at least half of the flights; however we also observe deviations that could be interpreted as saccadic realignments (see the period between frames 150 and 200 in Figure 4.17, 50th percentile panel). The flights on the high end of the distribution (95th and 100th percentile) suggest that for a few flights the birds may be fixating on other features during the flight. For example, the flight with the maximum deviation from the ideal case (the 100th percentile) suggests the bird initially fixated on the edge of the obstacles that was furthest from the turn (shown in blue), and then fixated on a feature between the edges of the obstacles (note that the inside edge, shown in red, is not exactly constant in longitude angle in the period between 100 and 200 frames after takeoff). The mean RMSE over all flights is 3.4°, which is comparable to the mean results for the flights grouped by bird (see Table 4.8). We additionally computed the mean position of the midpoint of the obstacles' inside edge in the bird's visual field, in latitude and longitude, for the range of frames of interest. The results are included for reference in Appendix C.10.

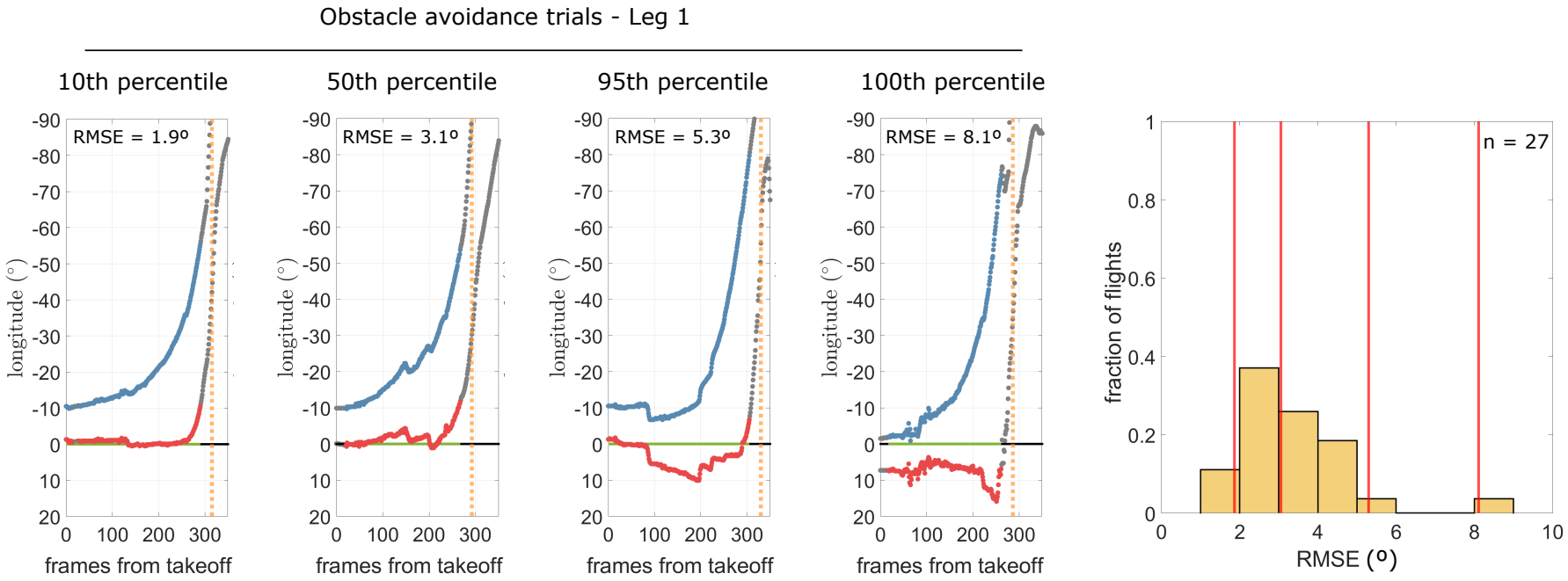


Figure 4.17: Deviation from an ideal alignment with the obstacles' inside edge. We inspected how much each flight deviated from the ideal case in which the bird's sagittal plane is always aligned with the obstacles' inside edge. We characterised the deviation with the root-mean-square error (RMSE); the distribution of its value across the $n = 27$ flights analysed is shown (right panel). The four panels from the left show the minimum (blue) and maximum (red) longitude angles registered for the obstacles per frame, for flights with a RMSE value corresponding to the 10th, 50th, 95th and 100th percentile. The inside edge of the obstacles in an ideal edge-fixation case is also shown (green). Only the samples within the subset of frames of interest are in colour; the samples outside this range are shown in grey for the data derived from the semantic maps, and black for the ideal case. The frame at which the obstacles are passed in the y direction is shown (yellow vertical dashed line). The location of the selected percentiles on the histogram is marked with red vertical lines. The sampling frequency is 200 Hz.

Table 4.8: Mean root-mean-square error (RMSE) between the longitude angle of the obstacles' inside edge and the ideal case, over all flights per bird. We computed, for each flight, the RMSE between the longitude angle of the obstacles' inside edge, and the ideal case in which the bird's sagittal plane is always aligned with the obstacles' inside edge. The table shows the mean RMSE values over all obstacle avoidance flights executed in leg 1, grouped by bird. The standard deviation σ and the total number of flights per group n is also reported.

Bird	Obstacle avoidance trials - Leg 1
Drogon	$3.5°$ ($\sigma = 2.1°$, $n = 8$)
Charmander	$2.7°$ ($\sigma = 1.5°$, $n = 5$)
Ruby	$3.3°$ ($\sigma = 0.7°$, $n = 8$)
Toothless	$3.9°$ ($\sigma = 0.8°$, $n = 6$)

4.4.4.2 Estimation of the loss of fixation point

To estimate the point in time at which the birds no longer frontally fixate the obstacles' edge, we inspect for each flight when the inside edge of the obstacles substantially deviated from a constant longitude angle. For this analysis, we consider the range of frames from takeoff until the obstacles are passed in the y direction (marked with a yellow vertical dashed line in Figure 4.17), excluding the frames in which the head pose was interpolated.

We use the M-estimator sample consensus (MSAC) algorithm, an iterative method for model-fitting in the presence of outliers. The algorithm determines the maximum number of samples, or 'inliers', that are within a certain distance of a model. In our case, we use a constant longitude angle model and a maximum distance of $1.5°$. This maximum distance was selected based on trial-and-error over a subset of flights, comparing the results obtained to the point we would assign from visual inspection. We determine the instant of loss of fixation as the sample with the largest frame number within the computed inliers. We explore this approach as a potential option to investigate fixations in the bird's visual field, as it has the advantage of imposing that a fixation is sustained in time, and it does not require the fixations to be closer than a certain threshold to the bird's sagittal plane. However, for 89% of the flights, the mean longitude angle of the detected inliers per flight is within $3°$ from the estimated sagittal plane (see Figure 4.18).

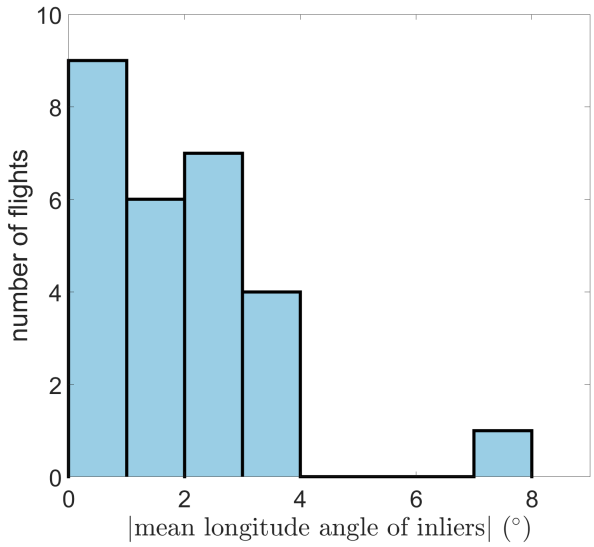


Figure 4.18: Mean longitude angle of the MSAC inliers' across all flights (in absolute value). The MSAC algorithm computes for each flight the maximum number of samples in the obstacles' inside edge longitude angle curve that are within 1.5° from a constant longitude angle model. The figure shows the distribution of the inliers' mean longitude angle per flight, in absolute value, for $n = 27$ flights.

Figure 4.19 shows the results for the loss of fixation points estimated per flight, in terms of longitude angle and frames before the obstacles are passed. As in previous sections, the data from the required flights is mirrored relative to the y-axis of the plot for clarity. Figure 4.20 depicts their location along the trajectory, in a coordinate system fixed to the obstacles. The obstacles' coordinate system is defined per flight, with its origin at the midpoint of the obstacles' top markers and parallel to the motion capture coordinate system. Table 4.9 reports the mean perpendicular distance to the obstacles and the mean time period ahead of passing the obstacles at the loss of fixation point, grouped by bird. The large variability across birds may suggest that the method is not suitable for the birds' different strategies. We observed, for example, that Drogon often reoriented its head around 50 frames before passing the obstacles, potentially resulting in an estimated loss of fixation point that appears 'too late' (see Figure 4.19). On average across all birds, the loss of fixation point occurs at a distance of 1.51 m, and 271.7 ms (54.3 frames) before passing the obstacles.

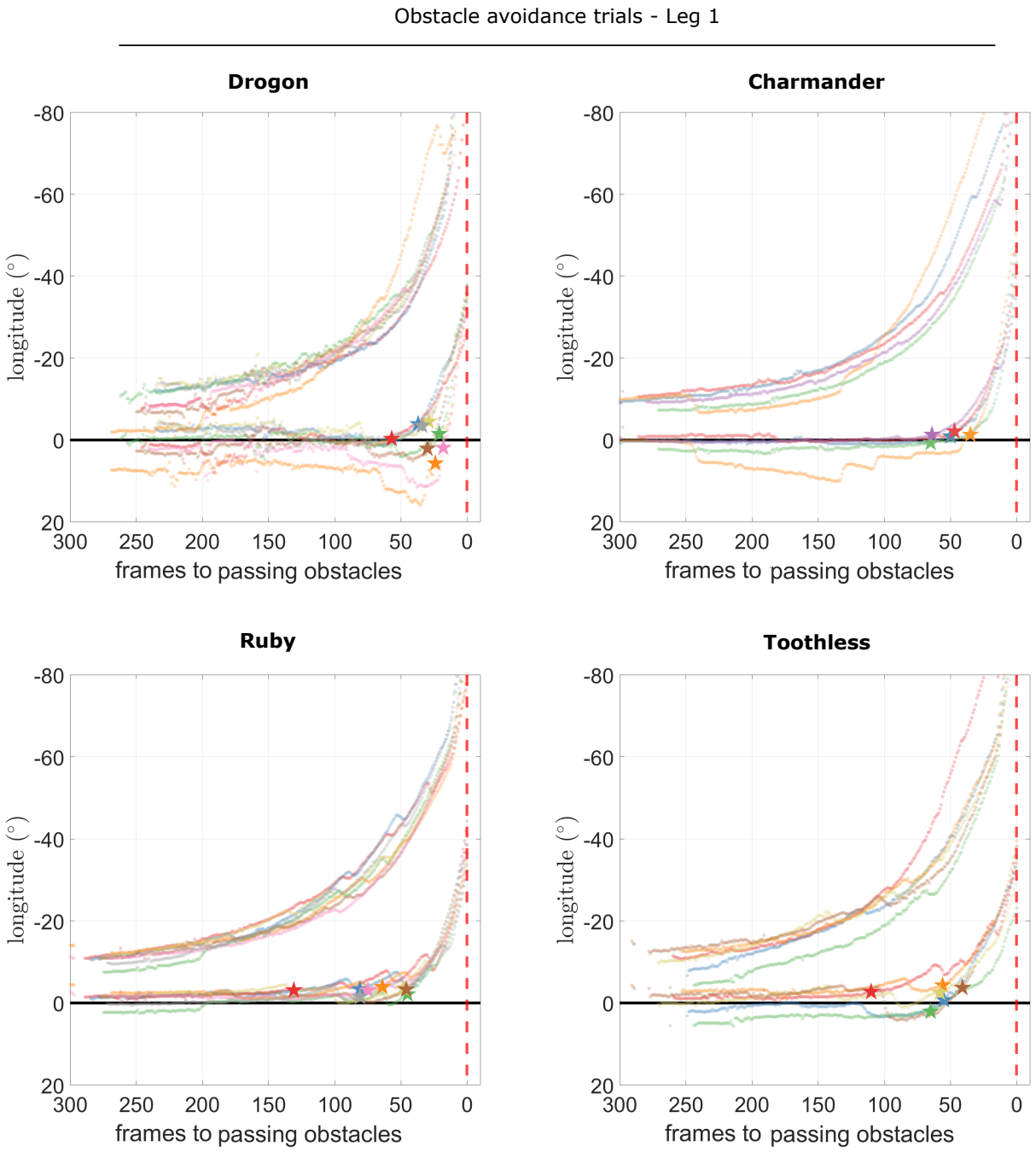


Figure 4.19: Estimated loss of fixation point per flight. The longitude angles of the inside and outside edges of the obstacles are represented for each flight executed in leg 1 of the obstacle avoidance trials, against the number of frames until the obstacles are passed. The flights are grouped by bird. The estimated loss of fixation point per flight is shown (star markers) on the inside edge curves, and the point of passing the obstacles is highlighted (red dashed vertical line). The sampling frequency is 200 Hz.

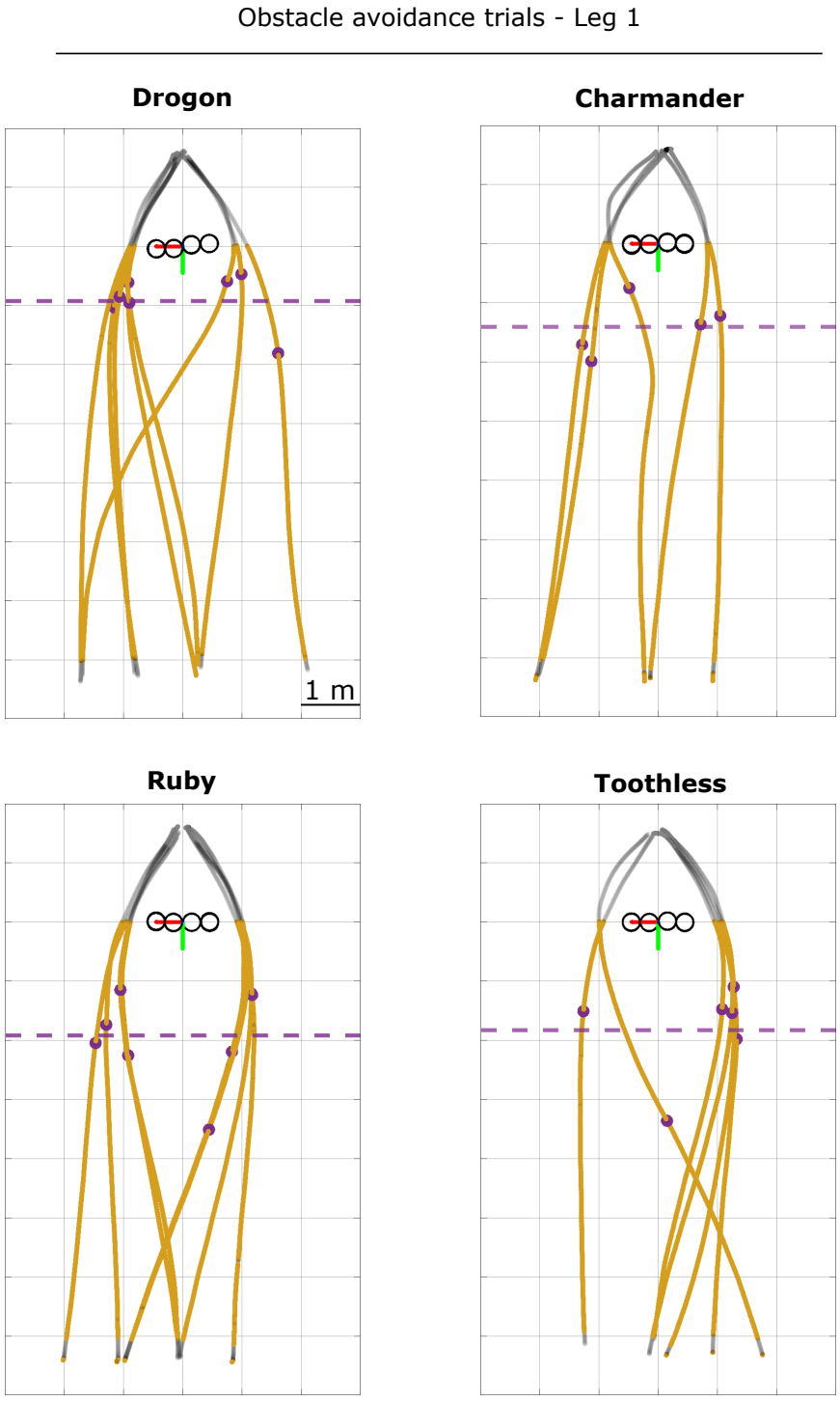


Figure 4.20: Loss of fixation point on the birds' trajectories. The trajectories of each bird around the obstacles are represented in the obstacles' coordinate system (x-axis shown in red, y-axis shown in green). The estimated loss of fixation point per flight is highlighted (purple marker) as well as their average y-coordinate (purple dashed line). The range of frames of interest is shown in yellow, and samples outside this range are shown in grey. The grid is represented at 1 m spacing. The obstacles' positions (black white-filled circles) are included for reference.

Table 4.9: Mean perpendicular distance to the obstacles, and mean time period before passing the obstacles at the estimated loss of fixation point, over all flights per bird. For all flights executed in leg 1 of the obstacle avoidance trial we estimated the loss of fixation point. The table shows the mean perpendicular distance to the obstacles $\bar{y}_{LOF}$ and the mean time period ahead of passing the obstacles $\bar{\Delta}t_{LOF}$ at the estimated points of loss of fixation (LOF), across all flights per bird. The mean values across all flights are also included. The standard deviation σ and the total number of flights per group n is reported.

Bird	$\bar{y}_{LOF}$ (m)	$\bar{\Delta}t_{LOF}$ (ms)	n
Drogon	0.93 ($\sigma = 0.42$)	156.3 ($\sigma = 61.0$)	8
Charmander	1.40 ($\sigma = 0.47$)	261.0 ($\sigma = 62.8$)	5
Ruby	1.92 ($\sigma = 0.79$)	356.9 ($\sigma = 142.6$)	8
Toothless	1.83 ($\sigma = 0.80$)	320.8 ($\sigma = 118.9$)	6
All flights	1.51 ($\sigma = 0.74$)	271.7 ($\sigma = 129.3$)	27

4.5 Discussion

In this chapter, we carried out a study of the gaze strategy used by Harris' hawks when executing perching and obstacle avoidance manoeuvres. We used detailed semantic descriptions of the visual scene the birds experience while executing these manoeuvres, based on motion capture data from experiments with birds. The results from the behavioural analysis showed how the generated synthetic data can provide a unique insight into the birds' behaviour. Additionally, the focus on a subset of the data allowed us to consolidate the methods of analysis, and pinpoint specific hypotheses to test with the full dataset in the future.

Head stabilisation

First, our results showed that the birds stabilised their heads in flight. The spherical histograms in Figure 4.6 ad Appendix C.5 clearly illustrate this, by contrasting the visual coordinate system, which represents the bird's visual field orientation, and the trajectory coordinate system, which did not consider the bird's head orientation. Head stabilisation is a well-known mechanism in birds (Land et al., 2012; Theunissen et al., 2017; Warrick et al., 2002; Pete et al., 2015), and it is key for the extraction of information from optic flow, such as depth or object segmentation (Eckmeier et al., 2008; Eckmeier et al., 2013; Raudies et al., 2012; Ravi et al., 2022). In fast

manoeuvres such as the ones studied here, birds are thought to heavily rely on visual cues derived from optic flow (Davies et al., 1990; Lee et al., 1993; Bhagavatula et al., 2011; Eckmeier et al., 2008) over other mechanisms such as stereopsis.

Distinct gaze strategies around landing perch and obstacles

Second, we found that the birds used distinct gaze strategies when approaching the landing perch and the obstacles. Birds primarily centred the landing perch relative to their sagittal planes, thereby producing a bilateral looming stimulus in their visual fields. In contrast, birds frontally fixated the obstacles' edge when turning around them, thereby leading to unilateral expansion in their visual fields. This suggests that the birds adapted their gaze strategy for goals (or elements to intercept) and obstacles (or elements to avoid collision with), as has been suggested in humans (Raudies et al., 2012; Rothkopf et al., 2009; Rothkopf et al., 2007), and presumably to aid visual computation during navigation (Eckmeier et al., 2013; Raudies et al., 2012; Rothkopf et al., 2009).

Landing perch centring

Perch centring in the final landing phase is thought to aid computing perch proximity (Kress et al., 2015; Lee et al., 1993; Davies et al., 1990). This behaviour induces visual cues that have been reported to support a controlled landing: the size and the rate of expansion of the landing perch on the retina have been linked to the control of braking (Lee et al., 1993), to the timing of foot extension (Davies et al., 1990) and the timing of downward tail pitching (Kress et al., 2015). Electrophysiological measurements on pigeons found evidence of a subpopulation of neurons responding selectively to looming objects, and maximally so when the ratio between the rate of expansion of the object and its retinal size was constant (Wang et al., 1992). This ratio, sometimes called τ or RREV (relative retinal expansion velocity), provides a first-order approximation to the time-to-contact and is thought to be used by many animals to control approach manoeuvres (Wagner, 1982; van Breugel et al., 2012; Ravi et al., 2022), including landing in Harris' hawks (Davies et al., 1990).

In this chapter, we built upon this well-studied behaviour to estimate the visual coordinate system of the bird in its final approach to the perch. The data suggest this is a solid assumption (as shown in Figure 4.4), but it does introduce a circularity issue: we will inevitably find the bird centres the perch with respect to its sagittal plane just before landing. However, the fact that this perch-centering behaviour was also observed in the period before the final landing phase, and especially the fact that the perch was held level throughout the flight, suggests that our method of calibrating the bird's visual coordinate system is reasonably accurate.

Frontal fixation of the obstacles' edge

However, stronger support of our visual coordinate system calibration is that birds frontally fixated the inside edge of the obstacles during flight; similar behaviours around obstacles have been reported in humans walking in virtual environments (Rothkopf et al., 2009; Rothkopf et al., 2007), in lovebirds (Kress et al., 2015), in zebra finches (Eckmeier et al., 2013) and in bumblebees (Ravi et al., 2022). This strategy is thought to aid obstacle segmentation based on motion contrast: by fixating the edge of the obstacle as the animal approaches it, the contrast between the motion of the background and the obstacle in the retina increases (Raudies et al., 2012; Ravi et al., 2022; Eckmeier et al., 2013). Neural recordings of zebra finches seem to support this hypothesis: when zebra finches watched an optic flow reconstruction of an obstacle avoidance flight, individual neurons in the nucleus rotundus peaked when the point of expansion of the optic flow field matched the obstacle's position (Eckmeier et al., 2013). The authors also identified neurons that increased firing when the obstacle was laterally passed. However, in both cases, the neurons identified were sparse and the majority did not respond to the obstacles, which suggests that further research is needed. Frontally-fixating the edge of an obstacle to turn around it may also serve to estimate the curvature of the turn, as has been suggested for the case of humans steering a car. In this situation, humans consistently fixate on the inside edge of the bend, specifically on the stationary point of the curve (Land et al., 1994; Mole et al., 2021), presumably to predict

the curvature of the road ahead from the angle between the velocity vector and the gaze direction. This could be comparable to the behaviour we observed in the birds. The visual expansion of the obstacles unilaterally in the bird's visual field could also be involved in mediating the steering away response, similarly to the collision-avoidance responses observed in fruit flies (Tammero et al., 2002).

In this chapter, we have focused on the differences between the birds' gaze strategies around the landing perch and the obstacles based on how their extension in longitude angle varies along the trajectory. However, if the birds are fixating the top edge of the landing perch, as the results suggest (see Figure 4.7), they may be doing so to facilitate the perch's segmentation from the background, in line with the motion contrast hypothesis. We did not analyse this further in the current chapter, but it would be a relevant hypothesis to explore next with the optic flow data generated.

Deviations from perch centring

We also found that the birds deviated further and more frequently from a perch-centring strategy in the period prior to landing, when obstacles were present (1.5 m in front of the takeoff perch), than when obstacles were absent. On average, birds also spent more time aligned with one of the landing perch's edges, rather than its centre, in the obstacle flights, and the fixated edge was most frequently the closest to the sagittal plane when the perch was first visible. However, the results suggest there may be differences across birds; this should be inspected with a larger dataset. The behaviour of fixating on the edge of the perch closest to the sagittal plane prior to landing has been previously reported in lovebirds making sharp turns (Kress et al., 2015). Indeed, the difference we observe between flights with and without obstacles suggests that the strategy may be used by the birds when the trajectory requires more turning. As suggested by Rothkopf et al. (2009), fixating your gaze on a target you aim to steer to may serve to simplify the control strategy, by reducing it to servoing relative to the fixated frame; the birds may be doing this in steps using the edge of the perch as an intermediate target.

Limitations

The results of this chapter have allowed us to pinpoint specific hypotheses on the gaze strategies used by the birds, but the study has some limitations. The reduced number of flights per bird limits the power of most statistical tests. Extending the analysis to the full dataset would allow us to verify the results obtained with more confidence. For example, a larger dataset would allow us to inspect the behaviour of fixating on the landing perch's edge further, and test whether the differences observed between flights with and without obstacles are related to the turning demands of the trajectory, as we hypothesised. Note as well that in the analysis of the gaze strategy during perching, the flights with and without obstacles cover slightly different periods of the flight due to the differences in visibility of the landing perch along the trajectory (see Figure 4.9) and due to the exclusion of the frames used to estimate the visual coordinate system (a larger fraction of frames from the flights without obstacles was used to fit $\vec{v}_{gaze}$). Although we do not expect this to change the main conclusions of the results, because the obstacle flights include a larger period in which the bird is more likely to be centring the perch, it could be an aspect to revisit in follow-up analyses. Statistical tests on a larger dataset would also allow us to inspect whether the strategies across birds are significantly different.

Another important limitation of this study is the aforementioned circularity, introduced by calibrating the gaze direction against the midpoint of the perch. Even though the data supported the underlying assumption (see Figure 4.4), this calibration method limits what can be concluded about the birds' behaviour during perching. An alternative gaze calibration approach might be considered in future studies, potentially using a dedicated set of trials.

One of the aims of this study was to explore different methods to analyse gaze fixations. To study the fraction of frames in which the bird was aligned with the perch's edge or centre, we simply considered which of those features was closest to the sagittal plane of the bird per frame while also being within $\pm 3°$ from it; we selected this threshold based on a rough estimate of the range of eye

motion in the birds. This produced reasonable results for most flights, but the $\pm 3°$ threshold seemed broad in some cases; indeed we would expect eye movements in flight to be below this estimated maximum value (Ros et al., 2017b; Kano et al., 2018). In the same analysis, we did not require the fixations on the perch's edge or centre to be sustained in time either, but we did explore this option in the analysis of the fixation on the obstacles' edge. A recommended next step would be to consider a method that combines both requirements, for example by rejecting the fits provided by the MSAC algorithm whose intercept is beyond a certain threshold from the estimated sagittal plane.

The optic flow data generated per flight could also be used to inspect other possible points of fixation. In the analysis presented here, we focused on specific features that were highlighted in the spherical histograms across flights (i.e., the landing perch's centre and the obstacles' edge), but inspecting the minima in the optic flow field may point to other features fixated on by the birds. We carried out some preliminary analyses in this direction for a single flight, in which we found that the optic flow minima appeared most frequently aligned with the bird's estimated gaze direction; this was not the case when we considered the data from the trajectory coordinate system, which provides additional support to the estimated visual coordinate system (see further details in Appendix C.11). If the optic flow data is used in posterior analyses, it may be worth considering a smoothing algorithm for the head pose, since any computations based on differences across frames will likely increase the noise in the resulting signal. This could be achieved with a low-pass filter based on the SLERP algorithm for quaternion interpolation (Shoemake, 1985), as we did in Chapter 3 for the pursuit flight (see Appendix B.2.2). A total variation denoising method could also be useful to inspect saccades, as it preserves the sharp step-like changes in the signal.

The method used to estimate the loss of fixation point on the obstacles' edge also presents opportunities for improvement. We found that, for some birds, it matched the point we would estimate from visual inspection (see for example Charmander's flights in Figure 4.19), but less so for others (see for example Drogon's flights in

Figure 4.19). This may indicate that the birds are using slightly different strategies. One way to deal with this could be to define several different criteria to estimate the loss of fixation point and inspect the distribution of estimates obtained. Additionally, from the work by Eckmeier et al. (2013), it seems relevant to inspect when the focus of expansion of the optic flow field falls on the obstacles, and compare these instants to the loss of fixation point estimated.

Next steps

The results presented here open up several new avenues for follow-up studies. An interesting option would be to inspect whether changes in the gaze strategy reflect changes in the trajectory. We have already suggested some points in this direction, such as investigating whether the turning required by the trajectory relates to the edge-fixating behaviour in perching, or inspecting the angle between the gaze direction and the velocity vector. For these analyses, models such as the attractor-repeller one presented in Appendix E of this thesis (inspired by the work of Fajen et al., 2003), or the pursuit models used in Brighton et al. (2019) could be especially useful.

Incorporating the optic flow data in the analysis may also lead to relevant insights. For example, combining the optic flow and semantic data generated, we can investigate the motion contrast between an obstacle and its immediate background, as in Ravi et al. (2022). We could further inspect if this strategy is also used to facilitate the detection of the perch. Additionally, we could inspect how often the focus of expansion of the optic flow field coincides with the obstacles, following the conclusions from Eckmeier et al. (2013).

The data analysed in this chapter has provided a unique insight into the birds' behaviour and helped consolidate possible methods of analysis, but it only represents a small part of the potential of the dataset. Combining our results with the optic flow and depth data already generated, we will be able to inspect the gaze strategies used in flight in further detail. Moreover, the rendering pipeline could easily be extended to the full dataset of over 200 trials, for which the motion capture

data has been labelled and postprocessed. Further understanding the principles behind the birds' gaze strategies in flight may not only reveal important aspects about their sensory-motor decisions, but also inspire new vision-based strategies for autonomous flying systems.

References

Bhagavatula, P. S., Claudianos, C., Ibbotson, M. R., and Srinivasan, M. V. (2011). "Optic flow cues guide flight in birds". In: *Curr. Biol.* 21.21, pp. 1794–1799. DOI: 10.1016/j.cub.2011.09.009.

Billington, J., Webster, R. J., Sherratt, T. N., Wilkie, R. M., and Hassall, C. (2020). "The (under)use of eye-tracking in evolutionary ecology". In: *Trends Ecol. Evol.* 35.6, pp. 495–502. DOI: 10.1016/j.tree.2020.01.003.

Blender Online Community (2021). *Blender - a 3D modelling and rendering package*. Stichting Blender Foundation, Amsterdam.

Boeddeker, N., Dittmar, L., Stürzl, W., and Egelhaaf, M. (2010). "The fine structure of honeybee head and body yaw movements in a homing task". In: *Proc. R. Soc. B Biol. Sci.* 277.1689, pp. 1899–1906. DOI: 10.1098/rspb.2009.2326.

Brighton, C. H. and Taylor, G. K. (2019). "Hawks steer attacks using a guidance system tuned for close pursuit of erratically manoeuvring targets". In: *Nat. Commun.* 10.1, pp. 1–28. DOI: 10.1038/s41467-019-10454-z.

Brighton, C. H., Thomas, A. L., and Taylor, G. K. (2017). "Terminal attack trajectories of peregrine falcons are described by the proportional navigation guidance law of missiles". In: *Proc. Natl. Acad. Sci.* 114.51, p. 201714532. DOI: 10.1073/pnas.1714532114.

Cellini, B. and Mongeau, J. M. (2020). "Active vision shapes and coordinates flight motor responses in flies". In: *Proc. Natl. Acad. Sci. U. S. A.* 117.37, pp. 23085–23095. DOI: 10.1073/pnas.1920846117.

Davies, M. N. O. and Green, P. R. (1990). "Optic flow-field variables trigger landing in hawk but not in pigeons". In: *Naturwissenschaften* 77.3, pp. 142–144. DOI: 10.1007/BF01134481.

Eckmeier, D., Geurten, B. R., Kress, D., Mertes, M., Kern, R., Egelhaaf, M., and Bischof, H. J. (2008). "Gaze strategy in the free flying zebra finch (Taeniopygia guttata)". In: *PLoS One* 3.12. DOI: 10.1371/journal.pone.0003956.

Eckmeier, D., Kern, R., Egelhaaf, M., and Bischof, H. J. (2013). "Encoding of naturalistic optic flow by motion sensitive neurons of nucleus rotundus in the zebra finch (Taeniopygia guttata)". In: *Front. Integr. Neurosci.* 7.SEP, pp. 1–17. DOI: 10.3389/fnint.2013.00068.

Egelhaaf, M., Boeddeker, N., Kern, R., Kurtz, R., and Lindemann, J. P. (2012). "Spatial vision in insects is facilitated by shaping the dynamics of visual input through behavioural action". In: *Front. Neural Circuits* 6.December, pp. 1–23. DOI: 10.3389/fncir.2012.00108.

Fajen, B. R. and Warren, W. H. (2003). "Behavioral Dynamics of Steering, Obstacle Avoidance, and Route Selection". In: *J. Exp. Psychol. Hum. Percept. Perform.* 29.2, pp. 343–362. DOI: 10.1037/0096-1523.29.2.343.

Hayhoe, M. M. (2017). “Vision and Action”. In: *Annu. Rev. Vis. Sci.* 3, pp. 389–413. DOI: 10.1146/annurev-vision-102016-061437.

Hazan, Y., Kra, Y., Yarin, I., Wagner, H., and Gutfreund, Y. (2015). “Visual-auditory integration for visual search: a behavioral study in barn owls”. In: *Front. Integr. Neurosci.* 9, pp. 1–12. DOI: 10.3389/fnint.2015.00011.

Kane, S. A., Fulton, A. H., and Rosenthal, L. J. (2015). “When hawks attack: animal-borne video studies of goshawk pursuit and prey-evasion strategies”. In: *J. Exp. Biol.* 218.2, pp. 212–222. DOI: 10.1242/jeb.108597.

Kane, S. A. and Zamani, M. (2014). “Falcons pursue prey using visual motion cues: new perspectives from animal-borne cameras”. In: *J. Exp. Biol.* 217.2, pp. 225–234. DOI: 10.1242/jeb.092403.

Kano, F., Walker, J., Sasaki, T., and Biro, D. (2018). “Head-mounted sensors reveal visual attention of free-flying homing pigeons”. In: *J. Exp. Biol.* 221.17, pp. 1–13. DOI: 10.1242/jeb.183475.

Kress, D., Van Bokhorst, E., and Lentink, D. (2015). “How lovebirds maneuver rapidly using super-fast head saccades and image feature stabilization”. In: *PLoS One* 10.6, pp. 1–24. DOI: 10.1371/journal.pone.0129287.

Land, M. F. and Lee, D. N. (1994). “Where we look when we steer”. In: *Nature* 369.6483, pp. 742–744. DOI: 10.1038/369742a0.

Land, M. F. (2006). “Eye movements and the control of actions in everyday life”. In: *Prog. Retin. Eye Res.* 25.3, pp. 296–324. DOI: 10.1016/j.preteyeres.2006.01.002.

Land, M. F. and Nilsson, D. E. (2012). *Animal eyes*. 2nd ed. Oxford: Oxford University Press, p. 288. DOI: 10.1093/acprof:oso/9780199581139.001.0001.

Lappi, O. (2022). “Gaze strategies in driving–an ecological approach”. In: *Front. Psychol.* 13.March. DOI: 10.3389/fpsyg.2022.821440.

Lee, D. N., Davies, M. N., Green, P. R., and Van Der Weel, F. R. (1993). “Visual control of velocity of approach by pigeons when landing”. In: *J. Exp. Biol.* 104, pp. 106–131. DOI: 10.4324/9780203936672.

Lev-Ari, T. and Gutfreund, Y. (2018). “Interactions between top-down and bottom-up attention in barn owls (Tyto alba)”. In: *Anim. Cogn.* 21.2, pp. 197–205. DOI: 10.1007/s10071-017-1150-2.

Lin, H. T., Ros, I. G., and Biewener, A. A. (2014). “Through the eyes of a bird: modelling visually guided obstacle flight”. In: *J. R. Soc. Interface* 11.96, pp. 1–12. DOI: 10.1098/rsif.2014.0239.

Martin, G. R. (2007). “Visual fields and their functions in birds”. In: *J. Ornithol.* 148.Suppl. 2, S547–S562. DOI: 10.1007/s10336-007-0213-6.

Matthis, J. S., Muller, K. S., Bonnen, K. L., and Hayhoe, M. M. (2022). “Retinal optic flow during natural locomotion”. In: *PLOS Comput. Biol.* 18.2, pp. 1–30. DOI: 10.1371/journal.pcbi.1009575.

Matthis, J. S., Yates, J. L., and Hayhoe, M. M. (2018). “Gaze and the control of foot placement when walking in natural terrain”. In: *Curr. Biol.* 28.8, 1224–1233.e5. DOI: 10.1016/j.cub.2018.03.008.

Mitkus, M., Potier, S., Martin, G. R., Duriez, O., and Kelber, A. (2018). “Raptor vision”. In: *Oxford Res. Encycl. Neurosci.* March. Oxford University Press, pp. 1–38. DOI: 10.1093/acrefore/9780190264086.013.232.

Mole, C., Pekkanen, J., Sheppard, W. E., Markkula, G., and Wilkie, R. M. (2021). “Drivers use active gaze to monitor waypoints during automated driving”. In: *Sci. Rep.* 11.1, pp. 1–18. DOI: 10.1038/s41598-020-80126-2.

O'Rourke, C. T., Hall, M. I., Pitlik, T., and Fernández-Juricic, E. (2010). "Hawk eyes I: Diurnal raptors differ in visual fields and degree of eye movement". In: *PLoS One* 5.9, pp. 1–8. DOI: 10.1371/journal.pone.0012802.

Pekkanen, J. and Lappi, O. (2017). "A new and general approach to signal denoising and eye movement classification based on segmented linear regression". In: *Sci. Rep.* 7.1, pp. 1–12. DOI: 10.1038/s41598-017-17983-x.

Pelz, J. B. and Rothkopf, C. A. (2007). "Oculomotor behavior in natural and man-made environments". In: *Eye Movements*. Elsevier, pp. 661–676. DOI: 10.1016/B978-008044980-7/50033-1.

Pete, A. E., Kress, D., Dimitrov, M. A., and Lentink, D. (2015). "The role of passive avian head stabilization in flapping flight". In: *J. R. Soc. Interface* 12.110, p. 20150508. DOI: 10.1098/rsif.2015.0508.

Potier, S., Bonadonna, F., Kelber, A., Martin, G. R., Isard, P.-F., Dulaurent, T., and Duriez, O. (2016). "Visual abilities in two raptors with different ecology". In: *J. Exp. Biol.* 219.17, pp. 2639–2649. DOI: 10.1242/jeb.142083.

Raudies, F., Mingolla, E., and Neumann, H. (2012). "Active gaze control improves optic flow-based segmentation and steering". In: *PLoS One* 7.6. DOI: 10.1371/journal.pone.0038446.

Ravi, S., Siesenop, T., Bertrand, O. J., Li, L., Doussot, C., Fisher, A., Warren, W. H., and Egelhaaf, M. (2022). "Bumblebees display characteristics of active vision during robust obstacle avoidance flight". In: *J. Exp. Biol.* 225.4. DOI: 10.1242/jeb.243021.

Rivers, T. J., Sirota, M. G., Guttentag, A. I., Ogorodnikov, D. A., Shah, N. A., and Beloozerova, I. N. (2014). "Gaze shifts and fixations dominate gaze behavior of walking cats". In: *Neuroscience* 275, pp. 477–499. DOI: 10.1016/j.neuroscience.2014.06.034.

Ros, I. G., Bhagavatula, P. S., Lin, H. T., and Biewener, A. A. (2017a). "Rules to fly by: pigeons navigating horizontal obstacles limit steering by selecting gaps most aligned to their flight direction". In: *Interface Focus* 7.1, p. 20160093. DOI: 10.1098/rsfs.2016.0093.

Ros, I. G. and Biewener, A. A. (2017b). "Pigeons (C. livia) follow their head during turning flight: Head stabilization underlies the visual control of flight". In: *Front. Neurosci.* 11.DEC, pp. 1–12. DOI: 10.3389/fnins.2017.00655.

Rothkopf, C. A., Ballard, D. H., and Hayhoe, M. M. (2007). "Task and context determine where you look". In: *J. Vis.* 7.14, pp. 1–20. DOI: 10.1167/7.14.16.

Rothkopf, C. A. and Ballard, D. H. (2009). "Image statistics at the point of gaze during human navigation". In: *Vis. Neurosci.* 26.1, pp. 81–92. DOI: 10.1017/S0952523808080978.

Schilstra, C. and van Hateren, J. (1998). "Stabilizing gaze in flying blowflies". In: *Nature* 395.October, p. 1998. DOI: 10.1038/27114.

Shoemake, K. (1985). "Animating rotation with quaternion curves". In: *Proc. 12th Annu. Conf. Comput. Graph. Interact. Tech.* Vol. 19. 3, pp. 245–254. DOI: 10.1145/325165.325242.

Tammero, L. F. and Dickinson, M. H. (2002). "Collision-avoidance and landing responses are mediated by separate pathways in the fruit fly, Drosophila melanogaster". In: *J. Exp. Biol.* 205.18, pp. 2785–2798. DOI: 10.1242/jeb.205.18.2785.

The Mathworks Inc. (2020). *MATLAB version 9.9.0.1538559 (R2020b) Update 3*. Natick, Massachusetts.

Theunissen, L. M., Reid, T., and Troje, N. F. (2017). "Pigeons use distinct stop phases to control pecking". In: *J. Exp. Biol.* 220.3, pp. 437–444. DOI: 10.1242/jeb.147850.

Van Breugel, F. and Dickinson, M. H. (2012). "The visual control of landing and obstacle avoidance in the fruit fly Drosophila melanogaster". In: *J. Exp. Biol.* 215.11, pp. 1783–1798. DOI: 10.1242/jeb.066498.

Volgushev, M., Nguyen, C. T., Iyer, G. S., and Beloozerova, I. N. (2022). "When cats need to see to step accurately?" In: *J. Physiol.* 600.1, pp. 75–94. DOI: 10.1113/JP282255.

Wagner, H. (1982). "Flowfield trigger landing in flies". In: *Nature* 297.May, pp. 147–148.

Wang, Y. and Frost, B. J. (1992). "Time to collision is signalled by neurons in the nucleus rotundus of pigeons". In: *Nature* 356.6366, pp. 236–238. DOI: 10.1038/356236a0.

Warrick, D. R., Bundle, M. W., and Dial, K. P. (2002). "Bird maneuvering flight: blurred bodies, clear heads". In: *Integr. Comp. Biol.* 42.1, pp. 141–148. DOI: 10.1093/icb/42.1.141.

Wilkie, R. M., Wann, J. P., and Allison, R. S. (2008). "Active gaze, visual look-ahead, and locomotor control". In: *J. Exp. Psychol. Hum. Percept. Perform.* 34.5, pp. 1150–1164. DOI: 10.1037/0096-1523.34.5.1150.

Windsor, S. P. and Taylor, G. K. (2017). "Head movements quadruple the range of speeds encoded by the insect motion vision system in hawkmoths". In: *Proc. R. Soc. B Biol. Sci.* 284.1864. DOI: 10.1098/rspb.2017.1622.

Yorzinski, J. L. (2019). "Conjugate eye movements guide jumping locomotion in an avian species". In: *J. Exp. Biol.* 222.20. DOI: 10.1242/jeb.211565.

Yorzinski, J. L., Patricelli, G. L., Platt, M. L., and Land, M. F. (2015). "Eye and head movements shape gaze shifts in Indian peafowl". In: *J. Exp. Biol.* 218.23, pp. 3771–3776. DOI: 10.1242/jeb.129544.

Zubair, H. N., Chu, K. M., Johnson, J. L., Rivers, T. J., and Beloozerova, I. N. (2019). "Gaze coordination with strides during walking in the cat". In: *J. Physiol.* 597.21, pp. 5195–5229. DOI: 10.1113/JP278108.

5

Conclusions and future work

In this thesis, I have presented a novel approach to investigate the gaze strategies of birds in manoeuvring flight. I used motion capture data from experiments with birds, and tools from computer vision, to characterise in detail the visual scene that the birds experience in flight, considering their full visual fields. I applied the method to the case of Harris' hawks executing perching and obstacle avoidance manoeuvres, and showed how this approach can provide unique insights into the birds' behaviour. In the previous chapters, I discussed in detail the findings and implications of each of the pieces of work. In this chapter, I summarise the main conclusions of the research, discuss the general limitations of the approach, and provide suggestions for future directions.

5.1 Main conclusions

The research presented in this thesis covered three key challenges of studying gaze strategies of birds in flight.

First, I explored the challenge of accurately tracking an animal's gaze direction. Motion capture systems are currently the most accurate species-agnostic method to track 3D movements in humans and animals. However, the commercial algorithms that accompany these systems to track the identity of the markers over time are

often optimised for human motion, and underperform when applied to animals. As a result, animal studies that make use of motion capture frequently require additional postprocessing work to identify the markers over time (i.e., label the markers), which in turn limits the possibilities of high-throughput studies.

In **Chapter 2**, I presented a solution that deals with this problem by design, deriving an optimal rigid arrangement of markers that maximises the performance of a per-frame labelling algorithm. With experiments, I verified the aims of the design and assessed the performance of the proposed labelling approach, comparing it to the commercially available solution from the motion capture system. The proposed method proved more robust to reduced sampling frequencies and higher rotational speeds, and did not carry mislabellings from one frame to subsequent ones, thanks to its reliance on geometrical features of the marker pack that can be defined per frame. I focused on the case of tracking head movements of birds in flight in a motion capture system; however, the design was parametrised to facilitate its application to other animals and other marker-based 3D motion tracking setups. Although I focus on a rigid arrangement of markers with no marker dropout, the method is still valuable as a first bulk-labelling step within a more sophisticated labelling pipeline. The proposed method supports the collection of large and accurate 3D motion datasets in animals, which can be key to improve the performance of video-based methods for 3D pose estimation, or to explore data-driven models of animal motor control.

In **Chapter 3**, I explored the problem of capturing the visual scene that the animal experiences, focusing on the case of Harris' hawks in flight. I proposed to solve this computationally, using a synthetic reconstruction of the environment that the birds fly in and an approximation of their point of view during flight. Specifically, I used motion capture data from the birds' head movements to estimate a coordinate system representative of the animals' visual fields, which I then used to define virtual cameras that represented the birds' perspectives in flight. To model the lab environment, I explored a hybrid approach that combined basic geometric primitives defined using the motion capture data, and dense 3D meshes captured with a SLAM application. In this way, I minimised the accumulated drift and

noise typical of large 3D dense meshes, while also reducing the modelling effort required for the most intricate shapes. With the proposed method, it is possible to generate a set of synthetic visual inputs that characterise the information that is likely available to the bird in flight: RGB data, semantic maps, depth maps, and optic flow data. These are defined for every pixel in the bird's visual field, for every timestep. I used three sample bird flights, of a pursuit and an obstacle avoidance manoeuvre, to demonstrate the unique advantages of the method: (1) the possibility of defining alternative gaze scenarios for hypothesis testing, (2) the ability to consider the full visual field of the bird, and (3) the possibility of adapting the 3D model of the environment to the required level of detail. The results from the sample flights showed that comparable analyses would be exceedingly challenging using alternative approaches such as head-mounted cameras, or would be much more limited if relying on point estimates of the bird's gaze direction.

In **Chapter** 4, I investigated the challenge of using the generated synthetic data to answer behavioural questions. Specifically, I generated semantic descriptions of the visual scene that the birds experience in flight, to characterise their gaze strategy during perching and obstacle avoidance manoeuvres in $n = 97$ flights. The results showed that all birds actively stabilised their heads in flight, to maintain the landing perch and the obstacles in specific areas of their visual fields. This was clearly illustrated by contrasting the results obtained in the visual coordinate system, which represents the bird's visual field orientation, and the trajectory coordinate system, which translates like the bird's head but does not account for the bird's head rotations. I also found that the birds used distinct gaze strategies around the landing perch and the obstacles. While the perch was centred relative to the bird's sagittal plane for the majority of the perching manoeuvre, the inside edge of the obstacles was fixated frontally when the birds had to turn around them. The fact that this edge-fixating behaviour around obstacles has been reported previously in humans (Rothkopf et al., 2009; Rothkopf et al., 2007; Raudies et al., 2012), lovebirds (Kress et al., 2015) and bumblebees (Ravi et al., 2022) provides independent, if circumstantial, support to the proposed method of estimating the

bird's visual coordinate system. I discussed why this strategy of fixating on the edge of an obstacle may be advantageous for a fast moving animal, focusing on the motion contrast hypothesis that has been previously suggested. I highlighted this point as one of the main hypotheses to inspect with the full dataset in the future. The results of this study also suggest that when approaching the perch, the birds deviated more frequently from a centring strategy when obstacles were in place (1.5 m in front of the takeoff perch), than when no obstacles were present. With obstacles in place, the birds also spent more time on average aligned with one of the landing perch's edges. I suggested this behaviour relates to the amount of turning required by the trajectory, in line with similar observations of lovebirds executing sharp turning manoeuvres (Kress et al., 2015); this would be another key hypothesis to test with the full dataset.

5.2 Limitations and recommendations

In each of the previous chapters I discussed the specific limitations of each part of the research and suggested directions of improvement. Here, I comment on the broader limitations of the thesis.

The method that I used to estimate the visual coordinate system of the birds proved to be a reasonable approach: across all birds, the data showed a good agreement with our assumptions (see Figures 4.4 and 4.5) and the results from the behavioural analysis were supported by the previous literature. However, there is also room for improvement. One of the current limitations is the reliance on the *a priori* assumption that the bird fixates its gaze on the landing perch upon landing. Although this proved to be a solid assumption *a posteriori*, it may not be easily translated to other manoeuvres. When perching is not involved in the manoeuvre of interest, one option could be to calibrate the bird's gaze direction using data from a dedicated calibration trial, with an equivalent gaze-fixation assumption. I explored this option in the sample pursuit flight presented in Chapter 3, for which I calibrated the bird's gaze direction using a trial in which a piece of meat with

a marker was presented to the animal. However, this approach produced a much lower number of samples appropriate for the calibration than the method based on the perching assumption. We also noticed that after several weeks it became harder to attract the birds' attention with this method. An alternative option could be to record a set of perching trials before recording the manoeuvre of interest, exclusively to calibrate the bird's visual coordinate system. However, this approach would limit the number of trials of interest that can be recorded, as the reward provided in the gaze calibration flights would count towards the total food intake of the birds per day. Additionally, we would still have a problem of circularity when analysing the birds' gaze behaviour during perching (see Section 4.5); one option to limit this issue could be to use a perch of substantially different geometry in the gaze calibration trials, such as a falcon-type perch, which resembles a small stool.

However a better solution, which would likely improve the accuracy of the estimated visual coordinate system, was suggested in Chapter 3: it consists of integrating a pair of calibrated stereo cameras with the motion capture system. Using video-based motion tracking tools, it would be possible to extract during a calibration trial the 3D coordinates of identifiable features on the bird, such as its bill and eyes, and express them in the headpack coordinate system; a similar approach is described as an annotation tool in Naik et al. (2020). In this way, the bird's head pose could be directly estimated relative to the headpack coordinate system. Even if this method still assumes the bird's eyes are fixed relative to the head, it would allow us to model more accurately the bird's visual coordinate system, mainly by reducing the misalignment between the estimated gaze direction and the true forward direction of the bird's head. In this thesis, the bird's visual coordinate system is estimated using a set of assumptions (see Section 3.3.3) and its origin is placed at the headpack's baseplate, which is < 20 mm offset from the midpoint between the bird's eyes (see Sections 3.3.3.1 and 3.3.3.2). This offset is expected to introduce a small angular deviation of about 2.3° at 0.5 m distance, between the true line of sight through the bird's eyes and the estimated gaze direction. Using the suggested stereo pair approach, we would be able to better approximate the midpoint between

the bird's eyes in the headpack coordinate system and reduce this error (note that the midpoint between the bird's eyes is the ideal origin of the sphere representing its visual field, see Martin, 2007). We would also be able to measure precisely the stereo baseline and use it to define the bird's view as a binocular system in Blender. More importantly, this approach would allow us to have a common calibration method independent of the recorded manoeuvre, and if properly automated, it would allow us to scale up the analyses to much larger datasets.

Throughout this thesis I have considered the forward direction of the bird's head as a good proxy of the animal's gaze direction. Although this is a very frequent approach used in the literature (Kano et al., 2018; Brighton et al., 2017; Ros et al., 2017b; Kress et al., 2015; Kane et al., 2015; Eckmeier et al., 2008), it relies on the assumption that eye movements relative to the bird's head are small. The degree of eye motion in Harris' hawks has not been experimentally measured and we can only rely on estimates from birds of the same family. These suggest a range of eye motion of around 6° (see Chapter 4, Section 4.3.2.1 and Wallman et al., 1985). However, the visual systems of birds of prey show large variation, and there may be relevant differences within orders, families or even individuals (Mitkus et al., 2018). Specifically, the magnitude of a bird's eye movements and their converged/diverged state will modify the extension of the binocular, monocular and blind regions in the animal's visual field (Potier et al., 2020; Tyrrell et al., 2018), and the estimated positions of the frontal and lateral-facing foveas relative to the forward direction of the head. However, since the range of eye motion is estimated around 6°, we don't expect eye movements to significantly affect the main conclusions of the work in this thesis, such as the stabilisation of the target in the binocular area in Chapter 3. Nevertheless, precise estimates of the degree of eye motion in Harris' hawks could be very informative for the method I present here, for example to help define appropriate thresholds when analysing the animals' gaze strategies.

A further improvement to the gaze direction estimate would be to track the eye movements of the birds in flight. This would allow us to inspect how well the small eye movements assumption holds for Harris' hawks in flight, and would also open up

a range of new studies. For example, we could explore changes in the bird's visual field configuration in flight (Tyrrell et al., 2018), inspect how the birds make use of the high-resolution areas in their visual field (Potier et al., 2016), or investigate whether the hawks can track targets simultaneously and independently with each eye, as it has been shown for grackles (Yorzinski, 2021). However, this seems currently very challenging. Although eye-tracking in free-moving mice has proven possible (Holmgren et al., 2021), the method is not currently easily translatable to a larger flying animal. Indeed, eye-tracking equipment in birds has only been used in relatively large species doing terrestrial tasks (Yorzinski et al., 2013; Yorzinski et al., 2014; Yorzinski et al., 2015; Yorzinski, 2019).

The methods I presented here to precisely track a bird's head movements (Chapter 2), and to generate synthetic data that characterise its visual scene in flight (Chapter 4) were conceived to support the collection of large amounts of data across several individuals. The data analysed in Chapter 4 represents a large quantity of data in terms of the number of frames tracked and rendered, but was limited in the number of flights per bird. Indeed, the size and richness of the data created its own practical complications. Studying a subset of the full dataset was an academic choice we made, after the pandemic disruption and due to time constraints. This is also the reason why I focused on analysing the semantic data in Chapter 4. However, I did explore methods of analysing the optic flow data, and some of the preliminary results are included in Appendix C. Additionally, in Appendix D I include a description of a larger dataset of pursuit and obstacle avoidance flights of Harris' hawks that I collected at the start of my DPhil. This proved to be too noisy to reliably extract the birds' head pose (the motion capture system was not fully optimised for our experiments at the time), but could be used to analyse the birds' head trajectories in future studies.

Finally, the methods described in this thesis are primarily designed around a motion capture system. Although this reduces the applicability of the approach in many aspects, some of the techniques discussed could easily be adapted to equivalent setups. For example, the marker labelling approach described in Chapter 2 could

be directly applied to marker-based video methods, as long as the problem is to label the 3D trajectories of unidentified markers. The basic geometric model of the environment described in Chapter 3 could alternatively be obtained with photogrammetry or laser-scanning methods, similarly to the work in Holmgren et al. (2021). However, the convenience of a centralised and synchronised system, and the accuracy provided by the motion capture system may be harder to replicate with other setups. Using a motion capture system also meant the birds were recorded in a lab environment, which is not representative of their natural habitat.

5.3 Future directions

The research presented here has aimed to facilitate the collection of large amounts of data across many individuals, and as discussed, some key steps have been taken towards this aim. However, these are only the first steps in a much larger endeavour, and further automation of the complete pipeline would allow us to explore a completely different approach to animal behaviour: one based on models driven by data. This would be an exciting avenue to explore next, and this thesis sets a solid basis for it.

With analytical models of behaviour, we typically start by defining a model based on our prior knowledge or assumptions about the animal's behaviour. The model is then fitted to the empirical data, and if the fit is good, we conclude that the model's underlying principles also apply to the organism (Im et al., 2020). For example, this is the aim with which the model in Appendix E was developed, and the approach followed by most models used in animal behaviour (see for example Brighton et al., 2017; Brighton et al., 2019; Fabian et al., 2022). With data-driven models, frequently based on machine-learning techniques, the methodology is slightly different: the data is fitted using complex function approximators, like deep neural networks, that have large ability to capture a wide variety of relations (Goodfellow et al., 2016). However, these models are difficult to interpret (Im et al., 2020; Heras et al., 2019). But even with this caveat, there are aspects of data-driven models that

make them interesting for behavioural modelling. The models' complexity can allow us to better capture the behaviour of the animal, and potentially reveal aspects beyond our prior assumptions (Im et al., 2020). These models also facilitate the definition of objective metrics to compare behaviours, and support the automated analysis of very large data sets (Eyjolfsdottir et al., 2017).

Some examples of data-driven models of animal behaviour exist in the literature: for example, the work by Zeiler et al. (2009), in which a generative model of the head-bobbing motion of a pigeon is explored, or the work by Boots et al. (2013), in which an imitation learning approach is used to model the navigation strategy of a hawkmoth. More recent works include the study by Heras et al. (2019), in which a modular deep neural network is used to model the collective behaviour of fish, or the work by Eyjolfsdottir et al. (2017), in which the sensory-motor relation of a fly is modelled using recurrent neural networks trained on real flies' data. In a follow-up work (Im et al., 2020), the authors inspected further how to ensure these machine learning models are accurate representations of the animal's behaviour, focusing on how well they perform on long-term predictions. In the fields of robotics and animation there is also a growing interest in similar approaches, with examples of imitation learning methods used with animal data to learn quadruped locomotion (Zhang et al., 2018; Peng et al., 2020). Works such as these show how machine learning approaches have large potential to capture the richness of animal behaviour, and offer new insights into the algorithms animals use to sense and act in their environment.

Another relevant development from this thesis would be to translate the approach presented here to experiments in the field. There is a growing interest within the fields of ecology and conservation in exploring methods to reconstruct an animal's experience of its environment (Tuia et al., 2022; McClune, 2018; Haalck et al., 2020), at much larger scales than the ones I explored here. To adapt the proposed method to the field, bio-loggers combining GPS and IMU units could be used to track the bird's head position and orientation (Kano et al., 2018). Note, however, that differential GPS technology or a similar approach would likely be required to

improve location accuracy (Keshavarzi et al., 2021; Sachs, 2016). To reconstruct the natural environment, laser scanners, structure-from-motion or photogrammetry techniques may be used (Tuia et al., 2022; Stürzl et al., 2015; Stuerzl et al., 2016; Schulte et al., 2019). Consumer-level handheld mapping devices like the one I used in Section 3.3.4.2, or more recent models, such as the iPhone 12 Pro or the iPad Pro 2021, could be considered; these have been shown to reconstruct forest environments with reasonably accurate results (Tatsumi et al., 2022; Gollob et al., 2021). In fact, a 3D model of a section of the Wytham Woods forest (Oxford, UK) covering 559 trees already exists[1] (Calders et al., 2018). This location is very close to the field station in which the Oxford Flight group is based.

On the other hand, there are still interesting research questions to address in the lab with the approach described here. For example, cylindrical obstacles like the ones I used are frequently considered in studies of obstacle avoidance in animals (Ravi et al., 2022; Lin et al., 2014; Ros et al., 2017a) and humans (Fajen et al., 2003; Rothkopf et al., 2009; Rothkopf et al., 2007), but it would be interesting to explore how the observed gaze strategies adapt to objects of different shapes and structure. For example, it is unclear how the birds would adapt an edge-fixation strategy around the obstacles if these had slanted or diagonal edges. Obstacles that are representative of the animal's natural environment, such as small trees or bushes, would also be relevant to explore; in Chapter 3, I showed some preliminary results in this direction (see Figure 3.18). A set of lab experiments with a 'simulated forest' could be useful for comparison with experiments in the field.

Finally, the method presented here could also be used in electrophysiological studies, to investigate the visumotor control of birds in flight (Altshuler et al., 2018). In Eckmeier et al. (2013), naturalistic stimuli similar to the data generated here were presented to head-restrained zebra finches, while recording from single neurons in the nucleus rotundus. Although the neural response was complex and sparse, their results seemed in agreement with the motion contrast hypothesis for the edge-fixation behaviour around obstacles. A synthetic dataset generated with

[1] https://bitbucket.org/tree_research/wytham_woods_3d_model/src/master/

the method presented in this thesis could be used in a similar approach, and help elucidate questions related to the visual cues used for flight control.

Overall, this thesis opens up several avenues for innovative research in the fields of sensory ecology and neuroethology, and sets the path for data-driven models of animal behaviour. Extending this method to the field may provide unique insights into how animals perceive and interact with their environment, and potentially help with problems of anthropogenic impact, such as bird collisions with wind turbines. Approaches like behavioural cloning or imitation learning applied to synthetic datasets like the ones discussed here may also help us unravel how birds use vision to navigate clutter, and inspire new strategies for visually-guided autonomous drones.

References

Altshuler, D. L. and Srinivasan, M. V. (2018). "Comparison of visually guided flight in insects and birds". In: *Front. Neurosci.* 12, p. 157. DOI: 10.3389/fnins.2018.00157.

Boots, B. and Fox, D. (2013). "Learning Dynamic Policies from Demonstration". In: *NIPS Work. Adv. Mach. Learn. Sensorimotor Control*, pp. 1–8.

Brighton, C. H. and Taylor, G. K. (2019). "Hawks steer attacks using a guidance system tuned for close pursuit of erratically manoeuvring targets". In: *Nat. Commun.* 10.1, pp. 1–28. DOI: 10.1038/s41467-019-10454-z.

Brighton, C. H., Thomas, A. L., and Taylor, G. K. (2017). "Terminal attack trajectories of peregrine falcons are described by the proportional navigation guidance law of missiles". In: *Proc. Natl. Acad. Sci.* 114.51, p. 201714532. DOI: 10.1073/pnas.1714532114.

Calders, K., Origo, N., Burt, A., Disney, M., Nightingale, J., Raumonen, P., Åkerblom, M., Malhi, Y., and Lewis, P. (2018). "Realistic forest stand reconstruction from terrestrial LiDAR for radiative transfer modelling". In: *Remote Sens.* 10.6, pp. 1–15. DOI: 10.3390/rs10060933.

Eckmeier, D., Geurten, B. R., Kress, D., Mertes, M., Kern, R., Egelhaaf, M., and Bischof, H. J. (2008). "Gaze strategy in the free flying zebra finch (Taeniopygia guttata)". In: *PLoS One* 3.12. DOI: 10.1371/journal.pone.0003956.

Eckmeier, D., Kern, R., Egelhaaf, M., and Bischof, H. J. (2013). "Encoding of naturalistic optic flow by motion sensitive neurons of nucleus rotundus in the zebra finch (Taeniopygia guttata)". In: *Front. Integr. Neurosci.* 7.SEP, pp. 1–17. DOI: 10.3389/fnint.2013.00068.

Eyjolfsdottir, E., Branson, K., Yue, Y., and Perona, P. (2017). "Learning recurrent representations for hierarchical behavior modeling". In: *5th Int. Conf. Learn. Represent. ICLR 2017 - Conf. Track Proc.*, pp. 1–12. arXiv: 1611.00094.

Fabian, S. T., Sumner, M. E., Wardill, T. J., and Gonzalez-Bellido, P. T. (2022). "Avoiding obstacles while intercepting a moving target: a miniature fly's solution". In: *J. Exp. Biol.* 225.4. DOI: 10.1242/jeb.243568.

Fajen, B. R. and Warren, W. H. (2003). "Behavioral Dynamics of Steering, Obstacle Avoidance, and Route Selection". In: *J. Exp. Psychol. Hum. Percept. Perform.* 29.2, pp. 343–362. DOI: 10.1037/0096-1523.29.2.343.

Gollob, C., Ritter, T., Kraßnitzer, R., Tockner, A., and Nothdurft, A. (2021). "Measurement of forest inventory parameters with apple ipad pro and integrated lidar technology". In: *Remote Sens.* 13.16, pp. 1–35. DOI: 10.3390/rs13163129.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. MIT press.

Haalck, L., Mangan, M., Webb, B., and Risse, B. (2020). "Towards image-based animal tracking in natural environments using a freely moving camera". In: *J. Neurosci. Methods* 330.108455. DOI: 10.1016/j.jneumeth.2019.108455.

Heras, F. J. H., Romero-Ferrero, F., Hinz, R. C., and de Polavieja, G. G. (2019). "Deep attention networks reveal the rules of collective motion in zebrafish". In: *PLOS Comput. Biol.* 15.9, pp. 1–23. DOI: 10.1371/journal.pcbi.1007354.

Holmgren, C. D., Stahr, P., Wallace, D. J., Voit, K.-M., Matheson, E. J., Sawinski, J., Bassetto, G., and Kerr, J. N. D. (2021). "Visual pursuit behavior in mice maintains the pursued prey on the retinal region with least optic flow". In: *Elife* 10, pp. 1–34. DOI: 10.7554/eLife.70838.

Im, D. J., Kwak, I., and Branson, K. (2020). "Evaluation metrics for behaviour modeling". In: *arXiv Prepr. arXiv2007.12298*, pp. 1–17. arXiv: 2007.12298.

Kane, S. A., Fulton, A. H., and Rosenthal, L. J. (2015). "When hawks attack: animal-borne video studies of goshawk pursuit and prey-evasion strategies". In: *J. Exp. Biol.* 218.2, pp. 212–222. DOI: 10.1242/jeb.108597.

Kano, F., Walker, J., Sasaki, T., and Biro, D. (2018). "Head-mounted sensors reveal visual attention of free-flying homing pigeons". In: *J. Exp. Biol.* 221.17, pp. 1–13. DOI: 10.1242/jeb.183475.

Keshavarzi, H., Lee, C., Johnson, M., Abbott, D., Ni, W., and Campbell, D. L. M. (2021). "Validation of real-time kinematic (RTK) devices on sheep to detect grazing movement leaders and social networks in merino ewes". In: *Sensors* 21.3, p. 924.

Kress, D., Van Bokhorst, E., and Lentink, D. (2015). "How lovebirds maneuver rapidly using super-fast head saccades and image feature stabilization". In: *PLoS One* 10.6, pp. 1–24. DOI: 10.1371/journal.pone.0129287.

Lin, H. T., Ros, I. G., and Biewener, A. A. (2014). "Through the eyes of a bird: modelling visually guided obstacle flight". In: *J. R. Soc. Interface* 11.96, pp. 1–12. DOI: 10.1098/rsif.2014.0239.

Martin, G. R. (2007). "Visual fields and their functions in birds". In: *J. Ornithol.* 148.Suppl. 2, S547–S562. DOI: 10.1007/s10336-007-0213-6.

McClune, D. W. (2018). "Joining the dots: reconstructing 3D environments and movement paths using animal-borne devices". In: *Anim. Biotelemetry* 6, p. 5. DOI: 10.1186/s40317-018-0150-6.

Mitkus, M., Potier, S., Martin, G. R., Duriez, O., and Kelber, A. (2018). "Raptor vision". In: *Oxford Res. Encycl. Neurosci.* March. Oxford University Press, pp. 1–38. DOI: 10.1093/acrefore/9780190264086.013.232.

Naik, H., Bastien, R., Navab, N., and Couzin, I. D. (2020). "Animals in virtual environments". In: *IEEE Trans. Vis. Comput. Graph.* 26.5, pp. 2073–2083. DOI: 10.1109/TVCG.2020.2973063.

Peng, X. B., Coumans, E., Zhang, T., Lee, T. W. E., Tan, J., and Levine, S. (2020). "Learning agile robotic locomotion skills by imitating animals". In: *arXiv*. DOI: 10.15607/rss.2020.xvi.064. arXiv: 2004.00784.

Potier, S., Bonadonna, F., Kelber, A., Martin, G. R., Isard, P.-F., Dulaurent, T., and Duriez, O. (2016). "Visual abilities in two raptors with different ecology". In: *J. Exp. Biol.* 219.17, pp. 2639–2649. DOI: 10.1242/jeb.142083.

Potier, S., Mitkus, M., and Kelber, A. (2020). "Visual adaptations of diurnal and nocturnal raptors". In: *Semin. Cell Dev. Biol.* 106.May, pp. 116–126. DOI: 10.1016/j.semcdb.2020.05.004.

Raudies, F., Mingolla, E., and Neumann, H. (2012). "Active gaze control improves optic flow-based segmentation and steering". In: *PLoS One* 7.6. DOI: 10.1371/journal.pone.0038446.

Ravi, S., Siesenop, T., Bertrand, O. J., Li, L., Doussot, C., Fisher, A., Warren, W. H., and Egelhaaf, M. (2022). "Bumblebees display characteristics of active vision during robust obstacle avoidance flight". In: *J. Exp. Biol.* 225.4. DOI: 10.1242/jeb.243021.

Ros, I. G., Bhagavatula, P. S., Lin, H. T., and Biewener, A. A. (2017a). "Rules to fly by: pigeons navigating horizontal obstacles limit steering by selecting gaps most aligned to their flight direction". In: *Interface Focus* 7.1, p. 20160093. DOI: 10.1098/rsfs.2016.0093.

Ros, I. G. and Biewener, A. A. (2017b). "Pigeons (C. livia) follow their head during turning flight: Head stabilization underlies the visual control of flight". In: *Front. Neurosci.* 11.DEC, pp. 1–12. DOI: 10.3389/fnins.2017.00655.

Rothkopf, C. A., Ballard, D. H., and Hayhoe, M. M. (2007). "Task and context determine where you look". In: *J. Vis.* 7.14, pp. 1–20. DOI: 10.1167/7.14.16.

Rothkopf, C. A. and Ballard, D. H. (2009). "Image statistics at the point of gaze during human navigation". In: *Vis. Neurosci.* 26.1, pp. 81–92. DOI: 10.1017/S0952523808080978.

Sachs, G. (2016). "In-flight measurement of upwind dynamic soaring in albatrosses". In: *Prog. Oceanogr.* 142, pp. 47–57. DOI: 10.1016/j.pocean.2016.01.003.

Schulte, P., Zeil, J., and Stürzl, W. (2019). "An insect-inspired model for acquiring views for homing". In: *Biol. Cybern.* 113.4, pp. 439–451. DOI: 10.1007/s00422-019-00800-1.

Stuerzl, W., Zeil, J., Boeddeker, N., and Hemmi, J. M. (2016). "How wasps acquire and use views for homing". In: *Curr. Biol.* 26, pp. 470–482. DOI: 10.1016/j.cub.2015.12.052.

Stürzl, W., Grixa, I., Mair, E., Narendra, A., and Zeil, J. (2015). "Three-dimensional models of natural environments and the mapping of navigational information". In: *J. Comp. Physiol. A Neuroethol. Sensory, Neural, Behav. Physiol.* 201.6, pp. 563–584. DOI: 10.1007/s00359-015-1002-y.

Tatsumi, S., Yamaguchi, K., and Furuya, N. (2022). "ForestScanner: A mobile application for measuring and mapping trees with LiDAR-equipped iPhone and iPad". In: *bioRxiv*, pp. 95–100. DOI: 10.1101/2021.12.11.472207.

Tuia, D., Kellenberger, B., Beery, S., Costelloe, B. R., Zuffi, S., Risse, B., Mathis, A., Mathis, M. W., van Langevelde, F., Burghardt, T., Kays, R., Klinck, H., Wikelski, M., Couzin, I. D., van Horn, G., Crofoot, M. C., Stewart, C. V., Berger-Wolf, T., et al. (2022). "Perspectives in machine learning for wildlife conservation". In: *Nat. Commun.* 13.1, pp. 1–15. DOI: 10.1038/s41467-022-27980-y.

Tyrrell, L. P., Goller, B., Moore, B. A., Altshuler, D. L., and Fernández-Juricic, E. (2018). "The orientation of visual space from the perspective of hummingbirds". In: *Front. Neurosci.* 12.16, pp. 1–9. DOI: 10.3389/fnins.2018.00016.

Wallman, J. and Pettigrew, J. D. (1985). "Conjugate and disjunctive saccades in two avian species with contrasting oculomotor strategies". In: *J. Neurosci.* 5.6, pp. 1418–1428. DOI: 10.1523/jneurosci.05-06-01418.1985.

Yorzinski, J. L., Patricelli, G. L., Babcock, J. S., Pearson, J. M., and Platt, M. L. (2013). "Through their eyes: selective attention in peahens during courtship". In: *J. Exp. Biol.* 216.16, pp. 3035–3046. DOI: 10.1242/jeb.087338.

Yorzinski, J. L. and Platt, M. L. (2014). "Selective attention in peacocks during predator detection". In: *Anim. Cogn.* 17.3, pp. 767–777. DOI: 10.1007/s10071-013-0708-x.

Yorzinski, J. L. (2019). "Conjugate eye movements guide jumping locomotion in an avian species". In: *J. Exp. Biol.* 222.20. DOI: 10.1242/jeb.211565.

Yorzinski, J. L. (2021). "Great-tailed grackles can independently direct their eyes toward different targets". In: *Exp. Brain Res.* 239.7, pp. 2119–2126. DOI: 10.1007/s00221-021-06122-8.

Yorzinski, J. L., Patricelli, G. L., Platt, M. L., and Land, M. F. (2015). "Eye and head movements shape gaze shifts in Indian peafowl". In: *J. Exp. Biol.* 218.23, pp. 3771–3776. DOI: 10.1242/jeb.129544.

Zeiler, M. D., Taylor, G. W., Troje, N. F., and Hinton, G. E. (2009). "Modeling pigeon behaviour using a conditional restricted Boltzmann machine". In: *ESANN 2009 Proceedings, 17th Eur. Symp. Artif. Neural Networks - Adv. Comput. Intell. Learn.*, pp. 391–396.

Zhang, H., Starke, S., Komura, T., and Saito, J. (2018). "Mode-adaptive neural networks for quadruped motion control". In: *ACM Trans. Graph.* 37.4, pp. 1–11. DOI: 10.1145/3197517.3201366.

Appendices

A

Appendices for Chapter 2: *Optimal marker pack design for a time-independent labelling algorithm*

Contents

A.1 Triangle inequality condition 190
A.2 Cayley-Menger determinant condition 191
A.3 Optimisation procedure 192
A.4 Golden solutions 193
A.5 Design plane l_{min}, l_{max} 193
A.6 Computing 3D coordinates of vertices from edges . . . 195
A.7 Design considerations for 3D-printed headpack for birds 195
A.8 Design considerations for 3D-printed two-markers supports 197
A.9 Drill speed measurements with tachometer 197
A.10 Computation of ground-truth labels 198
A.11 Definition of templates in Nexus 202
A.12 Histograms of residuals per trial 202
References 202

A.1 Triangle inequality condition

The triangle inequality states that given three face edges l_a, l_b and l_c, they will shape a triangle if they satisfy the following:

$$l_a + l_b > l_c \quad \cap \quad l_a + l_c > l_b \quad \cap \quad l_b + l_c > l_a \tag{A.1}$$

where $\cap$ represents logical intersection. These three inequalities can be more succinctly expressed as:

$$2max(l_a, l_b, l_c) < l_a + l_b + l_c \tag{A.2}$$

The triangle inequality condition must be satisfied for each of the sets of face edges in a non-degenerate tetrahedron. However, it is a necessary but not sufficient condition for a set of edges' lengths to form a tetrahedron. For example, a set of five edges of length 4, and one edge of length 7 fulfils the triangle inequalities, but cannot form a non-degenerate tetrahedron: if we take the five edges of length 4 as defining two equilateral triangles that make consecutive faces, it is only possible to define a tetrahedron with positive volume if the length of the remaining edge is below $4\sqrt{3}$ (example from Wirth et al., 2009, Figure 1).

A.2 Cayley-Menger determinant condition

The Cayley-Menger determinant condition requires the determinant D to be positive, with D defined as follows:

$$D = \begin{vmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & {l_a}^2 & {l_b}^2 & {l_c}^2 \\ 1 & {l_a}^2 & 0 & \tilde{l_c}^2 & \tilde{l_b}^2 \\ 1 & {l_b}^2 & \tilde{l_c}^2 & 0 & \tilde{l_a}^2 \\ 1 & {l_c}^2 & \tilde{l_b}^2 & \tilde{l_a}^2 & 0 \end{vmatrix} > 0 \tag{A.3}$$

where l_a, l_b, l_c are three non-opposing edges defining a vertex and $\tilde{l_a}$, $\tilde{l_b}$ and $\tilde{l_c}$ are the corresponding opposite edges. If D is positive, its value relates to the volume of the tetrahedron V as $D = 288V^2$ (Wirth et al., 2009). If $D = 0$ and the face edges satisfy the triangle inequalities, we have a degenerate tetrahedron (i.e., a tetrahedron with zero volume).

Note that the determinant D is in principle different for a tetrahedron and its dual. This is clearer if we analyse the extended expression for D:

$$\begin{aligned} D = \quad & 2{l_a}^2\tilde{l_a}^2({l_b}^2 + \tilde{l_b}^2 + {l_c}^2 + \tilde{l_c}^2 - {l_a}^2 - \tilde{l_a}^2) \\ & + 2{l_b}^2\tilde{l_b}^2({l_c}^2 + \tilde{l_c}^2 + {l_a}^2 + \tilde{l_a}^2 - {l_b}^2 - \tilde{l_b}^2) \\ & + 2{l_c}^2\tilde{l_c}^2({l_a}^2 + \tilde{l_a}^2 + {l_b}^2 + \tilde{l_b}^2 - {l_c}^2 - \tilde{l_c}^2) \\ & + ({l_a}^2 - \tilde{l_a}^2)({l_b}^2 - \tilde{l_b}^2)({l_c}^2 - \tilde{l_c}^2) \\ & - ({l_a}^2 + \tilde{l_a}^2)({l_b}^2 + \tilde{l_b}^2)({l_c}^2 + \tilde{l_c}^2) \end{aligned} \tag{A.4}$$

The penultimate term on the right-hand side will change sign for a tetrahedron and its dual. The determinant will be the same for both tetrahedrons only if that anti-symmetric term is zero, which occurs if at least one pair of opposite edges has the same length (Wirth et al., 2009).

A.3 Optimisation procedure

Our goal is to find a set of six edges' lengths that define a non-degenerate tetrahedron, and maximise the minimum distance between the vertices' representations in EE space.

We parametrised the lengths of the six edges in the set using their maximum and minimum values, l_{min} and l_{max}. This makes two edges' lengths, which we call l_e and l_f, equal to these values: $l_e = l_{max}$ and $l_f = l_{min}$. The four remaining edges' lengths l_a, l_b, l_c, l_d are the independent variables to optimise. We assume $l_a \leqslant l_b \leqslant l_c \leqslant l_d$.

For each of the 15 possible pairings of opposite edges' lengths we proceeded as follows:

1. **Select a set of non-opposite edges that define a vertex.** Once this is defined, the geometric interpretation for the rest of the edge sets of the tetrahedron (vertex edges and face edges) is fully defined. Note that even though it seems that can select from eight different triples of non-opposite edges, they all define the vertices or faces of just two tetrahedra, which are dual of each other. We will see that for this first stage of the optimisation, both dual tetrahedra give the same result, and therefore we can focus on analysing one of them only.

2. **Compute for each vertex its coordinates in EE space**, following the expression on Equation 2.1.

3. **Compute the square-distances between the vertices' representations in EE space.** For any two vertices i and j, the square-distance function will have the form:
$$d_{ij}^2 = f(l_{min}, l_{max}; l_a, l_b, l_c, l_d) \tag{A.5}$$
where the semicolon separates the parameters from the independent variables. Note that a tetrahedron and its dual will have the same expressions for the square-distance between their vertices in EE space.

4. **Compute the minimum square-distance h between any two vertices in EE space.** We define h as the minimum square-distance between any two vertices:
$$\begin{aligned} &h(l_{min}, l_{max}; l_a, l_b, l_c, l_d) = \\ &\quad min(d_{12}^2, d_{13}^2, d_{14}^2, d_{23}^2, d_{24}^2, d_{34}^2) \end{aligned} \tag{A.6}$$
The function h is piecewise, since at different regions of the l_a, l_b, l_c, l_d space a different square-distance function d_{ij}^2 will be the minimum one.

5. **Find the values of l_a, l_b, l_c, l_d such that $h = h_{max}$** The function h is piecewise and its derivative is not smooth across the different regions of the l_a, l_b, l_c, l_d space. We solve for the points of maximum h with a grid search, sweeping over a range of values for l_a, l_b, l_c, l_d. We first ran a coarser sweep and then refined over the optimal points found; further details are explained below. The maximum value of h, h_{max} has the following form:
$$h_{max} = \alpha(l_{max} - l_{min})^2 \tag{A.7}$$
where α has a different value for each mapping between opposite edges.

The results for each pairing between opposite edges are shown in Table 2.1. As we have mentioned, at this stage it is enough to consider only one of the dual tetrahedrons for each possible pairing of opposite edges, because the square-distance functions d_{ij}^2 (and consequently the h function) are the same for a tetrahedron and its dual. For most pairings of opposite edges, we found several solutions, that is, several sets of optimal edges lengths ${l_a}^*, {l_b}^*, {l_c}^*, {l_d}^*$ where h reached its maximum value. The edges' lengths for each solution are shown in Table 2.2. From all the optimal solutions, we select those with highest α and inspect them further.

In the optimisation procedure, we used for convenience the adimensional variables for the edges' lengths a, b, c, d, as defined in Equation 2.2. When performing the grid search, we first swept over a coarse grid of values to obtain a set of candidate optimal solutions. In this coarse sweep, the values for a, b, c, d ranged from 0 to 1 in steps of 0.01. We then ran a finer sweep around the candidate optimal solutions found, considering values within ± 0.1 of them, in steps of 0.001. We use these adimensional variables to represent the solutions in Table 2.2.

A.4 Golden solutions

We refer to the solutions with largest α in Table 2.1 as *golden solutions* because they have an interesting relationship between their edges' lengths and the golden ratio. First, for all these solutions exactly two of the four edges lengths in ${l_a}^*, {l_b}^*, {l_c}^*, {l_d}^*$ correspond to either l_{max} or l_{min} (see Table 2.2). Second, the two remaining edges (named l_1 and l_2 below) have the following relation with the golden ratio ϕ:

$$\begin{aligned} \frac{l_1 - l_{min}}{l_{max} - l_{min}} &= \frac{1}{\phi}, \\ \frac{l_2 - l_{min}}{l_{max} - l_{min}} &= \frac{1}{\phi^2}, \end{aligned} \tag{A.8}$$

with $\phi = \frac{1+\sqrt{5}}{2}$.

A.5 Design plane l_{min}, l_{max}

We can build 20 different tetrahedra from the set of 10 golden solutions, since each of them can be constructed as a certain tetrahedron or its dual. Given any of these solutions, we would typically select the values for the l_{min} and l_{max} parameters such that h_{max} is as large as possible. However, there are some restrictions to the values of l_{min} and l_{max} we can select: there are constraints for the set of edges to define a non-degenerate tetrahedron, and there are constraints due to our particular experimental application. To inspect the space of feasible values of l_{min} and l_{max}, we represent the h_{max} function along with the constraints to l_{min} and l_{max} in the l_{min}, l_{max} plane.

Since the h_{max} function for the golden solutions has the form $h_{max} = 0.58(l_{max} - l_{min})^2$, we can represent it in the l_{min}, l_{max} plane as a series of isolines, along which the h_{max} value is constant. For convenience, we will represent $\sqrt{h_{max}}$ instead, so that its units are more easily interpretable.

The constraints to the values of l_{min} and l_{max} define a region in the l_{min}, l_{max} plane, which is different for each of the 20 tetrahedra that can be built from the golden solutions. By definition, we have that $l_{min} < l_{max}$, and therefore the region of feasible designs will always be in the semiplane defined by that constraint. The rest of the constraints imposed to the l_{min} and l_{max} values are:

1. **Triangle inequality condition**. Each set of face edges for a solution tetrahedron needs to fulfil the triangle inequality condition. This leads to at most three conditions of the form:

$$l_{min} < K l_{max}. \tag{A.9}$$

   with $K > 1$. In the l_{min}, l_{max} plane, a triangle inequality condition restricts the space to a semiplane with a linear boundary. The triangle conditions are, in principle, different for a tetrahedron and its dual.

2. **Cayley-Menger determinant condition.**. The requirement of the Cayley-Menger determinant to be positive leads to one constraint of the form:

$$\begin{aligned} K_0 {l_{min}}^6 + K_1 {l_{min}}^5 l_{max} + K_2 {l_{min}}^4 {l_{max}}^2 \\ + K_3 {l_{min}}^3 {l_{max}}^3 + K_4 {l_{min}}^2 {l_{max}}^4 \\ + K_5 l_{min} {l_{max}}^5 + K_6 {l_{max}}^6 > 0 \end{aligned} \tag{A.10}$$

   with $K_0, K_1, K_2, K_3, K_4, K_5, K_6 \in \mathbb{R}$. In the l_{min}, l_{max} plane, the Cayley-Menger determinant condition represents a region delimited by a sixth-degree polynomial, homogeneous and bivariate. As explained in Section A.2, this constraint is in principle different for a tetrahedron and its dual.

3. **Minimum edge length due to the motion capture system's resolution,** $l_{min,res}$. There is a minimum distance between two retroreflective markers, below which the motion capture system is unable to distinguish them as two separate points; we call this minimum distance $l_{min,res}$. This limitation imposes a constraint to the minimum value of l_{min} we can select for the tetrahedron of the form:

$$l_{min} > l_{min,res}. \tag{A.11}$$

4. **Maximum edge length due to size constraints,** $l_{max,size}$. There are limitations to the dimensions of the marker pack depending on the application. For example, if we are planning to attach the marker pack to the head of a bird, we are limited by an attachment area of approximately 50×50 mm^2. We reduce this limitation to a maximum edge length allowed by the application $l_{max,size}$, and express it in the l_{min}, l_{max} plane as a constraint of the form:

$$l_{max} < l_{max,size}. \tag{A.12}$$

The first two conditions derive from the requirements to build a non-degenerate tetrahedron and are different for every solution, and for a tetrahedron and its dual. The last two conditions refer to the construction of the tetrahedron as

a marker pack, and are common to all solutions and tetrahedrons as long as they have the same application.

Figure A.1 shows three different region plots for three tetrahedra derived from golden solutions. Note how two of them correspond to the same solution, but interpreted as a tetrahedron or its dual (Figure A.1a and b), and how each defines a different region. We selected from the set of golden solutions one with the largest feasible region, which in practice meant the least-restrictive Cayley-Menger determinant condition (Figure A.1c). We found six solutions with equally large feasible regions, and selected one of them at random.

A.6 Computing 3D coordinates of vertices from edges

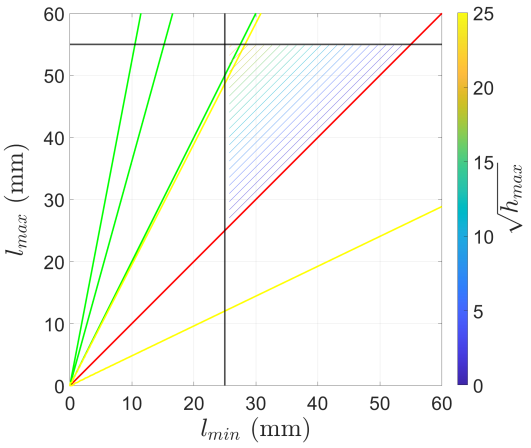
Once the design point is selected in the l_{min}, l_{max} plane, we can compute the optimal remaining edges ${l_a}^*, {l_b}^*, {l_c}^*, {l_d}^*$ with the expressions for the corresponding solution in Table 2.2 and with Equation 2.2. From the set of vertex edges we can derive the lengths of the edges that connect any two vertices, by determining the common element between the sets. Once these are known, we can solve for the vertices' coordinates in an arbitrary reference frame, for example by applying the following equations:

$$\begin{aligned}
&\vec{V_1} = [0,\, 0,\, 0] \\
&\vec{V_2} = [l_{1,2},\, 0,\, 0] \\
&V_{3,z} = 0 \\
&(V_{3,x} - V_{1,x})^2 + (V_{3,y} - V_{1,y})^2 = l_{1,3}^2 \\
&(V_{3,x} - V_{2,x})^2 + (V_{3,y} - V_{2,y})^2 = l_{2,3}^2 \\
&(V_{4,x} - V_{1,x})^2 + (V_{4,y} - V_{1,y})^2 + (V_{4,z} - V_{1,z})^2 = l_{1,4}^2 \\
&(V_{4,x} - V_{2,x})^2 + (V_{4,y} - V_{2,y})^2 + (V_{4,z} - V_{2,z})^2 = l_{2,4}^2 \\
&(V_{4,x} - V_{3,x})^2 + (V_{4,y} - V_{3,y})^2 + (V_{4,z} - V_{3,z})^2 = l_{3,4}^2
\end{aligned} \tag{A.13}$$

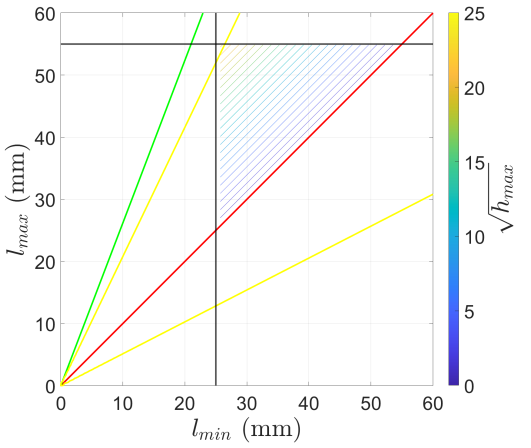
where $\vec{V_i} = [V_{i,x}, V_{i,y}, V_{i,z}]$ are the coordinates for vertex i, and $l_{i,j}$ is the edge length between vertex i and vertex j. In this case the reference frame is defined such that vertex 1 is at the origin, vertex 2 is along the x-axis in the XY plane, and vertex 3 is in the XY plane. Vertex 4 results from the intersection of three spheres centred at vertices 1, 2 and 3, with appropriate radii. We can import the vertices' coordinates into a CAD software to design the marker pack support for 3D-printing.

A.7 Design considerations for 3D-printed headpack for birds

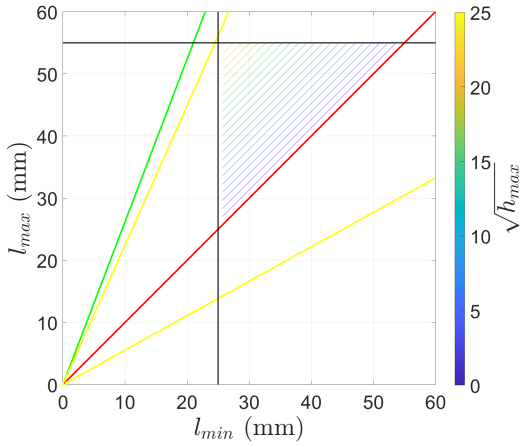
The main requirements for our headpack design is for it to be rigid, so that the markers stay at constant distance, and light, to minimise the weight on the birds'



(a) Pairing ID: 6, Solution ID: 1, tetrahedron A



(b) Pairing ID: 6, Solution ID: 1, tetrahedron B



(c) Pairing ID: 6, Solution ID: 2, tetrahedron A

Figure A.1: Examples of feasible regions in the l_{min}, l_{max} plane for three solutions. The design planes for three different tetrahedra are represented. The captions indicate the pairing between opposite edges and the solution they represent, as per Table 2.2. Tetrahedron A refers to the tetrahedron we can build for each solution, interpreting the first element of each pair of opposite edges from Table 2.1 as defining a vertex; tetrahedron B would be its dual. The regions delimited by the grey, yellow and red boundaries represent combinations of l_{min}, l_{max} parameters that produce non-degenerate tetrahedra for each selected solution. The following boundaries are represented: triangular faces condition (green); positive Cayley-Menger determinant (yellow); $l_{min,res}$ minimum distance between markers determined by the motion capture system (l_{min} constant grey line); $l_{max,size}$ maximum admissible distance between markers for a bird headpack (l_{max} constant grey line); and $l_{min} < l_{max}$ (red). The Cayley-Menger determinant condition (yellow) is always more restrictive than the triangular conditions per face (green). In cases (b) and (c), the triangular faces conditions for each face reduce to one boundary. Lines of constant $\sqrt{h_{max}}$ (mm) are shown, note that in some cases the isolines extend beyond the region of feasible designs.

heads. We accomplish this using an appropriate 3D printing material: we used DuraForm PA (polyamide nylon plastic) or ABS (acrylonitrile butadiene styrene) for the headpacks tested on the birds. We additionally required the headpack to disturb the birds' visual field as little as possible. For this we explored a few preliminary designs of the headpack, placing them at different orientations on a virtual mockup of the bird (see Figure 2.4b). We also ensured the headpack was easy to attach and detach, for which we collected feedback on preliminary designs from the falconers.

In our final design, we selected the vertices 2, 3 and 4 (as numbered in Equation A.13) as the ones defining the baseplate. This is where the marker pack was attached to the head of the bird with a piece of Velcro. The baseplate mimicked the shape of the Velcro strip on the bird's head, which approximated a circular sector. This was useful in guiding the placement of the headpack and making it as repeatable as possible. We added a small rectangular lip to the design to facilitate its detachment.

We used retroreflective markers of 6.4 mm diameter, and designed attachment points in the shape of truncated cones to attach them by pressure-fit. To ensure the markers' centres were at the computed vertex coordinates, we added a flat base to mark the surface at which the marker should be flush for its centre to be at the correct position. To ensure the pack was rigid while minimising the weight on the bird's head, we connected the markers with thin rods of 2 mm diameter. Note that the markers' in the pack won't be subject to strong stretching or compression, so no specific structural requirements were considered for the rods. All headpacks tested on the birds weighted 3 g or less, including the markers.

A.8 Design considerations for 3D-printed two-markers supports

We designed supports to hold two markers of 6.4 mm diameter, such that their centres would be placed at predefined distances. As in the headpack design, we used attachment points in the shape of truncated cones where the markers were fit by pressure. We added a flat base to mark the surfacee where the marker should be flush at, for its centre to be at the correct position. We joined the attachment points with a rod of 2 mm diameter.

A.9 Drill speed measurements with tachometer

For the drill experiments, we used a Bosch PSB 650 RE electronic drill (Robert Bosch Power Tools GmbH, Stuttgart, Germany). We selected four speed settings and measured their rotational speed using an optical tachometer (Mengshen Digital Tachometer TH-522; iMengsh, Wuhan, China). We used these values as external validation and for guidance when planning the experiments; however in the data analysis, we used the speed measurement derived from the motion capture data. We took three measurements with the tachometer at each speed setting, taking samples for 15 s and after running the drill for 45 s. The results are shown in Table A.1. The drill was fixed to a table using a 3-axis vice and turned on and off remotely.

Table A.1: Drill speed measurements with optical tachometer. We used an optical tachometer to measure the rotation speed of the drill at four different speed settings, named N1 to N4 in the table.

Speed setting	Measurement 1 (Hz)	Measurement 2 (Hz)	Measurement 3 (Hz)	Mean (Hz)
N1	16.6	16.8	16.8	16.7
N2	23.9	24	23.9	23.9
N3	36.5	36.9	36.9	36.8
N4	40.9	41.2	41.1	41.1

A.10 Computation of ground-truth labels

The following diagrams summarise the computation of the ground-truth labels for the drill trials (Figure A.2) and the bird trials (Figures A.3 and A.4). White boxes with black contour represent input data and blue boxes represent computation steps. Lighter blue tones are used to indicate substeps. Decision points are marked with yellow diamonds and output data is indicated by a red box.

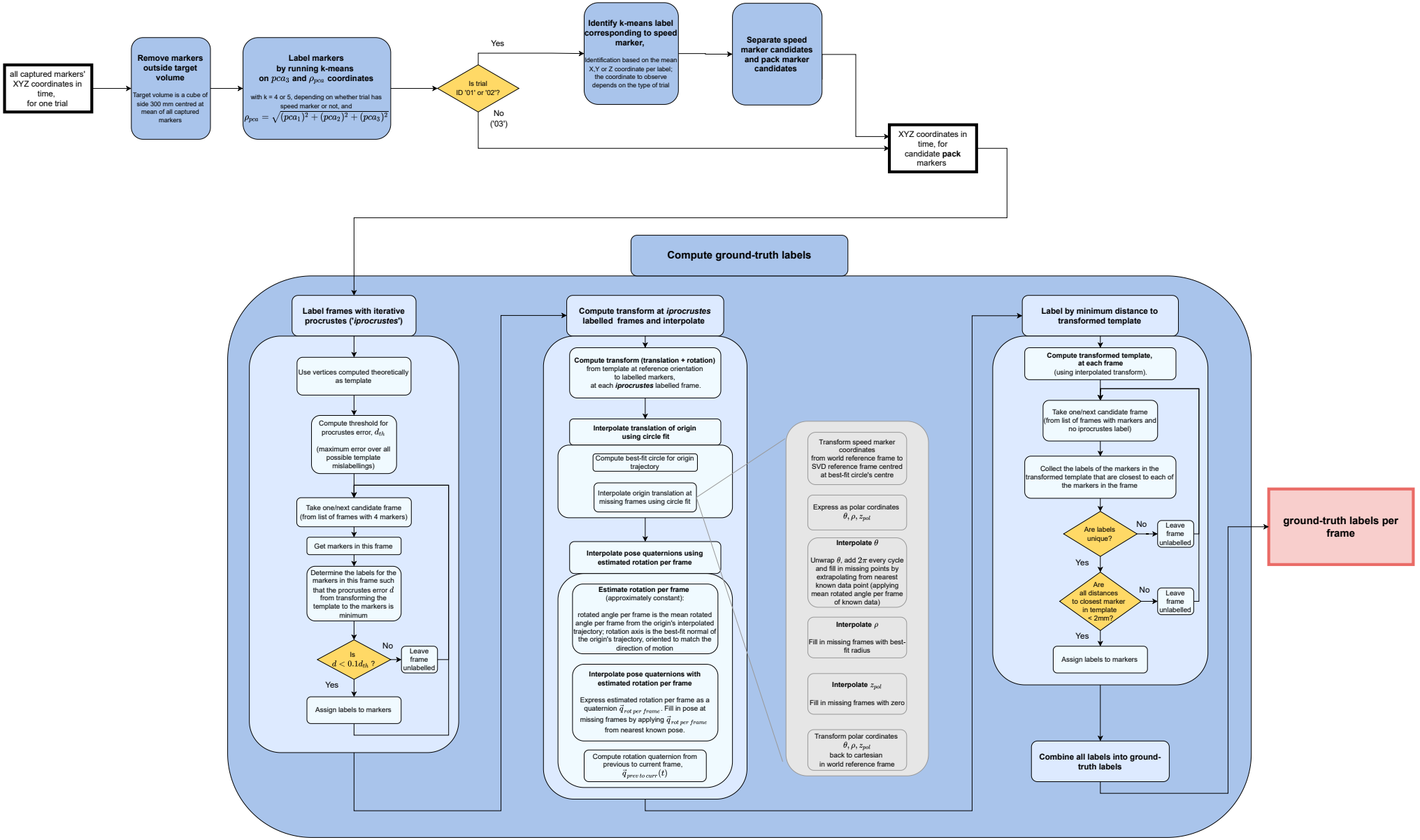


Figure A.2: Diagram for the computation of the ground-truth labels of the headpack markers in the drill trials.

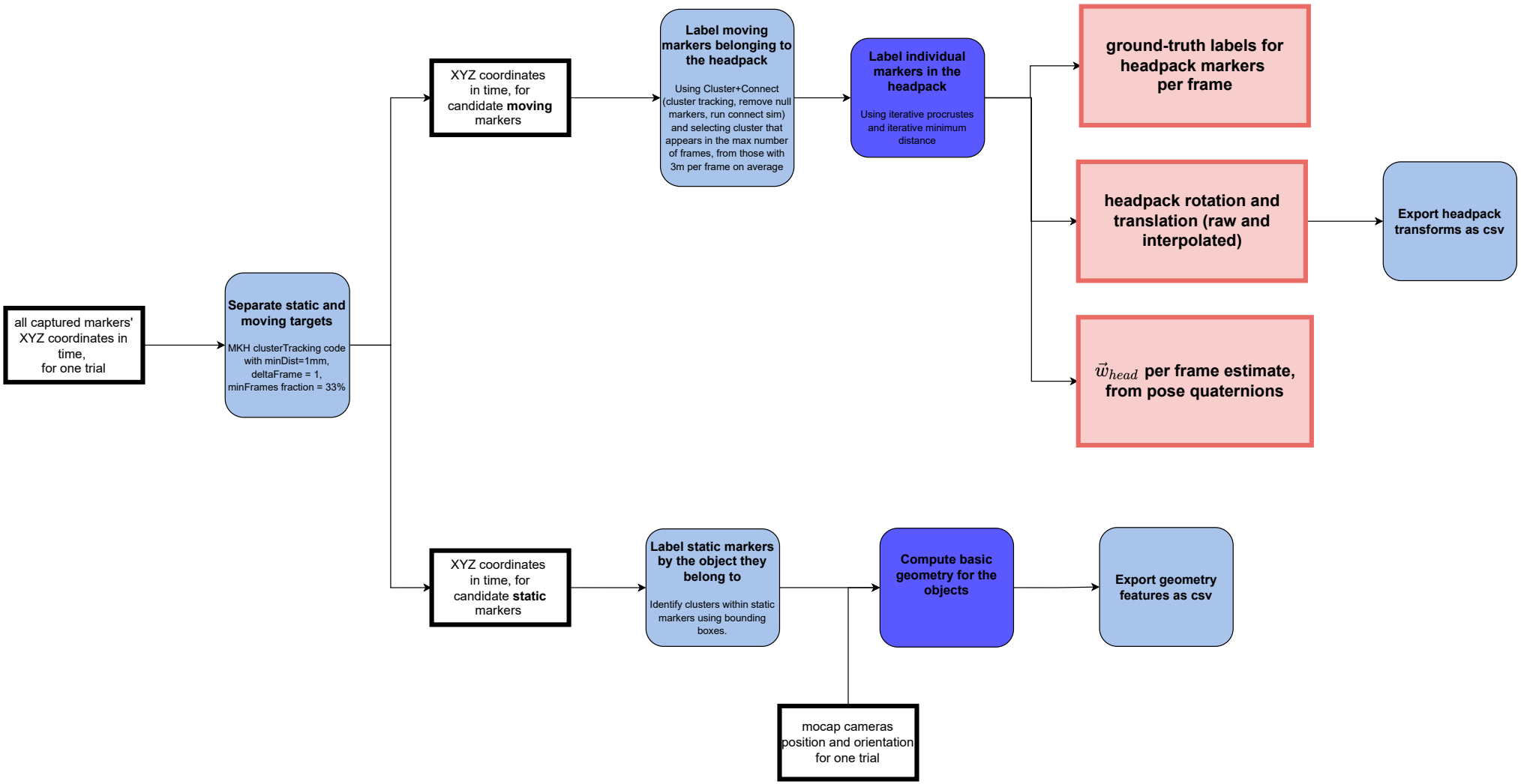


Figure A.3: Overview of the computation of the geometry and the ground-truth labels for the headpack markers in the bird trials.

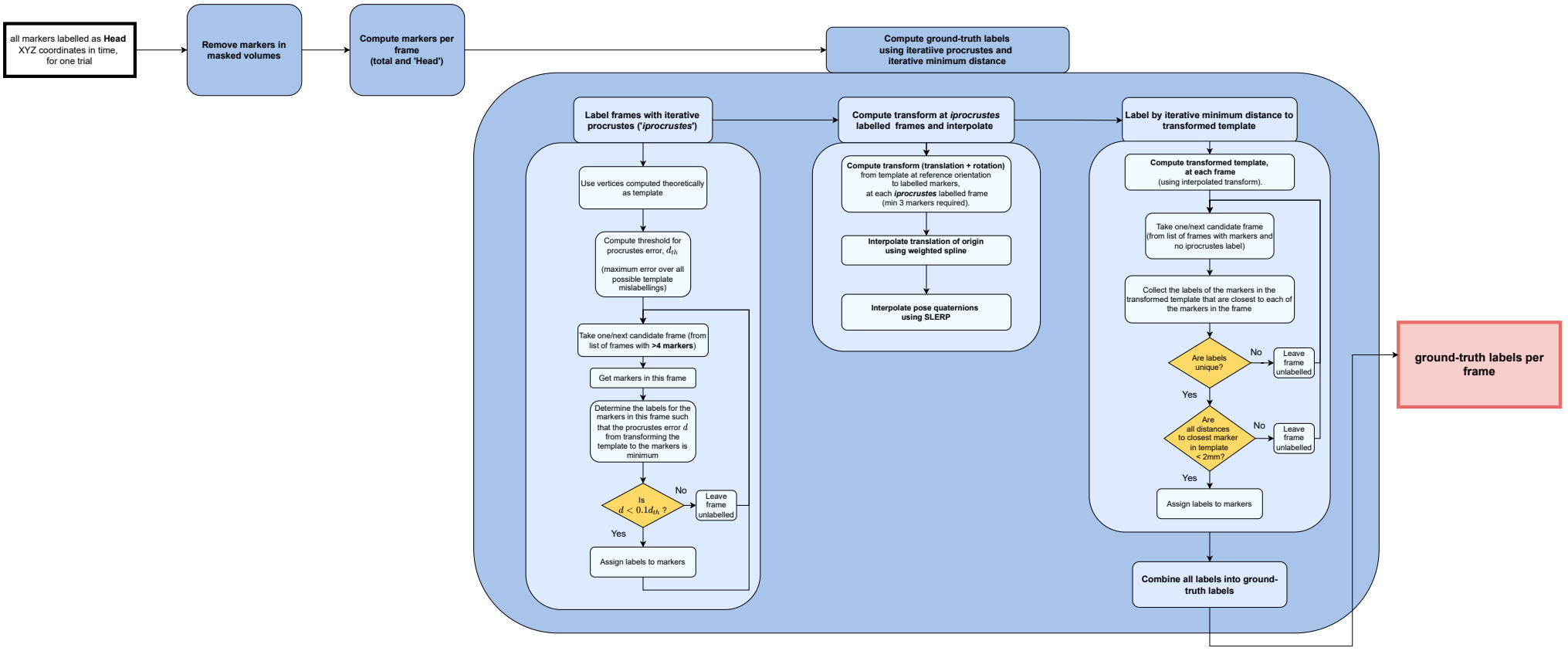


Figure A.4: Diagram for the computation of the ground-truth labels of the headpack markers in the bird trials.

A.11 Definition of templates in Nexus

We used one template for each headpack recorded in the drill trials (six templates), and one template for each bird-headpack pair in the bird trials (16 templates). We used the first recorded trial of each headpack or bird-headpack pair to define the template. For the drill trials, we selected the first frame when all four markers were reconstructed. For the bird trials, we chose the frame closest to the take-off jump in which all four markers were visible.

A.12 Histograms of residuals per trial

We inspected further why the drill trials recorded with $l_{max} = 45$ mm do not show the increased visibility effect when rotating around the z-axis. To do so we examine the residuals in the markers' 3D reconstructions per trial; these are represented in Figure A.5. We find three sets of trials with larger residuals than the rest: the drill trials recorded with $l_{max} = 45$ mm, the drill trials recorded with $l_{max} = 75$ mm and the bird trials recorded with Drogon and the suboptimal headpack. Larger residuals result in larger uncertainty in the markers' position, which results in an increased variance of the edges' lengths $\sigma^2_{i,k}$ and consequently a lower value of ρ. This could explain why the drill trials recorded with $l_{max} = 45$ mm don't show the increased visibility effect and why they have lower mean and median ρ values than the bird trials recorded with the same headpack. The same occurs with the Drogon trials recorded with the suboptimal headpack, which show lower ρ values than the trials recorded with the same headpack on the other birds (see Figure 2.12b, trials 0 to 53). The drill trials recorded with $l_{max} = 75$ mm also have large residuals, but their effect on ρ is likely diminished due to its clusters being more distant in EE space, compared to the $l_{max} = 45$ mm or the suboptimal headpack.

When inspecting the labelling performance of the EE algorithm across headpacks, we find an interesting result for the trials with large residuals. The EE labelling algorithm labels correctly 100% of the candidate frames in those trials when they use an optimal headpack ($l_{max} = 45$ mm and $l_{max} = 75$ mm on the drill). In contrast, in the suboptimal headpack trials on Drogon, the labelling performance is relatively poor (below 96.2%). This suggests a headpack design with a larger ρ value makes the EE labelling algorithm more robust against noisy marker reconstructions. Although this would match our expectations from the theoretical point of view, this effect should be tested directly.

References

Wirth, K. and Dreiding, A. (2009). "Edge lengths determining tetrahedrons". In: *Elem. der Math.* 64, pp. 160–170. DOI: 10.4171/em/129.

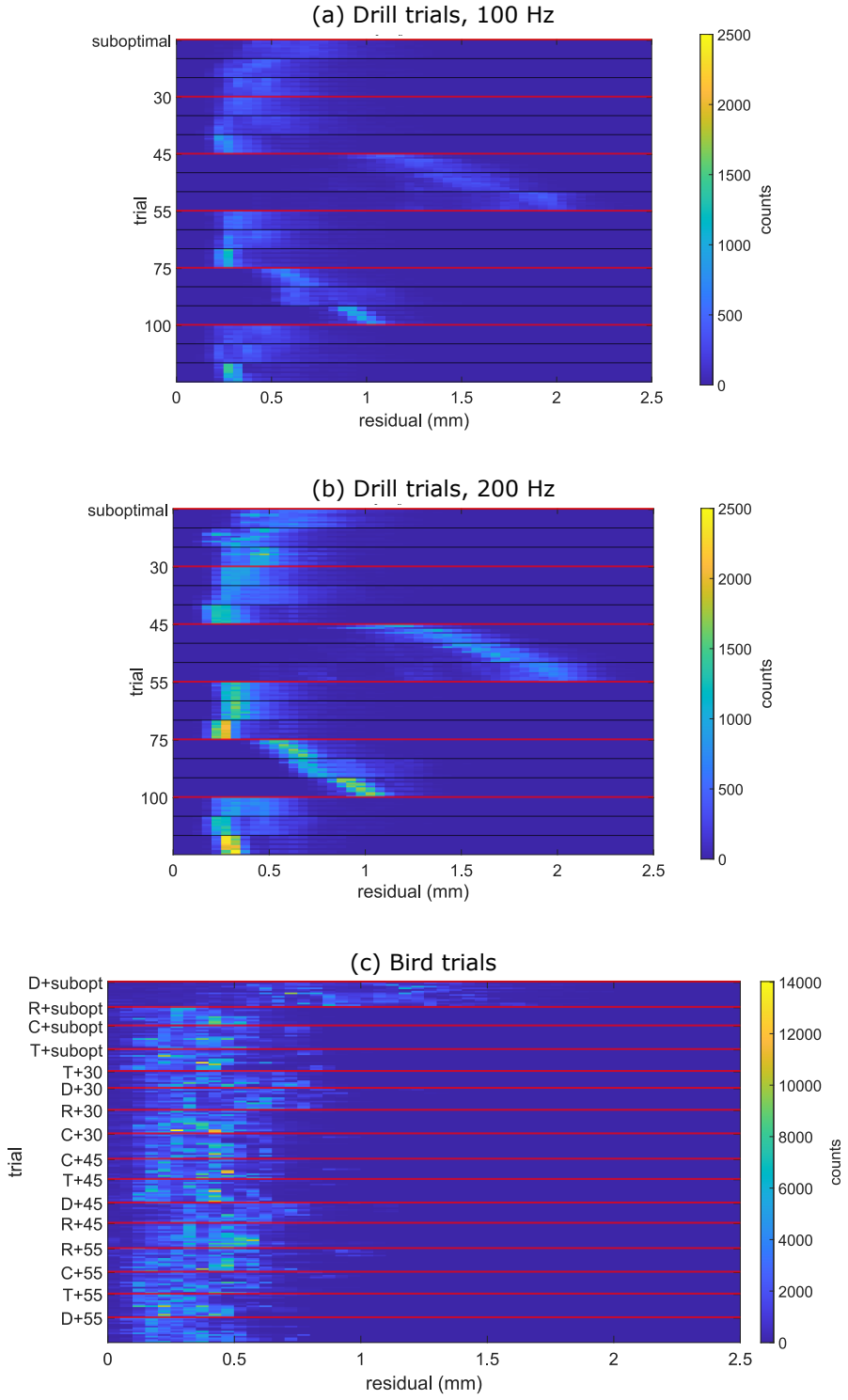


Figure A.5: Histograms of residuals for bird and drill trials. The residuals are computed by the motion capture system for the reconstructed 3D position of each marker. We represent for each trial its histogram of residuals using a bivariate plot, in which the x-axis represents the residuals' values, binned (bin width = 0.05), and the y-axis represents a list of trials. The colour at each x,y location corresponds to the number of residuals observed within a bin for a certain trial. Figures (a) and (b) show the residuals for the drill trials recorded at 100 Hz and 200 Hz respectively. Figure (c) show the residuals for the bird trials. The red lines indicate a change of headpack (in a and b) or a change of bird-headpack pair (in c). Black lines denote a change of rotation axis (in a and b).

B

Appendices for Chapter 3: *Through hawks' eyes: synthetically reconstructing the visual field of a bird in flight*

Contents

B.1 Bird flight experiments 205
B.1.1 Motion capture system 205
B.1.2 Placement of markers 205
B.1.3 Experimental procedure 206
B.2 Motion capture data postprocessing 207
B.2.1 Markers' reconstruction and labelling 207
B.2.2 Headpack transform per frame 209
B.2.3 Target transform per frame 210
B.2.4 Walls, floor and ceiling geometry 211
B.3 Modelling the bird's view 211
B.3.1 Fixation phase in gaze calibration trial 211
B.3.2 Distance between the origin and the midpoint between the eyes 212
B.3.3 Gaze direction estimate 214
B.3.4 Normal to the sagittal plane estimate 216
B.3.5 Retinal margins 217
B.3.6 Trajectory coordinate system 218
B.4 Hybrid model of the lab environment 218
B.4.1 Geometry of the pursuit flight 218
B.4.2 Geometry of the obstacle avoidance flights 218
B.4.3 Dense 3D map of the curtain 219
B.5 Supplementary videos 220
References 221

B.1 Bird flight experiments

B.1.1 Motion capture system

The motion capture lab is equipped with 22 infrared motion capture cameras (Vantage V16, Vicon Motion Systems Ltd, Oxford, UK; sampling rate 200 Hz) and four video cameras (Vue, Vicon Motion Systems Ltd, Oxford, UK; sampling rate 100 Hz, only used for reference). The cameras were mounted on a fixed scaffold from which camouflage netting was hung, to provide a natural-looking background texture and prevent the birds from perching on the lower parts of the scaffolding. The floor was carpeted with green astroturf.

The flights presented in this paper were recorded with two different captive-bred adult Harris' hawks: Drogon (0.6 kg) flew the obstacle avoidance flights, and Toothless (0.7 kg) the pursuit flight.

The pursuit flight is part of a dataset we recorded with four different birds, in which we collected between 3 and 4 flights per bird per day, during four weeks between December 2018 and January 2019, for a total of 251 trials. The number of flights per bird and day in the pursuit dataset was limited by the training approach: an extra food reward was required to get the bird off the target and back to the falconer's fist, which increased the food intake per trial. This reduced the amount of flights we could record while keeping the animals motivated to fly. Resetting the setup between trials also took more time than when recording perching or obstacle avoidance trials.

The obstacle avoidance flights are part of a dataset we recorded with four different birds (Drogon, Toothless, Charmander and Ruby), in which we collected 16 flights per bird per day for a total of two weeks in November 2020. Both periods include the training weeks for the birds.

B.1.2 Placement of markers

To track the birds' head movements, we used rigid supports of retroreflective markers for the birds to wear as a 'headpack'. In the pursuit flight, we used a headpack of 5 retroreflective markers, made of wooden thin rods glued to a plastic cross-shaped plate. In the obstacle avoidance flights, we specially designed 3D-printed rigid supports for 4 retroreflective markers.

The headpack was attached to the head of the bird by the falconer immediately before recording its set of flights, and removed at the end of the set. To ensure that the headpack did not move relative to the bird's head we recorded videos of the birds on the falconer's fist wearing the headpacks. For the flights analysed here, we did not detect visually any movement between the headpack and the head. The head movements when the bird is stationary are generally of larger amplitude than those observed in flight, so we consider these videos are a good indication that there was no significant relative motion between the bird's head and the headpack in flight.

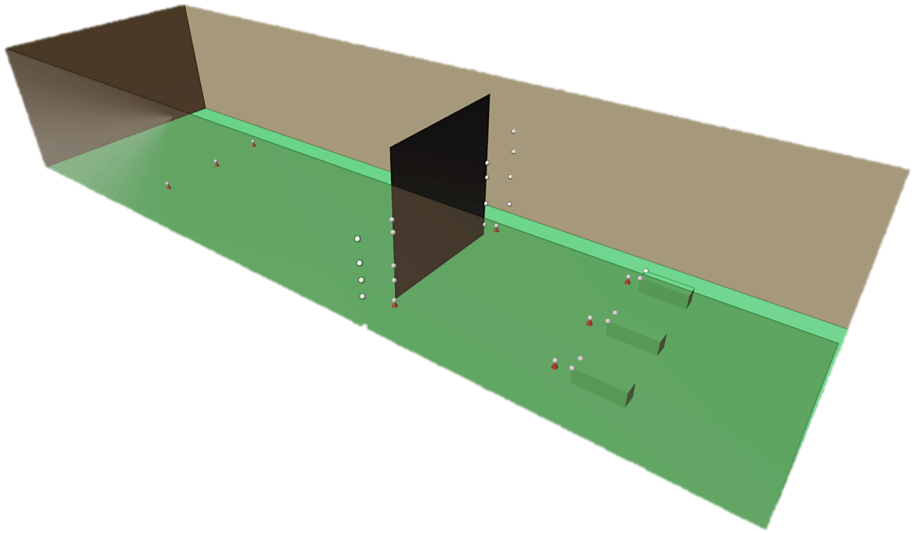


Figure B.1: Reference motion capture markers used in the pursuit trial. The median position of the reference markers in the pursuit trial is shown in a basic geometric model of the lab environment. The sidewalls are shown as partially transparent and the curtain is modelled as a plane, for clarity. The spheres representing the markers' positions are 5 times the markers' real size. Note that the walls' location was estimated from the motion capture cameras' positions and orientations, rather than from the markers placed at the curtain gaps.

In the pursuit flight, we fixed markers of 6.4-10 mm diameter to the pulleys' centres, to the top front vertices of the starting boxes, to the edges of the curtain (which were reinforced with a wooden rod) and to the camouflage netting at the curtain gap. Figure B.1 shows the median position of the reference markers used in the pursuit flight; note that the walls' positions were estimated with the motion capture cameras' position and orientation, rather than with the markers. To track the target's motion, we glued three markers (6.4 mm diameter) to its canvas cover. In the obstacle avoidance flight, we placed 14 mm diameter markers to the perches' edges and at the centre of the obstacles' tops.

B.1.3 Experimental procedure

In the pursuit dataset, the target was pulled using a line of 3 mm thick parachute-style cord, in green to minimise contrast with the astroturf carpet. To keep the target in tension, a long piece of line was attached to its back end and rolled around the first pulley of the target's trajectory. To minimise the risk of the bird predicting the target's trajectory from the line's placement, we placed dummy lines along the alternative paths and fixed them to the pulley's bases using small magnets. In the complete dataset, the takeoff position of the bird and the target were randomised between trials. The target's trajectory also varied between trials: we did this by randomising the pulleys the line would go through (between three options at the start, three options at the end, and two options at the curtain gaps). The curtain was made of two layers of flame-proofed cotton canvas of size 6.0 m $\times$ 3.5 m, and hang from a steel cable installed transversally across the motion capture room.

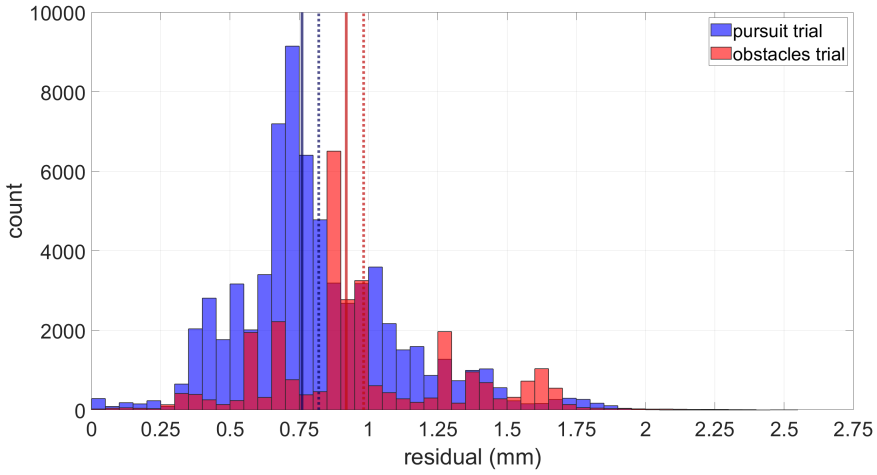


Figure B.2: Histogram of 3D point residuals. The residuals are computed by the motion capture system as part of the reconstruction step, in which the markers' 3D position is determined. The results for the rendered pursuit trial (blue) and obstacle avoidance trial (red) are shown. The obstacle avoidance trial consists of the two obstacle avoidance flights analysed in the main text. The mean (dashed line) and median (continuous line) values for both distributions are under 1 mm. Note that all markers and recorded frames per trial are considered. The bins' width is 0.05 mm.

The obstacle avoidance flights are part of a set in which we recorded the bird's head movements as it flew from the starting perch to the end perch, and back, with and without obstacles in place. For each trial we recorded two flights, that corresponded to the flight back and forth between the perches. Two falconers stood at either end of the room, to handle the bird and provide the food reward. The lateral position of the perches was randomised for every trial between three stations each. These were centred on the longitudinal axis of the room and distributed with 1 m spacing. Each obstacle was made up of two white expanded polystyrene cylinders of 1 m height, stacked on top of each other and bound together with white duct tape. The four obstacles were pushed together so that there were minimal gaps between them. For the trials with obstacles in place, we also randomised the side of the end perch from which the falconer would call the bird.

We calibrated the motion capture system using an active wand and following the manufacturer guidelines. In the obstacle avoidance set of trials, we calibrated before recording a set of trials with the same bird and headpack placement. In the pursuit dataset, we calibrated the system without the curtain in place, and before recording all trials per day. Since we recorded less trials in the pursuit dataset, we don't expect the calibration to deteriorate significantly (the flight presented here is recorded 1.5 hours after calibrating)

B.2 Motion capture data postprocessing

B.2.1 Markers' reconstruction and labelling

We used the commercial software Vicon Nexus 2.8.0, from the motion capture system manufacturers, to extract the unlabelled 3D coordinates of the retroreflective

markers per frame. This process is called 'markers' reconstruction'. The system provides a residual value for each reconstructed marker in 3D space to evaluate the measurement accuracy (Motion Lab Systems, 2021). Figure B.2 shows the histogram of residuals for the trials rendered in this paper. For the pursuit trial, the mean residual is 0.82 mm, the standard deviation is $\sigma = 0.29$ mm, and the median is 0.76 mm, for $n = 69537$ samples. For the obstacle avoidance trial, which consists of the two obstacle avoidance flights considered, the mean residual is 0.98 mm, the standard deviation is $\sigma = 0.34$ mm, and the median 0.92 mm, for $n = 29403$ samples. All markers and recorded frames per trial are included in the computation of the histograms (i.e., frames recorded before takeoff and after landing are also considered).

Nexus is designed for collection and annotation of 3D motion capture data, with a particular focus on human motion. We found that our experiments pushed the system to its limits in terms of tracking and labelling performance, and therefore considered alternative options to process our data. For a detailed description of the typical challenges of using marker-based motion capture on birds the reader is referred to Naik (2021).

For the pursuit flight, we labelled the individual headpack markers per frame using Nexus semi-automatic labelling tool. For the target, it was difficult to separate the individual identities of the markers with manual labelling, as the target twisted and turned during the trial. Instead we assigned a consistent set of three labels to the target markers, that allowed us to separate them easily from the rest. We exported the manual labels to MATLAB, and wrote custom scripts to label the remaining markers.

For the obstacle avoidance dataset, we wrote custom MATLAB scripts to separate headpack and object markers, and label them. To label the headpack markers, we used the ground-truth labelling approach described in Chapter 2 (Section 2.3.3.3 and Appendix A.10). This approach has three main steps:

- first, we labelled the frames with at least 4 markers by iteratively solving the orthogonal procrustes problem with the headpack template; we also determined the headpack's transform (position and orientation) at those frames.
- second, we interpolated the headpack's transform at the remaining frames; and
- third, we iteratively labelled the markers in the remaining frames, by assigning them the label of the closest marker of the corresponding interpolated headpack.

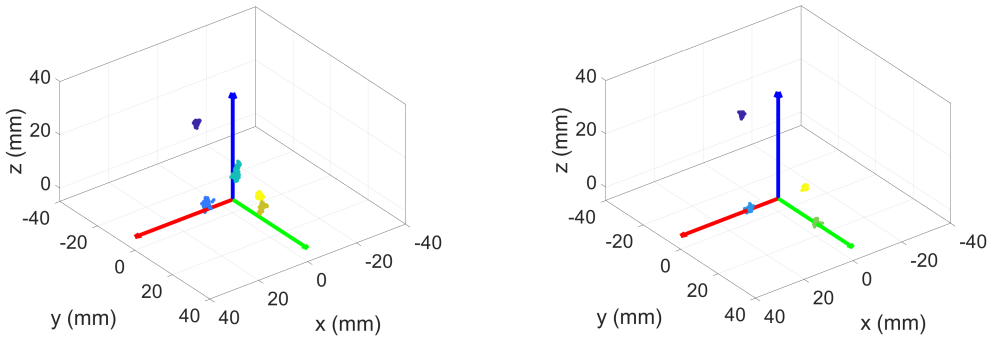
We iterated over the third step, keeping a constant minimum distance for accepting a label but updating the interpolated transform data for the headpack with the newly labelled frames every time. We stopped when there were no new frames labelled between two passes. To label the markers in the objects, we used predefined bounding boxes.

To assess the quality of our labelling of the headpack markers, we computed the maximum distance between a labelled marker and its corresponding position in the headpack template, in a local coordinate system. Results are shown in Figure B.3

Table B.1: Maximum distance d_{max} between labelled markers and corresponding markers in headpack template. We considered all rendered frames in which the headpack transform was not interpolated. Note that we used different headpack designs for the two trials, and therefore the same label does not correspond to the same marker. The obstacle avoidance trial consists of the two obstacle avoidance flights analysed in the main text.

	d_{max} (mm)	
Label	Pursuit trial	Obstacle avoidance trial
1	1.8	1.6
2	3.1	2.0
3	**4.6**	**2.8**
4	2.2	1.6
5	1.6	N.A.

and Table B.1. In line with our observations during manual labelling, the pursuit trial appears to have more jitter in the markers' positions (even though the residuals obtained from the system are not much larger, see Figure B.2). For the set of trials with the same bird and headpack placement as in the selected obstacle trial (which we used to estimate the visual coordinate system), the maximum distance registered across all trials was 3.7 mm.



(a) Pursuit trial (b) Obstacle avoidance trial

Figure B.3: Labelled markers per trial in a local coordinate system. The labelled markers for each of the pursuit (a) and the obstacle avoidance trials (b) are shown in a coordinate system that moves with the headpack. In (a) the visual coordinate system is used and in (b) the headpack coordinate system is shown; both correspond to the template used for labelling with procrustes in each trial. We considered all rendered frames in which the headpack transform was not interpolated. Note that we used different headpack designs for the two trials. The obstacle avoidance trial consists of the two obstacle avoidance flights analysed in the main text.s

B.2.2 Headpack transform per frame

To compute the headpack's transform per frame, we solved the procrustes problem on the labelled data. In the pursuit flight, we used a template computed from the

gaze calibration data. In the obstacle avoidance flights, we used the theoretical headpack design as template.

For the frames where there were not enough headpack markers reconstructed, or where these could not be reliably labelled, we interpolated the headpack's transform. In the pursuit flight, we used a smoothing cubic spline to interpolate the headpack's translation. In the obstacle avoidance flights, we used a weighted variation that takes into account the number of markers detected. In both cases we interpolated the headpack's rotation using the SLERP algorithm for quaternions, which assumes constant angular velocity (Shoemake, 1985). We assumed the short path between quaternions and constant pose in the extrapolation regions.

During manual labelling of the pursuit flight we found that some of the markers' reconstructions were jittery and noisy for certain sections, so we computed the procrustes transform on a subset of frames we determined as most reliable, and interpolated the translation and rotation of the headpack for the rest of the frames. The subset of frames considered most reliable is the set of frames with 3 or more markers reconstructed, excluding those frames in which the procrustes error is above a certain threshold, the markers are close to collinear, or the most prominent frontal marker is missing. We then applied a low-pass filter to the resulting head rotations, as a moving average filter of window size equal to 2 frames, using the SLERP algorithm:

$$q_f(t) = slerp(q_f(t-1), q(t), p), \tag{B.1}$$

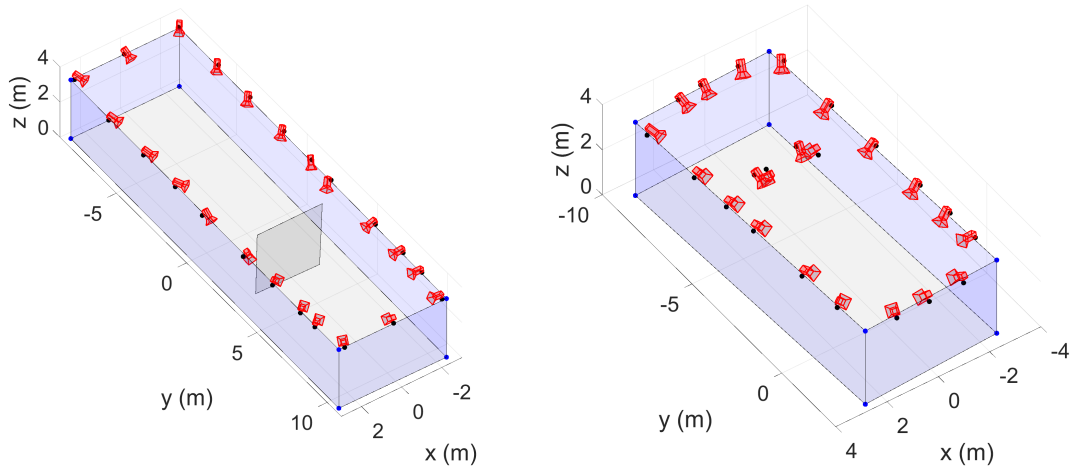
where $q_f(t)$ denotes the filtered quaternion at frame t, $q(t)$ denotes the quaternion data point, *slerp* denotes the implementation of the SLERP algorithm in MATLAB, and p is the interpolation parameter, which we set at 0.5. For the first frame $q_f(0) = q(0)$. We applied the same filter when computing the head rotations for the trajectory coordinate system.

B.2.3 Target transform per frame

For the pursuit flight, we modelled the target as a cylinder, with the three markers attached to it defining a right-angled triangle. We defined a coordinate system linked to the cylinder, whose origin is at the cylinder's centre, its y-axis is parallel to the longitudinal axis of the cylinder (pointing inwards), and its x and z-axis are two perpendicular radii of the cylinder.

We computed the translation of the cylinder as the trajectory of the circumcentre of the triangle. We estimated it using a cubic smoothing spline, that takes into account the number of markers reconstructed.

We define the rotation of the cylinder's coordinate system as follows: its y-axis is always parallel to the origin's velocity vector $\vec{v}_{origin}$, and its x-axis is always parallel to the floor plane of the lab. This mimics the actual motion of the target, which was in tension between two lines, one pulled by the linear motor and one looped around the initial pulley.



(a) Pursuit trial

(b) Obstacle avoidance trial

Figure B.4: Estimated location of the walls. We estimated the location of the walls from the motion capture cameras' positions and orientations for the pursuit flight (a) and the obstacle avoidance flights (b). Cameras are shown in red. From the dimensions of the cameras and their mounts, we estimated points in the scaffolding (black markers). We then used the scaffolding points to the define the walls' planes (blue transparent planes). The corners of the volume are highlighted with blue markers. A plane representing the curtain's location is shown for reference in (a), computed from the markers attached to it.

B.2.4 Walls, floor and ceiling geometry

The motion capture cameras were mounted on a fixed scaffold from which camouflage netting was hung. We estimated its location using the position and orientation of the motion capture cameras. These were computed during the calibration. We used the cameras' positions and orientations to estimate points on the scaffolding, taking into account the dimensions of the cameras and their mounts. We then fitted the estimated scaffolding points to a line for each wall section. This provided an estimate of the scaffolding rungs holding the netting and the cameras. We defined each wall as the plane perpendicular to the floor that contained the corresponding scaffolding rung, from which the camouflage netting was hung.

Figure B.4 shows the estimated walls and the motion cameras for the pursuit and the obstacle avoidance flights. The mean angle at the walls' corners was 90°, with standard deviation $\sigma = 0.2$ for the pursuit flight, and $\sigma = 0.5$ for the obstacle avoidance flights. We set the ceiling at the mean height of the estimated scaffolding points: 3.33 m for the pursuit flight and 3.25 m for the obstacle avoidance flights.

B.3 Modelling the bird's view

B.3.1 Fixation phase in gaze calibration trial

For the pursuit flight, we estimated the visual coordinate system of the bird using data from the gaze calibration trial. During this trial we recorded the bird's head movements while it was held on the falconer's fist and presented with a food reward

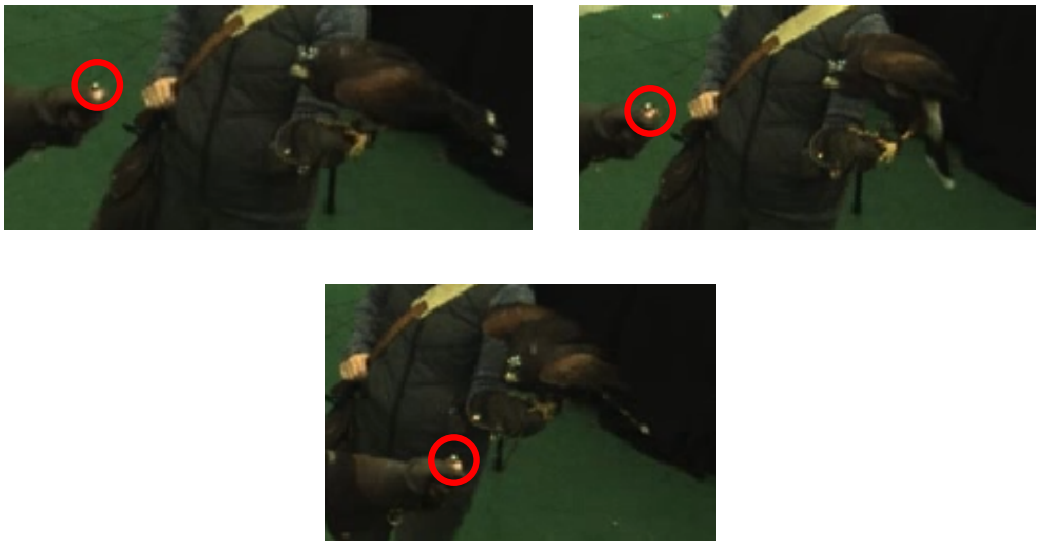


Figure B.5: Examples of the bird's behaviour in the fixation phase of the gaze calibration trial. For certain periods during the gaze calibration trial, the bird appeared to focus on the food reward presented to it, with a characteristic behaviour of lowering its head and preparing its wings for a downstroke. The marker on the food reward is highlighted with a red circle.

that had a marker attached to it. We used the data from the motion capture system's video cameras to identify a range of frames in which the bird is likely to be fixating on the food reward while keeping its head level, based on the animal's behaviour. We selected two sequences of 342 and 120 frames of motion capture data within the trial, in which the bird lowered its head and prepared its wings for a downstroke. Three samples of the typical posture of the bird in this phase are shown in Figure B.5.

B.3.2 Distance between the origin and the midpoint between the eyes

In the pursuit flight, the origin of the visual coordinate system is defined as the midpoint between the two lateral markers of the headpack (see Figure B.6a). In the obstacle avoidance flights, it is defined as the centroid of the headpack markers projected onto the plane defined by the headpack baseplate (see Figure B.8a; note that the headpack designs are not the same in the pursuit flight and the obstacle avoidance flights). We selected these points aiming to be close to the midpoint between the eyes. We estimated how much these points deviate from the midpoint between the eyes using snapshots of the videos of the birds wearing the headpacks, recorded before and after the trials.

The reference images used to estimate this offset for the pursuit flight are shown in Figures B.6 and B.7. Figure B.6a shows the selected origin lies between markers 2 and 5, and Figure B.6b shows it is closer to the sagittal plane (red dashed line) than to marker 5. We therefore estimate that the offset in the direction perpendicular to the sagittal plane is, at most, the distance between the origin and marker 5, which is 10 mm. To estimate the offset in the direction perpendicular

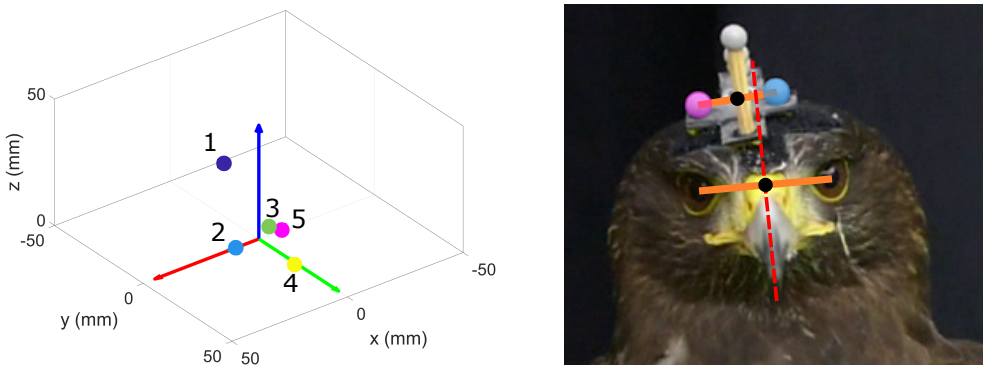


Figure B.6: Distance between selected origin and midpoint between the eyes for the pursuit flight, in the direction perpendicular to the sagittal plane. Figure (a) shows the disposition of the markers in the headpack coordinate system. The origin (which is the same for the headpack, visual and trajectory coordinate systems) lies at the midpoint between markers 2 (blue) and 5 (magenta). A front view of the bird wearing the headpack is shown in (b), with the orange segments highlighting the line between markers 2 (blue) and 5 (magenta), and the approximate line between the eyes. The midpoints of these segments are show as black dots. The trace of the sagittal plane (red dashed line) contains the midpoint between the eyes, and we estimate it based on the symmetry of the bill. We can see that the distance between the sagittal plane and the selected origin is smaller than the distance between the origin and marker 5 (10 mm).

to the headpack's baseplate, we selected three frames in which the baseplate plane was almost perpendicular to the camera plane. We estimated for each of them the real length of the yellow segment, based on the known real length of the green segment, and obtained a mean value of 20.1 mm.

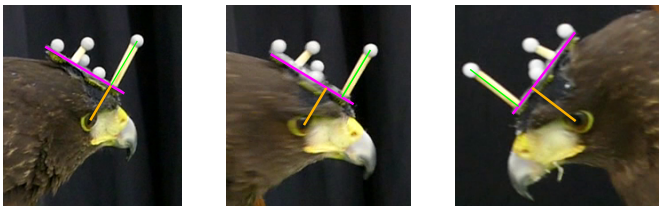


Figure B.7: Distance between selected origin and midpoint between the eyes for the pursuit flight, in the direction perpendicular to the headpack's plate. We selected three frames in which the headpack baseplate was almost perpendicular to the camera plane. We then estimated on each of them the real length of the yellow segment (the approximate distance from the headpack's baseplate to the bird's eye), based on the known length of the green segment (the height of marker 1 above the baseplate). We obtained a mean value of 20.1 mm.

We followed the same approach for the obstacle avoidance flights; the reference images used are shown in Figures B.8 and B.9. From these we estimate that the midpoint between the eyes is within 12 mm from the selected origin in the transverse direction (perpendicular to the sagittal plane), and 17.3 mm in the direction perpendicular to the headpack baseplate.

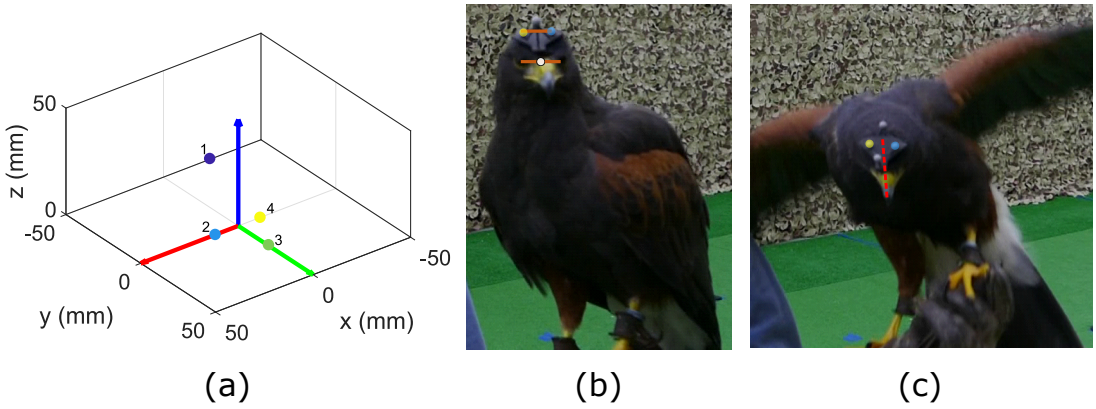


Figure B.8: Distance between selected origin and midpoint between the eyes for the obstacle avoidance flights, in the direction perpendicular to the sagittal plane. Figure (a) shows the disposition of the markers in the headpack coordinate system. The origin (which is the same for the headpack, visual and trajectory coordinate systems) lies approximately between markers 2 (blue) and 4 (yellow). A front view of the bird wearing the headpack is shown in (b), with the line between markers 2 (blue) and 4 (yellow) and the approximate line between the eyes. The midpoint of the line between the eyes (white dot) falls within the segment connecting markers 2 and 4. We therefore estimate that the distance between the midpoint between the eyes and the origin is at most, the maximum distance between the origin and one of the highlighted markers (12 mm). In (c), the red dashed line drawn following the bill's symmetry shows the approximate location of the sagittal plane, which contains the midpoint between the eyes.

B.3.3 Gaze direction estimate

To estimate $\vec{v}_{gaze}$, we computed the orthogonal regression line to the assumed points of fixation in the headpack coordinate system. We used the singular-value decomposition approach described in Söderkvist (2021):

$$USV^T = svd(X - \overline{X}), \tag{B.2}$$

$$\vec{v}_{gaze} = U(:, 1), \tag{B.3}$$

$$A_{gaze} = \overline{X} \tag{B.4}$$

where X denotes the array of sample points (size $3 \times N$, with N the number of sample points) and $\overline{X}$ is the mean of the sample points (size 3×1). The singular-value decomposition is represented by the operator $svd()$. The matrices U, S, V correspond to the matrix containing the left singular vectors in columns, the diagonal matrix containing the singular values, and the matrix containing the right singular vectors in columns, respectively. The estimated gaze direction unit vector $\vec{v}_{gaze}$ is computed as the first column of U. A_{gaze} is the centroid of the sample points, and a point in the fitted line. Table B.2 shows the complete fitting results for the flights considered in the main text.

The root mean square error (RMSE) in Table B.2 is computed with the error being the distance between the samples and the fitted line. We additionally computed the distance d, between the origin of the headpack coordinate system and the fitted line (Table B.2).

For both the pursuit and the obstacle avoidance trial, the distance d is larger than the distance we estimated between the origin of the headpack coordinate



Figure B.9: Distance between selected origin and midpoint between the eyes for the obstacle avoidance flights, in the direction perpendicular to the headpack's plate. We selected three frames in which the bird's sagittal plane was almost parallel to the camera plane. We then estimated on each of them the real length of the yellow segment (the approximate distance from the headpack's baseplate to the bird's eye), based on the known length of the green segment (the height of marker 1 above the baseplate). We obtained a mean value of 17.3 mm.

Table B.2: Results of the gaze direction fitting, for the pursuit and the obstacle avoidance flights. The estimated gaze direction $\vec{v}_{gaze}$ and the point in the fitted line A_{gaze} are both expressed in the headpack coordinate system. The distance between the fitted line and the origin of the headpack coordinate system is d. RMSE stands for root mean square error, where the error is the distance between the sample points and the fitted line. The number of samples per fit is denoted by n. For the pursuit flight, it corresponds to the number of frames in the fixation phase in the gaze calibration trial. For the obstacle avoidance flights, it corresponds to the number of frames in the final approach phase, across all the flights recorded with that bird and headpack placement.

Trial	Bird	$\vec{v}_{gaze}$	A_{gaze} (mm)	d (mm)	RMSE (mm)	n
Pursuit	Toothless	$[0.16, -0.94, 0.32]$	$[103.7, -572.3, 251.5]$	54.3	47.5	462
Obstacles	Drogon	$[0.13, -0.94, 0.31]$	$[131.4, -969.9, 247.0]$	68.4	65.6	1274

system and the midpoint between the eyes in Section B.3.2. Therefore, we don't expect either of these lines to closely represent a line that goes through the midpoint between the eyes. This could be due to the bird fixating on a different points than the ones we assume, due to eye movements that we are not accounting for, or due to existing movement of the headpack relative to the head. We would expect this last effect to be minimal, as we didn't find any evidence of it in the reference videos. In any case, we expect the estimated visual coordinate system to be more representative of the forward direction of the head and of the sagittal plane orientation than a coordinate system simply aligned to the headpack. The fact that our preliminary results seem in agreement with observations previously reported in the literature would support this, although it should be confirmed with more flights.

For the obstacle avoidance trial, we checked for every trial recorded with the

same bird and headpack placement, whether the slope of the fitted line per trial was similar across trials. The results are shown in Table B.3). We computed a deviation metric $\sigma_{\vec{v}_{gaze}}$ for the estimated unit gaze vector per trial, as the square-root of the trace of the covariance matrix:

$$\sigma_{\vec{v}_{gaze}} = \sqrt{\sigma^2_{v_{x,gaze}} + \sigma^2_{v_{y,gaze}} + \sigma^2_{v_{z,gaze}}}, \tag{B.5}$$

obtaining $\sigma_{\vec{v}_{gaze}} = 0.0463$. The mean gaze vector was $\overline{\vec{v}_{gaze}} = \begin{bmatrix} 0.11, & -0.95, & 0.29 \end{bmatrix}$, very similar to the gaze direction estimate considering all trials aggregated (second row in Table B.2). We also computed the distance from the origin of the headpack coordinate system to the fitted line per trial, which yielded a mean value of $\overline{d} = 65.56$ mm and a standard deviation of $\sigma_d = 22.88$ mm, showing the variability in the landing perch midpoint trajectories per trial.

Table B.3: Results of the gaze direction fitting per trial, for all trials with the same bird and headpack placement as the rendered obstacle avoidance trial. The estimated gaze direction $\vec{v}_{gaze}$ and the point in the fitted line A_{gaze} are both expressed in the headpack coordinate system. The distance between the fitted line and the origin of the headpack coordinate system is d. RMSE stands for root mean square error, where the error is the distance between the sample points and the fitted line. The number of samples n corresponds to the number of frames in the final approach phase, for each trial.

Trial	$\vec{v}_{gaze}$			A_{gaze} (mm)			d (mm)	RMSE	n
Drogon perching 1	0.10	−0.94	0.33	136.50	−1108.30	306.00	84.78	21.94	122
Drogon perching 2	0.13	−0.92	0.37	112.38	−1086.61	317.02	113.15	48.073	123
Drogon perching 3	0.17	−0.93	0.33	212.42	−1094.46	310.33	74.83	37.44	122
Drogon perching 4	0.12	−0.95	0.28	122.18	−1138.82	285.44	46.50	44.70	117
Drogon perching 5	0.13	−0.96	0.25	196.22	−1115.80	263.51	50.17	29.78	117
Drogon perching 6	0.12	−0.93	0.36	149.82	−1095.98	315.56	99.115	54.63	124
Drogon perching 7	0.09	−0.96	0.27	172.20	−1126.35	248.55	91.52	43.80	112
Drogon obstacles 1	0.12	−0.96	0.26	141.08	−686.26	172.10	59.95	62.09	57
Drogon obstacles 2	0.07	−0.95	0.29	65.88	−707.29	154.66	62.08	50.43	54
Drogon obstacles 3	0.09	−0.96	0.28	19.18	−711.67	158.51	65.75	48.99	56
Drogon obstacles 4	0.15	−0.96	0.26	117.33	−699.39	160.86	29.74	61.56	58
Drogon obstacles 5	0.11	−0.96	0.27	85.59	−709.25	155.57	44.30	15.51	53
Drogon obstacles 6	0.09	−0.96	0.27	32.18	−710.18	156.70	57.27	24.38	55
Drogon obstacles 7	0.12	−0.96	0.27	109.94	−697.48	153.10	50.87	73.59	55
Drogon obstacles 8	0.12	−0.95	0.28	83.00	−707.49	154.79	53.41	20.28	49

B.3.4 Normal to the sagittal plane estimate

Table B.4 shows the results of the $\vec{n}_{sagittal}$ estimation, for the pursuit and the obstacle avoidance trial.

Table B.4: Results of the normal to the sagittal plane fitting, for the pursuit and the obstacle avoidance trial. The estimated normal $\vec{n}_{sagittal}$ is expressed in the headpack coordinate system. RMSE stands for root mean square error, where the error is the residual of the least-squares problem solved. The number of samples per fit is denoted by n. For the pursuit trial, it corresponds to the number of frames in the fixation phase in the gaze calibration trial. For the obstacle avoidance trial, it corresponds to the number of frames in the final approach phase, across all the trials recorded with that bird and headpack placement.

Trial	Bird	$\vec{n}_{sagittal}$	RMSE	n
Pursuit	Toothless	$[0.98,\ 0.11,\ -0.16]$	0.04	462
Obstacles	Drogon	$[0.99,\ 0.14,\ -0.01]$	0.08	5056

B.3.5 Retinal margins

We computed the retinal margins per eye from the data reported for Harris' hawks by Potier et al. (2016). The authors measured the visual field experimentally, aligning the bird's sagittal plane with a visual perimeter and using an ophthalmoscopic reflex technique. They determined the degree of overlap between the retinal margins of the bird's eyes $\Delta\theta$, at several angles measured from the top of its head ϕ. We digitized the data from the paper (figures 5A, 5C and 6 from Potier et al., 2016) and fitted a smoothing spline periodic boundary conditions (8 polynomial pieces of order 2) to the relation $\Delta\theta$ vs ϕ (RMSE = 1.6°, Figure B.10). From the interpolated overlap $\Delta\theta$ we derived the retinal margins for each eye, θ_{left} and θ_{right}, assuming symmetry with respect to the sagittal plane. We also assumed our estimated gaze direction corresponds to the ϕ angle where the binocular overlap is widest ($\phi = 90°$) (in line with the point raised in the supplementary figure S1 from Potier et al., 2016).

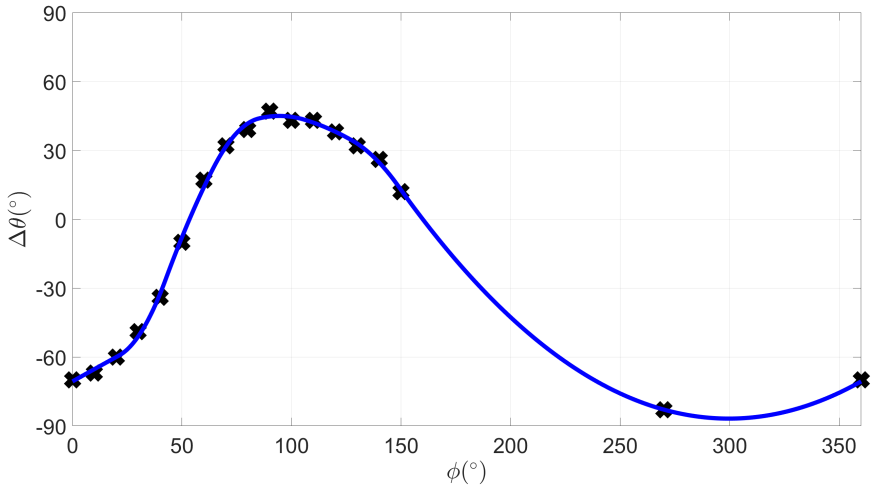


Figure B.10: Visual field of Harris' hawks. We digitized the data from Potier et al. (2016) defining the overlap between the retinal margins of the bird's eyes ($\Delta\theta$) at several angles measured from the top of its head (ϕ). We fitted the data (black crosses) to a smoothing spline (blue dots) with periodic boundary conditions (RMSE = 1.6°).

B.3.6 Trajectory coordinate system

We define the trajectory coordinate system as a coordinate system with its y-axis tangent to the forward direction of the trajectory and its x-axis parallel to the horizontal:

$$\vec{y}_T = -\frac{\vec{v}}{||\vec{v}||}, \tag{B.6}$$

$$\vec{x}_T = \frac{\vec{y}_T \times \vec{z}_{world}}{||\vec{y}_T \times \vec{z}_{world}||}, \tag{B.7}$$

$$\vec{z}_T = \vec{x}_T \times \vec{y}_T, \tag{B.8}$$

where $\vec{x}_T, \vec{y}_T, \vec{z}_T$ represent the versors of the trajectory coordinate system, $\vec{v}$ represents the velocity vector of the head, and $\vec{z}_{world}$ represents the world z-axis. We computed the velocity vector of the head $\vec{v}$ with a central differences scheme on the head interpolated trajectory using the *gradient* function in MATLAB. In the pursuit trial, we filtered the rotations of the trajectory coordinate system using the same approach applied to the visual coordinate system.

B.4 Hybrid model of the lab environment

B.4.1 Geometry of the pursuit flight

The pulleys used to guide the target's trajectory were diabolos of 6.5 cm radius and 15 cm length, glued to a metallic base of 1 cm width. We reduced their geometry in Blender to cones of 6.5 cm radius at the base, and height determined by the corresponding reference markers placed at the diabolos' axes. Some pulleys had additional markers placed on the contour but these were not used to model their geometry.

We modelled the starting boxes as cuboids of 1 m length, with their height and orientation defined by the reference markers placed at their top front vertices.

We modelled the target as a cylinder of 15 cm length and 2.54 cm diameter, based on reference images of the markers' location on the target and on the average distances measured between the markers over the whole trial (see Table B.5). In determining the target's width, we took into account the diameter of the markers attached to it (6.4 mm).

We rendered one trial in which we added a texture to the wall of the motion capture room, that mimicked the texture of the camouflage netting. We did this by tiling a picture of the netting with appropriate scaling. The results are presented in the Online Resources 1 and 2.

B.4.2 Geometry of the obstacle avoidance flights

We defined the geometry and locations of the perches and obstacles based on the position of their corresponding markers (placed at the perches' edges and at the centre of the obstacles' tops).

Table B.5: Mean distances between the markers fixed to the target. Computed from applying *k-means* clustering to the distances between markers labelled as belonging to the target. Length, width and diagonal refer to the dimensions of the cylinder with which we approximate the target's shape.

	mean value (mm)
length	91.2
width	31.8
diagonal	97.4

We reduced the perches to their top rungs and modelled them as cylinders of radius 4 cm (based on measurements of the actual perches). From reference images we estimated that a line between the centres of the markers on the perch's edges would be approximately tangent to the top rung cylinder.

We modelled the obstacles as cylinders of 0.3 m diameter. We estimated the offset between the larger markers' centre and the obstacles' tops by computing the mean deviation from 2 m height for these markers. We took this offset into account (14 mm, including the markers' base) when defining the obstacles' dimensions in Blender.

B.4.3 Dense 3D map of the curtain

The 3D mapping was carried out after collecting all bird trials of the day. To evaluate the error in the captured mesh, we recorded a motion capture trial just before carrying out the 3D mapping, that registered the static motion capture markers in the scene and on the ArUco calibration plate. We call this the mapping trial. We noticed the markers fixed to the curtain weren't clearly identifiable in the captured meshes, so we placed red tape around them to signal their approximate region prior to mapping. These pieces of tape were also in place during the bird trials.

We started the mapping procedure by setting up a local wireless network between the motion capture computer, the augmented reality smartphone and the laptop running *SemanticPaint*, making use of its streaming functionality. We then started the *SemanticPaint* application on the laptop and connected the smartphone's augmented reality app (*Tango Streamer*) to it. Once the connection was verified, the ArUco calibration plate was brought into camera view to compute the required coordinate system transformation. After visually confirming the computed coordinate system for the mesh was correct, the mapping was carried out.

To obtain the 3D map for the curtain, we cropped it from a partial mesh of the lab using MeshLab (Cignoni et al., 2008). We cleaned the resulting mesh by removing duplicate vertices and isolated pieces. We noticed that the floor plane in the captured mesh was slightly deviated from the motion capture coordinate system's floor so we corrected the mesh in MATLAB to match them.

To do this we first fitted a plane to the floor of the mesh with a reference normal equal to $[0, 0, 1]$, using *pcfitplane* in MATLAB. This is an implementation of the M-estimator sample consensus algorithm (MSAC), a variant of RANSAC. The plane was computed with a 99% confidence of finding the maximum number

of inliers, with a maximum distance between an inlier point and the plane of 10 cm, and a maximum angular distance between the normal vector of the fitted plane and the reference orientation $[0, 0, 1]$ of 2.5°. The resulting angle between the fitted plane and the reference orientation was 2.4° and the mean distance between inlier points and the fitted plane was 3.8 cm.

We rotated and translated the input mesh so that the fitted floor plane contained the origin of the motion capture system, and its normal was parallel to $[0, 0, 1]$. This implied the following transforms:

$$R = \begin{pmatrix} 1.00 & 0.00 & -0.04 \\ 0.00 & 1.00 & 0.005 \\ 0.04 & -0.005 & 1.00 \end{pmatrix}, \tag{B.9}$$

$$\vec{t} = \begin{bmatrix} 0.00, & 0.00, & 3.53 \end{bmatrix} \tag{B.10}$$

where R denotes the rotation matrix applied and $\vec{t}$ the translation vector, in centimetres.

Table B.6: Distances between the motion capture markers fixed to the curtain and the subset of red vertices in the curtain's mesh. The values shown represent the maximum, minimum and mean distance (in cm) between the median position of the curtain markers and the 20 nearest neighbours in the red vertices subset of the curtain's mesh.

	Pursuit trial		Mapping trial	
	Original mesh	Transformed mesh	Original mesh	Transformed mesh
Mean (cm)	9.7	9.3	9.1	8.9
Minimum (cm)	5.2	5.7	5.7	5.9
Maximum (cm)	16.4	16.6	13.0	13.3

We computed the deviation from the reference motion capture markers as follows. For the markers placed on the curtain edges, we computed their reference position as the median of its coordinates across all frames, for the pursuit trial and for the mapping trial. From all the points in the mesh, we selected a subset of 400 vertices, whose colour was closest to red (i.e., the vertices whose RGB vector had lowest sum-of-square error from the red vector, $[255, 0, 0]$). For each reference marker, we computed the mean distance to the 20 nearest vertices from the red subset (see Figure B.11). The mean, maximum and minimum distances for the original and the transformed mesh are shown in Table B.6.

B.5 Supplementary videos

Figure B.12 shows a representation of the orthographic projection used in the rendered videos. With an orthographic projection the distortion is less than if we use an equirectangular projection, but the most peripheral parts of the bird's

field of view are not included. Both projections are used to represent the RGB data in the supplementary videos. A description of the supplementary videos is summarised in Table 3.1.

For the pursuit trial, we exported the RGB channel in PNG format, and the rest of channels as part of a multilayer OpenEXR file. In the frame numbering shown in the video, the split between the two phases of the trial (before and after the curtain) is at frame 1132, and the interception frame corresponds to 1374.

For the obstacle avoidance trial, we exported all rendered passes as a mulitlayer OpenEXR file. For the RGB video, we used the *tonemap* function in MATLAB to read the HDR images. The analysis presented in the paper excludes the first and last 20 frames of the videos for each leg of the trial. In the frame numbering shown in the video, the landing perch becomes fully visible at frame 958 in leg 1, and frame 1997 in leg 2 of the trial. The observed saccades in the second leg of the trial occur at approximately frame 2029 and frame 2184. These are more evident in the orthographic projection.

We used the OpenEXR bindings available at https://github.com/skycaptain/openexr-matlab to read the OpenEXR files in MATLAB.

References

Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., and Ranzuglia, G. (2008). "MeshLab: an Open-Source Mesh Processing Tool". In: *Eurographics Ital. Chapter Conf.* Ed. by V. Scarano, R. D. Chiara, and U. Erra. The Eurographics Association. DOI: 10.2312/LocalChapterEvents/ItalChap/ItalianChapConf2008/129-136.

Motion Lab Systems (2021). *The C3D file format: a technical user guide.* Tech. rep. Baton Rouge, Louisiana: Motion Labs Systems, pp. 1–140.

Naik, H. (2021). "XR for all: closed-loop visual stimulation techniques for human and non-human animals". PhD thesis. Munich: Technische Universität München.

Potier, S., Bonadonna, F., Kelber, A., Martin, G. R., Isard, P.-F., Dulaurent, T., and Duriez, O. (2016). "Visual abilities in two raptors with different ecology". In: *J. Exp. Biol.* 219.17, pp. 2639–2649. DOI: 10.1242/jeb.142083.

Shoemake, K. (1985). "Animating rotation with quaternion curves". In: *Proc. 12th Annu. Conf. Comput. Graph. Interact. Tech.* Vol. 19. 3, pp. 245–254. DOI: 10.1145/325165.325242.

Söderkvist, I. (2021). "Using SVD for some fitting problems. Lecture notes for data analysis PhD course from Lulea University of Technology. Retrieved on 30 April 2021".

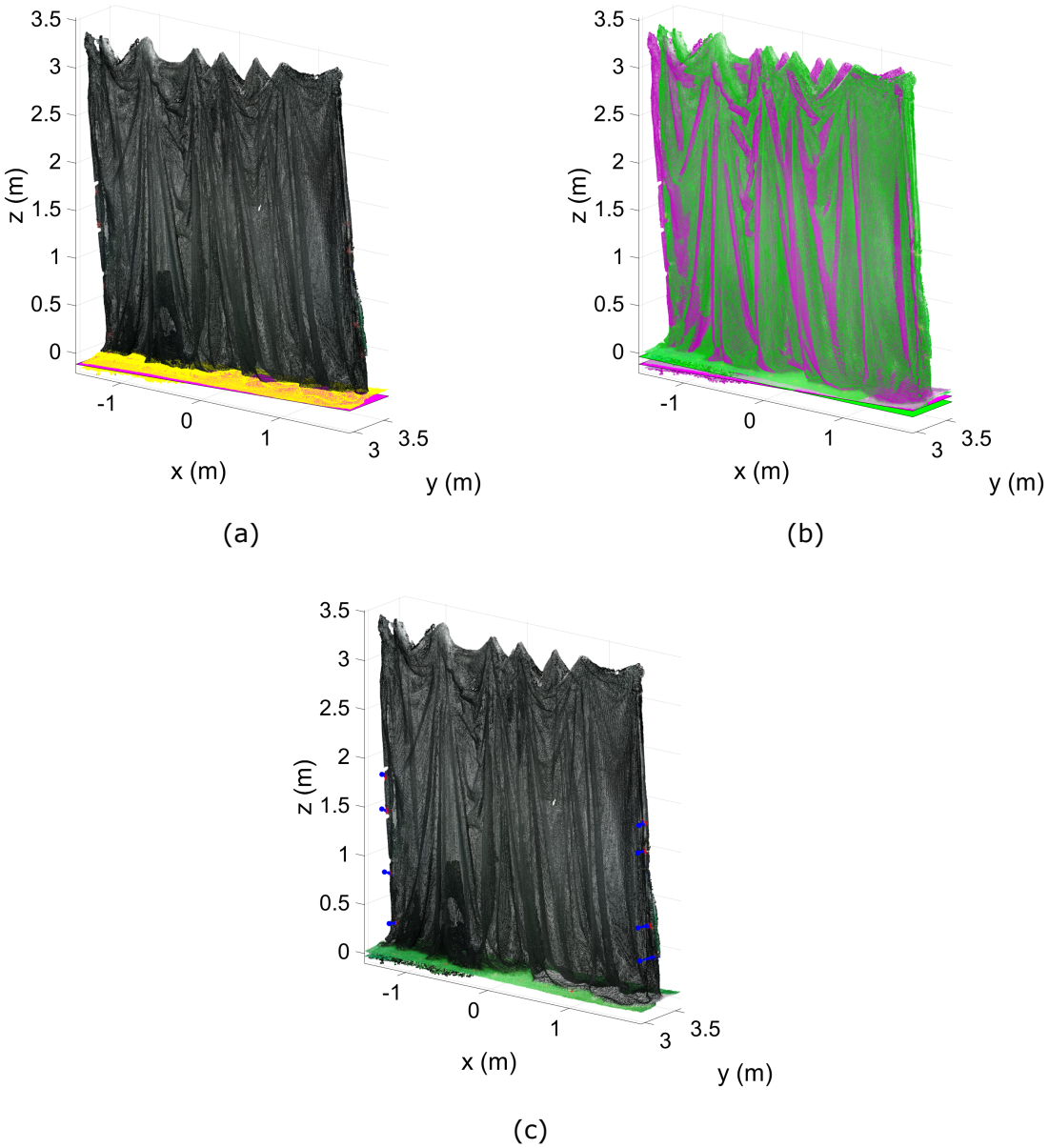


Figure B.11: Correction of the captured curtain mesh. We fitted a plane to the floor of the mesh; figure (a) shows the fitted plane (magenta) and the inlier points (yellow). We rotated the mesh so that its fitted floor plane was parallel to the floor plane of the motion capture coordinate system; figure (b) shows the original mesh and the fitted floor plane (magenta), and the rotated mesh and floor plane (green). We then translated the rotated mesh so that its floor plane contained the origin of the motion capture system (not shown). Figure (c) shows the median position of the markers fixed to the curtain during the pursuit trial (blue dots) and the segments (blue lines) connecting them to the nearest vertices from the red subset of the mesh (highlighted in red).

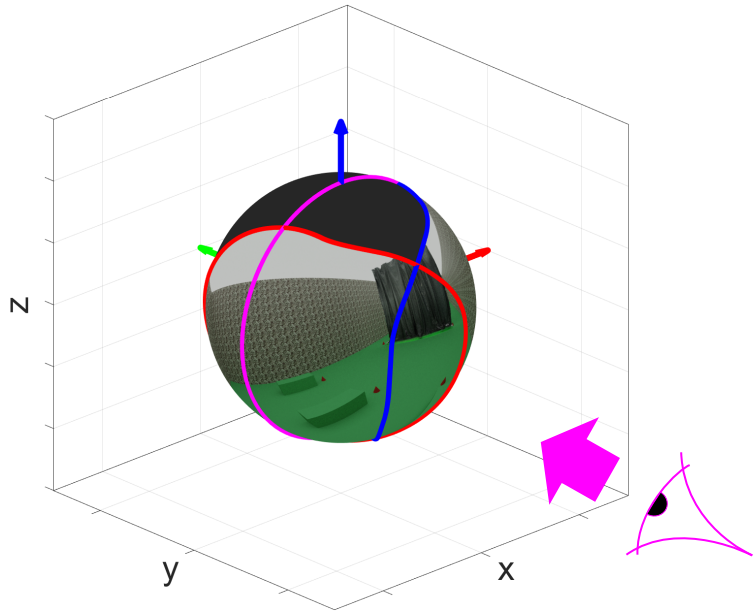


Figure B.12: Orthographic projection of the rendered output. The point of view of the rendered videos in orthographic projection is represented schematically. The rendered output (here, the RGB data) is represented in the visual field unit sphere of the bird. The retinal margins for the left (blue) and right (red) eyes are represented, as well as the blind area (black). The direction of the orthographic projection is represented with a magenta arrow and an eye icon. Note that when projecting the data as indicated by the arrow, the most peripheral parts of the field of view of the bird (beyond the magenta contour) are not included.

C

Appendices for Chapter 4: *Analysis of Harris' hawks' gaze strategies in perching and obstacle avoidance*

Contents

C.1 Motion capture experiments 224
C.2 Motion capture data postprocessing 225
C.3 Headpack coordinate system 226
C.4 Visual coordinate system 227
C.5 Spherical histograms 229
C.6 Latitudinal and longitudinal extension of the landing perch and the obstacles 238
C.7 Alignment with landing perch centre or bisector direction 238
C.8 Deviation from an ideal symmetric expansion of the landing perch 247
C.9 Mean position of the landing perch centre in the visual field 249
C.10 Mean position of the obstacle's edge in the visual field 249
C.11 Location of the minima of optic flow in the bird's visual field 252
References 252

C.1 Motion capture experiments

We worked with four captive-bred adult Harris' hawks, comprising three males (Drogon, Toothless, Charmander) and one female (Ruby). We recorded approxi-

mately 16 flights per bird per day, for a total of two weeks (including the training period for the birds) in November 2020. The data presented in this chapter are from flights recorded on one day of experiments, after six days of training. The birds wore different headpack designs because we used the collected data to validate a custom headpack tracking algorithm, described in Chapter 2 (see Figure C.1).



Figure C.1: Birds and headpacks used. We collected motion capture data from four Harris's hawks (*Parabuteo unicinctus*) wearing four headpack designs.

We recorded 279 trials over four days, across the four birds; however in this chapter we focused on a subset of 74 trials (148 flights; each trial recorded two flights) recorded on one day of experiments. From this subset, we discarded the trials in which any of the two flights comprising the trial had invalid take-offs or landings; we considered these invalid if the bird did not touch the perch with both legs. We manually determined the take-off and landing frames as the point at which the head was lowest when the bird was about to leave the perch or had just landed on it. The RGB rendered videos include a buffer of 20 frames before the take-off and landing frames used in the semantic data analysis.

For each recorded trial, two falconers stood at either end of the room, to handle the bird and provide the food reward. For the obstacle avoidance trials, we used a set of four styrofoam pillars. Each pillar was made up of two white expanded polystyrene cylinders stacked on top of each other and bound together with white duct tape. The four obstacles were pushed together so that there were minimal gaps between them. In the obstacle avoidance trials we also randomised the side of the end perch from which the falconer would call the bird.

We calibrated or recalibrated the motion capture system before each recording each bird's set of trials (i.e., four times per day). The headpack was attached to the head of the bird by the falconer immediately before recording the bird's set of trials, and it was removed at the end. For the first half of the trials, we recorded perching manoeuvres. For the second half, we recorded obstacle avoidance manoeuvres. To ensure that the headpack did not move during the flights, we recorded videos of the birds on the falconer's fist, wearing the headpacks, before recording their trials.

C.2 Motion capture data postprocessing

We labelled the headpack markers following the procedure described in Chapter 3 Appendix B.2, which in turn follows the ground-truth labelling method for bird

trials introduced in Chapter 2 (Section 2.3.3.3 and Appendix A.10). The method relies on first labelling the frames with 4 or more headpack markers by solving the orthogonal procrustes problem, and then labelling the remaining data using estimations of the headpack's position and orientation at the corresponding frames.

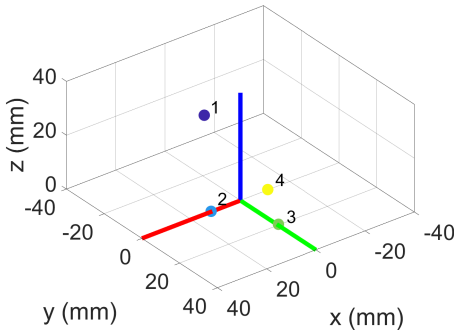
For the few frames where there were not enough headpack markers reconstructed, or where these could not be reliably labelled, we interpolated the headpack's translation using a weighted cubic spline that took into account the number of markers detected. To interpolate the headpack's rotation, we used the SLERP algorithm for quaternions (Shoemake, 1985), which assumes constant angular velocity. As a verification of the quality of the labelling, we computed the distance between labelled markers in the headpack coordinate system and their reference position in the template, for the frames in which the headpack transform was defined (i.e., not interpolated). For the 102 flights considered in the chapter across all birds (97 flights plus the 5 flights in which the birds flew over the obstacles), the labelled markers deviated less than 3.7 mm from their corresponding template position.

We followed the approach described in Chapter 3 Appendix B.2.4, to define the position of the walls per trial. We used the motion capture cameras' positions and orientations and the dimensions of their mounts to estimate points in the scaffold rungs in which the cameras were mounted. We used these points to define planes, modelling the netting that was hung from the scaffold during the experiments (see Appendix B.2.4). We set the ceiling at the mean height of the estimated scaffolding points. The motion capture cameras' positions and orientations were determined during the calibration, and therefore were the same for all trials recorded with the same bird. For all calibrations, the mean angle between consecutive walls was 90°, the standard deviation $\sigma = 0.5°$ and the ceiling mean position 3.25 m.

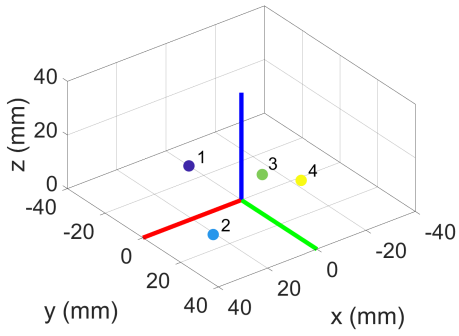
C.3 Headpack coordinate system

Figure C.2 shows the position of the markers for each of the headpack designs used in Chapter 4. The auxiliary coordinate system fixed to the headpack is defined as follows:

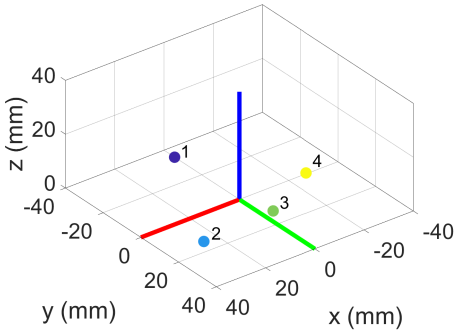
- the headpack baseplate plane is the plane defined by markers 1, 2 and 3 in the optimal headpacks, and by markers 2, 3 and 4 in the suboptimal one;
- the origin of the headpack coordinate system is the projection of the mean position of the markers (i.e., the centroid) on the headpack baseplate plane;
- the y-axis is defined as the unit vector from the origin to the projection of marker 1 on the baseplate plane (for the optimal headpacks, the projection of marker 1 coincides with the position of the marker, as it is contained in the baseplate plane);
- the x-axis is defined perpendicular to the y-axis and contained in the baseplate plane;
- the z-axis is the required one to form a right-hand coordinate system.



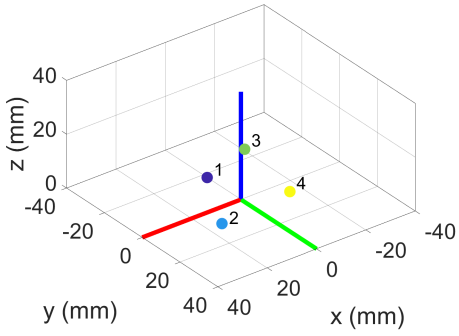
(a) Drogon, suboptimal headpack



(b) Charmander, optimal headpack
lmax=45 mm



(c) Ruby, optimal headpack
lmax=55 mm



(d) Toothless, optimal headpack
lmax=30 mm

Figure C.2: Headpack coordinate system for each design. The figure shows the position of the markers for each headpack design. The markers are coloured by their assigned label (show in text). The optimal headpacks are named based on their maximum distance between markers l_{max}. The headpack coordinate system x, y and z axes are shown in red, green and blue respectively.

C.4 Visual coordinate system

To estimate the visual coordinate system, we assume that the bird fixates its gaze on the perch midpoint during the final approach phase, and that it keeps its eyes level in flight. These assumptions are based on reports on previous literature (Brighton et al., 2019; Ros et al., 2017; Potier et al., 2016; Warrick et al., 2002) and on observations from our experiments; some examples are shown in Figure C.3.

We defined the final approach phase based on the distance between the bird and the perch: 0.5 to 2.0 m for perching trials and 0.5 to 1.0 m for obstacle trials. The thresholds are different for the two types of trials because we found gaze shifts up to approximately 1 m ahead of the perch in the obstacle avoidance trials; this is shown in Figure C.4. Note that the obstacles are 1.5 m ahead of the end perch.

Table C.1 includes additional results for the fitting of the gaze direction per bird, $\vec{v}_{gaze}$. The distance d between the origin of the headpack coordinate system and the fitted line is larger than the distance we expect between the origin of the headpack coordinate system and the midpoint between the bird's eyes. Therefore, we don't

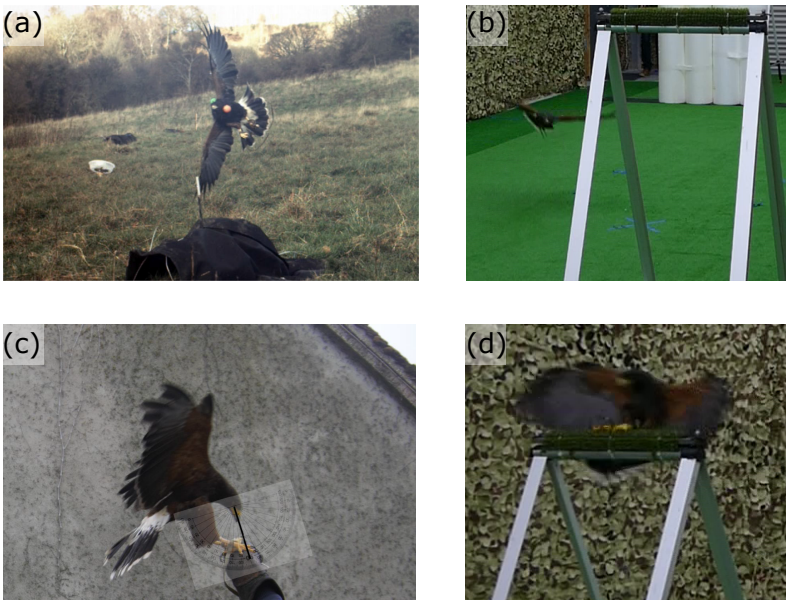


Figure C.3: Perch fixation and eyes-level assumption. Examples of a bird keeping its eyes level in flight, from literature Brighton et al., 2019 (a) and from our experiments (b). Examples of a bird fixating on the perch upon landing, from literature Potier et al., 2016 (c) and from our experiments(d).

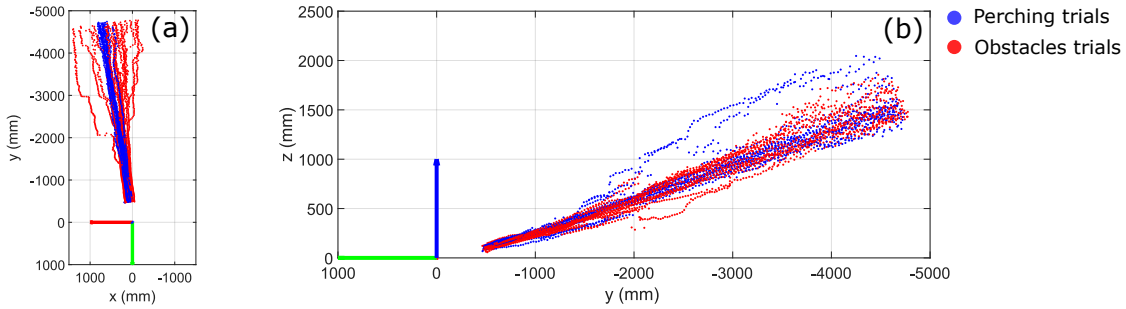


Figure C.4: Trajectories of the landing perch midpoint in the headpack coordinate system. Trajectories shown for all flights of one bird (Drogon), since the bird is 5 m away from the perch until it is 0.5 m from it. Figure (a) shows the projection onto the xy plane. Figure (b) shows the projection onto the yz plane. Note the head shifts (in the shape of steps), especially in the obstacle trials top-down view (Figure a, red curves).

expect the fitted line to closely represent a line that goes through the midpoint between the bird's eyes. We further inspected this by fitting a $\vec{v}_{gaze}$ vector per trial, for all of the trials executed by Drogon (see Appendix B.3.3, Table B.3). We found a small variation across trials in the fitted direction, but a larger variation in the distance d. This may result from the bird fixating a point relative to its head, but making small adjustments to locate this point in a relevant area of the visual field via eye movements. It could also be due to the bird fixating on a point that is close but not exactly at the perch's midpoint, or due to existing movement of the headpack relative to the head. However we expect this last effect to be minimal: we didn't find any evidence of it in the reference videos of the birds on the falconer's fist wearing the headpacks while moving their heads. The results from the behavioural analyses suggest that the estimated gaze direction is a reasonable estimate of the

forward direction of the bird's head, but without measurements of the eyes' motion, it is challenging to precisely determine the source of variation in the distance d.

Table C.1: Results of the gaze direction fits, considering all trials per bird. The estimated gaze direction $\vec{v}_{gaze}$ and the point in the fitted line A_{gaze} are both expressed in the headpack coordinate system. The distance between the fitted line and the origin of the headpack coordinate system is d. RMSE stands for root mean square error, where the error is the distance between the sample points and the fitted line. The number of samples n corresponds to the number of frames in the final approach phase, across all trials per bird.

Bird	$\vec{v}_{gaze}$	A_{gaze} (mm)	d (mm)	RMSE (mm)	n
Drogon	$[0.13, \ -0.94, \ 0.31]$	$[131.4, \ -969.9, \ 247.0]$	68.4	65.6	1274
Toothless	$[-0.07, \ -0.95, \ 0.31]$	$[-55.5, \ -969.6, \ 231.1]$	82.1	46.7	847
Charmander	$[0.004, \ -0.97, \ 0.23]$	$[36.7, \ -1046.4, \ 174.6]$	73.9	64.7	1328
Ruby	$[0.16, \ -0.96, \ 0.22]$	$[206.5, \ -966.0, \ 158.3]$	66.3	70.6	1353

C.5 Spherical histograms

Figure C.5 shows a schematic of how the spherical histograms were computed.

Figure C.6 shows two spherical histograms represented in the trajectory coordinate system. Note how the histograms reflect the shape of the bird's head trajectory relative to the objects of interest (i.e., the landing perch or the obstacles). This is easily seen for the histograms computed for the landing perch, in the perching trials (Figure C.6a). In them, the landing perch appears centred in longitude, but shows variation in latitude. This reflects the geometry of a perching manoeuvre in which the bird flies straight to the landing perch with a dive and a subsequent pitch-up. In the equivalent histogram for the obstacle avoidance trials (Figure C.6b), the contour of the landing perch additionally shows variation in longitude, reflecting a larger turn relative to the landing perch than in the perching trials. The comparison with the histograms computed from the visual coordinate system data show how the birds' gaze strategy unifies the view of the landing perch across different manoeuvres.

Figures C.7 to C.13 show the spherical histograms for the contour semantic data grouped by bird, type and leg of the trial, and coordinate system. Figure C.13 shows the results for the flights in which two birds flew over the obstacles, rather than around them.

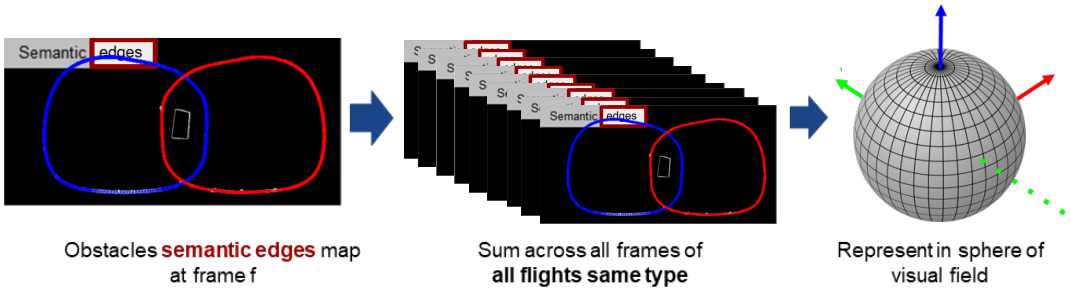


Figure C.5: Spherical histograms computation. The figure shows a schematic of how the spherical histograms were computed, focusing on the case of the obstacles. For each frame f, we set the value of the pixels defining the contour of the obstacles to 1, and the rest to 0 (left panel). Only pixels within the visible regions of the bird's visual field are considered; the retinal margins for the left (blue) and right (red) eye are shown for reference. We then add the results across the subset of frames considered and represent the results in the sphere of the bird's visual field.

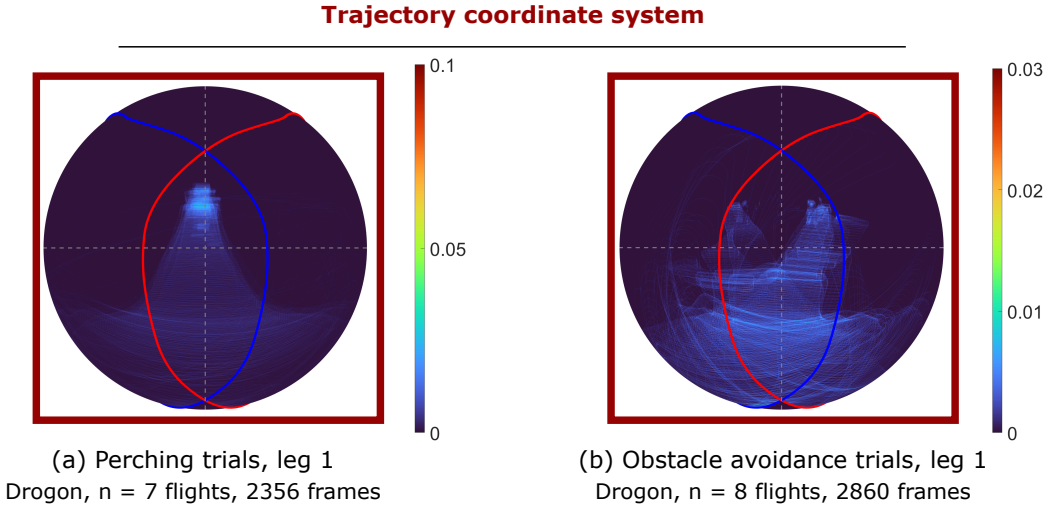


Figure C.6: Spherical histograms in trajectory coordinate system reflect the geometry of the trajectory. The figure shows the spherical histograms for the landing perch, for the flights executed by Drogon in leg 1 of the perching trials (a) and the obstacle avoidance trials (b). The visual field spheres are represented using an orthographic projection on the xz plane of the trajectory coordinate system; only the frontal hemispheres of the visual field are represented. The retinal margins from the right (red) and left (blue) eyes are shown for reference. The range of frames considered includes all frames from takeoff to landing, excluding the frames in which the head pose of the bird was interpolated (see number of flights n and total number of frames in subfigures' captions). Note the colourbars in (a) and (b) are not in the same range: the maximum values were selected to highlight the points of highest h metric per pixel in the corresponding representations in the visual coordinate system (see Figures C.7 and C.9). The planes at 0° latitude and 0° longitude are shown for reference (grey dotted lines).

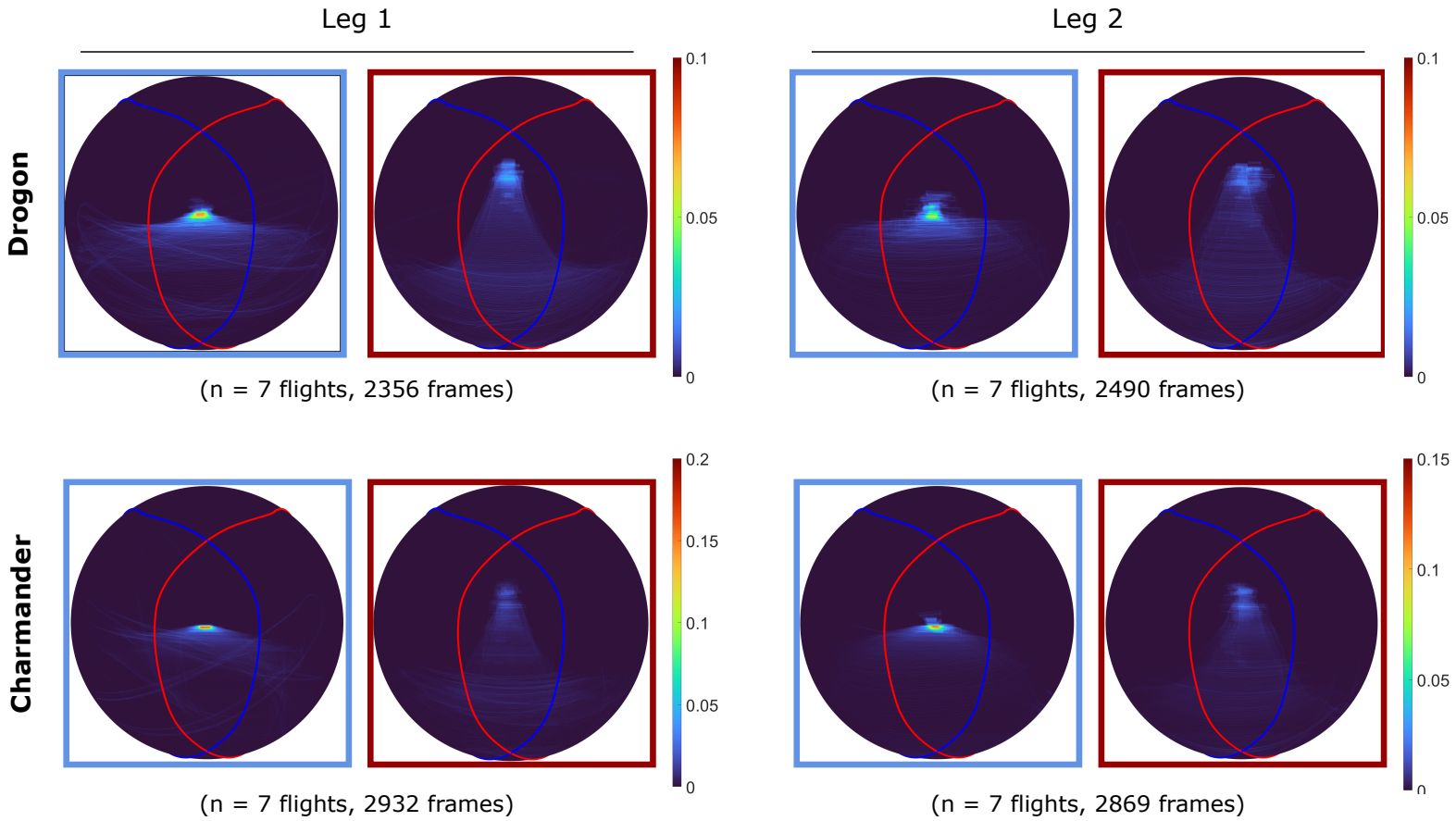


Figure C.7: Spherical histograms for the landing perch in the perching flights - Drogon and Charmander. The figure shows the spherical histograms for the landing perch in the perching flights of Drogon and Charmander, grouped by bird, leg of the trial and coordinate system. The visual field spheres are represented using an orthographic projection on the xz plane of the visual coordinate system (blue boxes) or the trajectory coordinate system (red boxes); only the frontal hemispheres of the visual field are represented. The retinal margins from the right (red) and left (blue) eyes are shown for reference. The range of frames considered includes all frames from takeoff to landing, excluding the frames in which the head pose of the bird was interpolated (see number of flights n and total number of frames in subfigures' captions). Note the colourbars across birds and legs of the trial are not in the same range: the maximum values were selected to highlight the points of highest h metric per pixel in the visual coordinate system representation.

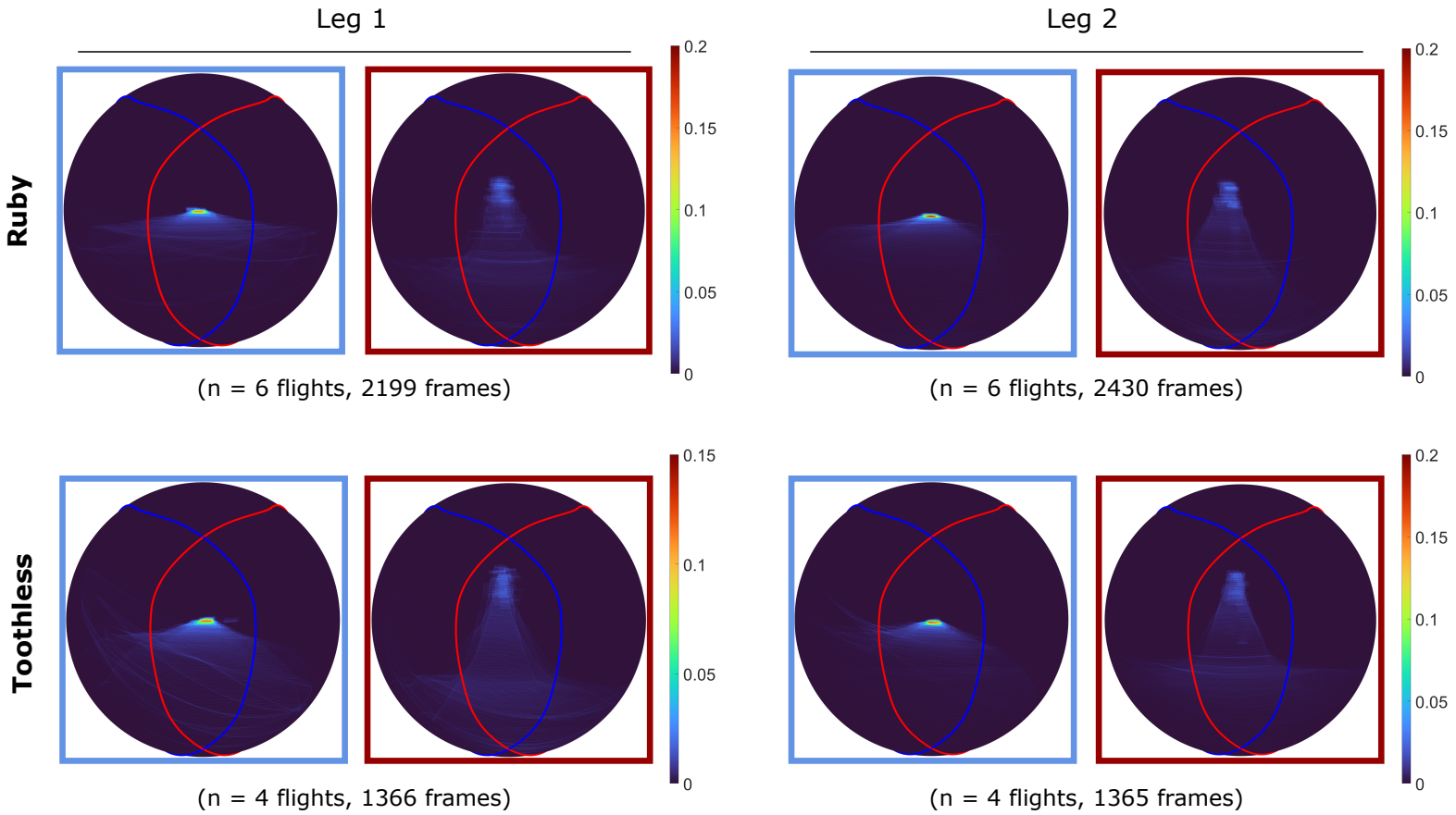


Figure C.8: Spherical histograms for the landing perch in the perching flights - Ruby and Toothless. The figure shows the spherical histograms for the landing perch in the perching flights of Ruby and Toothless, grouped by bird, leg of the trial and coordinate system. The visual field spheres are represented using an orthographic projection on the xz plane of the visual coordinate system (blue boxes) or the trajectory coordinate system (red boxes); only the frontal hemispheres of the visual field are represented. The retinal margins from the right (red) and left (blue) eyes are shown for reference. The range of frames considered includes all frames from takeoff to landing, excluding the frames in which the head pose of the bird was interpolated (see number of flights n and total number of frames in subfigures' captions). Note the colourbars across birds and legs of the trial are not in the same range: the maximum values were selected to highlight the points of highest h metric per pixel in the visual coordinate system representation.

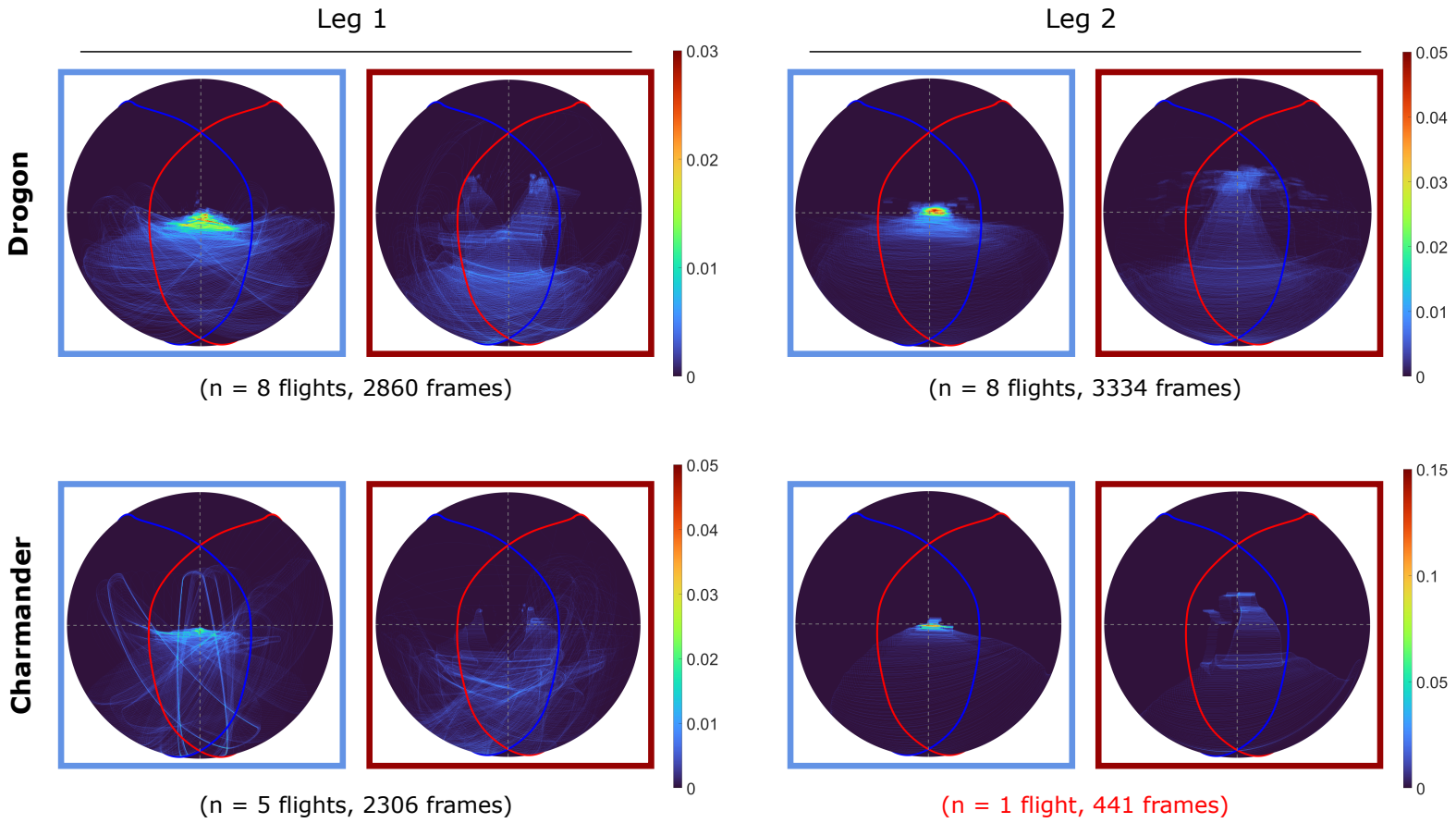


Figure C.9: Spherical histograms for the landing perch in the obstacle avoidance flights - Drogon and Charmander. The figure shows the spherical histograms for the landing perch in the obstacle avoidance flights of Drogon and Charmander, grouped by bird, leg of the trial and coordinate system. The visual field spheres are represented using an orthographic projection on the xz plane of the visual coordinate system (blue boxes) or the trajectory coordinate system (red boxes); only the frontal hemispheres of the visual field are represented. The retinal margins from the right (red) and left (blue) eyes are shown for reference. The range of frames considered includes all frames from takeoff to landing, excluding the frames in which the head pose of the bird was interpolated (see number of flights n and total number of frames in subfigures' captions). Note the colourbars across birds and legs of the trial are not in the same range: the maximum values were selected to highlight the points of highest h metric per pixel in the visual coordinate system representation. The figure caption in red highlights that some flights were excluded from the set because the bird flew over the obstacles rather than around them.

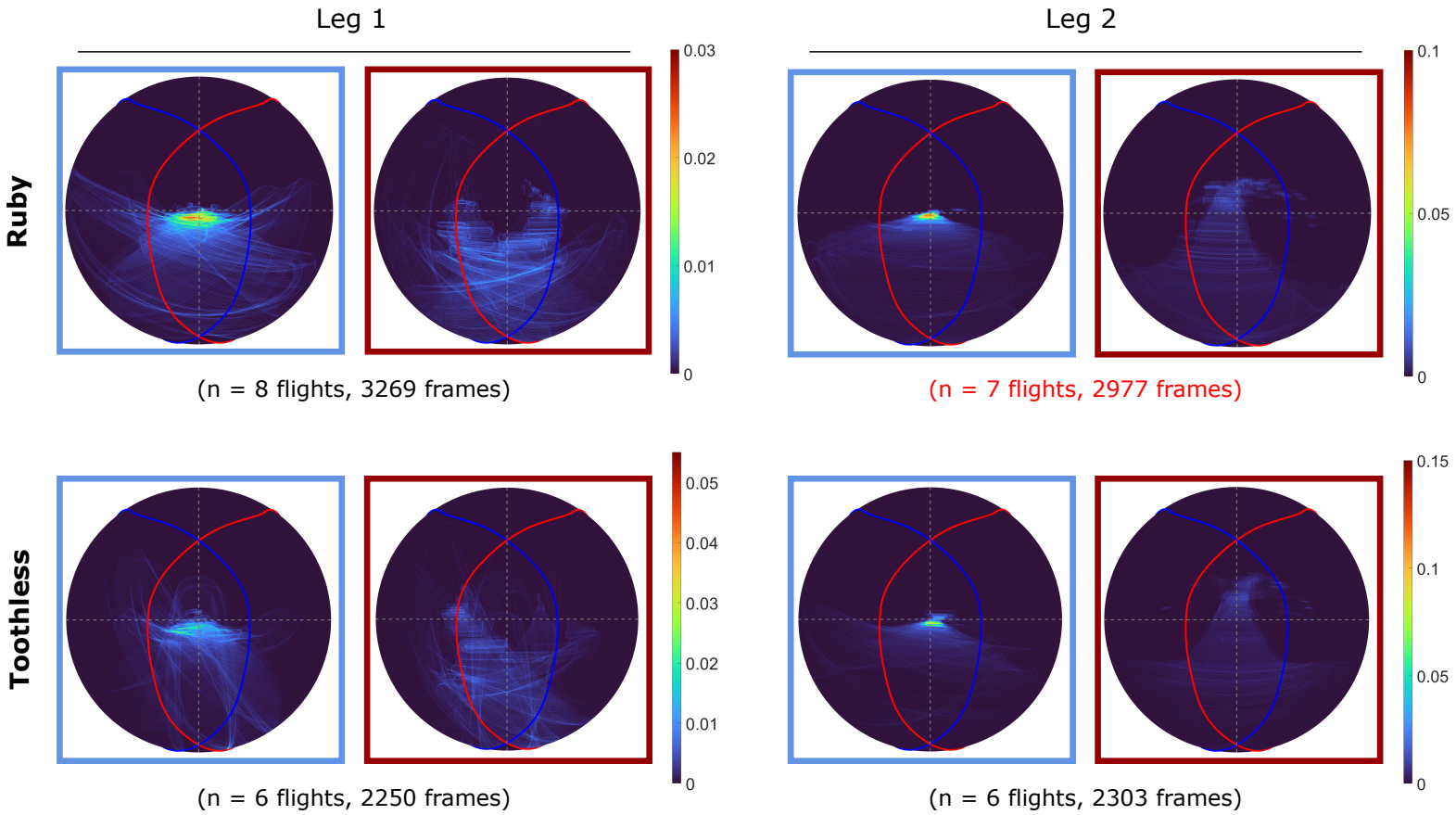


Figure C.10: Spherical histograms for the landing perch in the obstacle avoidance flights - Ruby and Toothless. The figure shows the spherical histograms for the landing perch in the obstacle avoidance flights of Ruby and Toothless, grouped by bird, leg of the trial and coordinate system. The visual field spheres are represented using an orthographic projection on the xz plane of the visual coordinate system (blue boxes) or the trajectory coordinate system (red boxes); only the frontal hemispheres of the visual field are represented. The retinal margins from the right (red) and left (blue) eyes are shown for reference. The range of frames considered includes all frames from takeoff to landing, excluding the frames in which the head pose of the bird was interpolated (see number of flights n and total number of frames in subfigures' captions). Note the colourbars across birds and legs of the trial are not in the same range: the maximum values were selected to highlight the points of highest h metric per pixel in the visual coordinate system representation. The figure caption in red highlights that some flights were excluded from the set because the bird flew over the obstacles rather than around them.

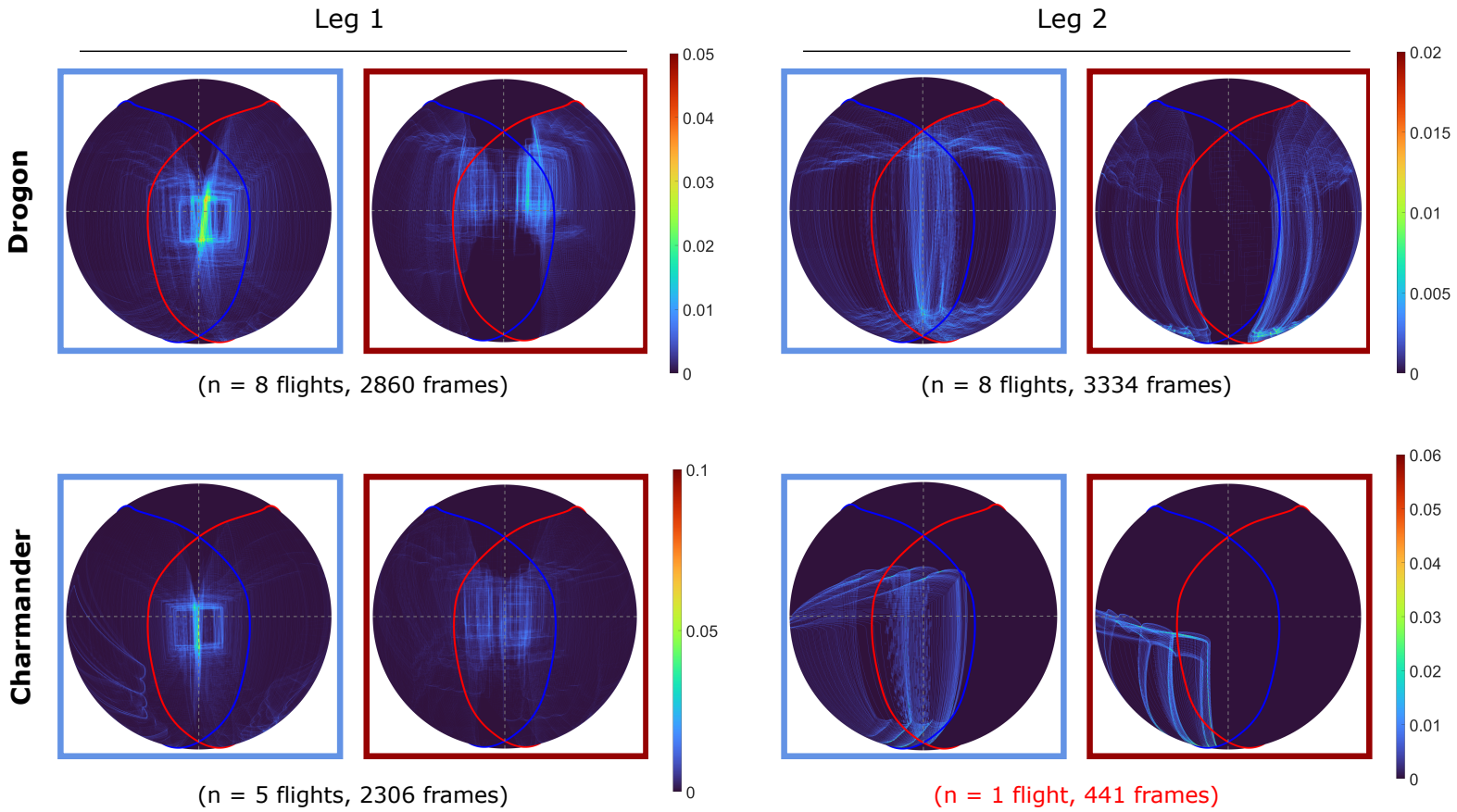


Figure C.11: Spherical histograms for the obstacles in the obstacle avoidance flights - Drogon and Charmander. The figure shows the spherical histograms for the obstacles in the obstacle avoidance flights of Drogon and Charmander, grouped by bird, leg of the trial and coordinate system. The visual field spheres are represented using an orthographic projection on the xz plane of the visual coordinate system (blue boxes) or the trajectory coordinate system (red boxes); only the frontal hemispheres of the visual field are represented. The retinal margins from the right (red) and left (blue) eyes are shown for reference. The range of frames considered includes all frames from takeoff to landing, excluding the frames in which the head pose of the bird was interpolated (see number of flights n and total number of frames in subfigures' captions). Note the colourbars across birds and legs of the trial are not in the same range: the maximum values were selected to highlight the points of highest h metric per pixel in the visual coordinate system representation. The figure caption in red highlights that some flights were excluded from the set because the bird flew over the obstacles rather than around them.

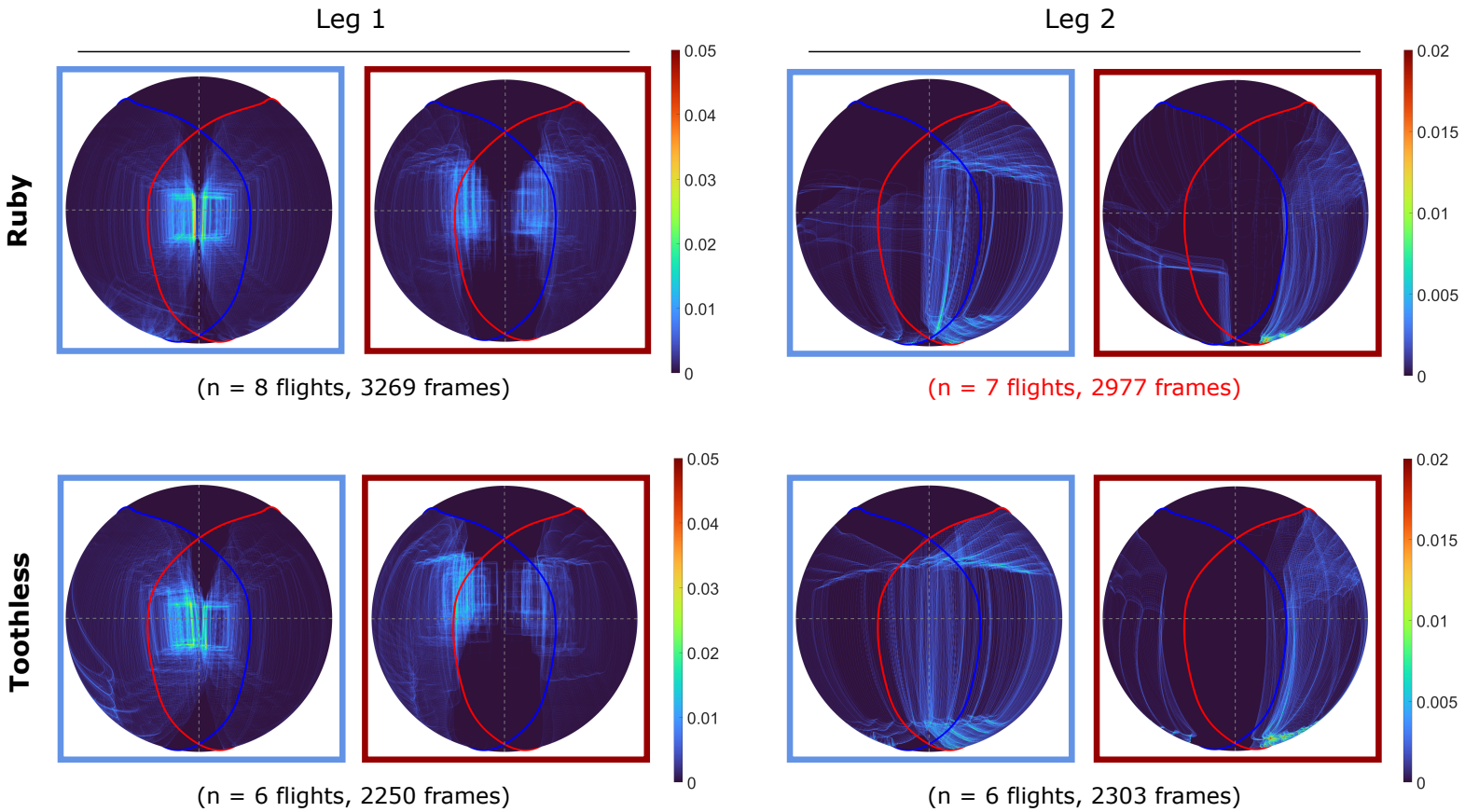


Figure C.12: Spherical histograms for the obstacles in the obstacle avoidance flights - Ruby and Toothless. The figure shows the spherical histograms for the obstacles in the obstacle avoidance flights of Ruby and Toothless, grouped by bird, leg of the trial and coordinate system. The visual field spheres are represented using an orthographic projection on the xz plane of the visual coordinate system (blue boxes) or the trajectory coordinate system (red boxes); only the frontal hemispheres of the visual field are represented. The retinal margins from the right (red) and left (blue) eyes are shown for reference. The range of frames considered includes all frames from takeoff to landing, excluding the frames in which the head pose of the bird was interpolated (see number of flights n and total number of frames in subfigures' captions). Note the colourbars across birds and legs of the trial are not in the same range: the maximum values were selected to highlight the points of highest h metric per pixel in the visual coordinate system representation. The figure caption in red highlights that some flights were excluded from the set because the bird flew over the obstacles rather than around them.

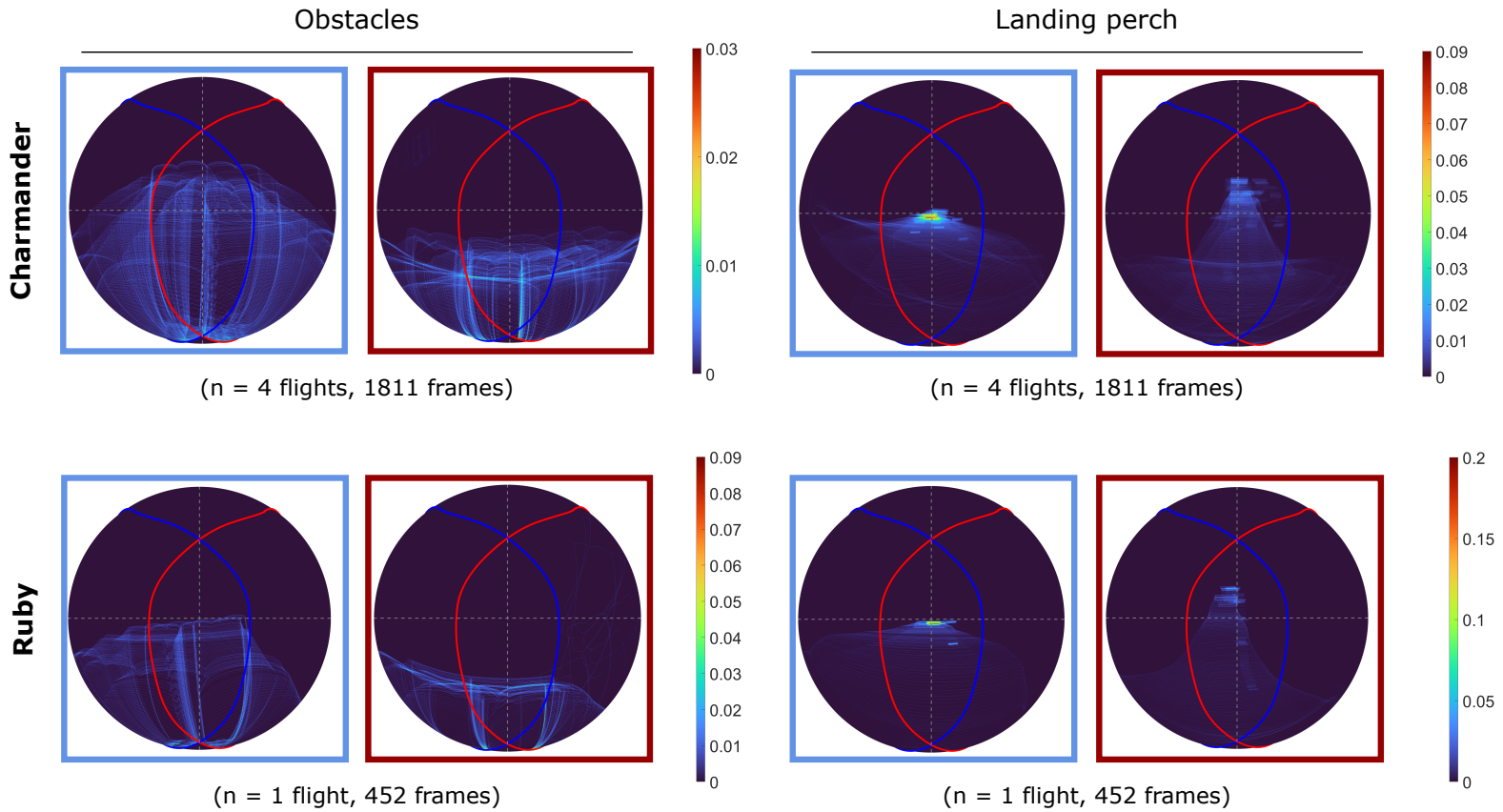


Figure C.13: Spherical histograms for the obstacles and the landing perch, for the flights over the obstacles. The figure shows the spherical histograms for the obstacles and the landing perch in the flights of Charmander and Ruby in which they flew over the obstacles, rather than around them. The data is grouped by bird, leg of the trial, coordinate system and object of interest (landing perch or obstacle). The visual field spheres are represented using an orthographic projection on the xz plane of the visual coordinate system (blue boxes) or the trajectory coordinate system (red boxes); only the frontal hemispheres of the visual field are represented. The retinal margins from the right (red) and left (blue) eyes are shown for reference. The range of frames considered includes all frames from takeoff to landing, excluding the frames in which the head pose of the bird was interpolated (see number of flights n and total number of frames in subfigures' captions). Note the colourbars across birds and legs of the trial are not in the same range: the maximum values were selected to highlight the points of highest h metric per pixel in the visual coordinate system representation.

C.6 Latitudinal and longitudinal extension of the landing perch and the obstacles

In this section, we include the curves for the extension in latitude of the landing perch along the trajectory, for the perching flights (Figures C.14 and C.15) and the obstacle avoidance flights (Figures C.16 and C.17). Additionally, in Figures C.19 and C.20 we represent the curves for the longitudinal extension of the obstacles, without mirroring the data relative to the y-axis of the plot.

C.7 Alignment with landing perch centre or bisector direction

We inspected the difference between assuming the bird is aligned with the landing perch's centre, and assuming it is aligned with the bisector of the angle subtended by the landing perch in longitudinal direction. To do so, we considered the triangle defined by the position of the bird's head and the landing perch's edges, as shown in Figure C.21. Specifically, we computed the difference in angle β between the vector along the median segment $\overrightarrow{OM}$, and the bisector vector $\overrightarrow{OB}$.

We first express the vectors from the bird's position O to the landing perch edges as:

$$\begin{aligned} \overrightarrow{OE_1} &= r_1(\cos\alpha, \sin\alpha), \\ \overrightarrow{OE_2} &= r_2(\cos\alpha, -\sin\alpha), \end{aligned} \tag{C.1}$$

where r_1 and r_2 are the distances from the bird to each of the edges. The vector from O along the median segment is by definition:

$$\overrightarrow{OM} = \frac{\overrightarrow{OE_1} + \overrightarrow{OE_2}}{2}. \tag{C.2}$$

From these we compute β as:

$$\begin{aligned} \tan\beta = \frac{OM_y}{OM_x} &= \frac{r_1\sin\alpha - r_2\sin\alpha}{r_1\cos\alpha + r_2\cos\alpha} \\ &= \left(\frac{r_1 - r_2}{r_1 + r_2}\right)\tan\alpha \end{aligned} \tag{C.3}$$

where OM_x and OM_y are the x and y components of the $\overrightarrow{OM}$ vector respectively, and α is half of the angle subtended by the perch.

We computed β across all 102 flights (including the flights in which the birds flew over the obstacles) using the motion capture data of the birds' head trajectories and the landing perch's edges. The results are shown in Figure C.22, where the contour of the data represents the maximum values observed at each value of $\hat{y}_{perch}$. The maximum value of β observed across all flights is under 1° from take-off until the bird is at $\hat{y}_{perch} = 0.14$ (1.26 m). Note that in the analysis of the deviation of the bird's gaze strategy with respect to an ideal symmetric expansion of the landing perch, we don't consider samples closer than $\hat{y} = 0.1$. The angle β only reaches 5° when the bird is closer than $\hat{y}_{perch} = 0.06$ to the perch (approximately 0.54 m).

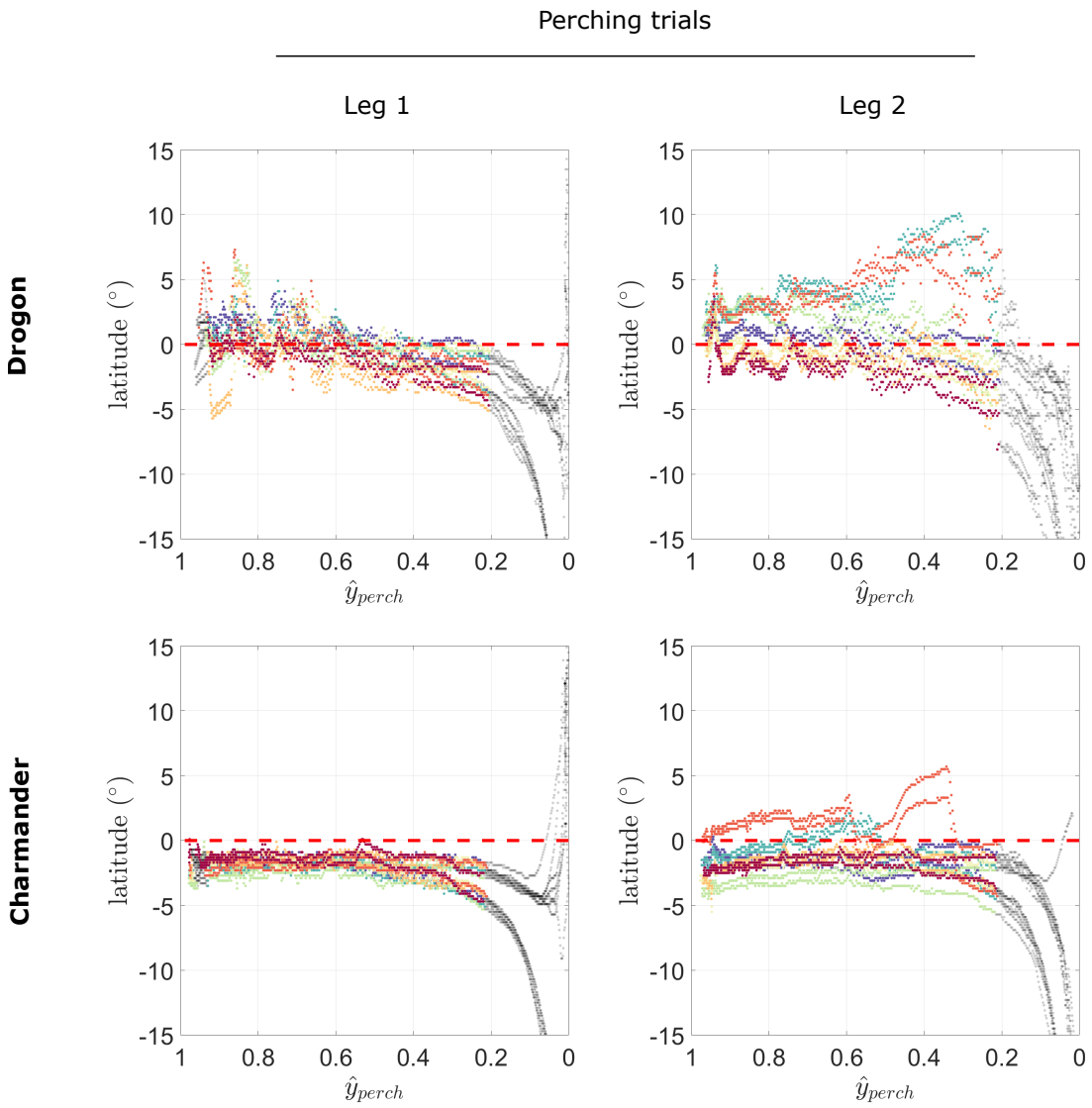


Figure C.14: Latitudinal extension of the landing perch in the bird's visual field along the trajectory - Drogon and Charmander perching flights. The latitudinal extension of the landing perch in the bird's visual field is represented against $\hat{y}_{perch}$, the normalised distance to the perch in y direction. The flights are grouped by bird and leg of the trial. We focus on a specific set of samples when analysing the deviation of a bird's gaze strategy from an ideal symmetric expansion of the landing perch in the visual field; this is the range of data points highlighted with coloured markers. The same colour is used for data points in the same flight and the samples outside of this range are shown in grey.

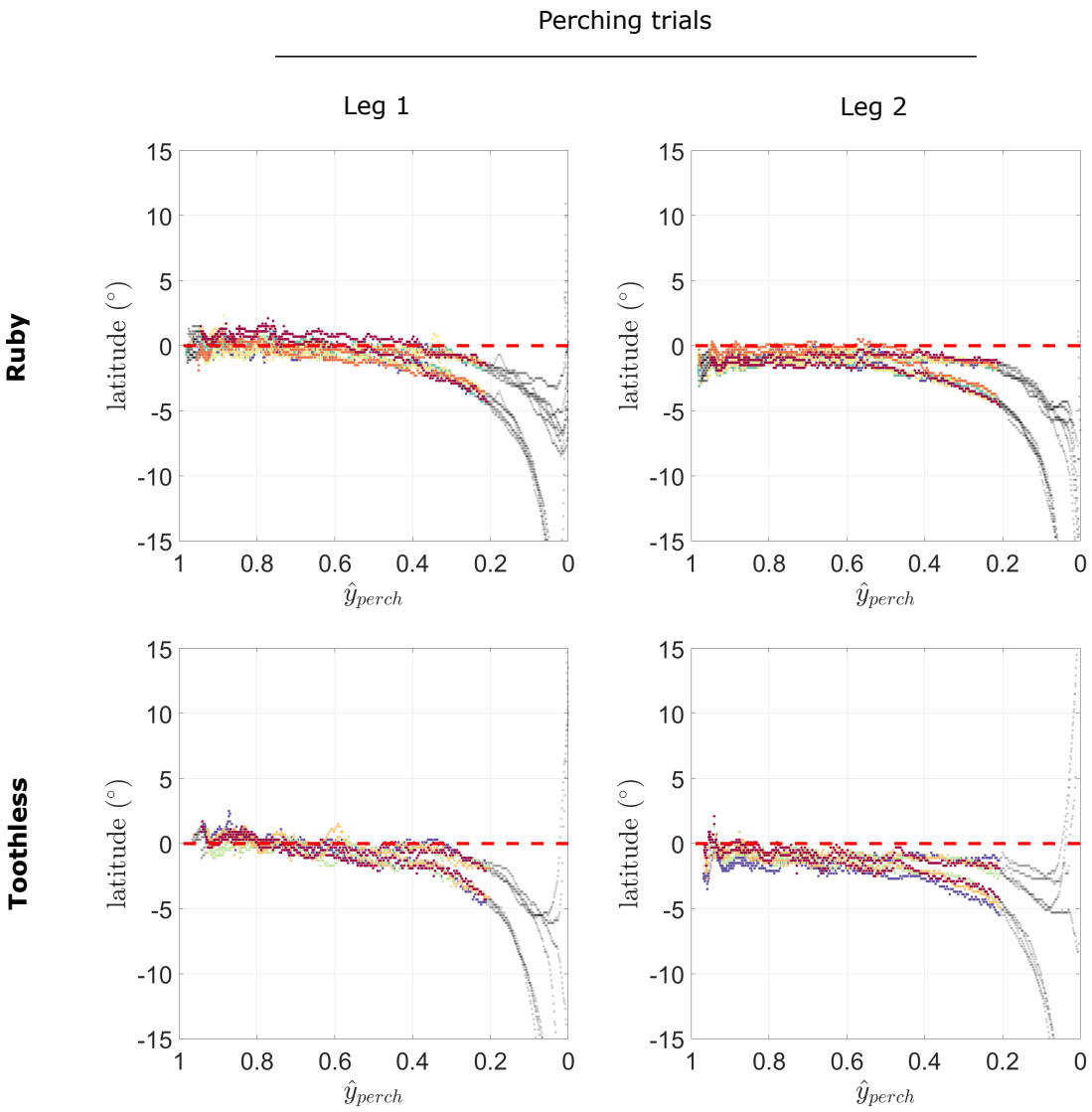


Figure C.15: Latitudinal extension of the landing perch in the bird's visual field along the trajectory - Ruby and Toothless perching flights. The latitudinal extension of the landing perch in the bird's visual field is represented against $\hat{y}_{perch}$, the normalised distance to the perch in y direction. The flights are grouped by bird and leg of the trial. We focus on a specific set of samples when analysing the deviation of a bird's gaze strategy from an ideal symmetric expansion of the landing perch in the visual field; this is the range of data points highlighted with coloured markers. The same colour is used for data points in the same flight and the samples outside of this range are shown in grey.

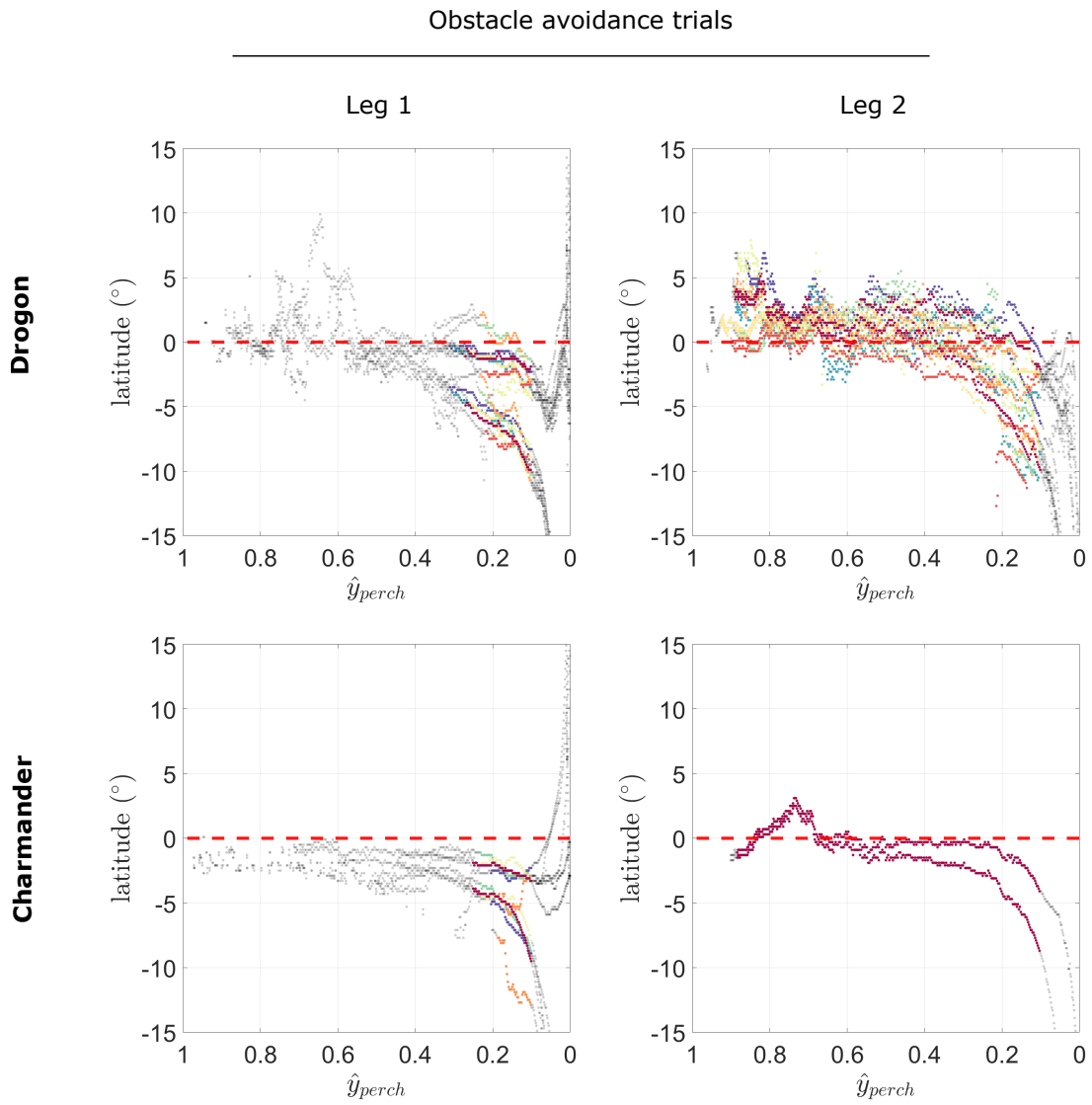


Figure C.16: Latitudinal extension of the landing perch in the bird's visual field along the trajectory - Drogon and Charmander obstacle avoidance flights. The latitudinal extension of the landing perch in the bird's visual field is represented against $\hat{y}_{perch}$, the normalised distance to the perch in y direction. The flights are grouped by bird and leg of the trial. We focus on a specific set of samples when analysing the deviation of a bird's gaze strategy from an ideal symmetric expansion of the landing perch in the visual field; this is the range of data points highlighted with coloured markers. The same colour is used for data points in the same flight and the samples outside of this range are shown in grey.

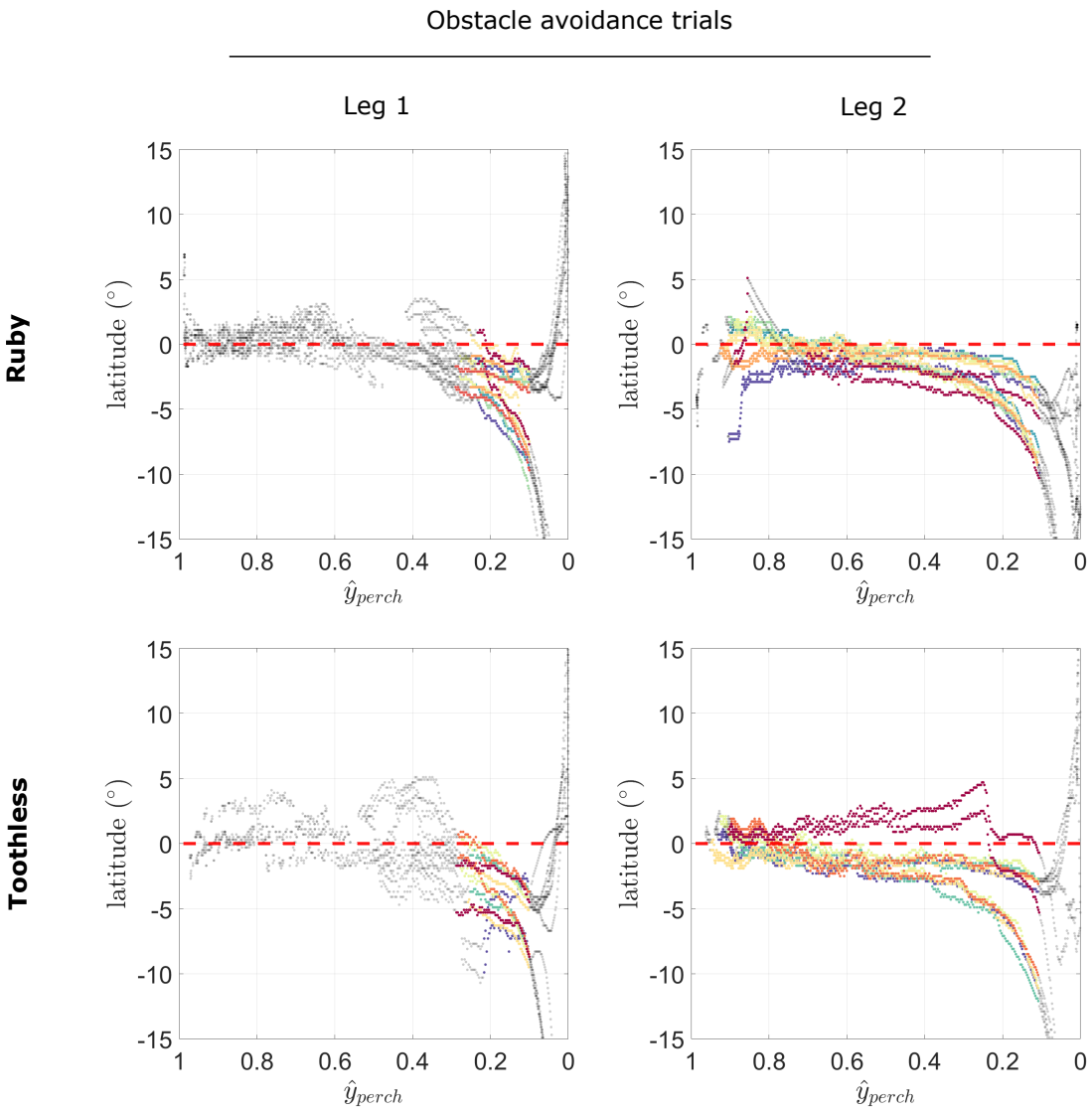


Figure C.17: Latitudinal extension of the landing perch in the bird's visual field along the trajectory - Ruby and Toothless obstacle avoidance flights. The latitudinal extension of the landing perch in the bird's visual field is represented against $\hat{y}_{perch}$, the normalised distance to the perch in y direction. The flights are grouped by bird and leg of the trial. We focus on a specific set of samples when analysing the deviation of a bird's gaze strategy from an ideal symmetric expansion of the landing perch in the visual field; this is the range of data points highlighted with coloured markers. The same colour is used for data points in the same flight and the samples outside of this range are shown in grey.

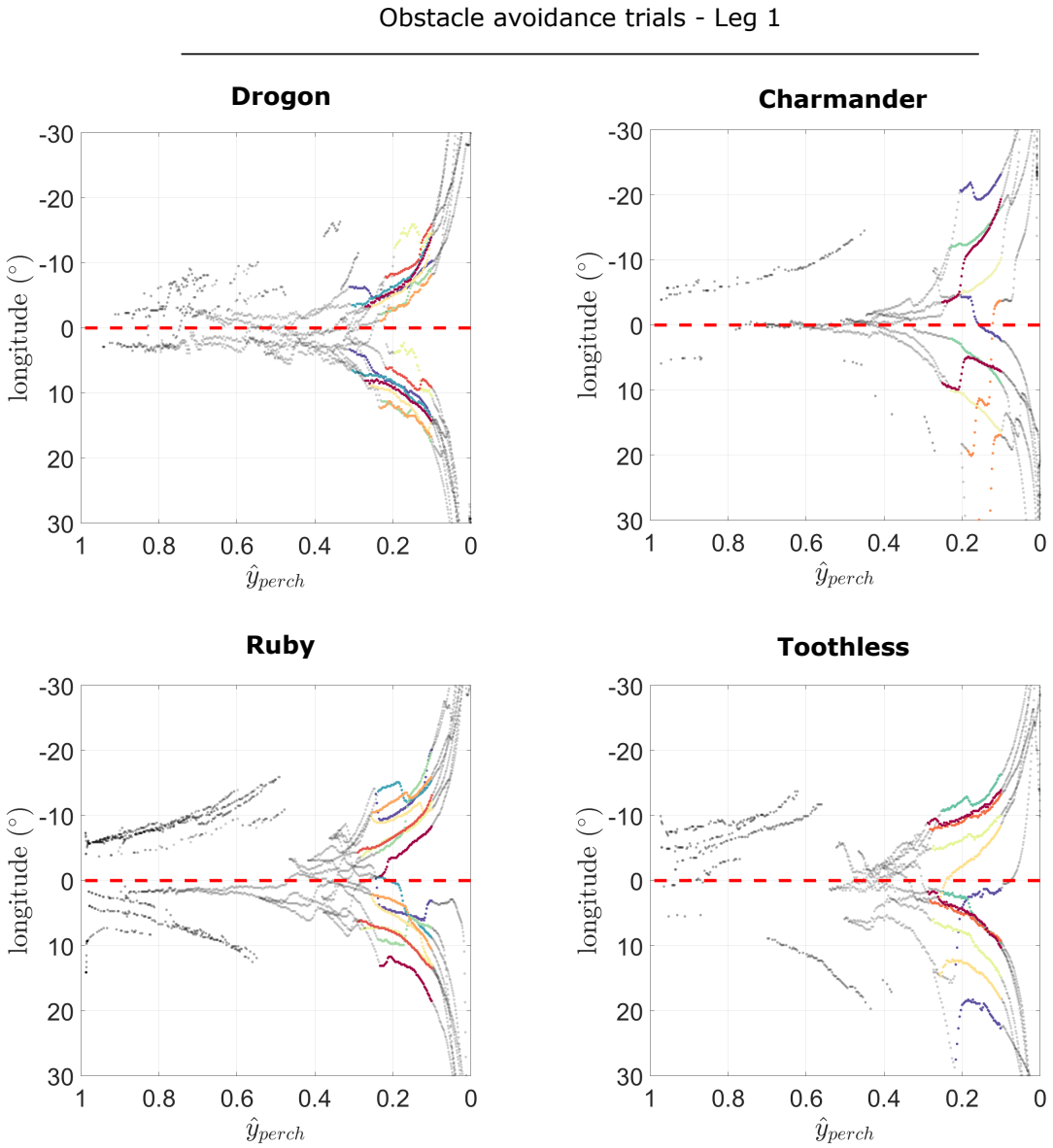


Figure C.18: Longitudinal extension of the landing perch in the bird's visual field along the trajectory - Obstacle avoidance leg 1 flights. The longitudinal extension of the landing perch in a bird's visual field is represented against $\hat{y}_{perch}$, the normalised distance to the perch in y direction. The flights are grouped by bird, type of trial and leg of the trial they are executed at. We focus on a specific set of samples when analysing the deviation of a bird's gaze strategy from an ideal symmetric expansion of the landing perch in the visual field; this is the range of data points highlighted with coloured markers. The same colour is used for data points in the same flight and the samples outside of this range are shown in grey. The perch is fully visible on the first sample show in colour; earlier samples either show a partially visible perch, or a thin stretch of the perch that is slightly visible through the gaps between the obstacles.

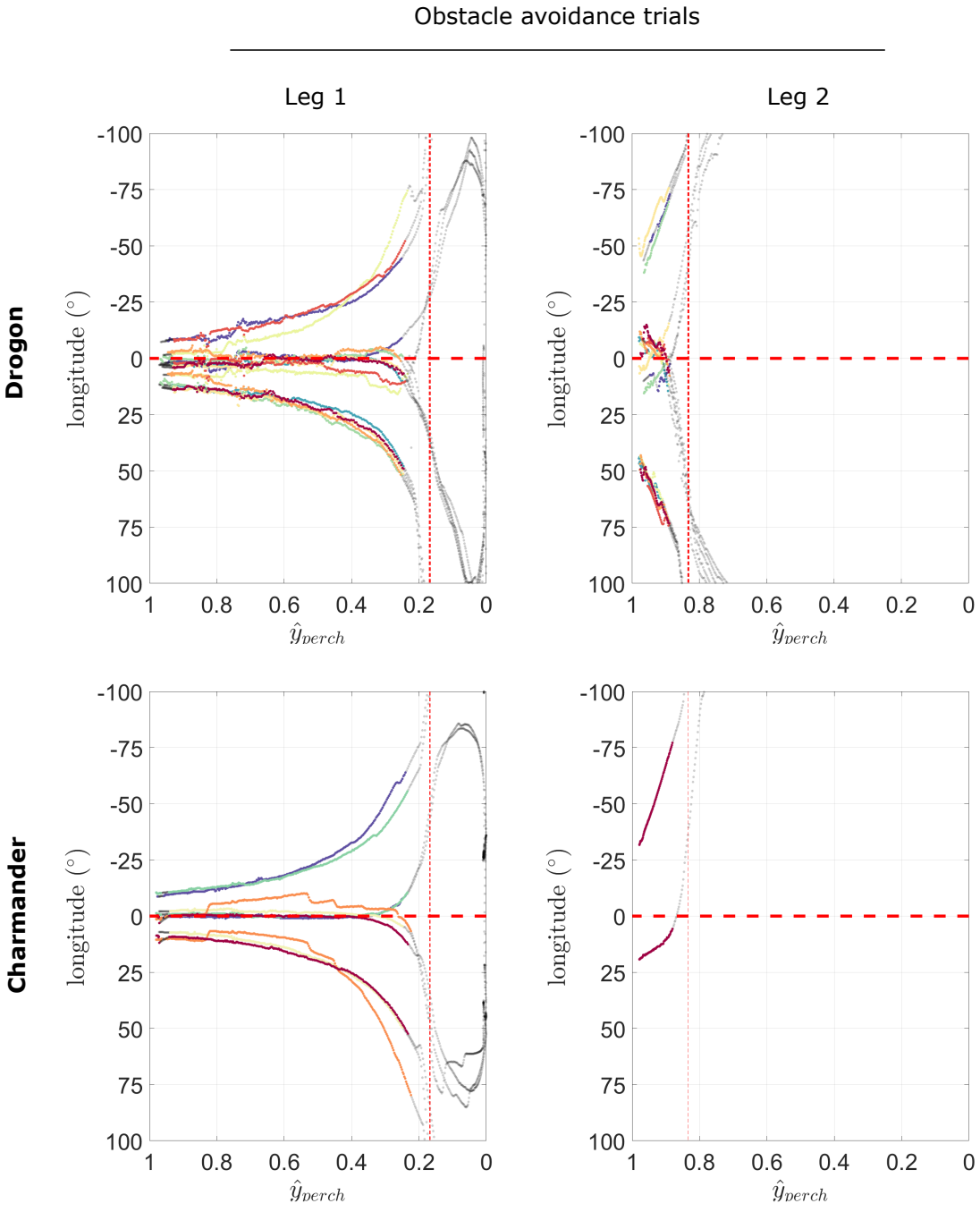


Figure C.19: Longitudinal extension of the obstacles in the bird's visual field along the trajectory (not mirrored) - Drogon and Charmander flights. The longitudinal extension of the obstacles in a bird's visual field is represented against $\hat{y}_{perch}$, the normalised distance to the perch along the y-axis of the motion capture coordinate system. The flights are grouped by bird and leg of the trial. We focus on a specific set of samples when analysing the deviation of the bird's gaze strategy from an ideal alignment with the inside edge of the obstacles; this is the range of data points highlighted with coloured markers. The same colour is used for data points in the same flight, and the samples outside of this range are shown in grey. For each trial, the point at which the bird's head trajectory passes the mean y-coordinate of the obstacles is marked with a $\hat{y} = constant$ line (red dashed lines).

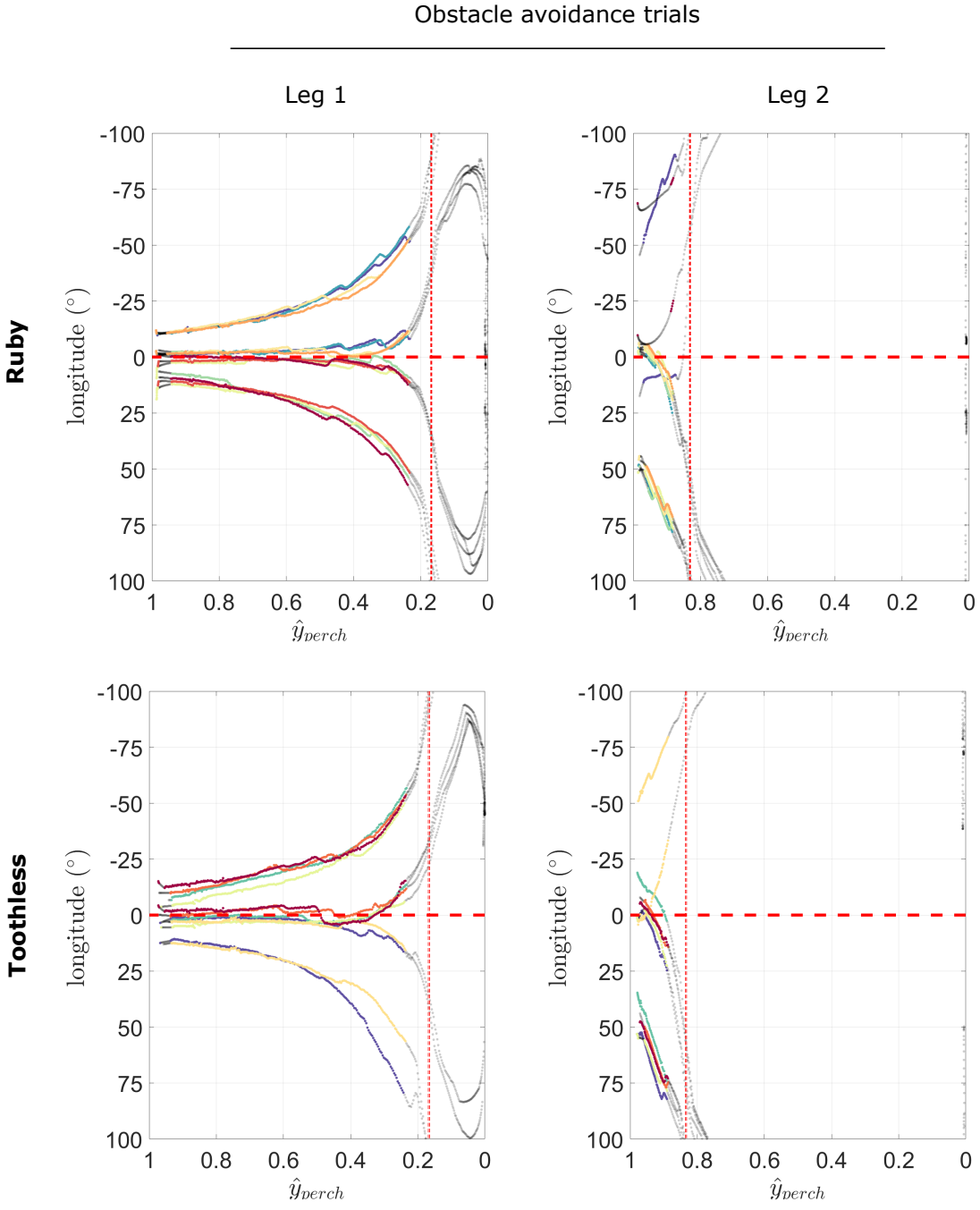


Figure C.20: Longitudinal extension of the obstacles in the bird's visual field along the trajectory (not mirrored) - Ruby and Toothless flights. The longitudinal extension of the obstacles in a bird's visual field is represented against $\hat{y}_{perch}$, the normalised distance to the perch along the y-axis of the motion capture coordinate system. The flights are grouped by bird and leg of the trial. We focus on a specific set of samples when analysing the deviation of the bird's gaze strategy from an ideal alignment with the inside edge of the obstacles; this is the range of data points highlighted with coloured markers. The same colour is used for data points in the same flight, and the samples outside of this range are shown in grey. For each trial, the point at which the bird's head trajectory passes the mean y-coordinate of the obstacles is marked with a $\hat{y} = constant$ line (red dashed lines).

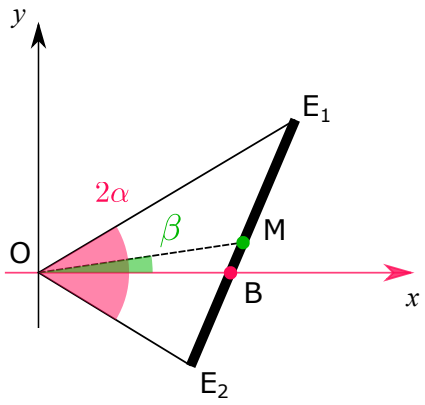


Figure C.21: Median and bisector of triangle defined by the bird's head and the landing perch's edges. The figure shows the geometry of the triangle defined by the bird's head (O) and the landing perch's edges (E_1 and E_2). The median of the triangle from the O vertex is the $\overrightarrow{OM}$ vector, where M is the midpoint between the perch's edges E_1 and E_2. The bisector of the angle subtended by the perch 2α (pink angle) is the vector $\overrightarrow{OB}$ (pink vector), which by definition coincides with the x-axis of the plot. The angle β (highlighted in green) is the angular difference between the vector along the median segment $\overrightarrow{OM}$ and the bisector vector $\overrightarrow{OB}$.

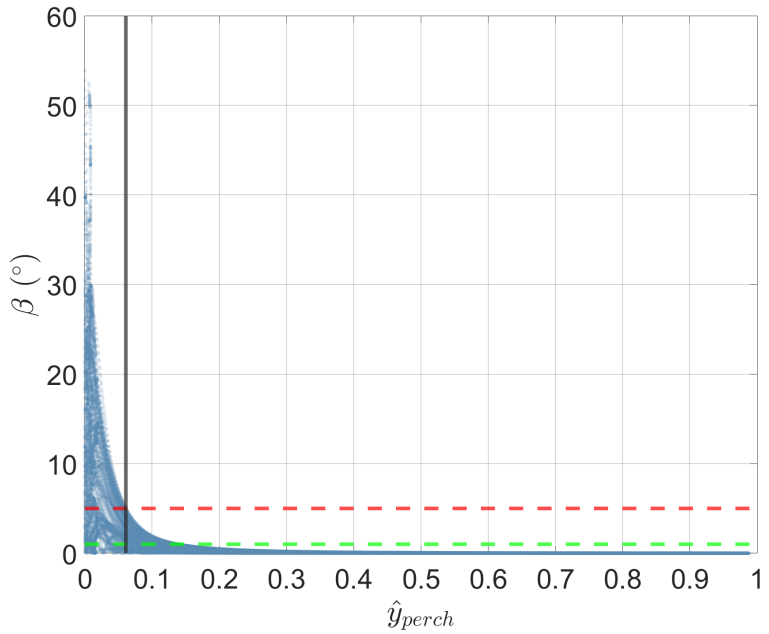


Figure C.22: Values of β angle per frame across all flights against the normalised distance to the landing perch $\hat{y}_{perch}$. The figure shows the values for the computed β angle per frame, across all flights, against the normalised distance to the landing perch $\hat{y}_{perch}$ (n=102 flights, including the flights in which the birds flew over the obstacles). Semi-transparent markers are used to reflect the density of the data. Reference lines at $\beta = 5°$ (red dashed line) and $\beta = 1°$ (green dashed line) are shown. When $\hat{y}_{perch} = 0.06$ (vertical black line), the maximum value of β observed is 5°.

C.8 Deviation from an ideal symmetric expansion of the landing perch

In the main text, we focused on leg 2 of the obstacle avoidance trials when analysing the gaze strategy around the landing perch. We made this choice because the landing perch in that leg was visible for longer (see Figure 4.9). For reference, we include here the results for the flights executed in the other leg of the obstacle avoidance trials (leg 1); Figure C.23 shows the results from the root-mean-square (RMSE) analysis.

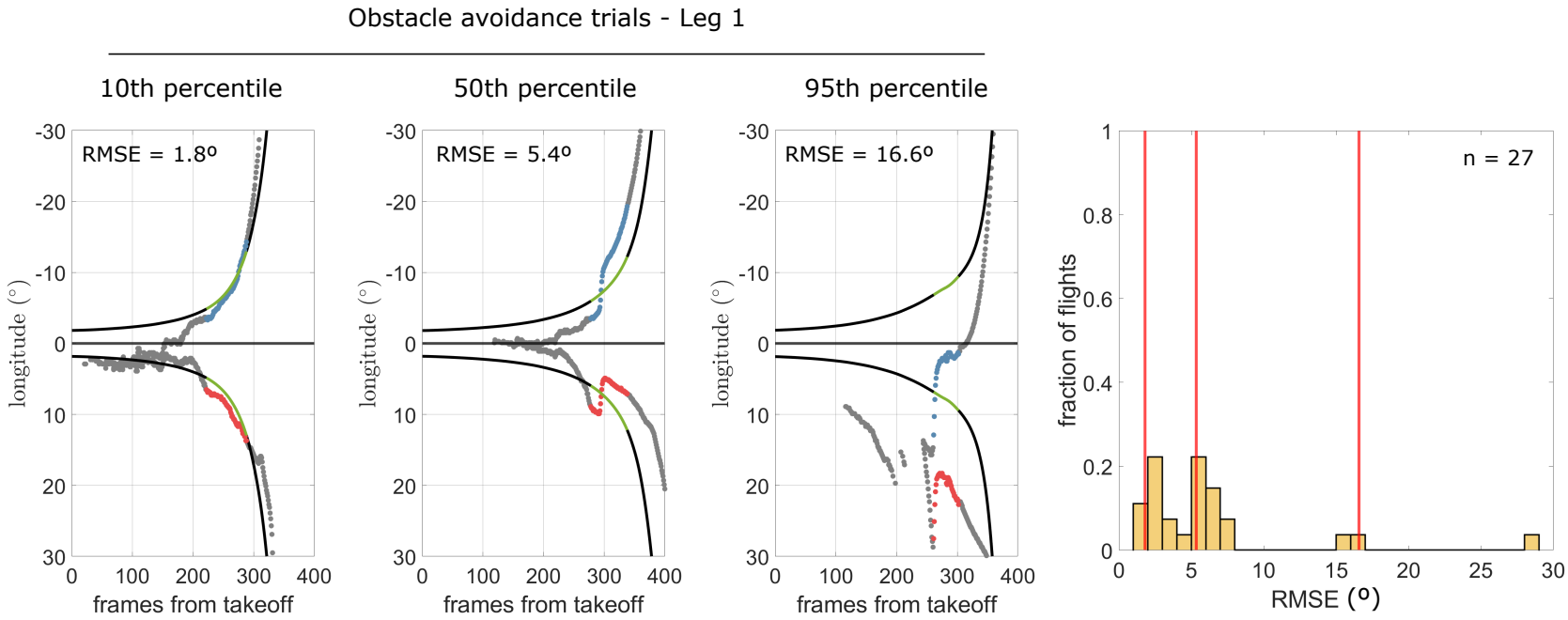


Figure C.23: Deviation from an ideal symmetric expansion of the landing perch in longitudinal direction in the bird's visual field - Obstacle avoidance leg 1 flights. We inspected for each obstacle avoidance flight executed in leg 1 of the trials, how much the longitudinal extension of the landing perch in the visual field deviated from an ideal symmetric expansion. We characterised the deviation with the root-mean-square error; the distribution of its value across the $n = 27$ flights analysed is shown (right panel). The three panels from the left show the longitudinal extension of the perch for flights with a root-mean-square error corresponding to the 10th, 50th and 95th percentile. The right and left edge of the landing perch are shown (red and blue respectively), as well as the edges in an ideal symmetric expansion case (green). Only the samples within the subset of frames of interest are shown in colour. The samples outside this range are shown in grey for the data derived from the semantic maps of the landing perch, and black for the ideal symmetric expansion. The location of the selected percentiles on the histogram is marked with red vertical lines.

C.9 Mean position of the landing perch centre in the visual field

We computed the mean position of the landing perch in the bird's visual field in latitude and longitude, for each flight. If in a given flight, the centre of the landing perch stays aligned with the estimated gaze direction for the majority of the frames, the mean position of the landing perch in the visual field for that flight will be close to 0° latitude and 0° longitude. If the centre of the landing perch is not aligned with the gaze direction for a large portion of the flight (for example, because another feature, such as the edge of the perch, is fixated instead), we expect the mean position of the landing perch to deviate from the estimated gaze direction.

To assess the landing perch's mean position, we computed the mean position of the midpoint of its top edge (the edge of maximum latitude in the landing perch's contour). The midpoint is computed in angle, as the bisector of the longitudinal angle subtended by the perch; for short we refer to it as simply the midpoint. We included all frames from take-off up to a normalised distance to the perch of $\hat{y}_{perch} = 0.1$, excluding the frames used in the estimation of $\vec{v}_{gaze}$ and the frames in which the head pose was interpolated (see colour markers in Figures 4.10, 4.11 and C.18). In the obstacle avoidance flights, we only considered the frames in which the landing perch is fully visible (that is, not partially occluded by the obstacles).

Figure C.24 shows the mean position of the landing perch centre for the perching flights, in the latitude and longitude plane. The value for each flight is coloured by leg of the trial, direction of the lateral offset between the takeoff and the landing perch (right, straight or left), and bird. The results suggest the landing perch midpoint is reasonably aligned with the estimated gaze direction throughout most of the perching flights.

Figure C.25 shows the mean position of the landing perch centre for the obstacle avoidance flights. The value for each flight is coloured by leg of the trial, side around the obstacles and bird. The results show larger variation in longitude than in the perching flights, and the flights executed in leg 1 of the trials seem biased to lower mean latitude values. This may be due to the landing perch being visible for very different parts of the trajectory in each leg of the trial. The flights in which the birds turn around the obstacles' right or left side also seem biased to positive and negative mean longitude angles respectively. Note that the side the birds turn around the obstacles also determines the side the landing perch appears from.

C.10 Mean position of the obstacle's edge in the visual field

We computed the mean position of the edge of the obstacles in the birds' visual field, for each flight. We considered the range of frames from takeoff up to 25 frames (0.125 s) before the obstacles are surpassed. The obstacles are surpassed when the y-coordinate of the bird's head is equal to the mean y-coordinate of the obstacles. We selected 25 frames prior to this point as an estimate of the time it takes the bird to fly a distance of one body length, at an approximate

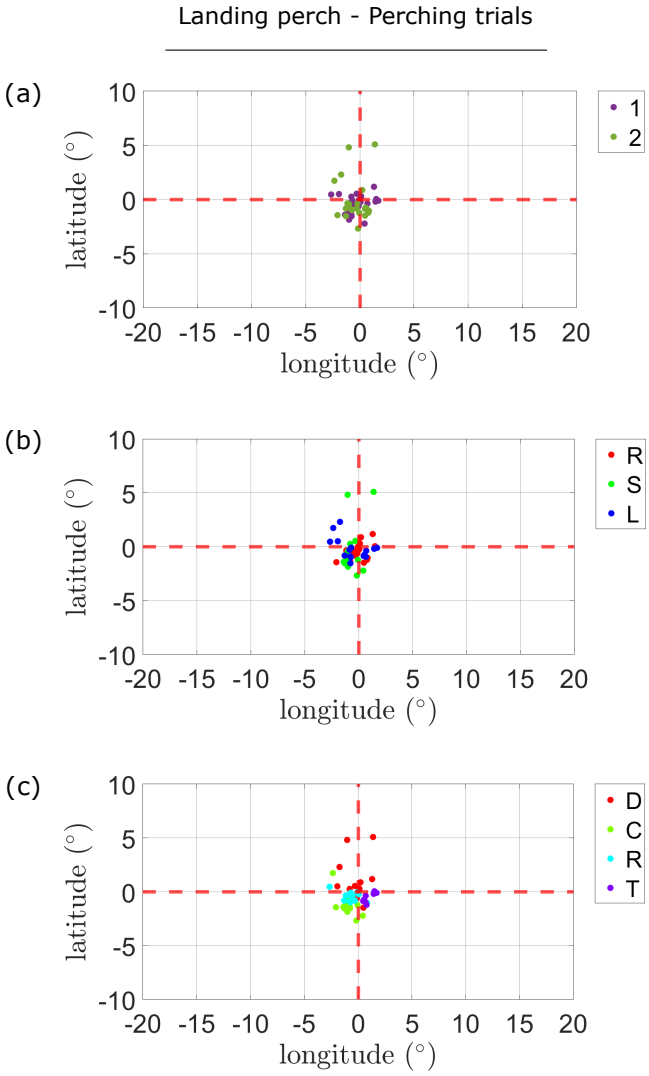


Figure C.24: Mean position of the landing perch centre in the bird's visual field per flight - Perching flights. The figure shows the mean position of the landing perch's centre per flight, in the latitude-longitude plane representative of the bird's visual field. We computed the landing perch's centre as the midpoint of the top edge. The values per flight are coloured by (a) leg of the trial; (b) direction of the lateral offset between takeoff and landing perch (Right, Straight or Left); and (c) bird (Drogon, Charmander, Ruby, Toothless). The 0° latitude and 0° longitude lines are highlighted for reference (red dashed lines).

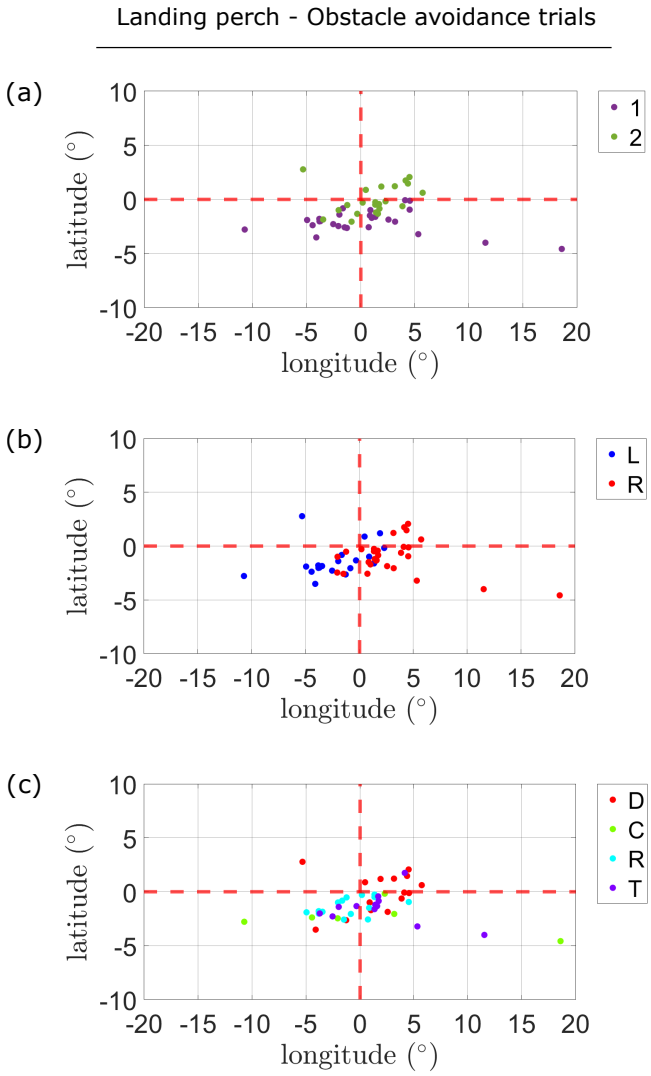


Figure C.25: Mean position of the landing perch centre in the bird's visual field per flight - Obstacle avoidance flights. The figure shows the mean position of the landing perch's centre per flight, in the latitude-longitude plane representative of the bird's visual field. We computed the landing perch's centre as the midpoint of the top edge. The values per flight are coloured by (a) leg of the trial; (b) side around the obstacles the bird flies (Right or Left); and (c) bird (Drogon, Charmander, Ruby, Toothless). The 0° latitude and 0° longitude lines are highlighted for reference (red dashed lines).

speed of 5 m/s. The range of frames considered per trial is shown with colour markers in Figures C.19 and C.20. For this range of frames, we computed the mean longitudinal position of the inside edge of the obstacles (relative to the bird's turn), and the mean mid-latitude of the obstacles. Note that we don't expect the midpoint of the obstacles in latitude to be aligned with the estimated gaze direction throughout the flight; we use its mean value to assess how stabilised the obstacles are in latitude in the range of frames considered.

Figure C.26 shows the mean position of the obstacles' edge midpoint in latitude and longitude. The value for each flight is coloured by leg of the trial, side around the obstacles and bird. In the flights executed in leg 1 of the trials, the mean position in longitude of the obstacles' inside edge is within $\pm 5°$ from the estimated gaze direction. In leg 2 of the trials there is larger variability. From the figure we don't see evident biases based on the side of the obstacles the bird turns around, but this should be checked with statistical tests.

C.11 Location of the minima of optic flow in the bird's visual field

For one of Drogon's obstacle avoidance trials, we investigated the most frequent location of the optic flow minima in the visual field. To do so, we computed spherical histograms representing the frequency with which an optic flow minimum appeared at each point in the visual field sphere. We defined an optic flow minimum per frame as a pixel whose backprojected 3D point displaced less than $10°/s$ from that frame to the next. We computed spherical histograms in the visual and the trajectory coordinate system; the results are shown in Figure C.27. We use the same metric h as in the main text, to represent the fraction of frames normalised by the pixel's size. We excluded the frames in which the bird's head pose was interpolated.

In the visual coordinate system, the minima of optic flow appear more frequently in the binocular area and close to the estimated gaze direction for both legs of the trial. In contrast, in the trajectory coordinate system, the minima of optic flow occur most frequently in areas outside the binocular region and not centred in the visual field. The optic flow data per frame in the visual coordinate system for the flight executed in leg 1 in Figure C.27 is included as a supplementary video `ESM_16.avi`.

References

Brighton, C. H. and Taylor, G. K. (2019). "Hawks steer attacks using a guidance system tuned for close pursuit of erratically manoeuvring targets". In: *Nat. Commun.* 10.1, pp. 1–28. DOI: 10.1038/s41467-019-10454-z.

Potier, S., Bonadonna, F., Kelber, A., Martin, G. R., Isard, P.-F., Dulaurent, T., and Duriez, O. (2016). "Visual abilities in two raptors with different ecology". In: *J. Exp. Biol.* 219.17, pp. 2639–2649. DOI: 10.1242/jeb.142083.

Ros, I. G. and Biewener, A. A. (2017). "Pigeons (C. livia) follow their head during turning flight: Head stabilization underlies the visual control of flight". In: *Front. Neurosci.* 11.DEC, pp. 1–12. DOI: 10.3389/fnins.2017.00655.

Shoemake, K. (1985). “Animating rotation with quaternion curves”. In: *Proc. 12th Annu. Conf. Comput. Graph. Interact. Tech.* Vol. 19. 3, pp. 245–254. DOI: 10.1145/325165.325242.

Warrick, D. R., Bundle, M. W., and Dial, K. P. (2002). “Bird maneuvering flight: blurred bodies, clear heads”. In: *Integr. Comp. Biol.* 42.1, pp. 141–148. DOI: 10.1093/icb/42.1.141.

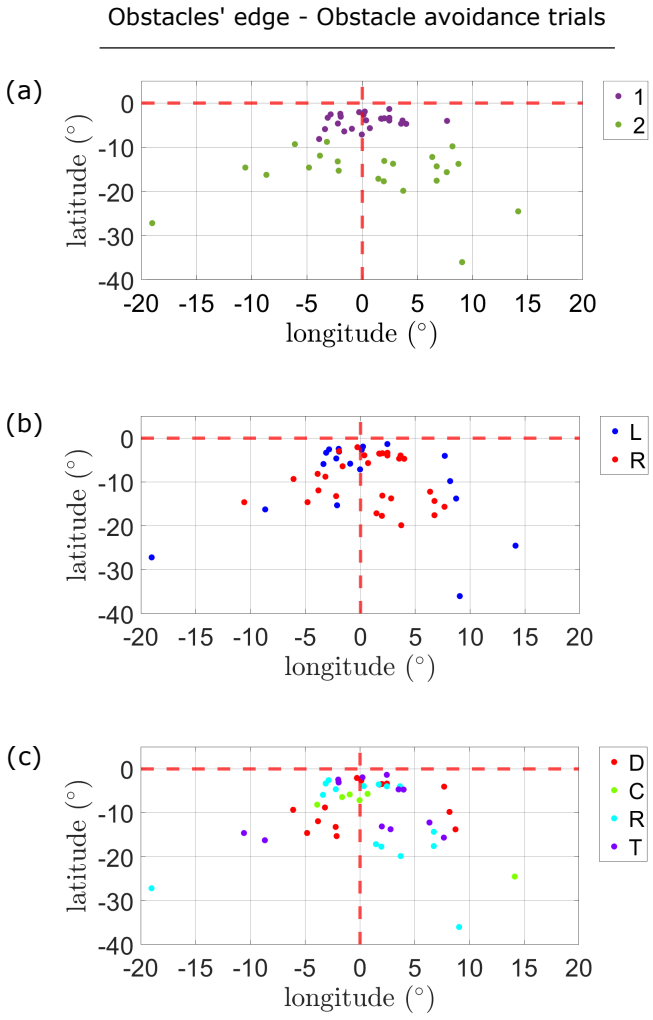


Figure C.26: Mean position of the obstacles' inside edge midpoint in the bird's visual field per flight. The figure shows the mean position of the midpoint of the obstacles' inside edge per flight, in the latitude-longitude plane representative of the bird's visual field. The values per flight are coloured by (a) leg of the trial; (b) side around the obstacles the bird flies (Right or Left); and (c) bird (Drogon, Charmander, Ruby, Toothless). The 0° latitude and 0° longitude lines are highlighted for reference (red dashed lines).

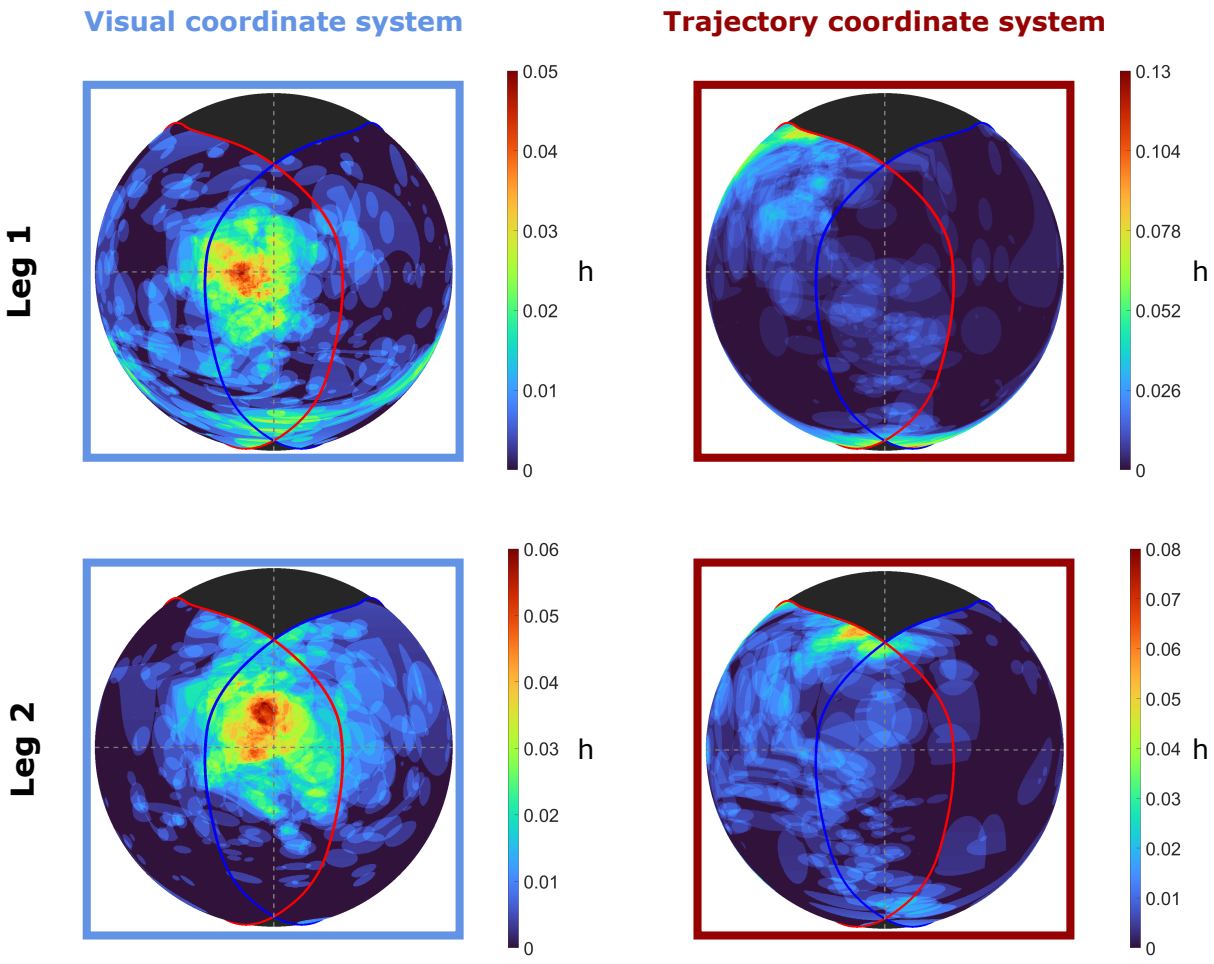


Figure C.27: Location of minima of optic flow in the visual field sphere, for two of Drogon's flights. The figure shows the spherical histograms for the minima of optic flow, for one of Drogon's obstacle avoidance trials. The results are represented in the visual coordinate system (blue box) and the trajectory coordinate system (red box), for both legs of the trial. The visual field spheres are represented using an orthographic projection on the xz plane the visual or the trajectory coordinate system. The retinal margins from the right (red) and left (blue) eyes are shown for reference and the caudal blind area is shaded in black. The maximum value for the colourbars is the rounded maximum value observed per histogram. The planes at 0° latitude and 0° longitude are shown for reference (grey dotted lines); the estimated gaze direction is at their intersection.

D

A dataset of obstacle avoidance flights in Harris' hawks, with static and moving goals

Contents

D.1 Introduction 256
D.2 Materials 257
D.2.1 Curtain 258
D.2.2 Perches 259
D.2.3 Pulleys and linear motor 259
D.2.4 Obstacles 259
D.3 Experimental procedure 260
D.3.1 Static goal trials 263
D.3.2 Moving goal trials 264
D.4 Summary of recorded flights 265
D.5 Labelling procedure 272
D.6 Randomisation based on stations 272
D.7 Randomisation based on grid 278
D.7.1 Static goal trials 278
D.7.2 Moving goal trials 282

D.1 Introduction

This appendix describes a set of experiments carried out between November 2018 and January 2019, in which we used motion capture to record head movements of

$N = 4$ Harris' hawks in obstacle avoidance flights. The purpose of this appendix is to provide a description of the experiments and the dataset collected, in case they are used in future studies.

We considered two types of trial: static and moving goal trials. In the static goal trials, the goal is a perch placed at one end of the room, opposite to the takeoff point. In the moving goal trials, the goal is the lure, a moving target that we dragged along the floor using a system of pulleys and a linear motor. We recorded in total 1217 static goal trials, and 202 moving goal trials (excluding familiarisation or 'warmup' trials). Each trial consists of one flight.

Apart from the perches in the static goal trials, and the lure in the moving goal trials, we added different elements to the environment to guide or alter the trajectory of the birds. We placed a curtain perpendicularly to the longitudinal axis of the room, leaving two gaps of around $1 - 1.1$ m each between the curtain edges and the walls, to force larger turns in the birds' trajectories. We also added obstacles in front and behind the curtain at randomised locations to interfere with the birds' expected trajectories. We considered two types of obstacles, pillars and monoliths. These were made of styrofoam, for ease of handling and to minimise the damage on the birds in case of unlikely impact.

In the static goal trials, we explored a variety of arrangements between the obstacles, and different options for randomisation schemes. In retrospect, we find this part of the dataset would have benefited from more consistency. However, we include details of all the variations used, in case subsets of the data can be used in later studies. The moving goal trials are more consistent in terms of obstacles and randomisation schemes used.

The experiments were planned and carried out by the author of this thesis in collaboration with DPhil student James Kempton, and supervised by Graham Taylor. The labelling of the motion capture data was carried out using a custom algorithm developed by the author, and an algorithm developed by postdoctoral researcher Dr. Marco Klein Hereenbrink.

D.2 Materials

The experiments were carried out in the Flight lab, in a volume of 19.2 m$\times$ 5.3 m$\times$ 3.3 m equipped with 22 motion capture cameras (Vicon Motion Systems Ltd., Oxford, UK). A schematic is shown in Figure D.1.

The walls of the room were covered with a camouflage net, which provided a natural-looking background texture and prevented the birds from landing on the motion capture cameras' scaffolding. The floor was carpeted with astroturf. Additional elements were placed in the flight lab for the experiments. The main setup consisted of a curtain and obstacles placed in front and/or behind it (see Figure D.1). In the static goal trials, we additionally placed two perches, each at one end of the room. In the moving goal trials, we added a system of pulleys and a linear motor. These elements are described in the next sections.

D.2.1 Curtain

The curtain was made of two layers of flame-proofed cotton canvas of size 6.0 m $\times$ 3.5 m. It was hung from a steel cable installed transversally across the room (see Figure D.3). It was extended leaving a gap on either side, between curtain and wall, of approximately 1 - 1.1 m in all trials (around one wingspan). The gap's width was exactly measured by placing between 1 and 4 markers on each side of the wall netting and the curtain edges. The longitudinal location of the curtain was kept constant for practical reasons.

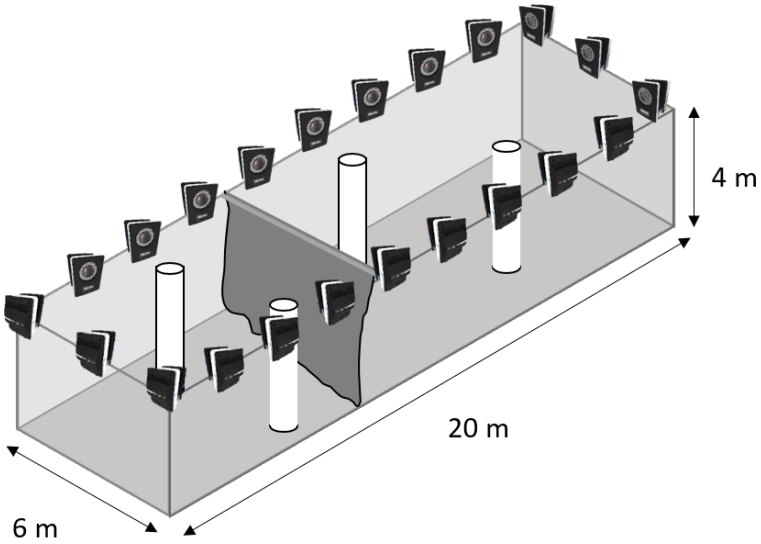


Figure D.1: Schematic of the flight lab, with the curtain and obstacles. The curtain was used to force turns. Obstacles made of individual pillars are shown. Camera icons from https://www.vicon.com/hardware/cameras/. Not to scale.



(a) Initial astroturf cover

(b) Final astroturf cover

Figure D.2: Flight lab with the two astroturf covers. Camouflage netting is used to cover the sidewalls. **(a)** For the first 8 weeks of the experiments (all static goal trials and three weeks of moving goal trials) the astroturf covered the central part of the room. **(b)** For the final three weeks it was extended to cover the width between the side wall netting



(a) Without reinforced edge



(b) With reinforced edge

Figure D.3: Curtain. Its extension was kept approximately constant for all experiments, with a gap width of 1 - 1.1 m. The gap width was registered for every trial by placing retroreflective markers in the curtain edge and wall netting.

D.2.2 Perches

In the static goal trials the birds were trained to fly from perch to perch, each one placed in opposite sides of the curtain. The perches were 60 cm wide and 134 cm high. We randomised the initial and end perch positions.

D.2.3 Pulleys and linear motor

In the moving goal trials, the birds were trained to follow a moving target, attached to a line and pulled by a linear motor. The target's line was passed through several pulleys (13 cm diameter, 16 cm high) placed on the floor. We used green lines to minimise contrast with the astroturf carpet and placed dummy lines along the alternative pulley paths, to minimise the risk of the bird predicting the target's trajectory from the lines' placement.

D.2.4 Obstacles

We used obstacles in both types of trial. In the static goal trials, we explored different types of obstacles, arrangements and randomisations for their locations. From there we consolidated a configuration for moving goal trials that was both practical (i.e., one we could rearrange quickly to capture as many flights as possible) and forced substantial turning in the birds' trajectories. In the moving goal trials, we used one type of obstacles and arrangement, which was an adaptation of the best configuration we found during the static goal trials. The types of obstacles and arrangements explored are described below.

Obstacle types

We used two types of obstacles, of simple geometry to facilitate their later synthetic reconstruction:

- monoliths or cuboid obstacles, of dimensions 2.4 m x 0.4 m x 0.225 m with a base of dimensions 0.5 m x 0.425 m x 0.075 m for stability. We used these in the static goal trials only.

- pillars or obstacles of cylindrical shape, of 2 m high and 30 cm wide. We used these in the static and the moving goal trials.

We trialled a different kind of obstacles (thinner poles) in some of the warmup (familiarisation) flights, but since they didn't force large turns in the birds' trajectories we discontinued their use.

Obstacle arrangements

We use the term *obstacle arrangement* to refer to a way of grouping obstacle elements to form an *obstacle unit*. For example, a pair of pillars is an obstacle arrangement that consists of grouping two pillars together to form one obstacle unit. We explored different obstacle arrangements, with the aim of inducing larger turning in the birds. We did that by using wider obstacles and/or by placing obstacles ahead of the curtain as well as behind it. We usually introduced obstacles in the back section first, following recommendations from the falconers.

The obstacle arrangements used are described in Table D.1 and shown in Figures D.4 to D.10. From the exploration during the static goal trials, we concluded the three pillars arrangement in the front and back sections was a good choice for our aims. We used this arrangement as a reference for the moving goal trials.



Figure D.4: Monoliths. Views from end perch. Arrangement used in static goal trials only.



Figure D.5: Individual pillars. Views from end perch. Arrangement used in static goal trials only.

D.3 Experimental procedure

In the static goal setup, each individual flew approximately 16 flights. For the first 6 out of 28 days we carried out 'warmup trials' (150 trials), to familiarise the birds with the setup.

Table D.1: Obstacle arrangements used in static and moving goal trials. Front and back sections refer to the areas ahead of and behind the curtain respectively. Except when stated, pillars are made up of two cylinders of 1 m height and 30 cm diameter stacked on top of each other.

Obstacle arrangement	Type of trial	Description
Monoliths	Static goal	• Placed in the front and back sections. • 369 recorded flights.
Individual pillars	Static goal	• Placed in the back section only or in the front and back sections simultaneously. • 219 recorded flights.
Pairs of pillars	Static goal	• Pillars grouped in pairs to form wider obstacles (60 cm wide). • Placed in the back section only. • 257 recorded flights.
Monoliths (front) and pairs of pillars (back)	Static goal	• Monoliths in the front section, and pairs of pillars in the back section. • 64 recorded flights.
Three pillars (front and back)	Static goal	• Pillars grouped in threes to form wider obstacles (90 cm wide). • Placed in the front and back sections simultaneously. • 155 recorded flights.
Pairs of high pillars (front) and three pillars (back)	Static goal	• Higher pillars (3 m high) by stacking three cylinders. • Placed pairs of high pillars in the front section and groups of three regular pillars in the back section. • 25 recorded flights.
Pillars in 2+1 configuration (front and back)	Moving goal	• Groups of three pillars in a 2+1 configuration. • A gap is left between a pair and a single one to allow the target to pass through. • Apart from the location, the side of the most voluminous obstacle is randomised. • Arrangement used in all moving goal trials. • 251 recorded flights.



Figure D.6: Pairs of pillars. Views from end perch. Arrangement used in static goal trials only.



Figure D.7: Monoliths (front) and pairs of pillars (back). Views from end perch. Arrangement used in static goal trials only.



Figure D.8: Three pillars (front and back). Views from end perch. Arrangement used in static goal trials only.



Figure D.9: High pillars (front) and three pillars (back). Views from initial perch (left) and end perch (right). Arrangement used in static goal trials only.

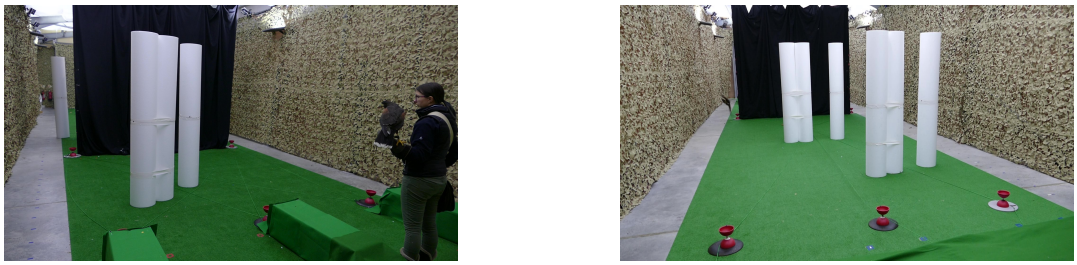


Figure D.10: Pillars in 2+1 configuration (front and back). View from initial perch (left) and end perch (right). This arrangement was used in all moving goal trials.

In the moving goal setup we collected maximum 4 flights per bird. The number of trials was lower than in the static goal case because of the training approach: a food reward ('pick-up piece') was required to get the bird off the lure, which meant larger food intake per trial. This limited the amount of flights we could carry out while keeping the animals motivated to fly. We used the trials on the first 6 out of 20 days as warmup trials (49 trials).

On any day of experiments, we first calibrated the motion capture system with the curtain removed. We then installed the curtain, generated the randomisation tables and prepared the setup for either the static or the moving goal trials. Before the flights, we carried out a gaze calibration recording, in which we recorded the pose of the head markers on the bird while a food reward with a marker was presented to it, at a range of $0.5 - 1.5$ m. We then recorded the bird's flights. After the flights across all birds were recorded, we carried out a 'coverage recording', in which we spun around the motion capture volume a cluster of markers at known distances between them. The aim was to later quantify how uniform the quality of the reconstruction was across the room. At the end of every day the lab environment was mapped using the SLAM application deployed in a virtual-reality smartphone.

The following sections provide further details on the experimental procedure followed in the static and the moving goal trials.

D.3.1 Static goal trials

In the static goal trials we randomised the initial and end perch positions and the obstacles positions. We collected 1217 trials, 1089 with obstacles in place and 128 without obstacles (warmup trials excluded).

Initial and end perches

We randomised the initial and end perch positions in two ways:

- Three discrete locations. For 848 trials, the perches positions were randomised between three possible options.

- Continuous initial perch. For a smaller set of trials (369 trials), the position of the initial perch was varied continuously in a range of 3.5 m, centred at the midpoint between the camouflage nets on the sidewalls. For 255 of these trials, the end perch position was fixed; for the remaining 114 trials the end perch was randomised in three possible locations. For all trials without obstacles (128 trials) the initial perch was varied continuously and the end perch was fixed.

The perches locations for static goal trials are represented schematically in Figure D.11. The distances in the Figure D.11 are indicative, and for gaze analysis and the reconstruction of the environment the motion capture data is always used as ground truth.

Obstacles for static goal trials

In 1089 of the 1217 static goal trials recorded we had obstacles in place. We explored two randomisations for their positions:

- For the monoliths we used a *randomisation based on stations*, for which we discretised the longitudinal position of the obstacles in several stations. We used variations of this approach for a total of 369 flights. Further details on this approach are included in Section D.6.
- For the pillars we used a *randomisation based on grid*, in which we defined a grid over the obstacles' space, and randomly picked a cell to determine the obstacles' position for each trial, with some restrictions. We used variations of this approach for the rest of static goal trials (720 flights). Further details on the variations of this approach are presented in Section D.7.1.

A schematic of these two approaches is shown in Figure D.12. In the grid randomisation, we restricted the possible locations of the obstacles to a certain set of cells for the majority of the trials. We again used the static goal trials to explore different ways of restricting the space, and then selected one for the moving goal trials. Schematics of the restrictions explored in static goal trials are shown in D.13 and further described in Section D.7.1.

D.3.2 Moving goal trials

In the moving goal trials we randomised the trajectory of the target by randomising the pulleys the line went through. We also randomised the obstacles' locations. We collected 202 flights (warmup trials excluded).

Pulleys

The location of the initial and end pulleys was randomised between three different positions each (see Figure D.15). The gap pulley, located at the curtain edges, was randomised between the left or right side. An additional pulley (called 'central pulley') was introduced between the curtain and the end pulley, and its longitudinal position was randomised. The addition of this central pulley was also randomised, distinguishing paths in diamond configuration (i.e., no central pulley) or in zigzag configuration (i.e., with central pulley). Figures D.14 and D.15 show top-down view schematics and pictures of the Flight lab for each configuration. We also randomised the falconer's position, who stood around 1 m behind of the initial pulleys, since the birds take off from their fist.

Obstacles for moving goal trials

In the moving goal trials, we used pillar obstacles in a 2+1 configuration. This configuration is based on the three pillar one from the static goal trials, but with a gap in between a pair and a single pillar. This was to prevent the bird from successfully avoiding the obstacles by simply closely following the target. With

a 2+1 configuration the target was able to go through the gap between them, but the bird was forced to turn.

We used a randomisation based on grid using reference trajectory restriction ('RTR', see Figure D.16). The reference trajectories for restricting the grid cells are possible target trajectories. We didn't use the actual initial and end positions of the target per trial, instead we randomised the reference trajectory for the grid restriction separately. This was done so that the bird couldn't predict from the obstacles' positions the trajectory of the target. Further details on the randomisation approach for the moving goal trials are presented in Section D.7.2.

D.4 Summary of recorded flights

Tables D.2 and D.3 summarise the number of recorded flights for each type of obstacle arrangement, for the static goal trials and the moving goal trials respectively.

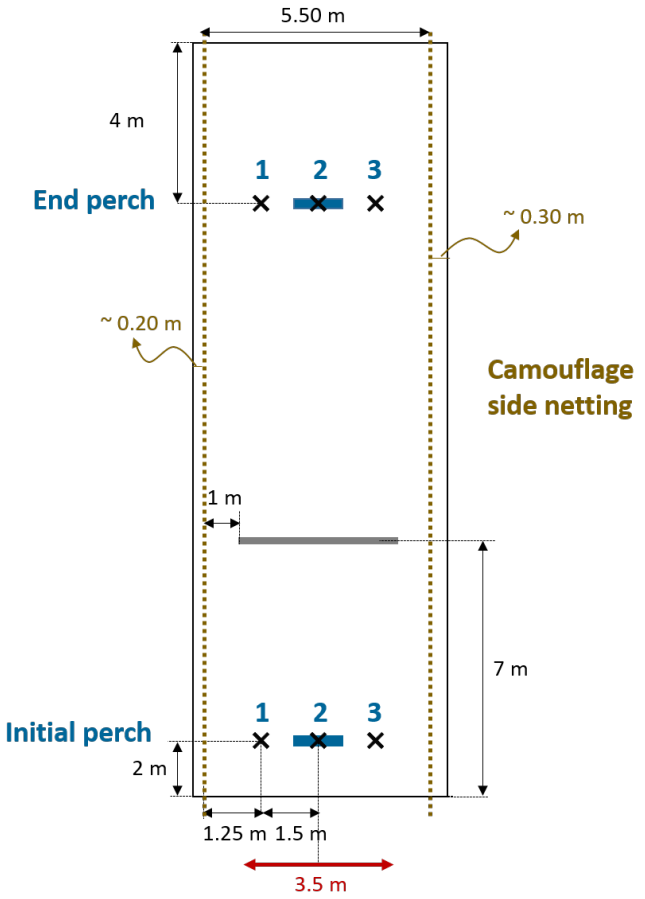
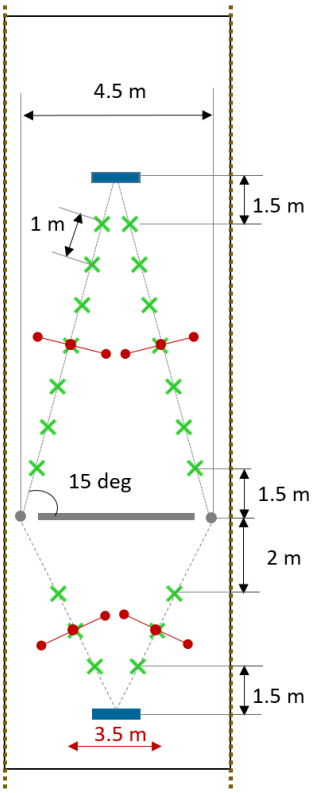
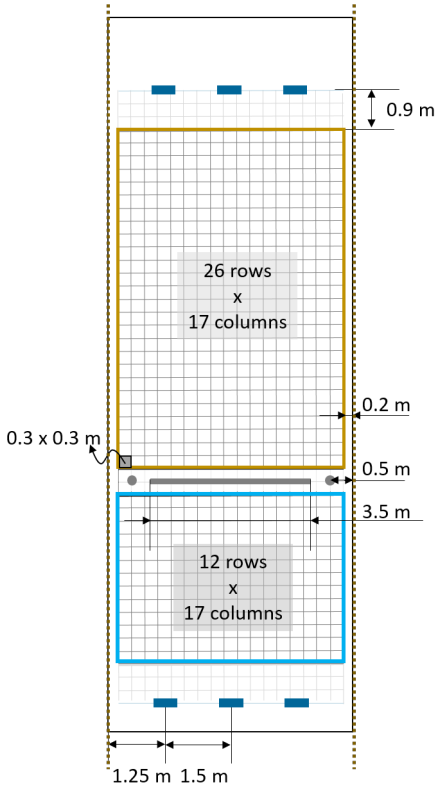


Figure D.11: Initial and end perches (static goal trials). The approximate locations of the perches (blue rectangles) and curtain (grey rectangle) are indicated in the figure. We considered three discrete positions for the initial and end perch (blue numbers). The black crosses represent the location of the perches' centres at these three positions. The red arrow indicates the range of positions for the initial perch when varied continuously. Note that the room is asymmetric in the way the side camouflage net is hung (khaki dotted line), so we centred the curtain and position 2 of the perches to the midpoint between the nets at the sidewalls. The perches' width is approximately 0.6 m. The direction of flight is from the bottom of the image to the top. Note we always use the motion capture data to determine the exact position of the static objects. Not to scale.

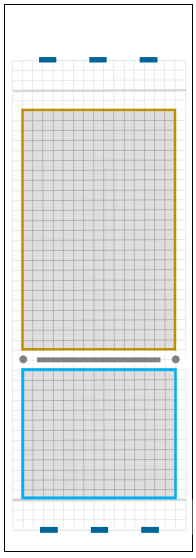


(a) Based on stations. Used for the monolith obstacles in the static goal trials. Approximate measurements shown. Not to scale.

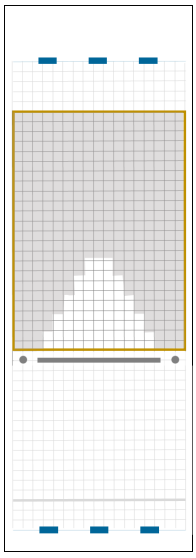


(b) Based on grid. Approach used for the pillar obstacles in the static and the moving goal trials. Not to scale.

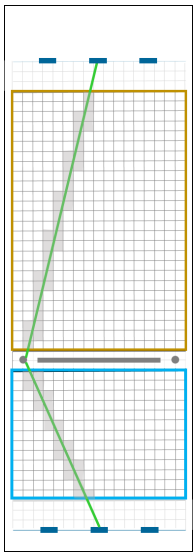
Figure D.12: Randomisation approaches (top-down view). The curtain (grey rectangle) and perches (blue rectangles) are shown, as well as the gaps' midpoints (grey dots). The lateral boundaries are the sidewalls netting (dotted khaki lines). Figure **(a)** shows the location of the different stations considered (green crosses) and their transversal deviations (red dots) along the reference line (grey dotted line). The initial perch was moved continuously (indicated by red arrow) in a range of 3.5 m, centred at the midpoint between nettings on the sidewalls. The end perch was fixed, except for the last 114 trials. Further details on the variations are presented on Table D.4. Figure **(b)** shows the general grid approach for static goal trials. The areas for the front obstacles (cyan box) and the back obstacles (yellow box) are indicated. The two rows ahead the initial and end perches were excluded following recommendations from the falconers (light grey cells). The distances shown were used for computing the reference trajectories.



(a) No restrictions, 15 columns, 12 front rows, 24 back rows.



(b) Triangle restriction, 17 columns, 24 back rows.



(c) Reference trajectory restriction (RTR)

Figure D.13: Restrictions for static goal trials. The possible obstacle locations are shown as grey cells. The perches (blue rectangles) and the curtain (grey rectangle) are also shown, as well as the gap midpoints (grey circles). The number of rows and columns is indicated when different from the general grid definition. For the case of reference trajectory restriction (RTR) **(c)**, a sample hypothetical reference trajectory is shown (green line) and the reference cells (cells that overlap with the trajectory) are shown in grey. A certain amount of cell deviations from these reference cells was considered (from 0 to 3 cells, see Tables D.5 and D.6)



(a) Front section.

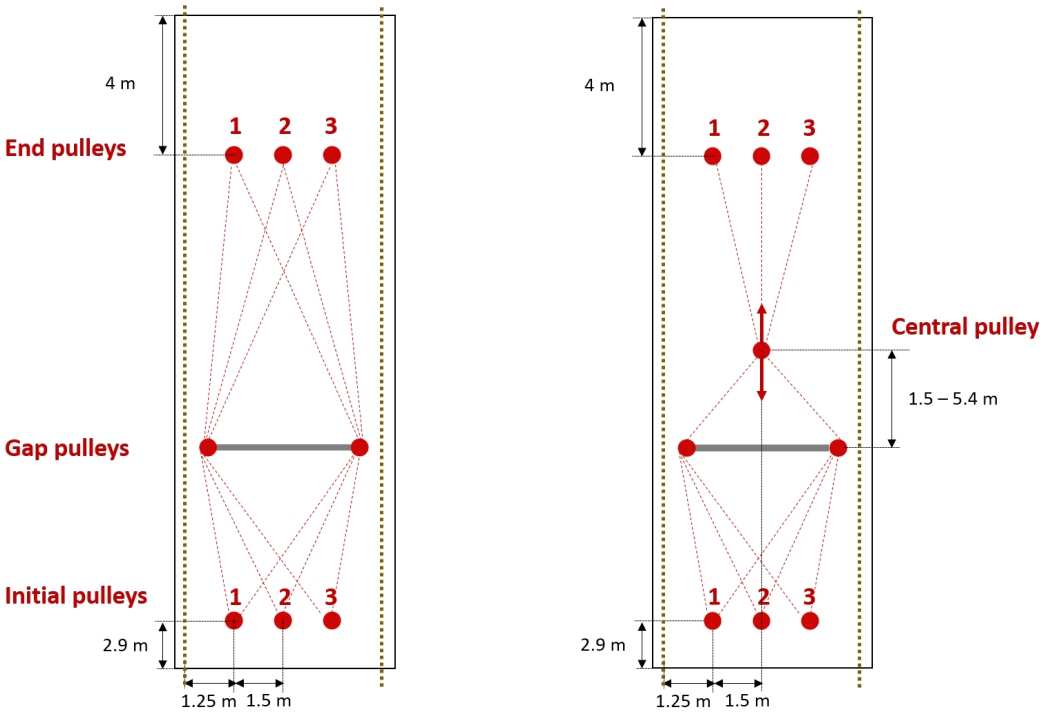


(b) Back section: diamond configuration.



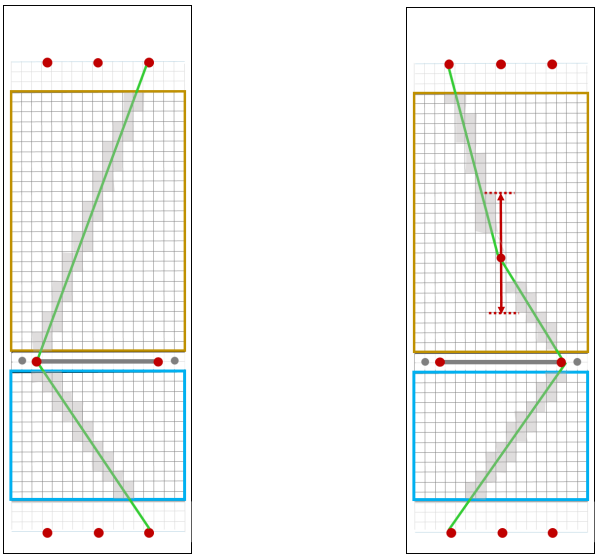
(c) Back section: zig-zag configuration.

Figure D.14: Pictures of the pulleys' configuration (moving goal trials). The pulleys define the trajectory points for the path of the target. **(a)** The initial pulleys and the boxes behind them that cover the target at the start are shown. The bird takes off from the falconer's fist around 1 m behind the row of initial pulleys. **(b)** End and gap pulleys are shown. **(c)** End, gap and central pulleys are shown.



(a) Diamond configuration **(b)** Zig-zag configuration

Figure D.15: Schematic of the pulleys' configuration (moving goal trials). The red circles represent the possible pulleys' locations and the red dotted lines the target's possible trajectories. Trajectories are obtained by randomising between three possible locations for the initial pulley, two possible locations for the gap pulley (left or right) and three locations for the end pulley. For the zig-zag configuration (b), we also randomised the longitudinal position of an additional central pulley, located between the curtain and the end pulley. The curtain's location is the same as in the static goal trials (repeated distances are omitted for clarity). The direction of flight is from the bottom of the image to the top. The birds take off from the falconers' fist, approximately 1 m behind the initial pulleys; this starting position was also randomised. The approximate location of the side netting is represented by the khaki dotted line. Schematic not to scale.



(a) Reference trajectory restriction - diamond

(b) Reference trajectory restriction - zigzag

Figure D.16: Restrictions for moving goal trials. The figure shows two sample hypothetical reference trajectories (green lines) and their reference cells (cells that overlap with the trajectory, in grey). The pulleys' locations are also shown (red circles). Note that a certain amount of cell deviations from these reference cells was considered (from 0 to 3 cells, diagonally or laterally, see Table D.7). The reference trajectories match the target's possible paths and thus go through the gap pulleys instead of the gap midpoints. Figure **(b)** shows the range of locations for the central pulley (red arrow). The two front rows ahead and behind the central pulley were excluded following recommendations from the falconers.

Table D.2: Types of obstacles and arrangements for the static goal trials. Percentages with respect to total recorded flights (in bold).

Type of obstacle	N flights	N flights no obstacles	N flights back obstacles only	N flights front obstacles only	N flights front and back obstacles
Monoliths	369 (27%)	5	45	56	263
Individual pillars	219 (16%)	0	102	0	117
Pair of pillars	257 (18%)	0	257	0	0
Monoliths (front) and pair of pillars (back)	64 (5%)	0	0	0	64
Three pillars	155 (11%)	0	0	0	155
Two high pillars (front) and three pillars (back)	25 (2%)	0	0	0	25
No obstacles	128 (9%)	128	0	0	0
Warmup (no obstacles or poles)	150 (11%)	54	72	8	16
Total (excluding warmup)	1217	138	400	56	623
Total	**1367**	192 (14%)	472 (35%)	64 (5%)	639 (47%)

Table D.3: Types of obstacles and arrangements for the moving goal trials. Percentages with respect to total recorded flights (in bold).

Type of obstacle	N flights	N flights no obstacles	N flights back obstacles only	N flights front obstacles only	N flights front and back obstacles
2+1 configuration	202 (80 %)	1	27	9	165
Warmup trials	49 (20 %)	49	0	0	0
Total raw	**251**	50 (20%)	27 (11%)	9 (4%)	165 (66%)

D.5 Labelling procedure

We developed custom algorithms to label the motion capture data. The method followed was similar to the one used to label the ground-truth data in the bird trials of Chapter 2.

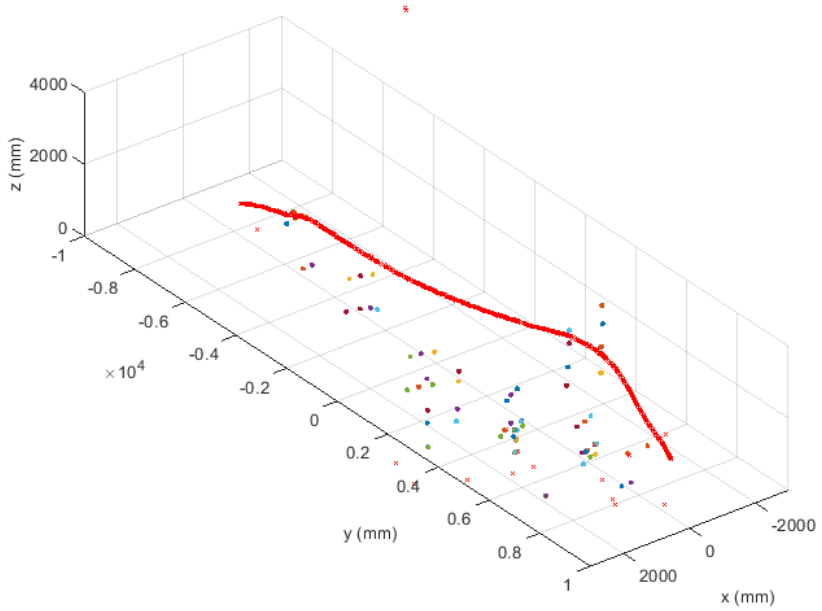
For each trial, we first separated static and moving markers. We used an approach based on bounding boxes to label the static markers based on the object they belonged to. Two examples of the object labelling are shown in Figures D.17b and D.18b, for a static and moving goal trial respectively. Figure D.18a shows an example of the bounding boxes used.

From the set of moving markers per trial, we identified the set of headpack markers using a custom clustering algorithm. We used a modified version of an algorithm developed by Marco Klein Hereenbrink. This algorithm defines markers as belonging to the same cluster if they are: (1) separated by less than a certain distance, (2) existing for a minimum number of frames, and (3) separated by a gap of a maximum number of frames. We manually tuned these parameters until the results were satisfactory. For the moving goal trials, we identified the target's markers following a similar approach, with additional constraints on the maximum z-coordinate observed. The identified cluster of headpack markers is shown in magenta for the sample trials in Figures D.17b and D.18b.

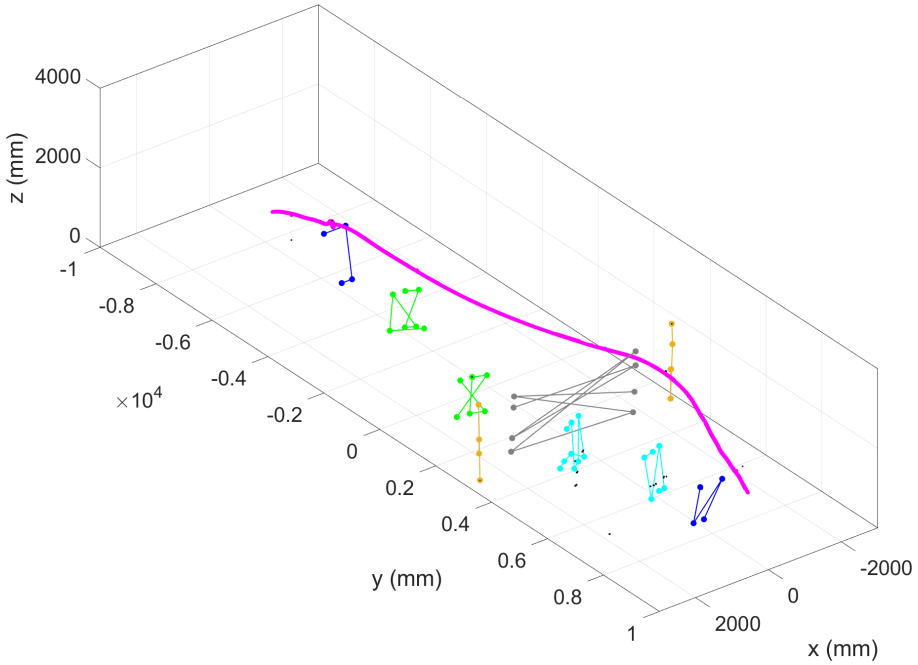
Once the set of headpack markers was identified, we used another custom algorithm to label the individual headpack markers per frame, and to extract the headpack translation and rotation per frame. This was only applied to the static goal trials. The approach is very similar to the one described in Chapter 2 to compute the ground-truth labels of the headpack markers on the bird trials; a schematic is shown in Figure D.19. The approach combines per-frame and temporal labelling methods. From an input template (or an estimated one from the data), a first set of labels is computed using geometric features of the headpack (in the figure, 'labelling in edges space'; note this is not the same as the 'EE' space defined in Chapter 2) and iteratively solving the procrustes problem per frame (in the figure, 'iprocrustes'). Then, the headpack transform per frame is computed in the labelled frames, and interpolated in the unlabelled frames. The unlabelled markers in the interpolated frames are labelled if they are close enough to the markers in the interpolated headpack, using a constant threshold ('iterative minimum distance', and 'minimum distance' in the figure). Finally, we combine all the labels from the different steps, define a final template using the data and extract the final headpack transform per frame. Figure D.20 shows an example flight in which the headpack pose was extracted after labelling the individual headpack markers.

D.6 Randomisation based on stations

We defined several longitudinal positions for the monolith obstacles, along the lines that join the midpoints of the curtain gaps and the central position of the perches (grey dotted line in Figure D.12a). The distance between stations is 1 m along this reference line. We explored six variations of this randomisation approach, which are summarised in Table D.4. Figure D.22 shows the area covered by the

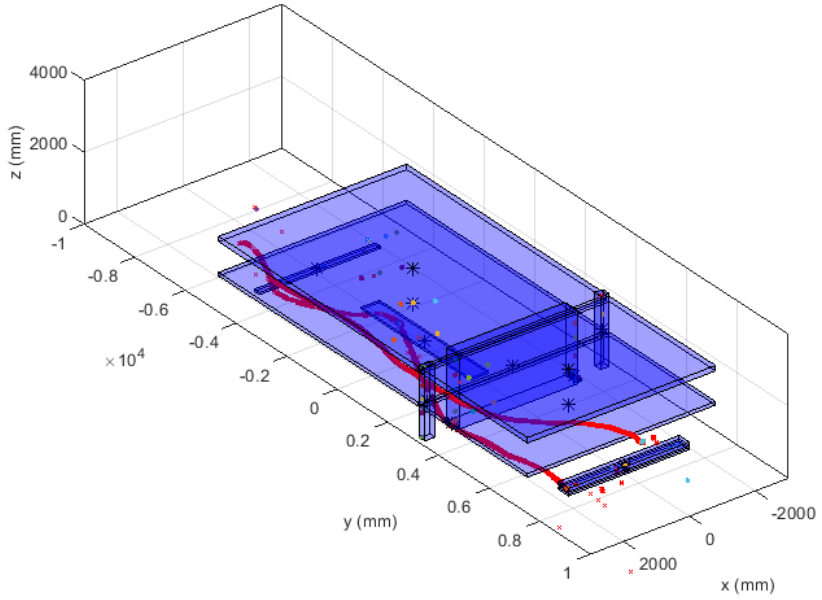


(a) Stationary and moving markers. Moving markers (red crosses) and stationary markers (colour dots) are shown.

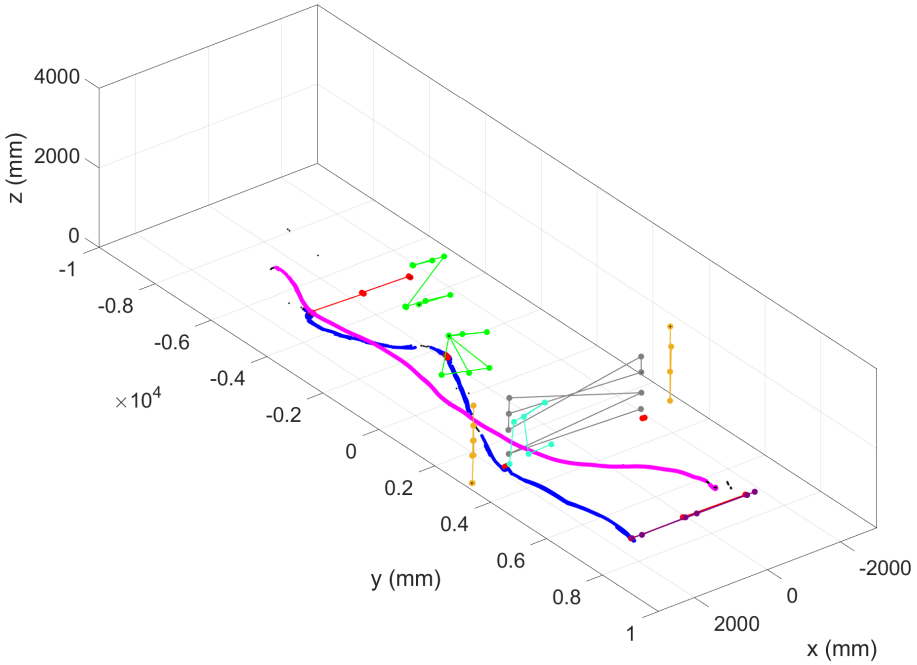


(b) Labelled objects for static goal trial. Markers are labelled according to the object they belong to. Objects are colour-coded: initial and end perch (navy), walls (yellow), curtain (grey), front obstacles (cyan) and back obstacles (green) are shown. Headpack markers are shown in magenta. Unlabelled markers are shown in black.

Figure D.17: Labelling of markers for a static goal trial.



(a) Stationary and moving markers. Moving markers (red crosses) and stationary markers (colour dots) are shown. Moving markers are mostly headpack markers and target markers. The bounding boxes for the different objects are shown in blue.



(b) Labelled objects for moving goal trial. Markers are labelled according to the object they belong to. Objects are colour-coded: starting (purple), pulleys (red), walls (yellow), curtain (grey), front obstacles (cyan) and back obstacles (green) are shown. Headpack markers are shown in magenta and lure markers in blue. Unlabelled markers are shown in black.

Figure D.18: Labelling of markers for a moving goal trial

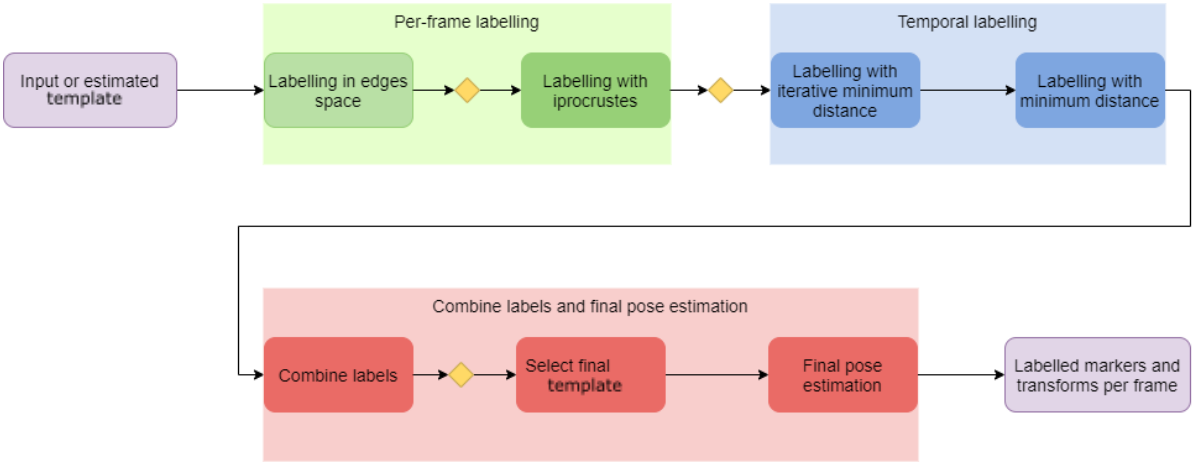


Figure D.19: Labelling approach followed for the individual markers on the headpack. The yellow diamond indicates the headpack template was updated using all the labelled data up to that point.

obstacles in each of these variations. For this set of trials only, when the trials were 'without obstacles', we placed the obstacles just in front or behind the curtain (away from the reference line but not hidden from view, see the area around the middle of the curtain in Figures D.22b to D.22f).

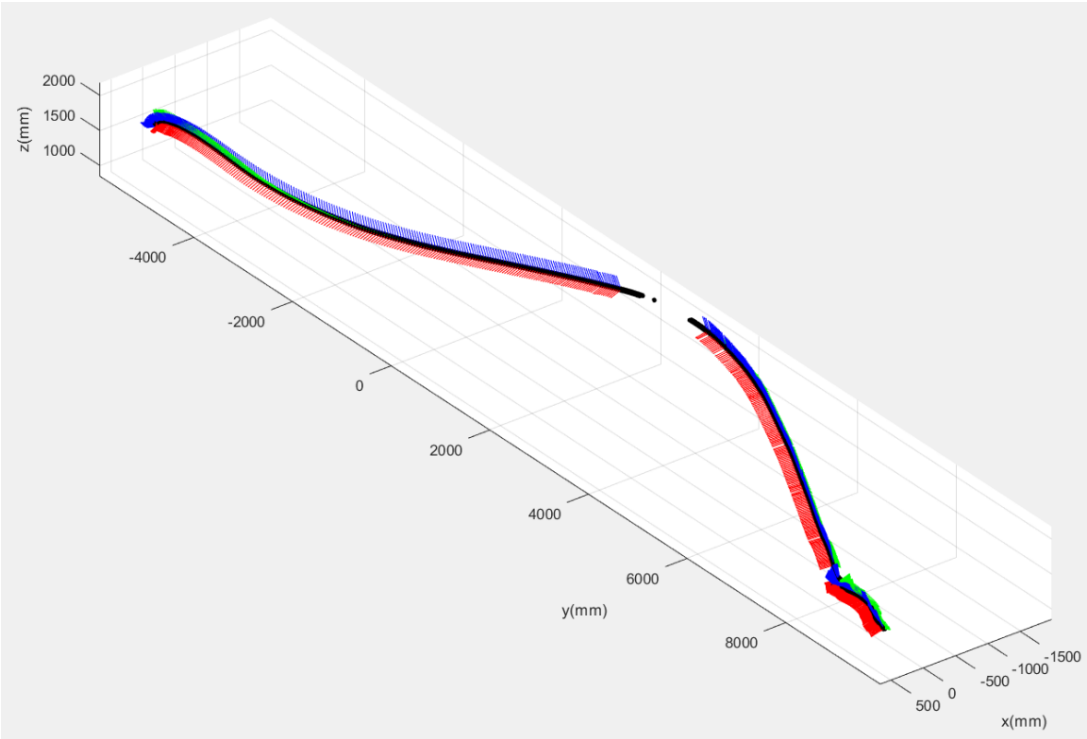


Figure D.20: Headpack pose for a sample static goal flight. The figure shows the headpack pose computed for a sample static goal flight. The red, green and blue vectors per frame represent the x, y and z-axis of a coordinate system fixed to the headpack. The headpack markers are plotted in black; note that for a portion of the flight no headpack markers were reconstructed. In the period before and after the missing markers there weren't enough markers to extract the headpack pose.

Table D.4: Randomisation based on stations. Summary of the variations of this approach. 'Cont.' stands for continuous

Randomisation variation	Number of obstacles	Initial perch	End perch	Front obstacles	Back obstacles
Longitudinal position	2 monoliths (front or back)	Cont.	Fixed	• 0-8 stations • 0 is 'no obstacles' • 1-7 are back obstacles stations • 8 is 'front obstacles only'	
Transversal position 1	4 monoliths (2 front, 2 back)	Cont.	Fixed	• 0-3 stations • Inner/outer	• 0-7 stations • Inner/outer
Transversal position 2	4 monoliths (2 front, 2 back)	Cont.	Fixed	• 0-3 stations • Inner/centre/outer	• 0-7 stations • Inner/centre/outer
Transversal asymmetry (back)	4 monoliths (2 front, 2 back)	Cont.	Fixed	• 0-3 stations • Inner/centre/outer	• 0-7 stations • Left back obstacle: Inner/centre/outer • Right back obstacle: Inner/centre/outer
Gap option (back)	4 monoliths (2 front, 2 back)	Cont.	Fixed	• 0-3 stations • Inner/centre/outer	• 0-7 stations • Back obstacle 1: left/right branch + inner/centre/outer • Back obstacle 2: left/right branch + inner/centre/outer
Gap option (front)	4 monoliths (2 front, 2 back)	Cont.	3 stations	• 0-3 stations • Front obstacle 1: left/right branch + continuous transversal • Front obstacle 2: left/right branch + continuous transversal position	• 0-7 stations • Back obstacle 1: left/right branch + continuous transversal position • Back obstacle 2: left/right branch + continuous transversal position

D.7 Randomisation based on grid

D.7.1 Static goal trials

We divided the obstacles' space into a grid and selected a cell of the grid at random with some restrictions, to define the obstacles' locations. Figure D.21 shows the grids defined for static goal trials and moving goal trials, and the possible obstacle locations prior to applying restrictions (grey cells). The differences between the grids' definitions owe to practical measures regarding the pulleys setup. In the static goal trials we explored different ways of restricting the space, and then selected one of the approaches for the moving goal trials.

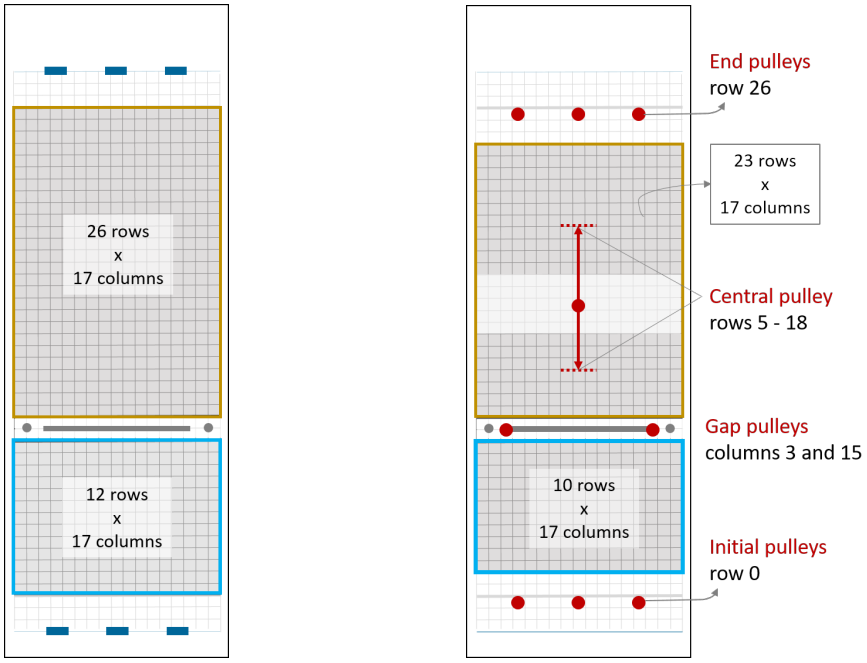
In the static goal trials we explored the following options (schematics of each case are shown in Figure D.13):

- *no restrictions (all cells).* For the first 152 flights we randomly assigned obstacles at one of the possible locations as shown in Figure D.13a. We used individual pillars in both back and front sections. Figures D.23a to D.23c show the obstacles' cover with this approach.

- *triangle restriction.* For the next 67 trials, we excluded the area in the wake of the curtain (see Figure D.13b). The area covered by the individual pillars in these trials is shown in Figure D.23d.

- *reference trajectory restriction (RTR).* For the majority of the trials we only considered cells that fell within a certain range of a hypothetical reference trajectory. This range (cell deviation) varied from 0 to 3 cells from the reference trajectory, diagonally or along the same row, depending on the set of trials. For the static goal trials, the reference trajectories started and ended at the actual initial and end perch positions of the given trial, but the curtain side they went through was randomly assigned (left or right). We first applied this approach on the back section only with a pair of pillars. Figure D.24 shows the obstacle cover for these trials. We then extended it to the front section as well, and used wider obstacles made of three pillars, or monoliths. Figure D.25 shows the obstacle cover for these.

The variations of these general approaches used in the static goal trials are summarised in Table D.5 and D.6.



(a) Grid for static goal trials. .

(b) Grid for moving goal trials.

Figure D.21: Grid definitions for static and moving goal trials. The curtain and gap midpoints are shown (grey rectangle and grey dots respectively), as well as the areas for front obstacles (cyan box) and back obstacles (yellow box). Cells omitted from the randomisation are shown as translucent; grey cells are potential obstacle locations before applying restrictions. The maximum area considered for obstacles is shown (see Tables D.5 and D.6 for particular cases). Figure **(a)** shows initial and end perches' positions (blue rectangles). Figure **(b)** shows the pulleys locations in the grid (red circles). Note that both initial and end pulleys are pushed inwards with respect to the perches' positions. Note the cells around these and the central pulley were excluded following recommendations from the falconers. The range of locations of the central pulley is indicated (red arrow). The boundaries of the obstacle areas used in the previous static goal trials are shown (dark thick row lines next to initial and end pulleys).

Table D.5: Randomisation based on grid for static goal trials: individual and pair of pillars.

Randomisation variation	Obstacles arrangement	Front obstacles?	Description	Number of flights
No restriction (all cells)	Individual pillars	Yes	• 15 columns, 12 rows front, 24 rows back	152
Triangle restriction	Individual pillars	No	• 17 columns, 24 rows back	67
Reference trajectory restriction + Biased + Transversal position 1	Pair of pillars	No	• 17 columns, 24 rows back • Ref. trajectory biased to right gap for initial position 2 • Transversal displacement along row (0-3 cells) • Side of second obstacle: closest to reference trajectory	64
Reference trajectory restriction + Biased + Transversal position 2	Pair of pillars	No	• 17 columns, 24 rows back • Transversal displacement along cell's diagonal (0-3 cells) • Maximum three cells deviation	64
Reference trajectory restriction + Unbiased + Spread	Pair of pillars	No	• 17 columns, 24 rows back • Ref.trajectory for initial position 2, randomly pick a curtain side • Two and one cell deviations • An error in the computed geometry led to further obstacle spread • Random position of second obstacle along the cell's diagonal • Added two columns symmetrically	64
Reference trajectory restriction + Unbiased + One cell deviation	Pair of pillars	No	• 17 columns, 26 rows back • One cell deviation • Added two rows at the back	65

Table D.6: Randomisation based on grid for static goal trials: obstacles in front and back

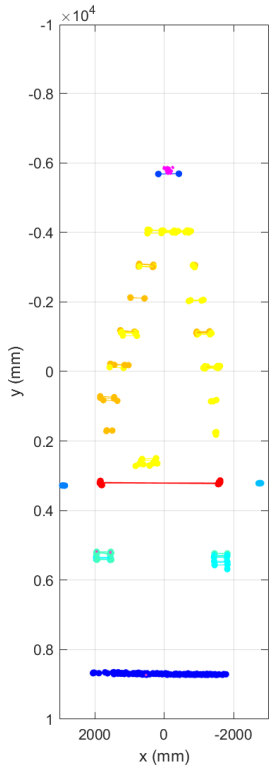
Randomisation variation	Obstacles arrangement	Front obstacles?	Description	Number of flights
Reference trajectory restriction + Unbiased + One cell deviation	Monoliths (front) and pair of pillars (back)	Yes	• 17 columns, 12 rows front, 26 rows back • Monolith as an obstacle unit for the front section	64
	Three pillars	Yes	• The randomised cell is the centre of the block of three pillars	131
Reference trajectory restriction + all trials with obstacles	Three pillars	Yes	• 17 columns, 12 rows front, 26 rows back • For each trial, an obstacle is placed using reference trajectories through both gaps	49

D.7.2 Moving goal trials

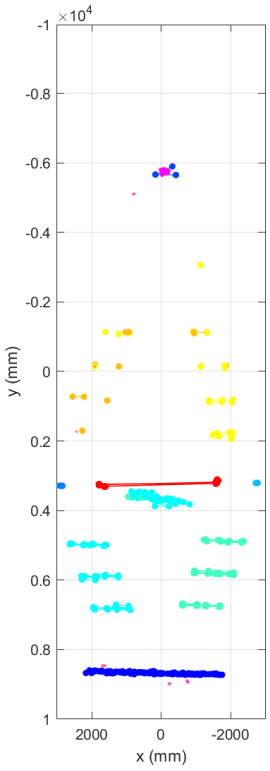
For the moving goal trials, we generated two hypothetical trajectories per flight, with a 1/4 possibility of having no obstacles (in both front and back). In the zig-zag configuration, we used the actual position of the central pulley (randomised from row 5 to 18 in the back) when generating the hypotheses for the reference trajectory. We randomised the lure speed between 7.5 and 8.5 m/s, in steps of 0.45 m/s after the first 16 trials, then capped the speed to 8 m/s after 48 trials. The specific variations on these approaches are summarised in Table D.7.

Table D.7: Randomisation based on grid for moving goal trials.

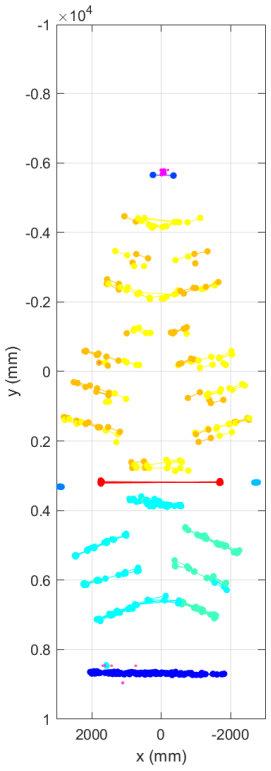
Randomisation variation	Obstacles arrangement	Description	Number of flights
Diamond - gap midpoints	2+1 configuration (2 back only)	• Initial pulleys at row 2 • 17 columns, 8 front rows, 23 back rows • Reference trajectory through gap midpoints • Cell randomisation indicates central pillar • Max one cell deviation transversally • Side of block of two always closest to ref. trajectory • Probability 1/4 of no obstacles for front and back	16
Diamond - varying lure speed	2+1 configuration (2 front, 2 back)	• Introduced lure speed variation (from $7.5m/sto8.5m/s$) • Initial pulleys at row 0 • 17 columns, 10 front rows, 23 back rows • Cell from randomisation indicates gap between pair and single pillar • Side of block of two is randomised • Max one cell deviation	16
Diamond - gap pulleys	2+1 configuration (2 front, 2 back)	• Reference trajectory through gap pulleys, assumed at curtain edges •No cell deviation	16
Diamond - capped lure speed	2+1 configuration (2 front, 2 back)	• Lure speed capped to $8m/s$ • No cell deviation	30
Zigzag	2+1 configuration (2 front, 2 back)	• Introduced randomisation of central pulley • No cell deviation	100
Combined	2+1 configuration (2 front, 2 back)	• We randomised per trial whether to do diamond or zigzag configuration (probability of 1/2 each)	32



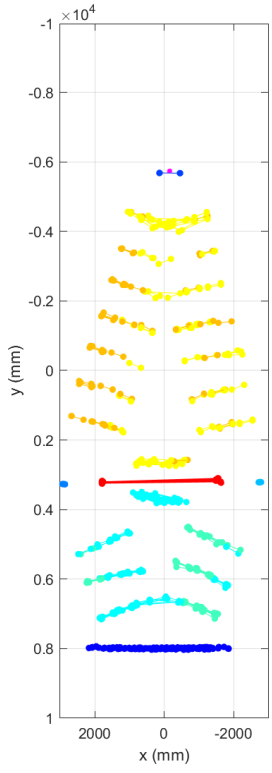
(a) Longitudinal position, $n = 64$



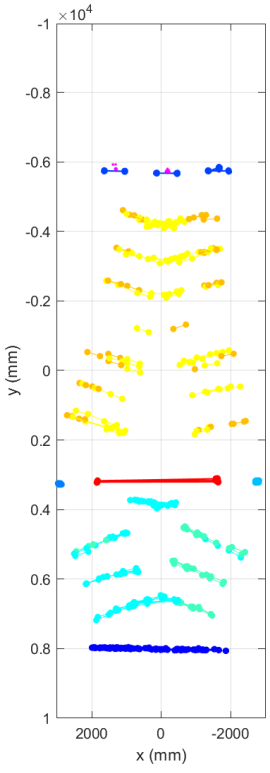
(b) Transversal position 1, $n = 64$



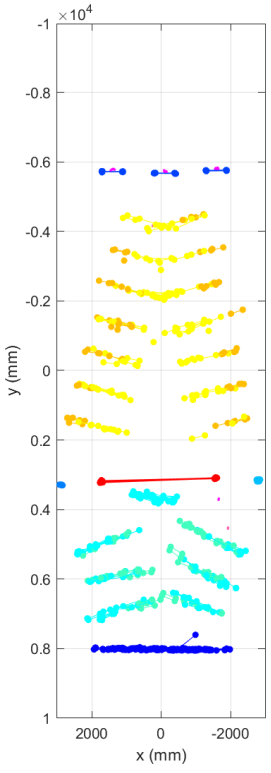
(c) Transversal position 2, $n = 63$



(d) Transversal asymmetry (back), $n = 64$

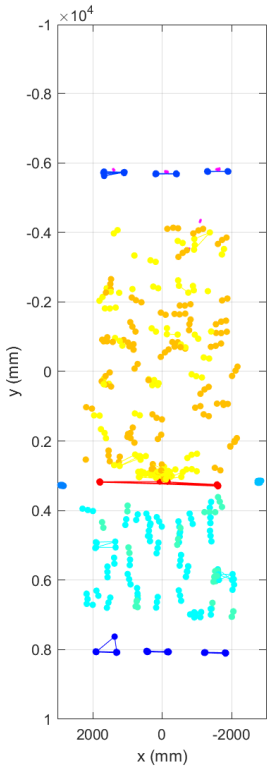


(e) Gap option (back), $n = 50$

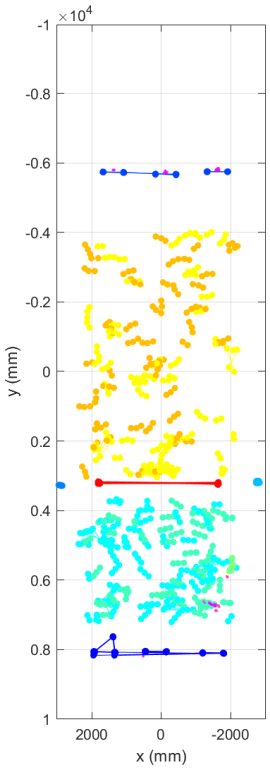


(f) Gap option (front), $n = 64$

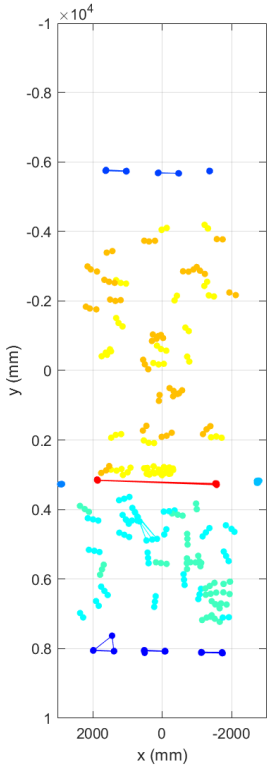
Figure D.22: Area covered by the obstacles for the static goal trials using the randomisation based on stations. Labelled object markers are shown for n trials. Figure titles refer to the variation used on the randomisation approach (see Table D.4). Markers belonging to initial and end perch (navy), walls (light blue), curtain (red), front obstacles (cyan shades) and back obstacles (yellow shades) are shown.



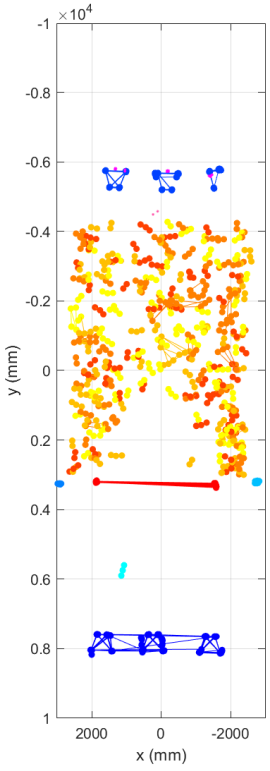
(a) No restriction, $n = 64$



(b) No restriction, $n = 64$

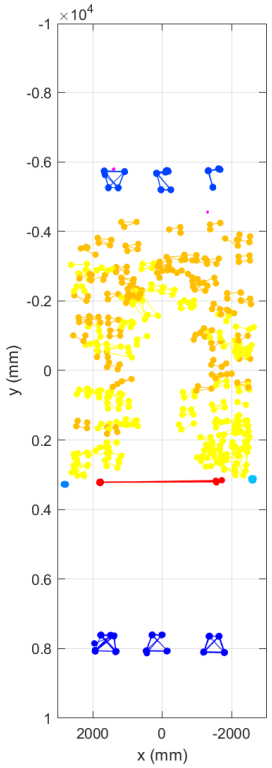


(c) No restriction, $n = 24$

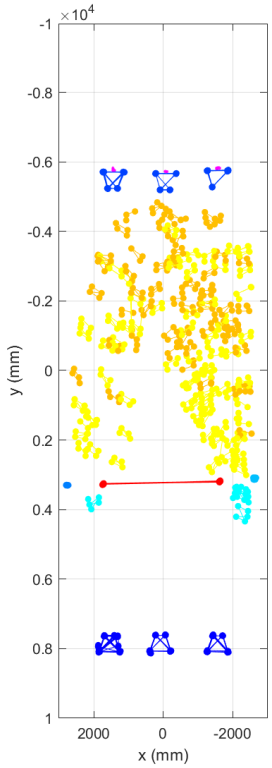


(d) Triangle restriction, $n = 67$

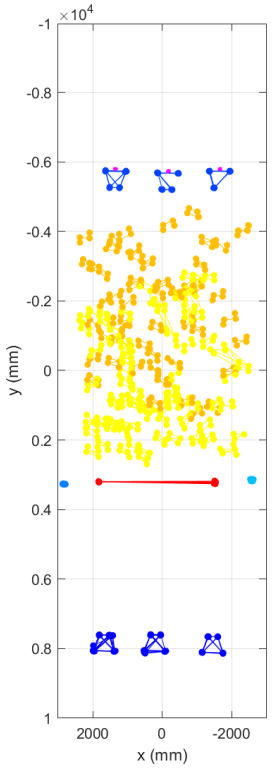
Figure D.23: Area covered by the obstacles in the static goal trials using the randomisation based on grid - Individual pillars. Labelled object markers are shown for n trials. Figure titles refer to the variation used on the randomisation approach (see Table D.5). Markers belonging to initial and end perch (navy), walls (light blue), curtain (red), front obstacles (cyan shades) and back obstacles (yellow shades) are shown.



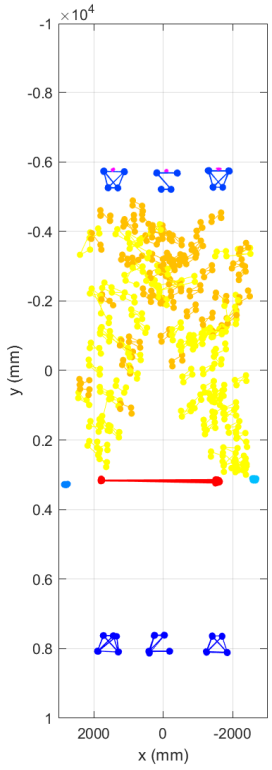
(a) RTR, biased and transversal position 1, $n = 64$



(b) RTR, biased and transversal position 2, $n = 64$

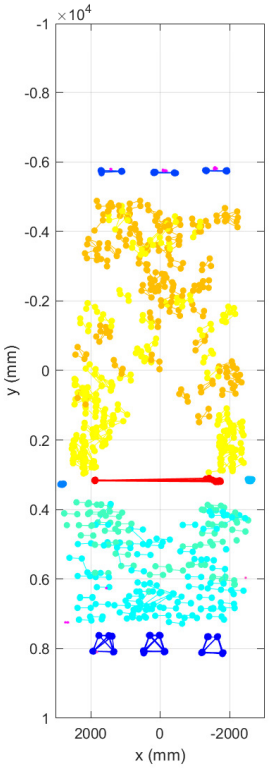


(c) RTR, unbiased and spread, $n = 64$

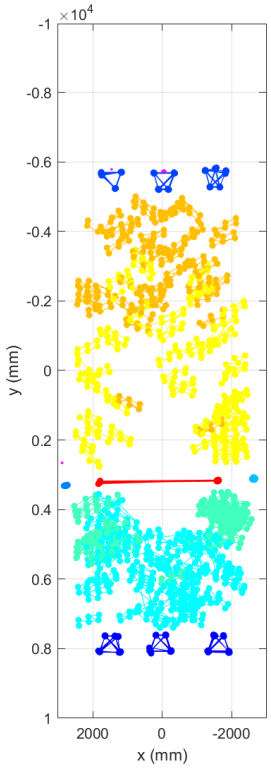


(d) RTR, unbiased, one cell deviation, $n = 65$

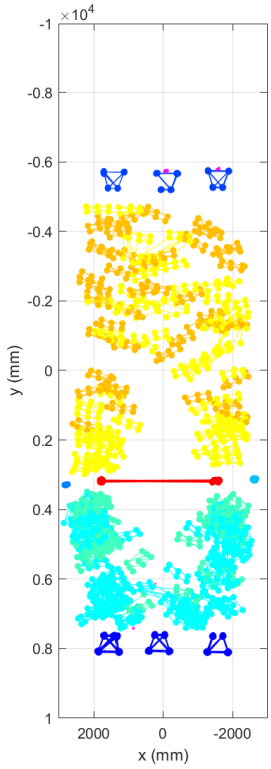
Figure D.24: Area covered by the obstacles in the static goal trials using the randomisation based on grid - Pair of pillars (back only). Labelled object markers are shown for n trials. Figure titles refer to the variation used on the randomisation approach (see Table D.5); 'RTR' stands for *reference trajectory restriction*. Markers belonging to initial and end perch (navy), walls (light blue), curtain (red), front obstacles (cyan shades) and back obstacles (yellow shades) are shown. Pink markers are noise (candidate perch markers). An error in the defined geometry in (c) led to further spread locations for a certain set of trials.



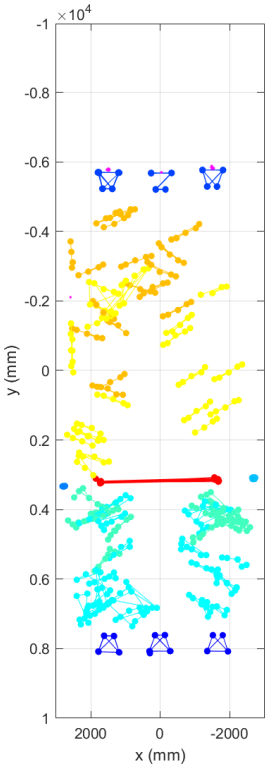
(a) Monoliths (front) and pair of pillars (back), $n = 64$



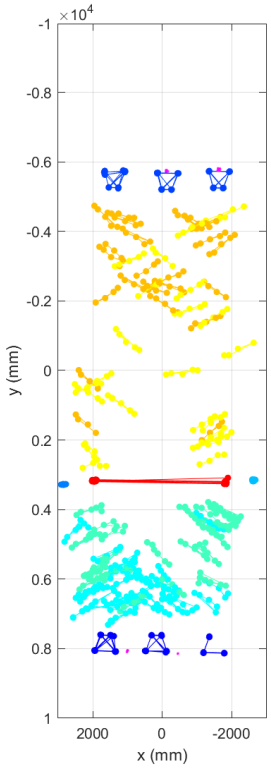
(b) Three pillars, $n = 48$



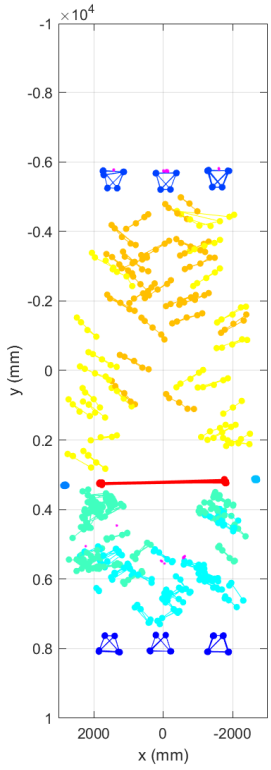
(c) Three pillars, $n = 59$



(d) Three pillars, $n = 24$

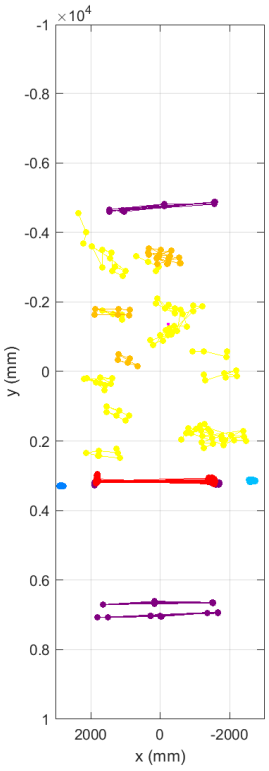


(e) Three pillars, $n = 24$

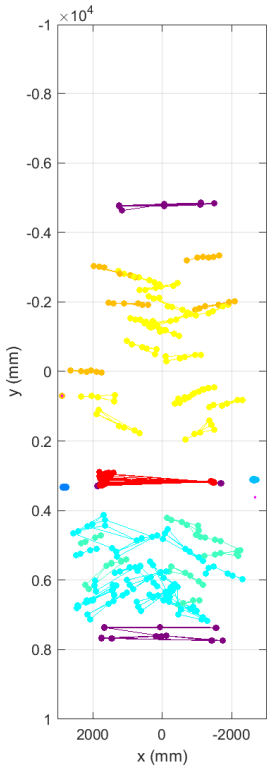


(f) Pair of high pillars (front) and three pillars (back), $n = 25$

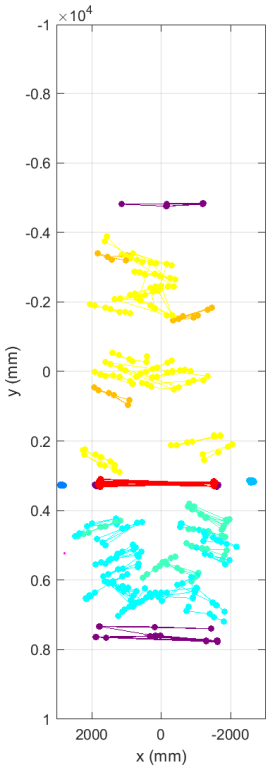
Figure D.25: Area covered by the obstacles in the static goal trials using randomisation based on grid - Obstacles front and back. Labelled object markers are shown for n trials. Figure titles refer to the obstacle configuration used; for the randomisation all use the variation RTR (reference trajectory restriction) - *unbiased, one cell deviation*, except **(c,d)** which use *all trials with obstacles* (see Table D.6). Markers belonging to initial and end perch (navy), walls (light blue), curtain (red), front obstacles (cyan shades) and back obstacles (yellow shades) are shown. Pink markers are noise (candidate perch markers).



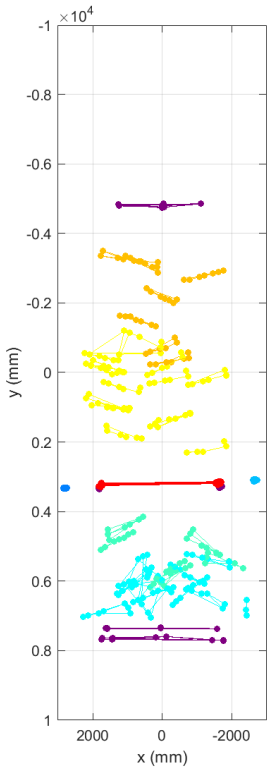
(a) Diamond - gap midpoint, $n = 16$



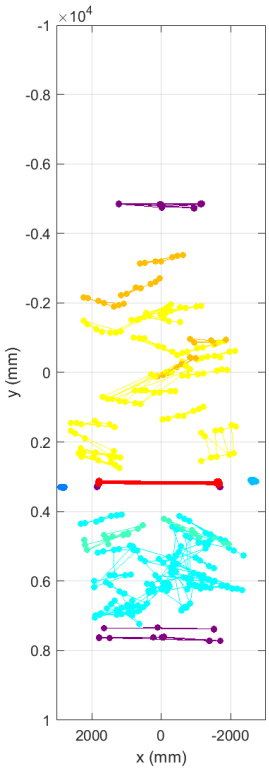
(b) Diamond - varying lure speed, $n = 16$



(c) Diamond - gap pulleys, $n = 16$

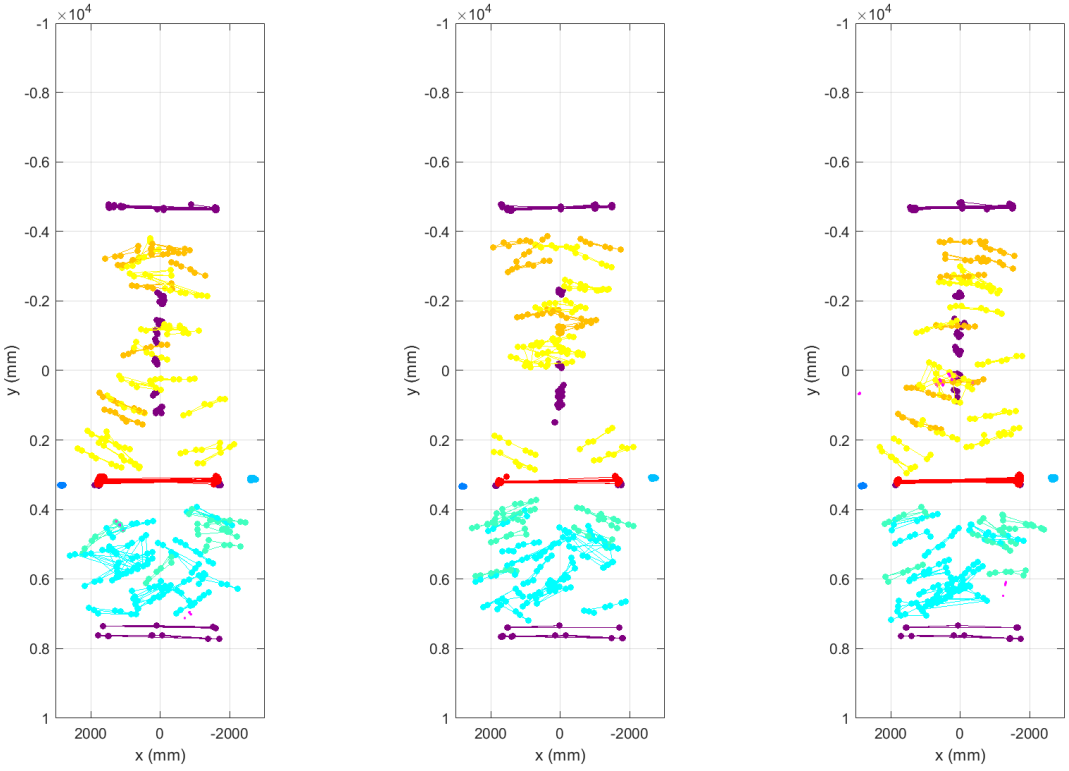


(d) Diamond - capped lure speed, $n = 14$

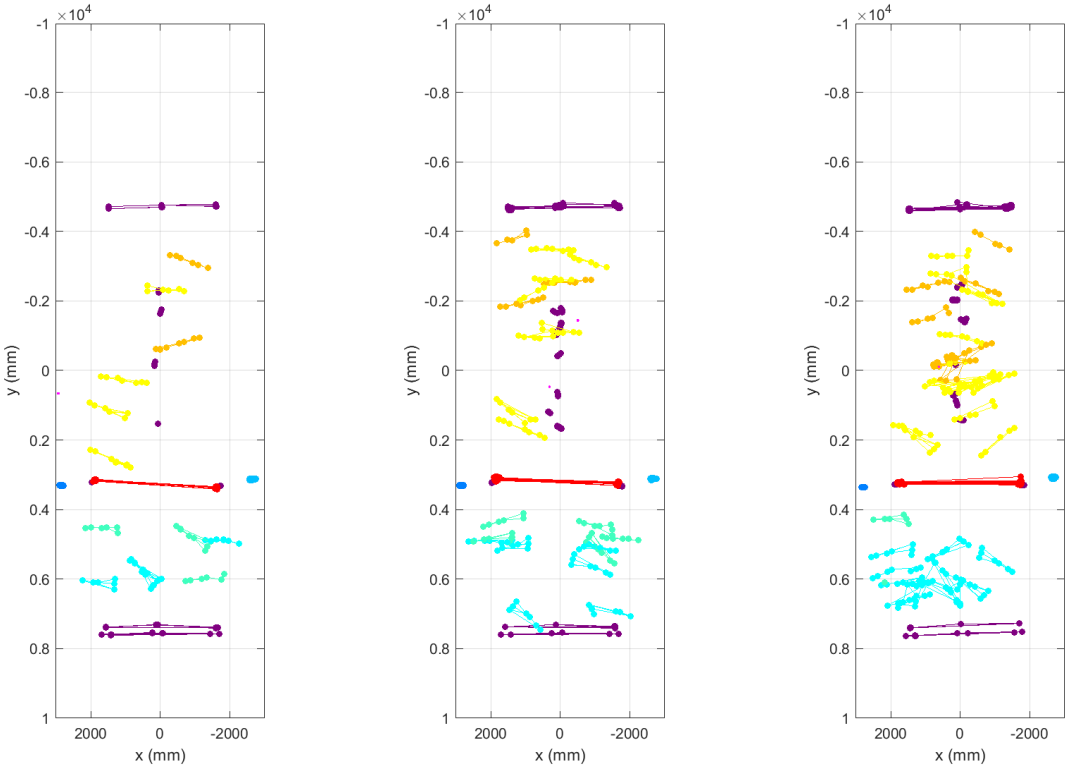


(e) Diamond - capped lure speed, $n = 16$

Figure D.26: Area covered by the obstacles in the moving goal trials - Diamond. Labelled object markers are shown for n trials. Figure titles refer to the variation used on the randomisation approach (see Table D.7). Markers belonging to the pulleys (purple), walls (light blue), curtain (red), front obstacles (cyan shades) and back obstacles (yellow shades) are shown.

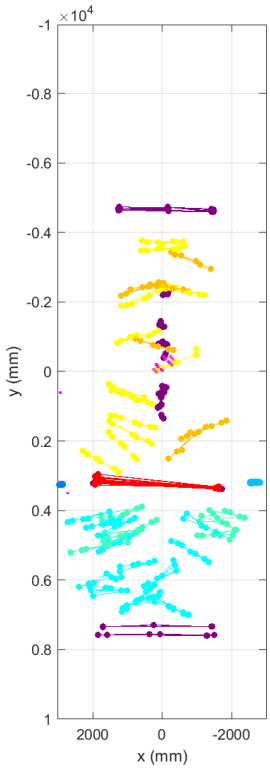


(a) Zigzag, $n = 16$ **(b) Zigzag**, $n = 15$ **(c) Zigzag**, $n = 16$

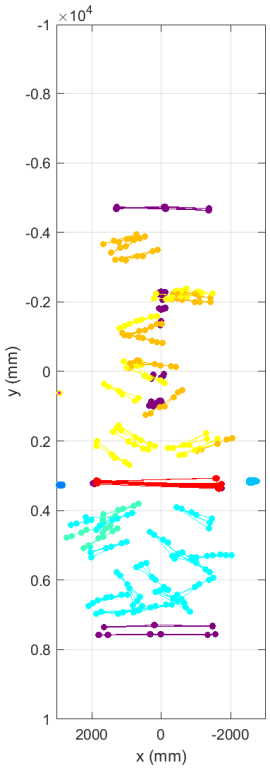


(d) Zigzag, $n = 4$ **(e) Zigzag**, $n = 9$ **(f) Zigzag**, $n = 12$

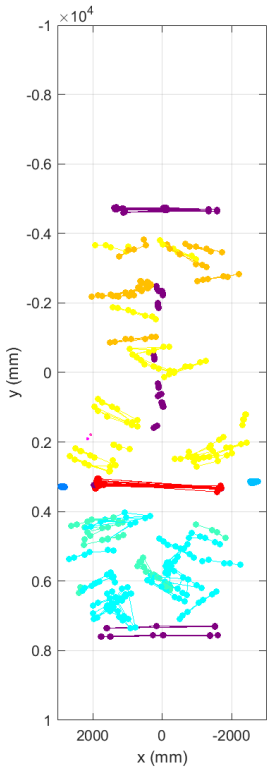
Figure D.27: Area covered by the obstacles in the moving goal trials - Zigzag. Labelled object markers are shown for n trials. Figure titles refer to the variation used on the randomisation approach (see Table D.7). Markers belonging to the pulleys (purple), walls (light blue), curtain (red), front obstacles (cyan shades) and back obstacles (yellow shades) are shown. Pink markers are candidate obstacle markers not clustered as a separate obstacle.



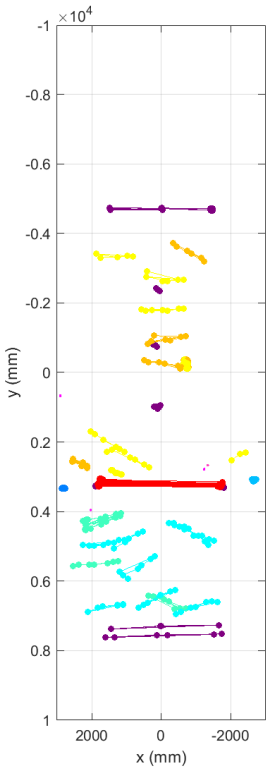
(a) Zigzag, $n = 14$



(b) Zigzag, $n = 14$



(c) Combined, $n = 16$



(d) Combined, $n = 16$

Figure D.28: Area covered by the obstacles in the moving goal trials - Zigzag and combined. Labelled object markers are shown for n trials. Figure titles refer to the variation used on the randomisation approach (see Table D.7). Markers belonging to the pulleys (purple), walls (light blue), curtain (red), front obstacles (cyan shades) and back obstacles (yellow shades) are shown. Pink markers are candidate obstacle markers not clustered as a separate obstacle.

E

An attractor-repeller model for the obstacle avoidance behaviour of Harris' hawks

Contents

E.1 Introduction 291
E.2 Model overview 292
E.3 Steering commands: normal acceleration 292
E.3.1 Scalar field s 293
E.3.2 $K_{s,i}$ and $K_{\phi,i}$ functions 294
E.4 Thrust commands: tangential acceleration 295
References 295

E.1 Introduction

The model presented here was conceived to analyse the data described in Appendix D (specifically, the static goal trials). Therefore, it refers to the setup of those experiments and the elements added to guide or alter the birds' trajectories. These elements were: an initial and an end perch, where the bird was trained to takeoff and land; a curtain, which we transversally across the room, at approximately one third of the total length of the room from the take-off position and leaving two gaps of around 1 m between the curtain edges and the walls; and different obstacles, in the shape of pillars or monoliths.

The goal of the model was to characterise the flight trajectories of Harris' hawks navigating through obstacles. In particular, we wanted to identify which

information about the environment is relevant for planning their manoeuvre through clutter, and when it is acquired.

This obstacle avoidance model is inspired in the work by Fajen et al. (2003) and is based on the attractor-repeller paradigm from robotics. This paradigm considers an agent (in our case, the bird) responding with different steering commands to the elements in the environment (such as a landing perch, an obstacle, the walls, the floor...). The steering commands depend on whether the environment element is classified as an attractor (or an element the bird will aim to fly towards, like the landing perch) or a repeller (or an element the bird will try to avoid, like the obstacles). In our model, steering commands additionally depend on the distance between the bird and the element, and on the angular position of the element relative to the bird's velocity vector. Apart from steering commands, our model also considers acceleration commands in the velocity vector, with simple predefined rules for accelerating, gliding and braking based on observations from the bird trajectories.

The landing perch of the experiments would typically be considered the 'goal' of the bird. We also consider the walls and the floor of the lab as elements the bird must avoid.

E.2 Model overview

The total acceleration on the bird at every timestep is the sum of the acceleration due to aerodynamic forces and the acceleration due to gravity:

$$\vec{a}_{bird} = \vec{a}_{aero} + \vec{g}, \tag{E.1}$$

where $\vec{a}_{aero}$ denotes the acceleration due to aerodynamic forces, and $\vec{g}$ the acceleration due to gravity, defined as $\vec{g} = -g\vec{k}$ with $\vec{k}$ positive in the direction opposite to gravity. We decompose the acceleration due to aerodynamic forces into a component perpendicular to the birds velocity $\vec{a}_n$, and a component parallel to the bird's velocity $\vec{a}_t$:

$$\vec{a}_{aero} = \vec{a}_n + \vec{a}_t \tag{E.2}$$

The normal acceleration $\vec{a}_n$ exclusively changes the direction of the velocity vector, and not its magnitude. Therefore, it can be seen as a steering command. The tangential acceleration $\vec{a}_t$ changes only the magnitude of the velocity vector, and not its direction, so it can be seen as a thrust (forward acceleration) command.

E.3 Steering commands: normal acceleration

The elements of the environment (goal, curtain, walls, obstacles) contribute to the normal acceleration $\vec{a}_n$:

$$\vec{a}_n = \vec{a}_{n,goal} + \sum_{j=1}^{N_{obstacles}} \vec{a}_{n,obstacle_j} + \vec{a}_{n,floor} + \vec{a}_{n,walls} \tag{E.3}$$

If we combine all attractor elements together (in this case, just the goal) and all repeller elements together (each individual obstacle, the curtain, the floor and the walls), we obtain a more compact expression:

$$\vec{a}_n = \sum_{k=1}^{N_{attractors}} \vec{a}_{n,k} + \sum_{j=1}^{N_{repellers}} \vec{a}_{n,j} \tag{E.4}$$

The magnitude of the steering contribution of an element i, $|\vec{a}_{n,i}|$, depends on two factors: the distance from the bird to the element, and the angular position of the element relative to the bird's velocity. This way, elements that are closer to the bird impose a larger steering command than those that are further, and elements most aligned with the bird's current trajectory will also have more effect on the bird's steering. The direction of the steering command of the ith element, $\vec{a}_{n,i}$, is a function of a desired steering direction $\vec{d}_i$, defined for each element, and which captures the repulsion or attraction behaviour of the element.

Mathematically, we define the steering command from the ith element as:

$$\vec{a}_{n,i} = -K_{s,i}K_{\phi,i}\vec{V} \times \frac{\vec{V} \times \vec{d}_i}{|\vec{V} \times \vec{d}_i|} \tag{E.5}$$

where $\vec{V}$ is the bird's velocity vector and $\vec{d}_i$ is the desired steering direction for the element i. K_s, i is a scalar function that depends on a distance-related metric s to the element i. K_ϕ, i is a scalar function that depends on the angle ϕ between the bird's velocity vector $\vec{V}$ and the desired direction vector $\vec{d}_i$. Except for the bird's velocity $\vec{V}$, all the other terms are a function of the distance-related metric s. We explain in further detail the terms in Equation E.5 in the following sections.

E.3.1 Scalar field s

The distance-related metric s is a scalar field defined for each element i of the environment. It captures the interaction of the ith element with the bird. We choose s to represent a distance metric to the element i. For example, for the curtain, a scalar field is used with s equal to the sum of distances to the curtain edges. The desired direction vector $\vec{d}$, the scalar function K_s and the scalar function K_ϕ are all functions of the scalar field s.

The isosurfaces of these scalar fields are the collection of points for which the scalar field s has a certain constant value. For example, for the curtain case, ellipses parallel to the floor plane with foci on the curtain edges have a constant s value. Similarly for the other elements of the environment, the isosurfaces are also associated to the element's shape.

The gradient of the scalar field s, ∇s, is a vector field whose vectors are always perpendicular to the isosurfaces of s, and always point in the direction of steepest increase. Because of the way the scalar field s is defined, the direction of the gradient vector ∇s can be seen as a direction of attraction or repulsion, to or

from the element; this idea is what guides the definition of a scalar fields s for each element in the first place.

The vector $\vec{d}$ is defined always parallel or antiparallel to ∇s, depending on whether the element is an attractor or repeller. The double cross product term between $\vec{V}$ and $\vec{d}$ in Equation E.5 results in the steering command being always perpendicular to $\vec{V}$ but as aligned as possible to the desired steering direction $\vec{d}$.

E.3.2 $K_{s,i}$ and $K_{\phi,i}$ functions

The $K_{s,i}$ and K_{ϕ}, i functions also depend on the scalar field s of the element. The $K_{s,i}$ function is an exponential function of s. It captures the decay of the steering command of an element as its distance to the bird (or more precisely, its value of s) increases. We define $K_{s,i}$ as a sum of two exponentials, each capturing the long range and the short range effect:

$$K_{s,i} = K_{SR,i} e^{-\frac{s}{c_{SR,i}}} + K_{LR,i} e^{-\frac{s}{c_{LR,i}}}. \tag{E.6}$$

The $c_{SR,i}$ and $c_{LR,i}$ parameters represent the area of influence for the short range and long range terms, respectively; these can be interpreted as the distance at which an element becomes relevant for the bird. The $K_{SR,i}$ and $K_{LR,i}$ parameters represent the magnitude of the response in the short and the long range regions, respectively.

The $K_{\phi,i}$ function captures the effect on the steering command of the angular position of an element i relative to the bird's velocity vector $\vec{V}$. We use different expressions for $K_{\phi,i}$, depending on whether the element i is attractor or a repeller.

For an attractor element, we define it as

$$K_{\phi,i} = \sin(\frac{\phi_i}{2}), \tag{E.7}$$

with ϕ_i denoting the angle between the bird velocity $\vec{V}$ and the desired steering direction of the element, $\vec{d_i}$. The variable ϕ_i is defined from 0° to 180°, so that $K_{\phi,i}$ increases as ϕ_i increases. Note that $K_{\phi,i}$ is maximal when the bird is at $\phi_i = 180°$ from the desired direction. For a repeller we use the following expression:

$$K_{\phi,i} = \frac{\sin(\frac{\phi_i}{2})}{1 + e^{\frac{\cos\phi_i - \cos\phi_{th}}{c_\phi}}}, \tag{E.8}$$

The numerator of $K_{\phi,i}$ for repellers is the same as the $K_{\phi,i}$ expression for attractor elements (see Equation E.7). This makes sense since we would also like the bird to try to align with the desired direction $\vec{d_i}$. However, the repeller's expression includes an attenuation factor in the form of a sigmoid. This sigmoid is a function of the cosine of ϕ. It effectively attenuates the steering command of the element if its located at an angle ϕ_i beyond a certain threshold relative to the velocity vector $\vec{V}$. The threshold angle for attenuation is captured by the parameter ϕ_{th}. In this way ϕ_{th} can be interpreted as a measure of the relevant field of view for the bird. The parameter c_ϕ relates to the extent of the linear region of the sigmoid.

E.4 Thrust commands: tangential acceleration

The tangential acceleration captures the change in the bird's speed. From the empirical bird trajectories we observed the bird increases its mechanical energy approximately until the end perch becomes visible (when turning around the curtain). From then on, the bird dissipates its mechanical energy until landing, reaching the end perch at a fairly repeatable landing speed. With this in mind, we modelled the tangential acceleration as a piecewise function, with the visibility of the end perch as a decision parameter .

At the start of the flight, while the end perch is not visible, the bird accelerates. We model the tangential acceleration at this phase as an acceleration at constant power, parallel to the bird's velocity $\vec{V}$. This is a reasonable assumption for the bird's power strategy.

When the end perch becomes visible, the bird starts decelerating until it lands on it. We model the tangential acceleration at this phase as a constant acceleration, applied in the direction opposite to the velocity vector $\vec{V}$. This constant acceleration can be interpreted as a constant dissipative force (or aerodynamic drag) per unit mass. From the bird trajectories, we find it varies in the decelerating phase. We simplified the problem by distinguishing two regimes within the decelerating phase:

- At the start of the decelerating phase, the bird flies at approximately constant speed and constant dissipative force. We found the value of the force in this phase matched the wind tunnel measurements for the drag of a gliding Harris' hawk, flying at similar speed (Tucker et al., 1990). We call this first phase 'gliding phase' and use the wind tunnel drag measurements to derive the constant acceleration value for our model.

- After this gliding phase, we observed in the bird trajectories that the dissipative force increases sharply, before reducing again upon landing. We simplified this phenomena in our model by considering a different, but higher value for the constant acceleration when the bird is in the 'braking phase'.

We model when the decelerating bird should switch from gliding to braking based on a simple rule. At every timestep in the decelerating phase, we estimate the mechanical energy the bird needs to dissipate from its current state to reach the end perch at a certain landing speed. We additionally estimate the mechanical energy the bird would dissipate if decelerating in 'braking configuration' until the end perch (assuming a straight line trajectory). If the mechanical energy the bird needs to dissipate is larger than the estimated dissipated energy in braking configuration, then the bird will start braking. Otherwise, the bird will continue gliding.

References

Fajen, B. R. and Warren, W. H. (2003). "Behavioral Dynamics of Steering, Obstacle Avoidance, and Route Selection". In: *J. Exp. Psychol. Hum. Percept. Perform.* 29.2, pp. 343–362. DOI: 10.1037/0096-1523.29.2.343.

Tucker, V. A. and Heine, C. (1990). “Aerodynamics of gliding flight in a harris’ hawk, Parabuteo unicinctus”. In: *J. Exp. Biol.* 149, pp. 469–489.

Through hawks' eyes: synthetically reconstructing the visual field of a bird in flight

Sofía Miñano[1], Stuart Golodetz[2], Tommaso Cavallari[†] and Graham K. Taylor[1*]

[1]Department of Zoology, University of Oxford, 11a Mansfield Road, Oxford, OX1 3SZ, United Kingdom.

[2]Department of Computer Science, University of Oxford, Parks Road, Oxford, OX1 3QD, United Kingdom.

*Corresponding author(s). E-mail(s): graham.taylor@zoo.ox.ac.uk;
Contributing authors: sofia.minanogonzalez@biodtp.ox.ac.uk; stuart.golodetz@cs.ox.ac.uk; tommaso@tommasocavallari.com;

Abstract

Birds of prey rely on vision to execute flight manoeuvres that are key to their survival, such as intercepting fast-moving targets or navigating through clutter. A better understanding of the role played by vision during these manoeuvres is not only relevant within the field of animal behaviour, but could also have applications in autonomous drones. Here, we present a novel method that uses computer vision tools to analyse the role of active vision in bird flight, and demonstrate its use to answer behavioural questions. Combining motion capture data from Harris' hawks with a hybrid 3D model of the environment, we render RGB images, semantic maps, depth information, and optic flow outputs that characterise the visual experience of the bird in flight. Contrary to previous approaches, our method allows us to consider different camera models and alternative gaze strategies for the purposes of hypothesis testing, allows us to consider visual input over the complete visual field of the bird, and is not limited by the technical specifications and performance of a head-mounted camera light enough to attach to a bird's head in flight. We present pilot data from three sample flights: a pursuit flight, in which a hawk intercepts a moving target, and two obstacle avoidance flights. With this approach, we provide a reproducible method that facilitates the collection of large volumes of data across many individuals, opening up new avenues for data-driven models of animal behaviour.

Keywords: animal vision, visual field, reconstruction, bird, flight

Statements and Declarations

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 682501). SM's work was supported by funding from the Biotechnology and Biological Sciences Research Council (BBSRC) [grant number BB/M011224/1], via the Intersdisciplinary Bioscience Doctoral Training Partnership. SG's work was supported by Amazon Web Services as part of the Oxford-Singapore Human-Machine Collaboration Programme. This work has received approval

[†] Work carried out while at the Department of Engineering Science, University of Oxford.

from the Animal Welfare and Ethical Review Board of the Department of Zoology, University of Oxford, in accordance with University policy on the use of protected animals for scientific research, permit no. APA/1/5/ZOO/NASPA. The authors have no financial or proprietary interests in any material discussed in this article. The code to generate the model of the lab environment and define the corresponding virtual camera in Blender, and the datasets analysed during the current study will be made available upon publication.

1 Introduction

From intercepting moving targets to manoeuvring through clutter, birds use vision to coordinate flight manoeuvres with an agility and flexibility beyond the reach of current autonomous systems. Nevertheless, the links between their vision, guidance, and control are complex and poorly understood. Much more is known about the role of vision in insect flight (Taylor et al, 2008), because the size and sentience of birds complicates experimental identification of the input-output relationships characterising their visuomotor control (Altshuler and Srinivasan, 2018). To control the visual experience of a bird, previous works (Bhagavatula et al, 2011; Schiffner and Srinivasan, 2015; Dakin et al, 2016; Ros and Biewener, 2016) have investigated avian flight behaviour in abstract visual environments, such as corridors with vertically or horizontally striped walls (see Figure 1b). This has proven a useful first step in exploring how birds use visual self-motion cues, and in isolating its effects on flight control. For example, budgerigars flying through narrow corridors regulate flight speed in response to optic flow from sliding gratings projected onto the walls (Schiffner and Srinivasan, 2015). These methods are inherited from similar studies of insects (Baird et al, 2021; Tammero and Dickinson, 2002b,a; Altshuler and Srinivasan, 2018), so may oversimplify the rich visual input available to birds with their high visual acuity and complex neural organization (Altshuler and Srinivasan, 2018). Such abstract visual inputs are far from those experienced naturally in flight, so caution is warranted if using them to draw conclusions about the birds' strategies in the wild.

Here, we present a method for reconstructing the scene contained within the visual field of a bird flying through a structured environment, as a first step towards understanding how birds use visual information to guide and control their flight. Our synthetic reconstruction method is designed to simulate the physical stimulus to a bird's optomotor system, and avoids the main practical difficulties associated with making invasive physiological measurements of its response. In developing our method, we aim to support the analysis of large quantities of data across multiple individuals, using environments that may vary experimentally across trials. To achieve this, we focus on developing a non-invasive method for tracking the animal's self-motion, and on methods to reconstruct the geometry of the environment that allow us to capture different levels of complexity. Specifically, we combine high-speed motion capture data with a three-dimensional (3D) reconstruction of the laboratory environment, which we use to generate synthetic visual inputs modelling those that may be available to the bird in flight. These visual inputs characterise the bird's visual experience of its own self-motion in detail, opening up first several new avenues of research in bio-inspired computer vision and behavioural modelling.

In summary, we present a computer vision approach to understanding the visuomotor guidance and control of bird flight, and demonstrate the possibilities of using the data that we generate to answer mechanistic questions in behaviour. We present pilot data from three sample flights: one pursuit flight, in which a hawk intercepts a moving target pulled across the ground, and two obstacle avoidance flights, in which a hawk flies between two perches around a set of obstacles placed at varying distances from the take-off perch.

1.1 Active vision in birds

Bird vision is an active process, so understanding how birds interrogate their visual environment is key to unravelling the cognitive processes coordinating their flight manoeuvres. Birds mainly use head movements controlled by their neck motor system to look around the environment. This is because their eyes have a limited range of motion within the orbit, such that the largest eye movements driven by the oculomotor system are small compared to those made by the head (Yorzinski

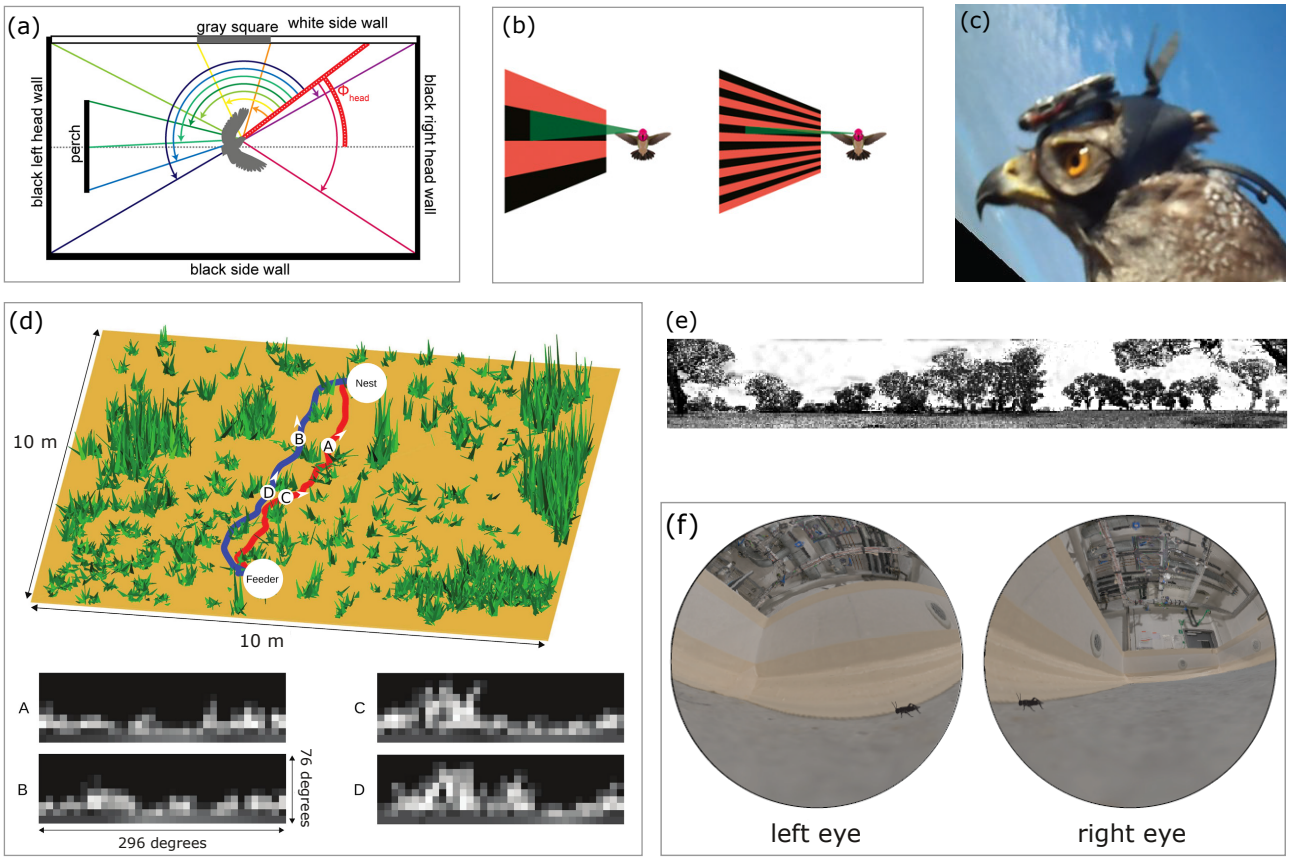


Fig. 1 Related work. (a) Overhead view of a lovebird turning in flight; the alignment of the forward direction of its head is compared against several features of the lab environment (Kress et al, 2015). (b) Hummingbirds flying along striped corridors (Dakin et al, 2016). (c) Northern goshawk with head-mounted camera (Kane et al, 2015). (d) Reconstruction of the environment around a nest of homing ants, and reconstructed views (A,B,C,D) at different instances of their recorded paths (Ardin et al, 2016). (e) Reconstruction of an ant's field of view, combining laser-scanned data and camera images (Stuerzl et al, 2015). (f) Reconstructed view from each eye of a mouse hunting a cricket, using laser-scanned data and texture from high-resolution images (Holmgren et al, 2021)

et al, 2015; Mitkus et al, 2018). The head movements of birds in flight display a characteristic saccade-and-fixate strategy recalling primate eye movements (Eckmeier et al, 2008; Ozawa, 2010; Kress et al, 2015; Ros and Biewener, 2017). This strategy supports the use of optic flow in guidance and control, because the rotational component of the optic flow field is eliminated during fixation, leaving only the translational component that containing depth information (Eckmeier et al, 2008). A small body of previous work has assessed how head movements are affected by the visual environment experienced in flight. For example, the frequency and amplitude of head saccades has been shown to vary in homing pigeons with the structure of the landscape they are flying through (Kano et al, 2018), the structure of the clutter they are negotiating (Ros and Biewener, 2016), and the presence of another individual when they are flying in pairs (Taylor et al, 2019).

Although head pose is the primary determinant of a bird's gaze direction, it can be measured most directly using eye-tracking cameras. These have so far been used to identify where birds look when assessing mates (Yorzinski et al, 2013), watching predators (Yorzinski and Platt, 2014; Yorzinski, 2021), or inspecting the environment (Yorzinski et al, 2015). Weight limitations have so far restricted this technology to terrestrial use cases, but whilst miniaturisation may be possible, eye-tracking cameras will always be challenging to use in flight, owing to the difficulty of keeping the camera steady without occluding the frontal field of view. Most studies therefore take head orientation as a proxy for gaze direction, which is

reasonable provided that the eyes move little relative to the head in flight, and is often sufficient to identify the features to which a bird is attending. For example, work on lovebirds flying to a perch (Kress et al, 2015) investigated the alignment of the bird's head with the edges of the perch and flight arena (see Figure 1a), whilst work on pigeons negotiating a forest of vertical poles (Lin et al, 2014) investigated the alignment of the bird's head with the obstacles it had to traverse. Both studies analysed the problem in two dimensions, however, focusing only on changes in head azimuth, and reducing these extended features of the environment to single points in the visual field. A complete understanding of the problem requires fuller knowledge of what a bird sees in flight, which is what motivates the visual reconstruction method of the present study.

1.2 Related work

Here we review the main approaches that have been used to reconstruct what animals see of their environment, as broader context for the contribution of the method that we present. These amount either to recording video using animal-borne cameras, or to reconstructing images synthetically using specialised hardware or rendering software, which is the approach that we adopt here.

1.2.1 Head-mounted cameras

Head-mounted video cameras can be used to sample the view that a bird experiences as a result of its self-motion through the environment (Figure 1c), and have been used so to analyse aerial attack behaviours in hawks and falcons (Kane et al, 2015; Kane and Zamani, 2014; Ochs et al, 2016). This approach can be especially informative regarding a bird's behaviour in its natural habitat, subject to the extreme limitations of pixel count, dynamic range, and field of view of any camera small enough to mount on the head. Payload is conventionally limited to $\leq 5\%$ of a bird's body mass on welfare grounds (Fair et al, 2010), but much more stringent limits may be required to ensure natural behaviour if the load is carried on the head (Kane and Zamani, 2014). The 20 g cameras that have been used previously (Kane et al, 2015; Kane and Zamani, 2014) are twice the weight of many small birds, so are only suitable for studying very large species such as raptors. Even so, it is simply not possible for a small camera to cover a bird's full field of view at an appropriate optical or sampling resolution. For example, the vertical field of view (31°) of the head-mounted cameras used to study Harris' hawks (Kane et al, 2015) compares unfavourably with the vertical extent of the binocular overlap (100°) in this species (Potier et al, 2016), and their 30 Hz frame rate is well below these birds' 81 Hz flicker fusion frequency (Potier et al, 2020). The possibilities for analysing head-mounted video data are further impacted by their low frame rates, and the motion blur associated with low shutter speeds and rolling shutters (Kane and Zamani, 2014). Finally, head-mounted cameras are prone to wobble unless surgically attached to the head (Lev-Ari and Gutfreund, 2018; Hazan et al, 2015), which is an undesirable intervention. Head-mounted video cameras therefore have less utility for studying visually-guided flight in birds than might first be imagined.

1.2.2 Synthetic reconstruction

The only obvious alternative to recording the scene using head-mounted cameras is to reconstruct it using specialist hardware (Stuerzl et al, 2010; Zhu et al, 2020) or rendering software (Holmgren et al, 2021). Rendering is attractive in offering complete control of the detail presented over the visual field (Holmgren et al, 2021), and is particularly well suited to scientific inference, because of the possibility of defining alternative views for hypothesis testing. State-of-the-art research on image-based rendering points to the possibility of reconstructing novel views given only a set of camera images with known pose (Tancik et al, 2022; Mildenhall et al, 2020), but current rendering applications require an explicit 3D model of the environment. Standard modelling approaches such as simultaneous localization and mapping (SLAM) suffer from accumulating noise and drift when covering large areas (Schonberger and Frahm, 2016), and automating the post-processing of the resulting meshes can be challenging (Risse et al, 2018; Stuerzl et al, 2015). Nevertheless, the quality of the dense maps that can now be captured with consumer-level handheld devices has improved greatly in the past few years, achieving results comparable to more expensive laser-scanning methods even in complex

natural environments (Tatsumi et al, 2022; Gollob et al, 2021).

Most synthetic reconstructions to date have focused on insects. Several studies have investigated ant navigation (Zeil et al, 2014; Ardin et al, 2015, 2016) using fully synthetic models of the natural environment (Figure 1d), and a basic geometric reconstruction of the laboratory environment has been used to investigate the role of optic flow in bee flight (Ravi et al, 2019, 2022). More detailed 3D models of the natural environment obtained using laser scanners, structure-from-motion (SfM), and photographic reconstruction techniques (Figure 1e) have been used to study homing flight behaviour in bees and wasps (Stuerzl et al, 2015, 2016; Schulte et al, 2019). This focus on insects reflects the fact that they are simpler to model than vertebrates, owing to their lower sampling resolution, and the fact that their eyes are rigidly fixed with respect to their head, but the same general approaches have more recently been extended to vertebrates. In particular, by tracking the animal's head and eye movements together, and by combining laser scanning and photographic texture to produce a high-resolution 3D reconstruction of its lab environment, one recent study was able to analyse the optic flow experienced by mice when pursuing prey within a 1 m^2 arena (Holmgren et al, 2021). The growing interest in reconstructing an animal's experience of its environment at the much larger spatial scales relevant to ecology and conservation (Tuia et al, 2022) has also led to demonstrations of mapping sensors worn by a dog (McClune, 2018), and a method for embedding an animal's track within an aerial view of its environment obtained by a mobile camera used for tracking (Haalck et al, 2020).

Fully synthetic reconstruction techniques have been used to study the effectiveness of movement displays of model lizards in naturalistic scenes under different lighting scenarios and wind conditions, assuming a realistic model of the lizard motion vision system (Bian et al, 2018, 2019, 2021). These are perhaps the only studies to date in which the environment has been rendered at a level of precision comparable to that experienced by the well-developed visual system of a bird in flight, albeit that they are rendered from the perspective of a stationary observer. Such fully synthetic renderings also lend themselves to being presented to an animal moving in a virtual reality (VR) environment. This approach has been used in tethered and free-flight studies of insect flight for some time (Kern et al, 2005; Taylor et al, 2008; Windsor and Taylor, 2017), but impressive VR environments have more recently been developed for freely-moving mice, flies, and zebra fish (Stowers et al, 2017), which enable stimulus presentation under open or closed loop experimental conditions (Naik et al, 2020). Such environments are currently limited to volumes of approximately 1 m^3, so have yet to find use for larger animals making larger-scale movements.

A key advantage of synthetic reconstruction is that it can be used to study the appearance of the visual environment in the context of an accurate model of the animal's visual system covering its entire field of view. This is important in capturing detail at spatiotemporal frequencies appropriate to the optical and sampling resolution of the eyes (Holmgren et al, 2021; Neumann, 2002), and the response properties of the motion vision system (Windsor and Taylor, 2017; Bian et al, 2018). Even so, rendering an RGB image that looks realistic to a human eye is not the same thing as reconstructing the stimulus that another animal's visual system receives, owing to variation in the number of photoreceptor types and their respective spectral response curves. For instance, whereas most humans are trichromatic, most birds are tetrachromatic. These details may not matter in the context of motion vision, however, because changes in luminance are always sufficient for motion detection, notwithstanding the possible importance of chromatic motion vision cues in humans (Ruppertsberg et al, 2003). Synthetic reconstruction therefore has considerable potential as an approach to studying visually-guided behaviours of all kinds.

1.3 Contribution

Here, we combine high-speed motion capture with 3D modelling of the laboratory environment, to render an approximation of the visual stimuli experienced by a bird in flight. By defining a virtual camera that replicates the bird's head pose within a 3D model of its environment, we reconstruct RGB renderings, semantic maps, depth maps, and optic flow over the complete visual field of the bird (see Figure 2). We make no attempt

to model the spectral response of the bird's visual system, although it would in principle be possible to use this to analyse the information contained in the RGB renderings given knowledge of the ambient lighting conditions. Analogous data have been generated at much lower spatial resolution for insects (Ravi et al, 2022, 2019; Schulte et al, 2019; Stuerzl et al, 2016, 2015), but this is the first time, to our knowledge, that such data have been produced for a bird in flight.

The method that we present allows for the possibility of defining a hybrid model of the lab environment, combining basic geometric primitives with dense 3D models. A representation based on geometric primitives alone may be sufficient to investigate how the bird's gaze strategy relates to the main elements in the scene (Minano and Taylor, 2021). However, if elements of complex shape or texture are present, then modelling them using geometric primitives may be out of the question, which limits the scope for high-throughput analysis. We address this by using a SLAM application on a consumer-level handheld mapping device to capture the more complex elements of the scene, which we combine with a more basic geometric representation of the environment. In this way, we limit the issues of drift and noise that are typical of the large, dense 3D maps produced by SLAM or SfM (Schonberger and Frahm, 2016), whilst also minimising the modelling effort.

We further facilitate the integration of the captured meshes with the motion capture data by transforming the mesh at acquisition time to the motion capture coordinate system using an ArUco fiducial marker (Romero-Ramirez et al, 2018; Garrido-Jurado et al, 2014). The novelty of our method therefore lies in its fusion of several different computer vision techniques, and in their application to the experimental and computational study of animal behaviour. Our aim is to provide a reproducible method that facilitates the collection and creation of large volumes of data across many individuals, with a view to future work using behavioural cloning and inverse reinforcement learning.

2 Methods

Here we describe the key details of our experimental method and reconstruction method. Further supporting detail on both aspects may be found in the Appendices.

2.1 Experimental method

We recorded several hundred flights from $n = 5$ Harris' hawks (*Parabuteo unicinctus*) flying in a large ($20 \times 6 \times 3.3$ m) motion capture lab, using 22 Vicon Vantage V16 motion capture cameras sampling at 200 Hz and a set of 4 Vicon Vue reference video cameras (Vicon Motion Systems Ltd., Oxford, UK). As our aim is only to describe and illustrate the reconstruction method, a complete description of these experiments will be provided separately elsewhere. Here we report on the results of three flights from two birds. One of these flights involved pursuit of a dummy target with food reward that was pulled along an erratic course; the others represent the two discrete legs of a back-and-forth flight involving obstacle avoidance and perching. Results from the two obstacle avoidance flights have also been presented in the preprint by Minano and Taylor (2021), using a slightly different analysis approach. Further details on the experimental procedure for these three flights are included in Appendix A.

2.1.1 Flight trials

For the pursuit trials, the bird chased a cylindrical artificial target with food reward (length: 0.15 m; diameter: 0.025 m) that was dragged in an unpredictable direction around a series of pulleys at an average speed of 5.6 m s^{-1}. To further challenge the bird's manoeuvring, we hung a black curtain across the room from floor to ceiling, leaving a gap of approximately one wingspan ($1.0 - 1.1$ m) to either side, through which the bird and target passed. For the obstacle avoidance flights, the bird flew back and forth between two perches set 9 m apart, with a set of four cylindrical styrofoam pillars (height: 2 m; diamater: 0.3 m) placed 1.5 m in front of one the perches. The walls of the laboratory environment were hung with camouflage netting, and other reconstruction cases that we trialled included placing small trees within this environment.

2.1.2 Motion capture

We tracked the bird's head, using a custom 'headpack' comprising a rigid arrangement of four

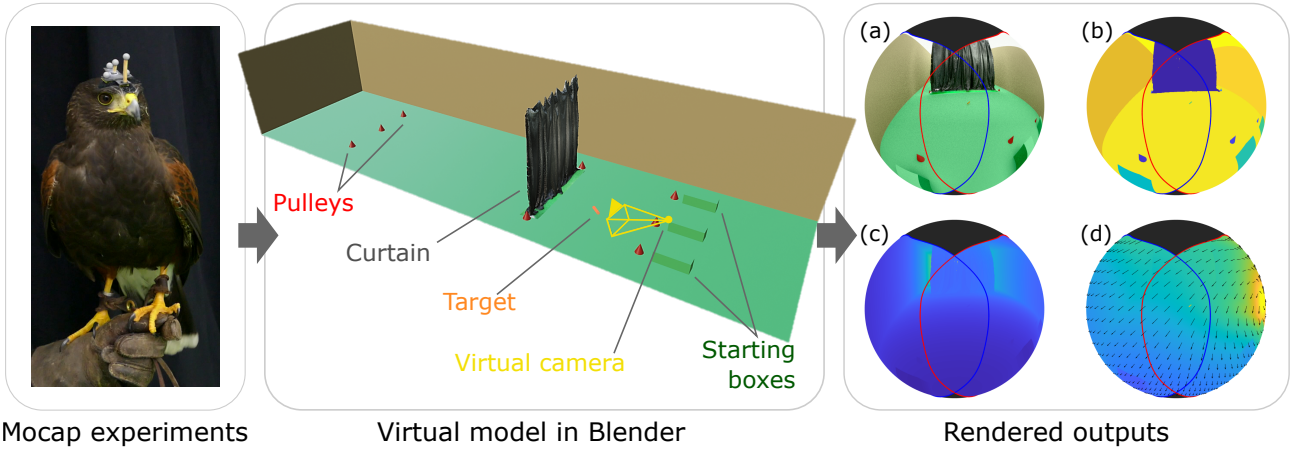


Fig. 2 Summary of method for reconstructing the visual information contained within the visual field of a bird in flight. The head movements of birds executing flight manoeuvres are recorded in a large motion capture lab (left panel). The lab environment is modelled in Blender, using geometric primitives and a dense 3D map to model objects with a more complex geometry (centre panel). The measured head pose is then used to define a virtual camera that is representative of the bird's visual field (right panel). With this approach we can generate detailed information describing the visual scene that the bird experiences in flight, including: (a) basic RGB renderings; (b) semantic maps; (c) depth maps; or (d) optic flow. In the virtual model of the lab shown (centre panel), the pulleys and target are displayed at twice their actual size for clarity. The spheres representing the bird's visual field in the right panel show the animal's view as it flies through the lab. The retinal margins of the left eye (blue line) and right eye (red line) are shown for reference, as well as the blind area caudal to the area sampled by the retina (black fill). Note that because the spheres are represented using an orthographic projection, not all the visual field of the bird is visible

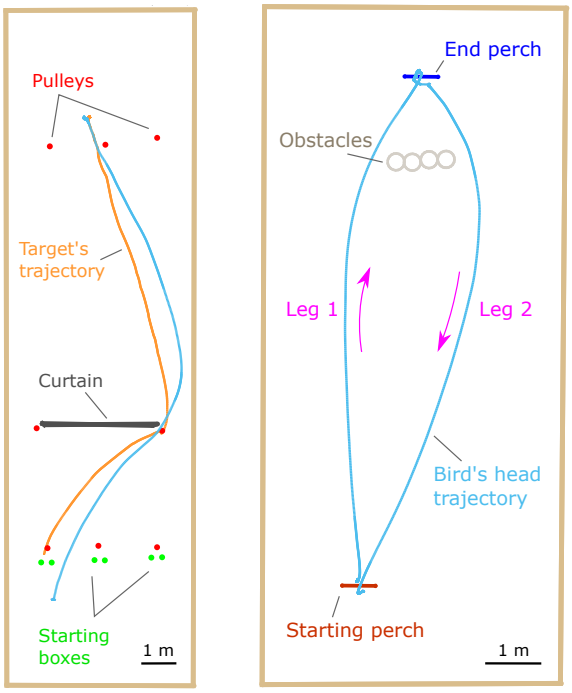
or five 6.4 mm diameter spherical retroreflective markers that we fixed to a Velcro patch glued to the bird's head. We tracked the target using three 6.4 mm diameter markers, and attached further 6.4 mm or 14 mm diameter markers to the main static elements of the scene, to allow us to locate these within the coordinate system of the motion capture measurements. We used Nexus v2.8.0 software (Vicon Motion Systems Ltd., Oxford, UK) to estimate the 3D positions of all the unlabelled retroreflective markers. For the pursuit flight, we labelled the headpack markers manually within Nexus, using its semi-automatic labelling functionality. For the obstacle avoidance flights, we labelled the markers automatically using custom scripts written in MATLAB R2020b (The Mathworks Inc., Natick, MA). These were designed to label stationary obstacle markers, to compute and interpolate the pose of the headpack and target, and to handle missing marker data. Further detail on these post-processing steps is presented in Appendix B.

2.1.3 Ethics statement

This work has received approval from the Animal Welfare and Ethical Review Board of the Department of Zoology, University of Oxford, in accordance with University policy on the use of protected animals for scientific research, permit no. APA/1/5/ZOO/NASPA, and is considered not to pose any significant risk of causing pain, suffering, damage or lasting harm to the animals. No adverse effects were noted during the trials.

2.2 Simulated visual input

We modelled the bird's view of the scene using Blender 2.93.5 (Blender Online Community, 2021) by defining a virtual camera whose field of view was matched that of the bird's eyes, and whose pose was matched to that of the bird's head. To relate the headpack coordinate system to the bird's visual coordinate system, we make three assumptions (Minano and Taylor, 2021). First, we assume that the bird's eye movements are negligibly small compared to those of its head (Kano et al, 2018; Brighton et al, 2017; Ros and Biewener, 2017; Kress et al, 2015; Kane and Zamani, 2014; Eckmeier et al, 2008). Second, we assume that the bird's gaze direction is known from first principles during calibration. In the pursuit case, we assume that the bird looks at food presented to it by the falconer; in the obstacle



(a) Pursuit trial (b) Obstacle avoidance trial

Fig. 3 Pursuit and obstacle avoidance flights. An overhead view of the motion capture trajectories collected on the pursuit and obstacle avoidance trials is shown. Note the different scale bar for each panel (bottom right). (a) The pursuit trial consists of one flight. The bird's head trajectory (light blue) is shown as it flies to intercept a moving target (orange). The bird takes off spontaneously from the fist of a falconer standing 1 m behind the starting boxes (green markers). The target is attached to a line that passes around a series of pulleys (red). It begins hidden within the leftmost starting box in the flight shown, and is pulled by a linear motor located at the top of the image (not shown). The curtain (black) is represented here as the best-fitting plane to the reference markers attached to its edges. (b) The obstacle avoidance trial consists of two flights, each corresponding to one leg of a back-and-forth flight (magenta arrows) between two perches, with a bank of tall cylindrical obstacles located 1.5 m ahead of the end perch. The walls (brown) are estimated as planes from the motion capture cameras' positions and orientations. The pursuit flight takes 2.5 s and the obstacle flights around 2 s each. Further details on the positions of the reference markers for the different objects can be found in Appendix D

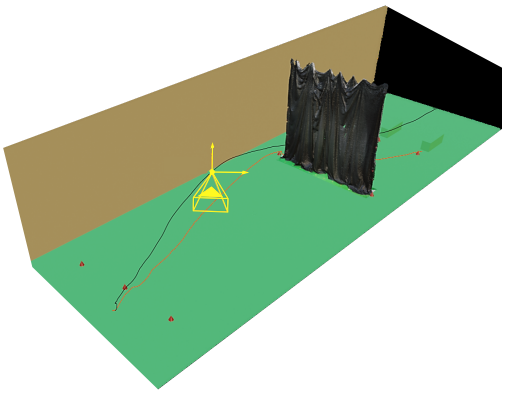
avoidance case, we assume that the bird looks at the perch centre upon landing (Potier et al, 2016; Kress et al, 2015). Third, we assume that the bird holds its eyes level during the calibration. This behaviour has been reported repeatedly in the bird flight literature (Brighton and Taylor, 2019; Ros and Biewener, 2017; Warrick et al, 2002), and is confirmed by the reference video. Further detail of these calibrations is provided in Appendix C.

2.2.1 Virtual camera

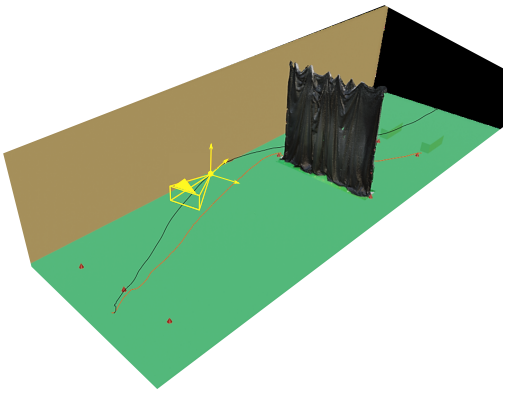
We modelled the scene viewed by the bird in Blender, using a 360° virtual camera whose translation and rotation per frame were matched to those of the estimated visual coordinate system fixed to the bird's head. As any vergence movements of the eyes are unknown, we modelled the binocular visual field as a monocular camera. We selected a resolution of 5 pixels per degree latitude and longitude, which results in pixels within the bird's visual field having a length and width approximately 10 times larger than the minimum resolution angle of the bird (Potier et al, 2016). For each flight, we rendered the view from this virtual camera in the bird's visual coordinate system using the Cycles rendering engine, and produced RGB, depth, semantic and optic flow data per frame. To test the effect of the bird's head movements, we also defined a trajectory coordinate system representing the pose of a horizon-levelled virtual camera whose optical axis was directed tangent to the bird's head trajectory (see Figure 4). Specifically, we set the y-axis of the trajectory coordinate system parallel to the birds' head velocity vector and its x-axis parallel to the floor plane, whilst fixing its origin to that of the visual coordinate system (see Appendix C.5). We again rendered the view from a virtual camera following the motion of this trajectory coordinate system for all three flights.

2.2.2 Hybrid model approach

We model the lab environment using a hybrid approach, which uses a combination of geometric primitives for the simple geometries in the scene, and dense 3D meshes for the most complicated ones. The dense 3D meshes are captured with a mobile device, and expressed at acquisition time in the same coordinate system as the motion capture trajectories. In this way we aim to minimise the modelling and postprocessing effort, while producing realistic representations of the environment. We demonstrate the use of a hybrid model of the lab for the pursuit flight, modelling the curtain using a dense 3D mesh, and the rest of the objects in the scene as geometric primitives (see middle panel in Figure 2). In the obstacle avoidance flights, we represented all the main elements of the scene using geometric primitives (see Figure 5). The synthetic scene can easily be made



(a) Visual coordinate system



(b) Trajectory coordinate system

Fig. 4 The two coordinate systems used for rendering. For each flight, we produce one set of renderings in which the virtual camera pose matches that of the visual coordinate system, and one set in which it matches the pose of the trajectory coordinate system (see text for definitions). The figure shows the pose of the virtual camera (yellow), the bird's head trajectory (black line) and the target's trajectory (orange line), for a selected frame in the final stages of the interception manoeuvre during the pursuit flight. In (a), the camera tracks the estimated visual coordinate system, which we expect to be representative of the pose of the bird's visual field. In (b), the camera tracks the trajectory coordinate system, which represents a horizon-level camera whose optical axis is tangent to the head's trajectory. Although the camera is represented schematically as a pyramid, a 360° virtual camera was used for all renderings

more photo-realistic as required: as an example, we provide an RGB rendering in which we added texture to the walls replicating the camouflage netting photographically (see Online Resource 1 and 2). The code to generate the model of the lab environment and define the corresponding virtual camera in Blender will be made available in GitHub upon publication.

2.2.3 Geometric primitives

We modelled the floor, ceiling and walls of the lab as planes. We determined the positions of the walls and ceiling from estimated points in the motion capture system's scaffolding, which we derived from the cameras' positions and orientations (see Appendix B.4). For the RGB renderings, we added texture to the walls representing the camouflage netting that was hung from the walls of the lab to provide a natural-looking texture and prevent the birds from perching. In the pursuit flight, we modelled the pulleys as cones and the boxes covering the target's initial position as cuboids (see middle panel in Figure 2). We modelled the target as a cylinder of 15 cm length and 2.54 cm diameter. We placed additional reference markers on the curtain's edges and on the wall netting at the curtain gap, but only used them to measure deviation from our modelled geometry (see Appendices A.2 and D.3). In the obstacle avoidance flights, we modelled the obstacles as vertical cylinders, and the perches as horizontal cylinders, thereby reducing each A-frame perch to the top rung on which the bird landed (Figure 5). The position, orientation and size of these geometric elements was determined from the retroreflective markers attached to the corresponding objects, and from measurements of the dimensions of the real objects. Further details on the definition of the geometric primitives are included in Appendix D.

2.2.4 Dense 3D map

To capture a dense 3D map of the curtain, we used the open-source *SemanticPaint* framework (Golodetz et al, 2015, 2018), which is built on top of InfiniTAM v3 (Prisacariu et al, 2017). We used the ASUS ZenFone AR smartphone as a mobile mapping sensor, and to perform visual-inertial odometry (ZenFone ZS571KL, ASUS, Taipei, Taiwan). To facilitate the integration of the mesh in the virtual model of the lab, we captured it in the motion capture coordinate system, using ArUco fiducial markers (Romero-Ramirez et al, 2018; Garrido-Jurado et al, 2014). The voxel size was set to 10 mm and the truncation distance to 40 mm (4× the voxel size).

The coordinate transformations applied to express the mesh in the motion capture coordinate system are summarised in Figure 6. By

default, the 3D mesh is captured in the SLAM world coordinate system, which is defined by the first camera pose. To compute the required transform from the SLAM world coordinate system to the motion capture coordinate system, we used an ArUco calibration plate. This consisted of an ArUco fiducial marker of size 28.8×28.8 cm fixed to an acrylic plastic sheet with three retroreflective markers (10 mm diameter) on three of its corners. When brought into camera view, the coordinates of the ArUco marker's corners are computed in the SLAM world coordinate system. Since we also placed retroreflective markers on these corners, their coordinates in the motion capture coordinate system are also known. By defining an auxiliary coordinate system with these three points, we can compute the transform from the SLAM world coordinate system to the motion capture one.

We used the open-source software MeshLab (Cignoni et al, 2008) to crop the mesh, remove duplicate vertices, and remove isolated pieces. We found that the floor plane of the mesh was slightly deviated from the motion capture system's xy-plane (2.4°, see Appendix D.3), likely due to drift. We used MATLAB's Point Cloud Processing functions to fit a plane to the floor of the mesh (mean error = 3.8 cm), and to compute the required transform to match the fitted plane with the motion capture's XY plane. The transformed mesh deviated on average by 0.093 m from the reference markers placed on the curtain's edges, as they were registered during the pursuit trial, and by 0.089 m from their position recorded just before capturing the mesh. Further details on the post-processing of the mesh and the deviation metrics are included in Appendix D.3.

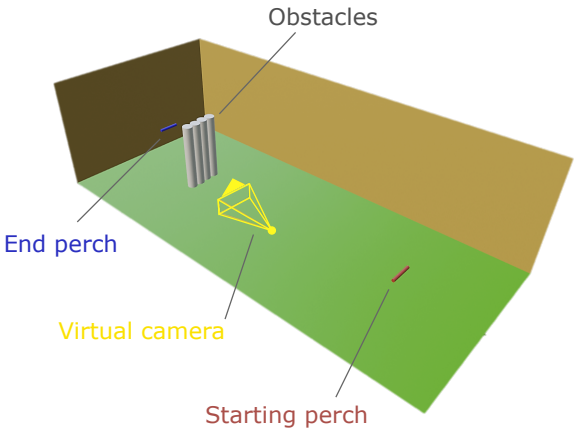


Fig. 5 Model of the lab environment for the obstacle avoidance flights. For the obstacle avoidance flights, we modelled the lab environment using basic geometric primitives to characterise the main elements in the environment. The walls, floor and ceiling were computed from the positions and orientations of the motion capture cameras and markers. The obstacles were modelled as vertical cylinders and the perches as horizontal ones, reducing the perches to their top rung. The positions of the perches and obstacles were determined using retroreflective markers attached to them. Their dimensions were derived from the motion capture markers, or from direct measurements of the objects

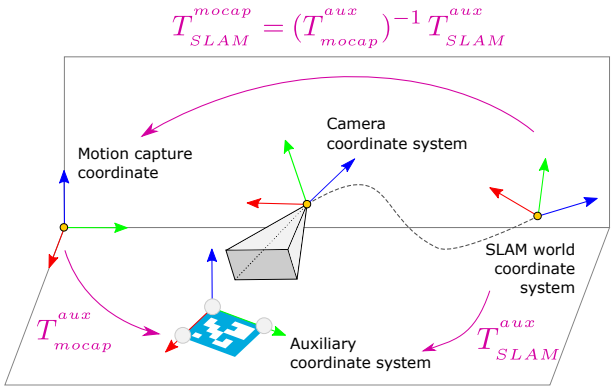


Fig. 6 Capturing the 3D dense map in the motion capture coordinate system. By default, the 3D mesh is captured in the SLAM world coordinate system, defined by the first camera pose. The camera's trajectory (dashed line) and pose in this coordinate system are computed via visual-inertial odometry. When the ArUco marker (shown in cyan) is brought into the camera's view, the coordinates of three of its corners are computed in the camera coordinate system. Thus, via visual-inertial odometry, they are also known in the SLAM world coordinate system. Since we placed three retroreflective markers on these corners (grey circles), their coordinates are also known in the motion capture coordinate system. The pose of an auxiliary coordinate system defined by these three points is therefore known in both the motion capture coordinate system and the SLAM world coordinate system, which allows us to compute the required transformation between them. The transforms between coordinate systems are shown in magenta, where T^B_A represents the transform from A to B

2.3 Comparison to previous approaches

Our method has several advantages over previous approaches to characterising the visual input experienced by a bird in flight (see Introduction). First, it allows us to consider visual input over the complete visual field of the bird, as opposed to confining this to the available field of view of a head-mounted camera (Ochs et al, 2016; Kane et al, 2015; Kane and Zamani, 2014) or to local features of the scene falling in the direction of the bird's gaze (Land, 1999; Eckmeier et al, 2008;

Yorzinski et al, 2013; Yorzinski and Platt, 2014; Yorzinski et al, 2015; Kress et al, 2015; Yorzinski, 2021). Second, it enables us to characterise this visual input without having to contain the animal within an abstract visual environment (Bhagavatula et al, 2011; Schiffner and Srinivasan, 2015; Dakin et al, 2016; Ros and Biewener, 2016). Third, we are not limited by the technical specifications and performance of a camera that is practical to attach to a bird's head in flight (Ochs et al, 2016; Kane et al, 2015; Kane and Zamani, 2014). Fourth, our method allows us to consider different camera models and alternative gaze strategies for the purposes of hypothesis testing, as we have done here in comparing views rendered in the visual coordinate system and trajectory coordinate system, respectively. Our method thereby opens up new avenues for research using data-driven models of animal behaviour, along the lines of those proposed to model quadruped locomotion for animation (Zhang et al, 2018), or motor control and perception in humanoids (Merel et al, 2020). As we now show, these data also offer unique insights into the bird's gaze strategy.

3 Results

For all three flights, we rendered the view from: (i) a virtual camera in the visual coordinate system, representing the translation and rotation of the bird's visual field inclusive of all head movements; and (ii) a virtual camera following the trajectory coordinate system, representing the translation and rotation of the bird exclusive of rotational head movements. Videos containing each of the rendered outputs per frame are included as Supplementary information (Online Resources 1 to 15); a brief description of them is included in the section Supplementary information, and further details on how they were computed is provided in Appendix E. As a demonstration of how these synthetic data can offer new insights into the bird's behaviour, we used the RGB and semantic data per frame to inspect the birds' gaze strategy, following a similar approach to the preliminary analysis presented by Minano and Taylor (2021).

3.1 Pursuit flight

The RGB renderings in the visual coordinate system for the pursuit trial show that the target remains within the bird's area of binocular overlap for almost the entire duration of the pursuit (see snapshots in Figure 7, and Online Resources 1 and 2). In contrast, the target is not held steady or centered within the trajectory coordinate system, for which the RGB renderings display pronounced pitch oscillations. This comparison confirms that the bird uses its rotational head movements to stabilize its gaze, and to keep the moving target reasonably well centered within its visual field. To analyse these properties quantitatively, we make use of the semantic data.

Figure 8 shows the frequency with which the target appears at each point in the visual field during the flight. For each pixel, the figure displays the metric:

$$h_i = \frac{n_i}{NA_i}, \tag{1}$$

where n_i denotes the number of frames over which the ith pixel saw the target, where N denotes the number of frames analysed, and where A_i denotes the solid angle covered by the ith pixel, normalised by the maximum solid angle that any pixel covers. Note that as the output is an equirectangular projection of a sphere, different pixels subtend different solid angles in the figure. We only consider those pixels falling within the visible areas of the bird's visual field, and exclude from the analysis any frames in which the head transform was interpolated, and any frames after interception (identified as the point at which the distance to the target first reaches a local minimum in the final approach phase). The results show that the target is held $\pm 10°$ longitude and from $-10°$ to $4°$ latitude in the visual coordinate system for most of the flight (Figure 8a). This contrasts with the results in the trajectory coordinate system, in which the target is not confined to this central area at all (Figure 8b). However, a limitation of these results is that they are affected by the apparent size of the target, and refer to data aggregated across the whole of the flight. How does the target's position in the visual field vary quantitatively through the flight?

Figure 9 plots the evolution of the target's edge contour in the visual coordinate system, where the visual field is shown in equirectangular projection and cropped close to the binocular area. The sections of flight before and after the curtain are plotted separately for clarity (Figure 9a and Figure 9b respectively). In the first section of the

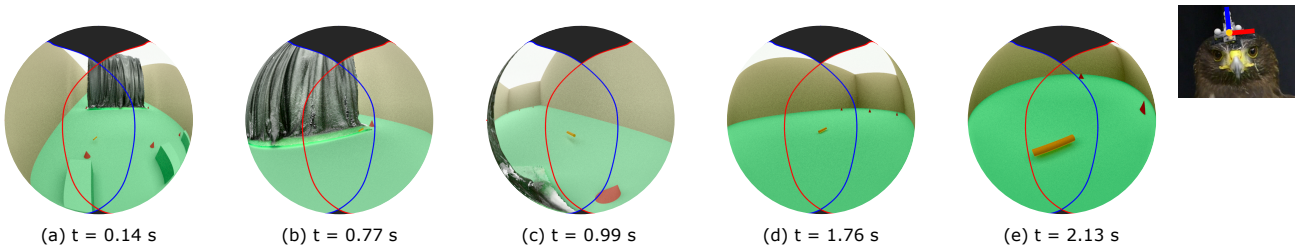


Fig. 7 Snapshots of the rendered pursuit flight. The spheres representing the bird's visual field show the animal's view as it flies through the lab; the retinal margins for the left eye (blue line) and right eye (red line) are shown for reference, as well as the blind area caudal to the visual field (black fill). Note that because the spheres are represented using an orthographic projection (right panel), not all of the visual field of the bird is visible. The inset shows the approximate orientation of the estimated visual coordinate system on the bird's head. The full flight takes 2.5 s

flight, the target begins drifting across the visual field, but then seems to be stabilised at approximately 10° longitude (Figure 9a). Target tracking seems to be lost temporarily as the bird turns around the curtain (green contours in Figure 9a), but is quickly recovered with the target now stabilised at −10° longitude (Figure 9b). Towards the end of the flight, the target gradually becomes centred in the visual field, looming until interception. The same evolution can be seen by inspecting the longitudinal position of the boundaries and midpoint of the target through time (Figure 9c). The target's boundaries also remain between −10° and 4° latitude for most of the flight, except at takeoff, when turning around the curtain, and at final interception (Figure 9d).

For comparison, we computed the equivalent path of the target as seen from the trajectory coordinate system (Figure 10). In the first part of the flight, the target shows considerably more oscillations in the vertical direction, likely reflecting wingbeat oscillations (Figure 10a). However just before turning around the curtain, the target appears to be aligned longitudinally with the bird's head velocity vector. In the second part of the flight (Figure 10b), the target is clearly not aligned with the head's velocity vector, drifting out of the central area of the trajectory coordinate system (Figure 10b). The comparison of Figures 10b and Figure 9b reflects how the estimated gaze direction diverges from the bird's velocity vector in the final phase of interception. In summary, although the moving target was not perfectly stabilised in the bird's visual field and sometimes appeared to be held at ±10° longitude, rather than being centered, the effect of the bird's head movements was to keep the target contained within the area of binocular overlap (Figure 9a).

3.2 Obstacle avoidance flights

The RGB renderings of the visual coordinate system for the obstacle avoidance trials show the obstacles located towards the centre of the visual field (see Online Resources 8 and 9 for leg 1, and Online Resources 12 and 13 for leg 2). This is not the case for the renderings of the same flights in the trajectory coordinate system, in which the obstacles are not centred, and where wingbeat oscillations are clearly visible (see Online Resources 10 and 11 for the flight corresponding to leg 1 of the trial, and Online Resources 14 and 15 for the flight corresponding to leg 2). This confirms that the bird actively stabilises its visual field against the pitch oscillations associated with its wingbeat, and also directs its gaze so as to keep the obstacles broadly centered. In this case, however, close inspection of the semantic data reveals a more subtle interpretation of how the bird is directing its gaze.

As the obstacles occupy a much larger portion of the field of view than does the target shown in the cropped equirectangular projections of Figure 9, we display the evolution of the obstacles' edge contour using an orthographic projection to reduce distortion in Figure 11. The obstacles' edge contour is represented from the point of takeoff until the point at which the landing perch is visible without occlusion. The data show that the nearside edge of the obstacles seen by the bird remains aligned longitudinally with the centre of the visual coordinate system in both flights (Figure 11, top row). This alignment does not appear for the semantic data rendered within

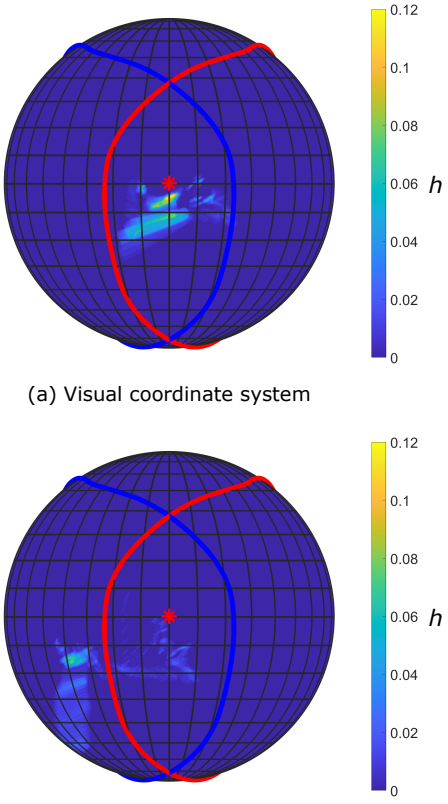


Fig. 8 Semantic heatmaps of the target for the pursuit flight. The heatmaps show the value of h (Eq. 1), representing the frequency with which the target appears at each pixel in the visual field throughout a flight, normalised by the relative area of the solid angle that each pixel subtends. Results are shown for a camera following the visual coordinate system (a), and the trajectory coordinate system (b). The virtual camera axis is marked by a red star: it corresponds to the estimated direction of the bird's gaze $\vec{v}_{gaze}$ in the visual coordinate system, and to the direction of the head's velocity vector in the trajectory coordinate system. In the virtual coordinate system (a), the target appears centred in the binocular area throughout the flight. In the trajectory coordinate system (b), it does not appear to be constrained to this region. We considered $n = 337$ frames, from the takeoff frame to the interception frame, and excluded those frames in which the head translation or rotation was interpolated. Note that the bird's visual field extends beyond what is shown in this projection (orthographic). The maximum value of the colormap was set to match the maximum value observed in the visual coordinate system data. The visual field is represented as seen from the bird's point of view

the trajectory coordinate system (Figure 11, bottom row), which reinforces the role of the bird's head movements in fixating the obstacles' nearside edge.

Figure 12 plots how the longitudinal positions of the nearside edge, the farside edge, and the midpoint of the obstacles and landing perch evolve through time. For the obstacle avoidance flight comprising the first leg of the trial, the nearside edge of the obstacles appears to remain approximately aligned with the centre of the visual field until the point at which the landing perch first becomes fully visible (Figure 12a). Beyond this point, the midpoint of the landing perch becomes the object most closely aligned with the centre of the visual field. For the obstacle avoidance flight comprising the second leg of the trial, the nearside edge of the obstacles is again aligned closely with the bird's estimated gaze direction until the point at which the landing perch becomes visible. The bird then appears to make a head saccade, such that its new gaze direction aligns with the nearside edge of the landing perch, which remains the case for approximately the next 0.8 s (Figure 12b). The bird then seems to make another head saccade, such that its estimated gaze direction becomes most aligned with the midpoint of the perch. Both head saccades can be seen in the corresponding RGB rendered video (Online Resources 12 and 13).

In summary, when avoiding the obstacles, the bird stabilized and centered the nearside edge of the obstacles in its visual coordinate system. The bird switched to fixating the perch once this was fully visible, looking at the nearside edge of the perch first in one of the flights, and then at its midpoint before landing. It is important to note that the latter statement is inevitable for the final set of frames that we used to calibrate the bird's gaze direction under the assumption that the bird would look at the midpoint of the perch when landing. However, the coincidence of the edge of the obstacles with the centre of the visual coordinate system provides strong internal support for the reliability of its calibration, and the midpoint of the perch was found to be centered within the visual coordinate system from well before the start of the set of frames that were used in the calibration (Figure 12).

4 Discussion

Here we have presented a method generating synthetic visual inputs characterising the visual experience of a bird whose head pose is recorded in flight using a motion capture system or other measurement system providing similar data (e.g. a head mounted inertial measurement unit). To our

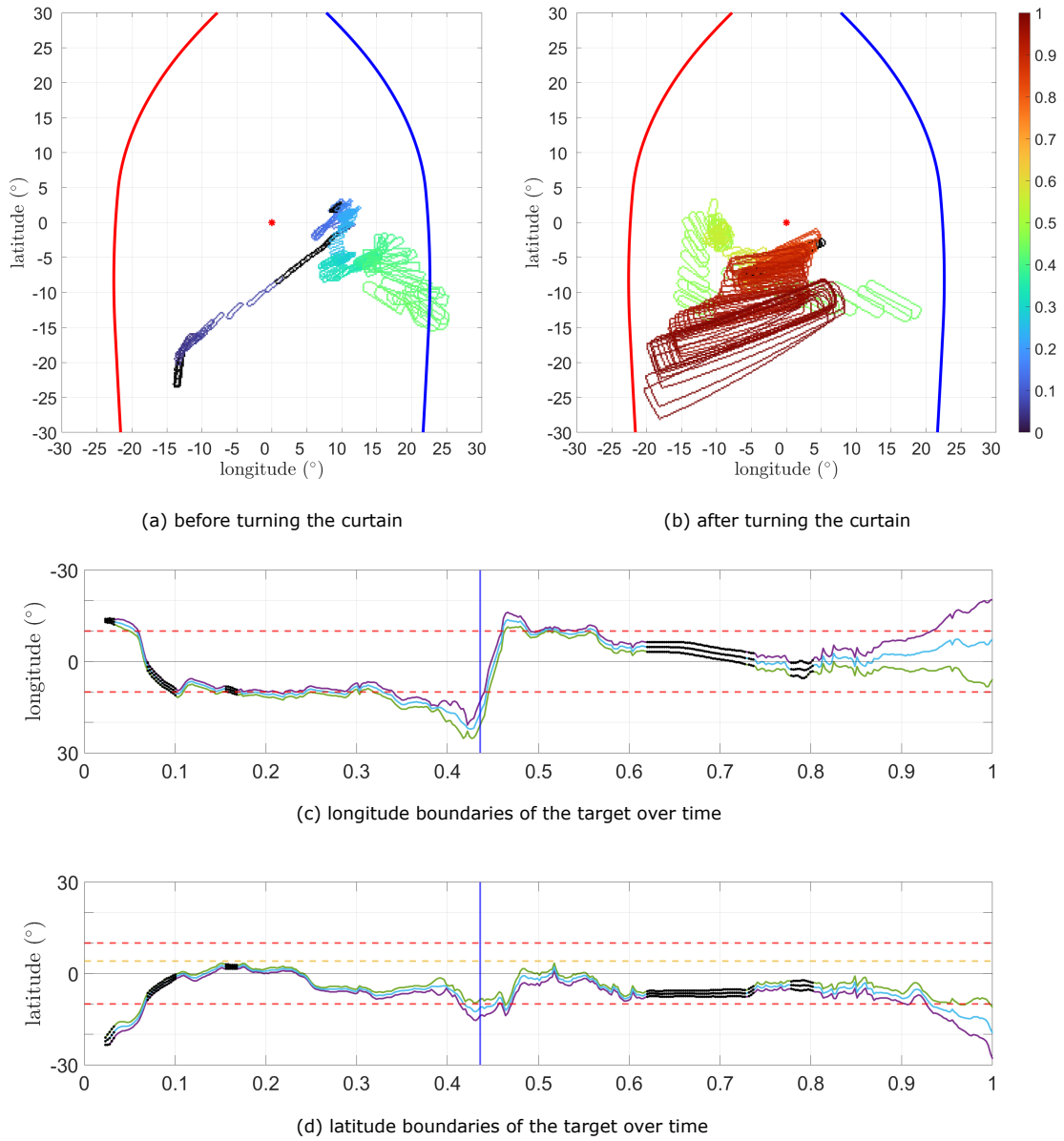
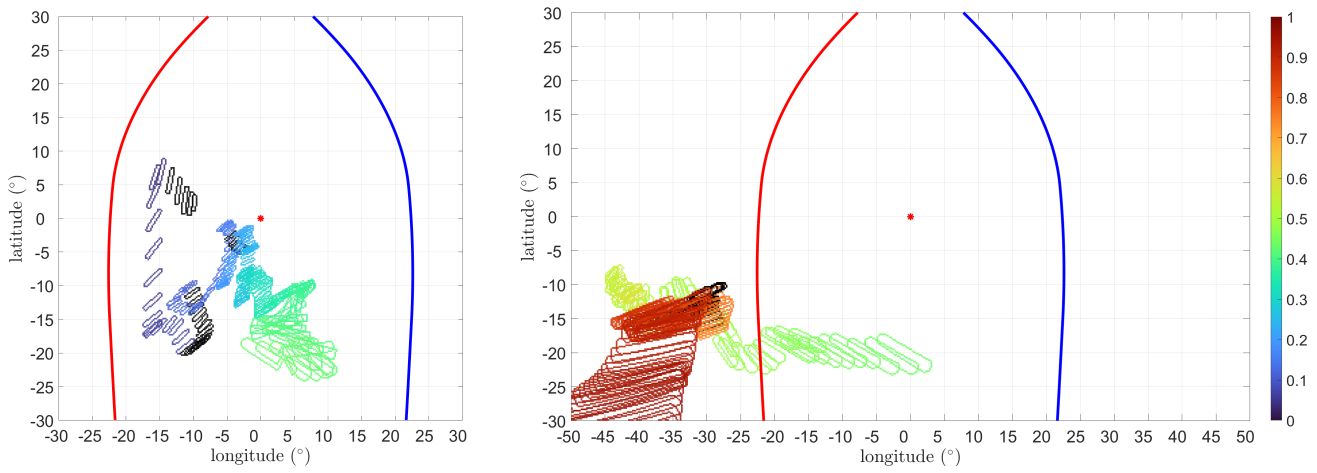


Fig. 9 Trajectory of the target in the bird's visual field. The edge contour of the target is represented for each frame of the pursuit flight in a cropped equirectangular projection of the area around the estimated gaze direction $\vec{v}_{gaze}$ (red star), for the frames before (a) and after (b) turning around the curtain. The colormap indicates time through the flight. The extension of the target in longitude and latitude over time is shown in (c) and (d) respectively, with the maximum (purple), minimum (green), and mean (blue) values represented per frame. For those frames in which the bird's head transform was interpolated, these values are shown in black. The blue vertical line in (c) and (d) indicates the frame that defines the data split before (a) and after (b) turning the curtain. Reference lines are shown in (c) and (d) at ±10° (red dashed lines) and at 4° latitude in (d) (yellow dashed line). The normalised time is 0 at the takeoff frame, identified when the bird dips its head just before the takeoff jump; note that it takes some frames for the target to become visible, but the linear motor pulling it was already triggered at this point. The normalised time is 1 at interception, when the bird and the target reach a local minimum of distance in the terminal phase of interception. The data shown cover 2.145 s

knowledge, this is the first time that such detailed data covering the complete visual field of an animal has been generated for non-human animals in general, and especially for birds in flight. Although we have limited our analysis of these data to presenting simple metrics for three flights as an illustration of the method, these are already sufficient to show the potential of using synthetic data generated using our method to investigate how birds



(a) before turning the curtain

(b) after turning the curtain

Fig. 10 Trajectory of the target in the trajectory coordinate system. The edge contour of the target is represented for each frame of the pursuit flight in a cropped equirectangular projection of the area around the camera axis (red star) in the trajectory coordinate system. The virtual camera's axis corresponds to the direction of the head's velocity vector. The extent of the retinal margins of the bird's right eye (red) and left eye (blue) is shown relative to the virtual camera's axis for reference, although they are not expected to be positioned correctly or consistently in this coordinate system. The data is split following the same criteria as in Figure 9, separating the frames before (a) and after (b) turning around the curtain. The colormap indicates time through the trial. Note that in the first part of the flight the target shows prominent pitch oscillations likely reflecting the reaction to the wingbeat. In the second part of the flight, the target is not constrained to the equivalent of the binocular area, showing that the velocity vector of the head trajectory is not aligned with the target

use vision to guide and control their flight. Comparable analyses would be exceedingly challenging using video data from head-mounted cameras, given the attendant limitations on payload, resolution, field of view, and motion blur, and would be much more limited if relying on point estimates of the bird's gaze direction in relation to prominent visual features in the environment. Our method allows us to define alternative gaze strategies, by comparing our renderings in the visual coordinate system to renderings in the trajectory coordinate system, in order to inspect counterfactual scenarios. Again, this would be difficult or impossible to do using any of the other reviewed approaches.

4.1 Key features of rendering method

Our rendering method uses a hybrid approach to model the lab environment, which combines basic geometric primitives defined using the motion capture data with dense 3D maps of features with more complex geometry. In this way, we avoid the accumulated drift and noise typical of large 3D maps, whilst reducing the modelling effort for the most intricate shapes, and making the most of the accurate motion capture data. To facilitate the integration of the dense meshes within the geometric model of the environment, they are expressed in the same coordinate system as the motion capture data at the point of acquisition. We have demonstrated the applicability of this hybrid approach for a simple example of reconstructing a curtain that was hung to act as an obstacle to the bird. However, the method would be especially relevant for modelling natural-looking environments with more complicated features, such as small trees. As an illustration of the viability of this approach, Figure 13 shows the 3D mesh of a set of trees that we placed around the lab and captured using *SemanticPaint*, in this case using a Kinect v1 as the mapping sensor.

4.2 Behavioural analyses: pursuit flight

Our analysis of the synthetic reconstructions demonstrate the potential of using these to provide further insight into the bird's behaviour. During the pursuit flight, we found that the target is held within the area of the bird's binocular overlap for most of the flight. Modelling a counterfactual gaze strategy, in which the virtual camera's principal axis is aligned with the bird's head velocity,

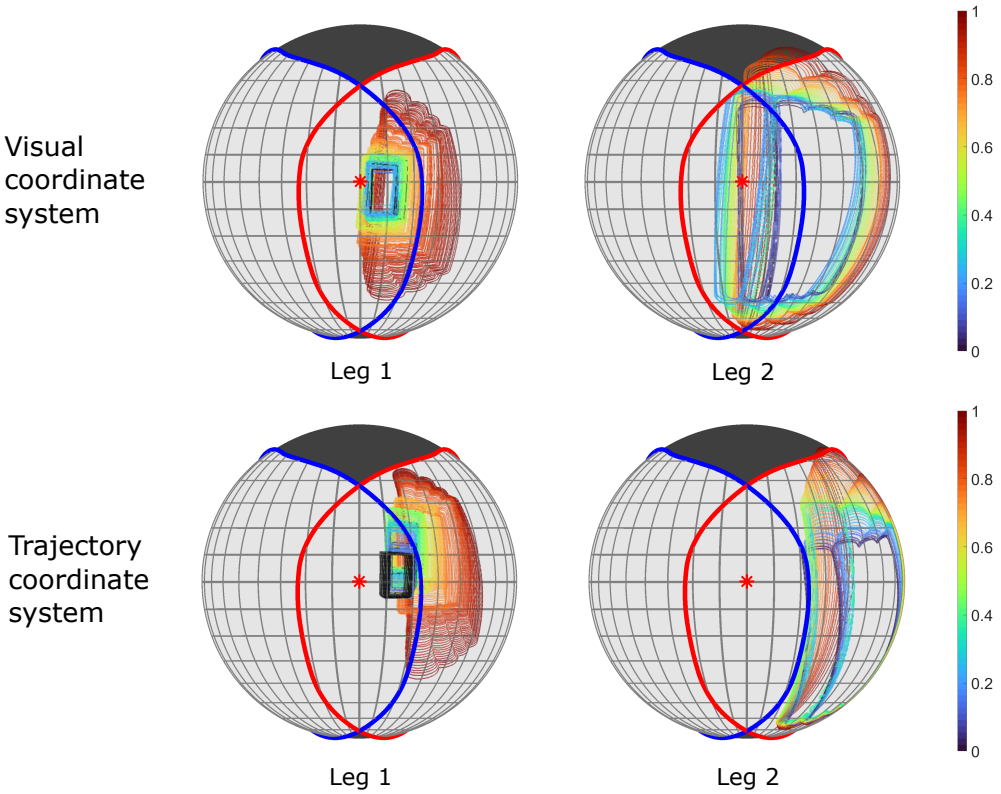


Fig. 11 Trajectory of the obstacles in the visual field sphere. The edge contour of the set of obstacles is represented in the virtual camera's visual field for each frame (coloured semi-transparent) of both obstacle avoidance flights, from the point of takeoff to the point at which the perch is fully visible (i.e. until the second blue dashed vertical line in Figure 12). The colormap indicates time through each flight. For the frames in which the camera's transform was interpolated (i.e. because not enough markers were reconstructed), the contour of the obstacles is shown in black. The retinal margins of are shown for the left eye (blue line) and right eye (red line), and the virtual camera's axis is shown for reference (red star). Note that the virtual camera's axis corresponds to the estimated gaze direction $\vec{v}_{gaze}$ in the visual coordinate system (top row), and to the direction of the head's velocity vector in the trajectory coordinate system (bottom row). In the visual coordinate system, the leftmost edge of the set of obstacles stays largely aligned with the estimated sagittal plane in both flights. In contrast, in the trajectory coordinate system, the obstacles are not stabilised in the vertical direction and they are not aligned with the head velocity vector either. The visible area in the bird's visual field extends beyond what is shown in this orthographic projection

corroborates our conclusion that the bird actively directs its gaze to keep the target in this region of the visual field. Inspecting the evolution of the target's position in the visual field through the flight, we find that the bird fixates the target at $\pm 10°$ longitude from the estimated gaze direction, and only centres it towards the terminal interception phase. In common with most other raptors, Harris' hawks have two points of high-resolution vision per retina: one projecting frontally and the other laterally (Mitkus et al, 2018; Potier et al, 2016). Whilst it would be premature to draw any firm conclusions from data for a single flight of a single bird, we hypothesise that these locations at $\pm 10°$ longitude in the visual field may correspond to the projections of the frontal foveae of the bird's left and right eyes. These are estimated to be located at between 9° and 16° longitude from the forward direction of the head in most raptors (Frost et al, 1990; Kane and Zamani, 2014; Tucker, 2000), and so unlike in humans, they may not project to a single point in the visual field—even when the eyes converge (Tyrrell et al, 2018). This being so, our results could indicate that the bird tracks the target with one or other of its frontal foveae throughout the flight, before centering it in the visual field prior to interception and producing a symmetric looming stimulus. It is also interesting that when the bird loses track of its target as it makes a sharp turn at the curtain, it appears to reacquire it with the other fovea.

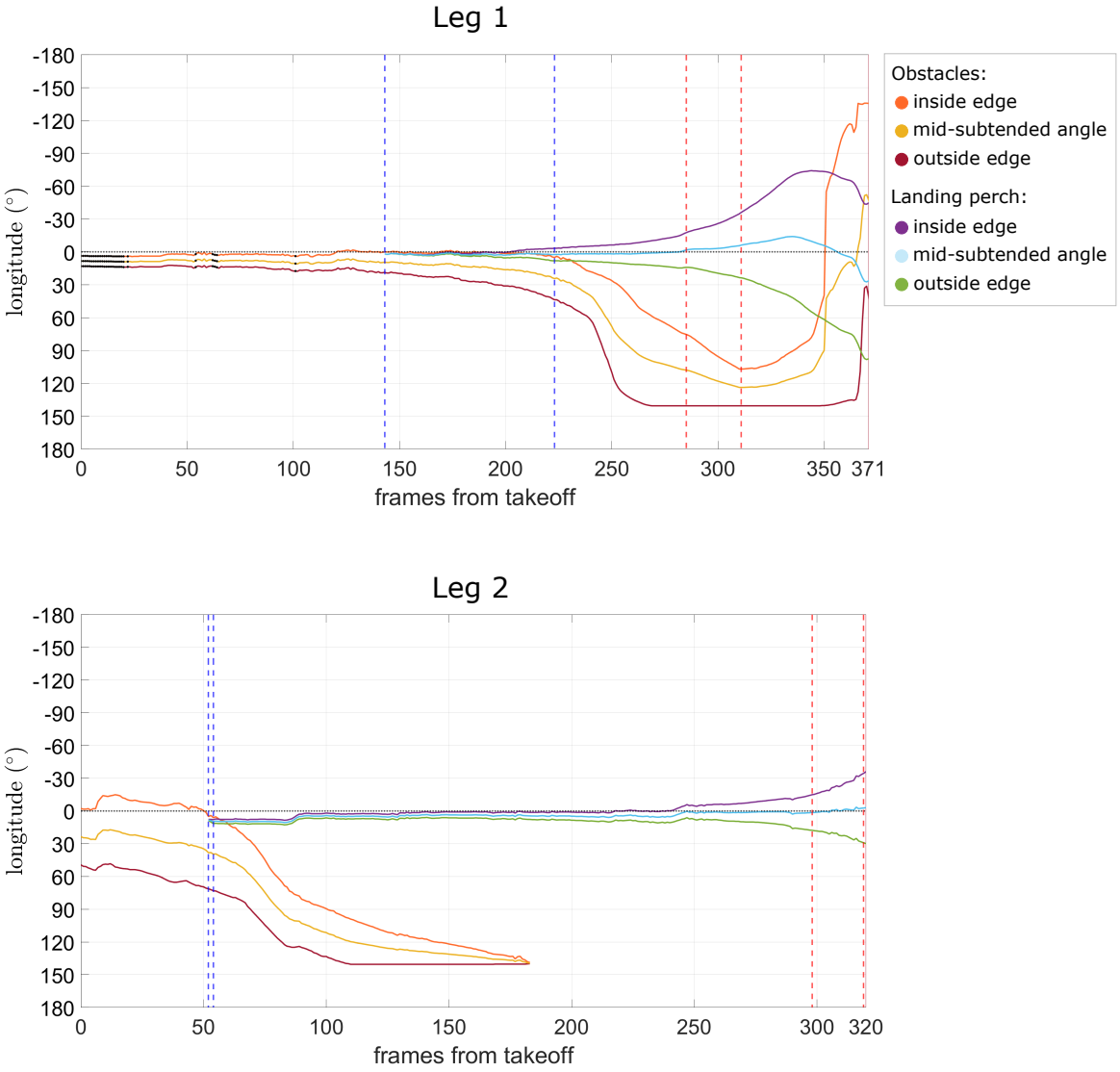


Fig. 12 Longitudinal extension of the set of obstacles and landing perch for the obstacle avoidance flights. The evolution through time of the longitudinal extension of the obstacles and landing perch is represented for the two obstacle avoidance flights, corresponding to leg 1 (top panel) and leg 2 (bottom panel) of the trial. The inside and outside edges of the obstacles and landing perch are labelled relative to the bird's turn, together with the midpoint of the angle subtended by the obstacles and landing perch. Black markers denote frames in which the bird's head transform was interpolated. The range of frames when the landing perch is partially occluded by the obstacles is marked between two vertical dashed lines. Note that because objects curve when they are close to the poles of the spherical virtual camera, it may be that the landing perch is fully visible but that its longitudinal extension overlaps with that of the obstacles (e.g. at around 225 frames from takeoff; see Online Resources 8 and 9). The first and last frames used to estimate the bird's gaze direction $\vec{v}_{gaze}$ are marked with vertical red dashed lines. In leg 1 of the flight, the obstacles appear within the central part of the visible field after 350 frames, but in reality the would have been occluded by the bird's body (see Online Resource 8). In leg 2, we can visually identify two potential head saccades: one at 87 frames after takeoff seems to align the estimated gaze direction with the left edge of the landing perch, as seen by the bird; the other at 242 frames after takeoff seems to align the estimated gaze direction with the centre of the perch (see Online Resources 12 and 13). The sampling rate is 200 Hz

4.3 Behavioural analyses: obstacle avoidance

In both obstacle avoidance flights that we analysed, we found that the inside edge of the obstacles was aligned with the longitudinal centre of the visual field for substantial portions of the flights in which the obstacles were visible. Again this was not observed in the counterfactual gaze strategy that we considered, in which the virtual camera's principal axis was aligned with the bird's head velocity. This is in accordance with previous

findings from other birds (Kress et al, 2015) and humans (Raudies et al, 2012; Rothkopf and Ballard, 2009), which both seem to fixate on the edges of objects than can be perceived as obstacles, and on the centre of objects perceived as goals. Moreover, in the flight corresponding to the second leg of the trial, the bird seemed to align its visual field first with the edge of the landing perch and then with its centre (see Online Resources 12 and 13, around frames 2029 and 2184 as numbered in the video). This is also in line with previous reports in lovebirds (Kress et al, 2015) and may reflect a strategy based on aiming at intermediate goals. The agreement with previous works indicates that our approach is reasonable. However, as with the pursuit flight, these hypotheses can only be confirmed or rejected by analysing the complete set of flights across different individuals.

4.4 Limitations and future work

We aimed to develop a reconstruction method that would enable the collection of large amounts of data from many individuals. Some key steps that we have taken towards this goal include our non-invasive tracking of head pose using marker-based motion capture, our programmatic definition of geometric primitives based on motion capture markers placed on objects in the environment, and our automated integration of the dense 3D maps on the motion capture coordinate system. The key limitation in our current pipeline is the need to calibrate the visual coordinate system against the coordinate system of the headpack, which we achieve using two alternative methods sharing common assumptions. A second potential limitation of the reconstructions method is its representation of a binocular system using a monocular visual field.

One way of overcoming both limitations would be to combine data from calibrated stereo cameras with the motion capture data. This could allow the pose of the head to be calibrated directly within the headpack coordinate system, also allowing measurement of the stereo baseline of the eyes. The potential for exploiting this approach is demonstrated by recent work by Naik (2021), who developed an annotation tool to obtain the 3D coordinates of identifiable features on a bird, such as its bill and eyes, with respect to its head. This will be a key focus of our future work, as it will allow us to streamline the process and obtain more accurate estimations of the visual coordinate system. Other potential improvements could involve automating the correction of the dense 3D maps, potentially using the reference motion capture markers in the mapped objects as fiducials. An improved ArUco plate would also be an easy way to improve floor alignment. Exploring the possibilities for tracking eye movements would also be of interest in providing better estimates of the range of the motions, or defining a more accurate visual coordinate system Holmgren et al (2021).

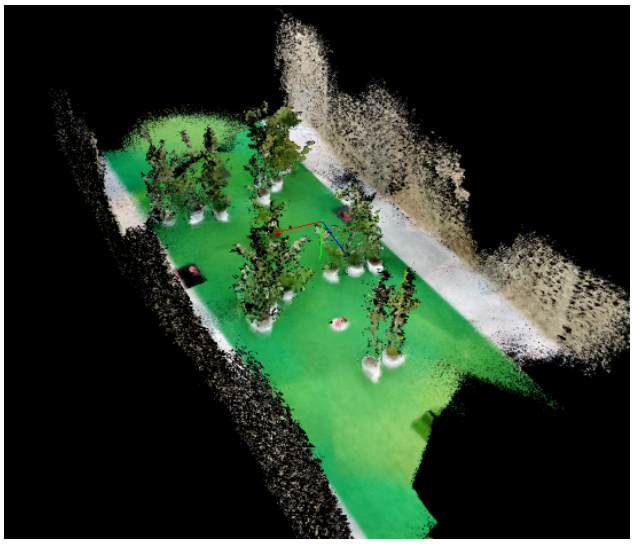


Fig. 13 3D model of a forest environment in the lab. We used a mix of 20 small laurel and bay trees of under 2 m height to recreate a forest environment in the motion capture lab. The 3D model was captured using *SemanticPaint* and a Kinect v1 as a mapping sensor. Note that in this case the world coordinate system of the captured map is not aligned with that of the motion capture coordinate system

There is still much to learn about how animals interact with their environment, and innovative methods such as the one we propose open a wide range of possibilities for behavioural analysis. Understanding the role of active vision in bird flight may reveal efficient processing strategies that could be translated to autonomous systems. Human-inspired active vision has already been explored for robotic applications (Seara and Schmidt, 2004; Seara et al, 2002, 2001), as active observers have been shown to solve basic vision problems more efficiently than passive ones, with many ill-posed problems becoming well-posed (Aloimonos et al, 1988). Visual-guidance

algorithms could also be tested directly using the data generated with our approach, similar to the method followed when studying navigation in homing ants by Ardin et al (2015). Finally, large datasets collected with this or similar methods could be combined with inverse reinforcement learning frameworks, leading to data-driven models offering new insights into the bird's gaze strategy in flight; similar approaches already exist that make use of human motion capture data to generate active-sensing behaviours in synthetic humanoids (Merel et al, 2020). In conclusion, we see many exciting opportunities for mutual collaboration between the animal behaviour and computer vision communities in all of these areas.

Supplementary information. We provide the rendered outputs as video supplementary material. For the pursuit flight, we include videos for the RGB, semantic, depth and optic flow synthetic data generated. For the obstacle avoidance flights, we include RGB videos. All videos are reproduced at 20Hz, (1/10 of the real speed), except the optic flow video which is reproduced at 5 Hz (1/40 of the real speed). The frame numbering shown in the videos follows the motion capture data system's numbering. A description of each of the video files included is presented in Table 1.

The RGB outputs are presented using two projections: equirectangular, in which the geometry appears distorted, but which shows the complete field of view of the bird; and orthographic, in which the distortion is reduced but doesn't include the most peripheral regions of the bird's visual field. The rest of the rendered outputs are only represented in orthographic projection. Note that for the orthographic case, the point of view is as if looking frontally to the bird's visual field sphere (see Figure 30). In all videos, a red contour around the figure indicates the head transform for that frame was interpolated. The retinal margins and the areas of the bird's visual field (blind, monocular and binocular) are overlaid on the rendered output.

In the optic flow video, the colormap represents for each frame the instantaneous angular speed per pixel, in degrees per second. The colorbar is in logarithmic scale and capped at $10°/s$ in the lower bound and $1000°/s$ in the upper bound. The vector field results from the transformation of the output data from pixel space to the surface of the unit sphere. Further details on the computation of the videos are included in Appendix E.

Acknowledgments. The authors would like to thank James Kempton who contributed in the data collection for the pursuit trial, and the falconers Helen Sanders, Lucy Larkman and Mark Parker for animal husbandry and handling during the experiments. We would also like to thank Marco KleinHeerenbrink for help with the markers' labelling custom algorithms.

Appendix A Motion capture experiments

A.1 Motion capture system

The motion capture lab is equipped with 22 infrared motion capture cameras (Vantage V16, Vicon Motion Systems Ltd, Oxford, UK; sampling rate 200 Hz) and four video cameras (Vue, Vicon Motion Systems Ltd, Oxford, UK; sampling rate 100 Hz, only used for reference). The cameras were mounted on a fixed scaffold from which camouflage netting was hung, to provide a natural-looking background texture and prevent the birds from perching on the lower parts of the scaffolding. The floor was carpeted with green astroturf.

The flights presented in this paper were recorded with two different captive-bred adult Harris' hawks: Drogon (0.6 kg) flew the obstacle avoidance flights, and Toothless (0.7 kg) the pursuit flight.

The pursuit flight is part of a dataset we recorded with four different birds, in which we collected between 3 and 4 flights per bird per day, during four weeks between December 2018 and January 2019, for a total of 251 trials. The number of flights per bird and day in the pursuit dataset was limited by the training approach: an extra food reward was required to get the bird off the target and back to the falconer's fist, which increased the food intake per trial. This reduced the amount of flights we could record while keeping the animals motivated to fly. Resetting the setup between trials also took more time than when recording perching or obstacle avoidance trials.

Table 1 Description of supplementary videos. We provide supplementary videos for the rendered output per trial. The description field in the table specifies the output data represented in the video, the coordinate system the virtual camera is tracking in the video and the projection used.

Trial		Filename	Description (output, coordinate system tracked, projection)
Pursuit		ESM_1.avi	RGB, visual coordinate system, orthographic
		ESM_2.avi	RGB, visual coordinate system, equirectangular
		ESM_3.avi	Depth, visual coordinate system, orthographic
		ESM_4.avi	Semantic, visual coordinate system, orthographic
		ESM_5.avi	Optic flow, visual coordinate system, orthographic
		ESM_6.avi	RGB, trajectory coordinate system, orthographic
		ESM_7.avi	RGB, trajectory coordinate system, equirectangular
Obstacle avoidance	Leg 1	ESM_8.avi	RGB, visual coordinate system, orthographic
		ESM_9.avi	RGB, visual coordinate system, equirectangular
		ESM_10.avi	RGB, trajectory coordinate system, orthographic
		ESM_11.avi	RGB, trajectory coordinate system, equirectangular
	Leg 2	ESM_12.avi	RGB, visual coordinate system, orthographic
		ESM_13.avi	RGB, visual coordinate system, equirectangular
		ESM_14.avi	RGB, trajectory coordinate system, orthographic
		ESM_15.avi	RGB, trajectory coordinate system, equirectangular

The obstacle avoidance flights are part of a dataset we recorded with four different birds (Drogon, Toothless, Charmander and Ruby), in which we collected 16 flights per bird per day for a total of two weeks in November 2020. Both periods include the training weeks for the birds.

A.2 Placement of markers

To track the birds' head movements, we used rigid supports of retroreflective markers for the birds to wear as a 'headpack'. In the pursuit flight, we used a headpack of 5 retroreflective markers, made of wooden thin rods glued to a plastic cross-shaped plate. In the obstacle avoidance flights, we specially designed 3D-printed rigid supports for 4 retroreflective markers.

The headpack was attached to the head of the bird by the falconer immediately before recording its set of flights, and removed at the end of the set. To ensure that the headpack did not move relative to the bird's head we recorded videos of the birds on the falconer's fist wearing the headpacks. For the flights analysed here, we did not detect visually any movement between the headpack and the head. The head movements when the bird is stationary are generally of larger amplitude than those observed in flight, so we consider these videos are a good indication that there was no significant relative motion between the bird's head and the headpack in flight.

In the pursuit flight, we fixed markers of 6.4-10 mm diameter to the pulleys' centres, to the top front vertices of the starting boxes, to the edges of the curtain (which were reinforced with a wooden rod) and to the camouflage netting at the curtain gap. Figure 14 shows the median position of the reference markers used in the pursuit flight; note that the walls' positions were estimated with the motion capture cameras' position and orientation, rather than with the markers. To track the target's motion, we glued three markers (6.4 mm diameter) to its canvas cover. In the obstacle avoidance flight, we placed 14 mm diameter markers to the perches' edges and at the centre of the obstacles' tops.

A.3 Experimental procedure

In the pursuit trial, a linear motor pulled the target on a line passing around a series of pulleys. The trajectory of the target was randomised by selecting for each trial a set of pulleys, from one of

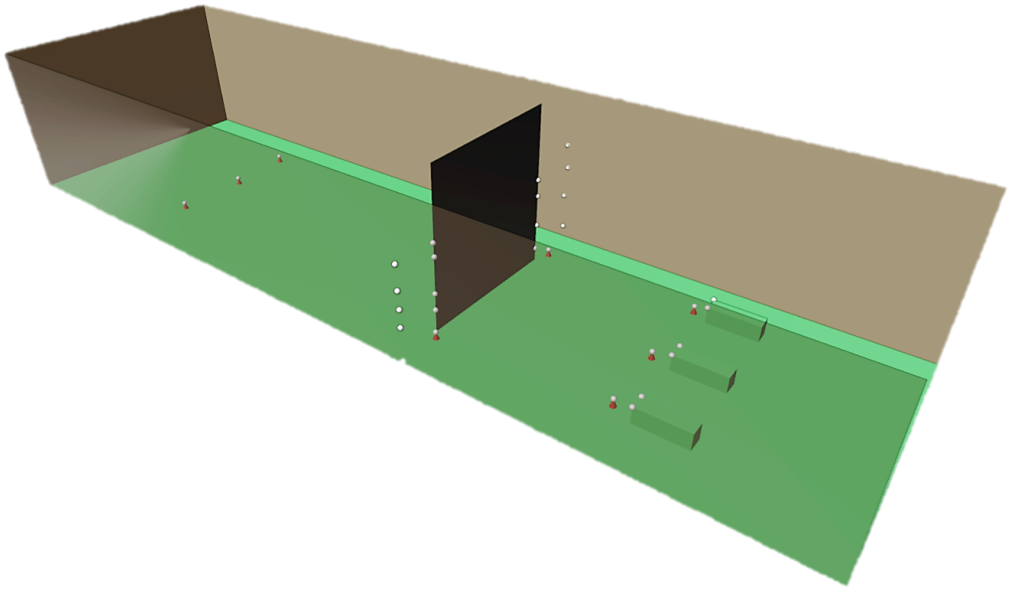


Fig. 14 Reference motion capture markers used in the pursuit trial. The median position of the reference markers in the pursuit trial is shown in a basic geometric model of the lab environment. The sidewalls are shown as partially transparent and the curtain is modelled as a plane, for clarity. The spheres representing the markers' positions are 5 times the markers' real size. Note that the walls' location was estimated from the motion capture cameras' positions and orientations, rather than from the markers placed at the curtain gaps

three starting positions to one of three end positions, via one of the two gaps around the curtain (Figure 3a). The curtain was made of two layers of flame-proofed cotton canvas of size 6.0 m $\times$ 3.5 m, and hang from a steel cable installed transversally across the motion capture room. The line was a 3 mm thick parachute-style cord, in green to minimise contrast with the astroturf carpet, and the average speed of the target was 5.6 m/s. To keep the target in tension, a long piece of line was attached to its back end and rolled around the first pulley of the target's trajectory. We also randomised the takeoff position of the bird and placed dummy lines along the alternative paths, to minimise the risk of the bird predicting the target's path. The bird began flying freely from its falconer's fist as soon as the target appeared from inside the starting box in which it was hidden. Before recording the pursuit trial, we recorded a 'gaze calibration trial' to estimate the bird's gaze direction within the headpack coordinate system. In it the falconer held the bird on their fist whilst we displayed a piece of food with a marker attached at a range of 0.5-1.5 m.

The obstacle avoidance flights are part of a set in which we recorded the bird's head movements as it flew from the starting perch to the end perch, and back, with and without obstacles in place. For each trial we recorded two flights, which corresponded to the trajectories back and forth between the perches. Two falconers stood at either end of the room, to handle the bird and provide the food reward. The lateral position of the perches was randomised for every trial between three stations each. These were centred on the longitudinal axis of the room and distributed with 1 m spacing. Each obstacle was made up of two white expanded polystyrene cylinders of 1 m height, stacked on top of each other and bound together with white duct tape. The four obstacles were pushed together so that there were minimal gaps between them. For the trials with obstacles in place, we also randomised the side of the end perch from which the falconer would call the bird. We used the complete set of flights from the same bird on the same day to calibrate the bird's gaze direction within the headpack coordinate system (see Section 2.2).

We calibrated the motion capture system using an active wand and following the manufacturer guidelines. In the obstacle avoidance set of trials, we calibrated before recording a set of trials with the same bird and headpack placement. In the pursuit dataset, we calibrated the system without

the curtain in place, and before recording all trials per day. Since we recorded less trials in the pursuit dataset, we don't expect the calibration to deteriorate significantly (the flight presented here is recorded 1.5 hours after calibrating)

Appendix B Motion capture data postprocessing

B.1 Markers' reconstruction and labelling

We used the commercial software Vicon Nexus 2.8.0, from the motion capture system manufacturers, to extract the unlabelled 3D coordinates of the retroreflective markers per frame. This process is called 'markers' reconstruction'. The system provides a residual value for each reconstructed marker in 3D space to evaluate the measurement accuracy (Motion Lab Systems, 2021). Figure 15 shows the histogram of residuals for the trials rendered in this paper. For the pursuit trial, the mean residual is 0.82 mm, the standard deviation is $\sigma = 0.29$ mm, and the median is 0.76 mm, for $n = 69537$ samples. For the obstacle avoidance trial, which consists of the two obstacle avoidance flights considered, the mean residual is 0.98 mm, the standard deviation is $\sigma = 0.34$ mm, and the median 0.92 mm, for $n = 29403$ samples. All markers and recorded frames per trial are included in the computation of the histograms (i.e., frames recorded before takeoff and after landing are also considered).

Nexus is designed for collection and annotation of 3D motion capture data, with a particular focus on human motion. We found that our experiments pushed the system to its limits in terms of tracking and labelling performance, and therefore considered alternative options to process our data. For a detailed description of the typical challenges of using marker-based motion capture on birds the reader is referred to Naik, 2021.

For the pursuit flight, we labelled the individual headpack markers per frame using Nexus semi-automatic labelling tool. For the target, it was difficult to separate the individual identities of the markers with manual labelling, as the target twisted and turned during the trial. Instead we assigned a consistent set of three labels to the target markers, that allowed us to separate them easily from the rest. We exported the manual labels to MATLAB, and wrote custom scripts to label the remaining markers.

For the obstacle avoidance dataset, we wrote custom MATLAB scripts to separate headpack and object markers, and label them. To label the headpack markers, we used the following approach:

- first, we labelled the frames with at least 4 markers by iteratively solving the orthogonal procrustes problem with the headpack template; we also determined the headpack's transform (position and orientation) at those frames.
- second, we interpolated the headpack's transform at the remaining frames; and
- third, we iteratively labelled the markers in the remaining frames, by assigning them the label of the closest marker of the corresponding interpolated headpack.

We iterated over the third step, keeping a constant minimum distance for accepting a label but updating the interpolated transform data for the headpack with the newly labelled frames every time. We stopped when there were no new frames labelled between two passes. To label the markers in the objects, we used predefined bounding boxes.

To assess the quality of our labelling of the headpack markers, we computed the maximum distance between a labelled marker and its corresponding position in the headpack template, in a local coordinate system. Results are shown in Figure 16 and Table 2. Although the ideal distance deviation without measurement error would of course be zero, the actual distance deviation is in all cases smaller than the diameter of the headpack markers, which implies that the markers are correctly labelled, albeit that their precise positioning is subject to measurement error. In line with our observations during manual labelling, the pursuit trial appears to have more jitter in the markers' positions (even though the residuals obtained from the system are not much larger, see Figure 15). For the set of trials with the same bird and headpack placement as in the selected obstacle avoidance trial (which we used to estimate the visual coordinate system), the maximum distance registered across all trials was 3.7 mm.

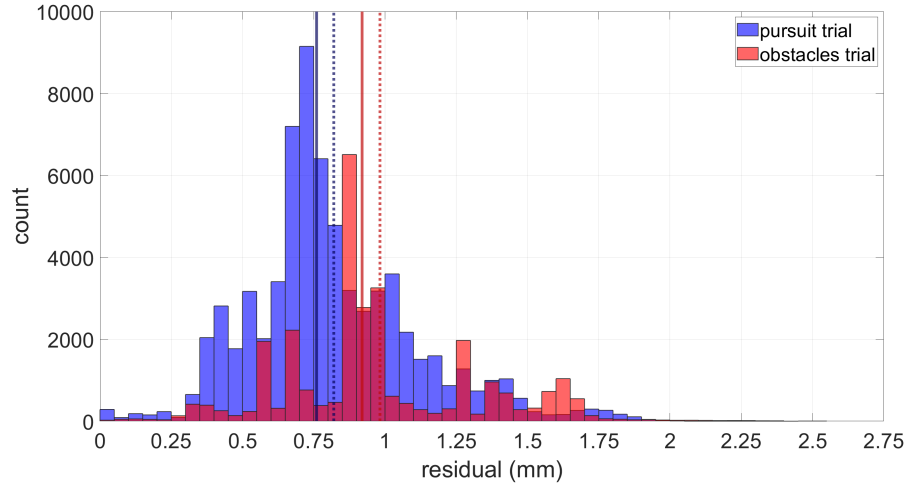


Fig. 15 Histogram of 3D point residuals. The residuals are computed by the motion capture system as part of the reconstruction step, in which the markers' 3D position is determined. The results for the rendered pursuit trial (blue) and obstacle avoidance trial (red) are shown. The obstacle avoidance trial consists of the two obstacle avoidance flights analysed in the main text. The mean (dashed line) and median (continuous line) values for both distributions are under 1 mm. Note that all markers and recorded frames per trial are considered. The bins' width is 0.05 mm

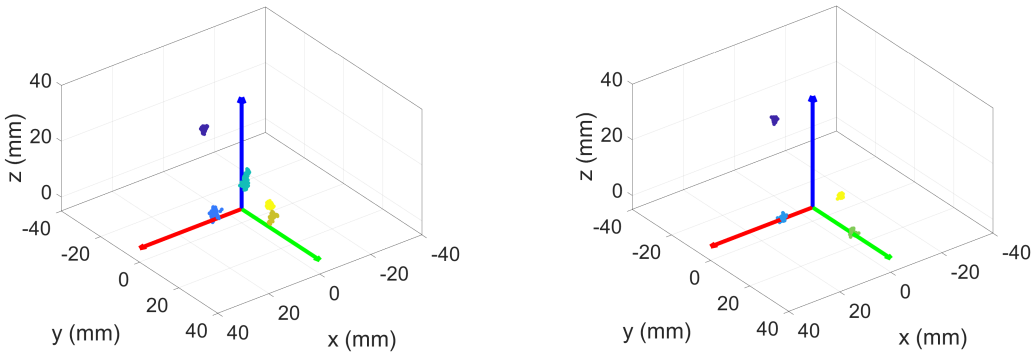


Fig. 16 Labelled markers per trial in a local coordinate system. The labelled markers for each of the pursuit (a) and the obstacle avoidance trials (b) are shown in a coordinate system that moves with the headpack. In (a) the visual coordinate system is used and in (b) the headpack coordinate system is shown; both correspond to the template used for labelling with procrustes in each trial. We considered all rendered frames in which the headpack transform was not interpolated. Note that we used different headpack designs for the two trials. The obstacle avoidance trial consists of the two obstacle avoidance flights analysed in the main text

B.2 Headpack transform per frame

To compute the headpack's transform per frame, we solved the procrustes problem on the labelled data. In the pursuit flight, we used a template computed from the gaze calibration data. In the obstacle avoidance flights, we used the theoretical headpack design as template.

For the frames where there were not enough headpack markers reconstructed, or where these could not be reliably labelled, we interpolated the headpack's transform. In the pursuit flight, we used a smoothing cubic spline to interpolate the headpack's translation. In the obstacle avoidance flights, we used a weighted variation that takes into account the number of markers detected. In both cases we interpolated the headpack's rotation using the SLERP algorithm for quaternions, which assumes constant angular velocity (Shoemake, 1985). We assumed the short path between

Table 2 Maximum distance d_{max} between labelled markers and corresponding markers in headpack template. We considered all rendered frames in which the headpack transform was not interpolated. Note that we used different headpack designs for the two trials, and therefore the same label does not correspond to the same marker. The obstacle avoidance trial consists of the two obstacle avoidance flights analysed in the main text

	d_{max} (mm)	
Label	Pursuit trial	Obstacle avoidance trial
1	1.8	1.6
2	3.1	2.0
3	**4.6**	**2.8**
4	2.2	1.6
5	1.6	N.A.

quaternions and constant pose in the extrapolation regions.

During manual labelling of the pursuit flight we found that some of the markers' reconstructions were jittery and noisy for certain sections, so we computed the procrustes transform on a subset of frames we determined as most reliable, and interpolated the translation and rotation of the headpack for the rest of the frames. The subset of frames considered most reliable is the set of frames with 3 or more markers reconstructed, excluding those frames in which the procrustes error is above a certain threshold, the markers are close to collinear, or the most prominent frontal marker is missing. We then applied a low-pass filter to the resulting head rotations, as a moving average filter of window size equal to 2 frames, using the SLERP algorithm:

$$q_f(t) = slerp(q_f(t-1), q(t), p), \tag{2}$$

where $q_f(t)$ denotes the filtered quaternion at frame t, $q(t)$ denotes the quaternion data point, *slerp* denotes the implementation of the SLERP algorithm in MATLAB, and p is the interpolation parameter, which we set at 0.5. For the first frame $q_f(0) = q(0)$. We applied the same filter when computing the head rotations for the trajectory coordinate system.

B.3 Target transform per frame

For the pursuit flight, we modelled the target as a cylinder, with the three markers attached to it defining a right-angled triangle. We defined a coordinate system linked to the cylinder, whose origin is at the cylinder's centre, its y-axis is parallel to the longitudinal axis of the cylinder (pointing inwards), and its x and z-axis are two perpendicular radii of the cylinder.

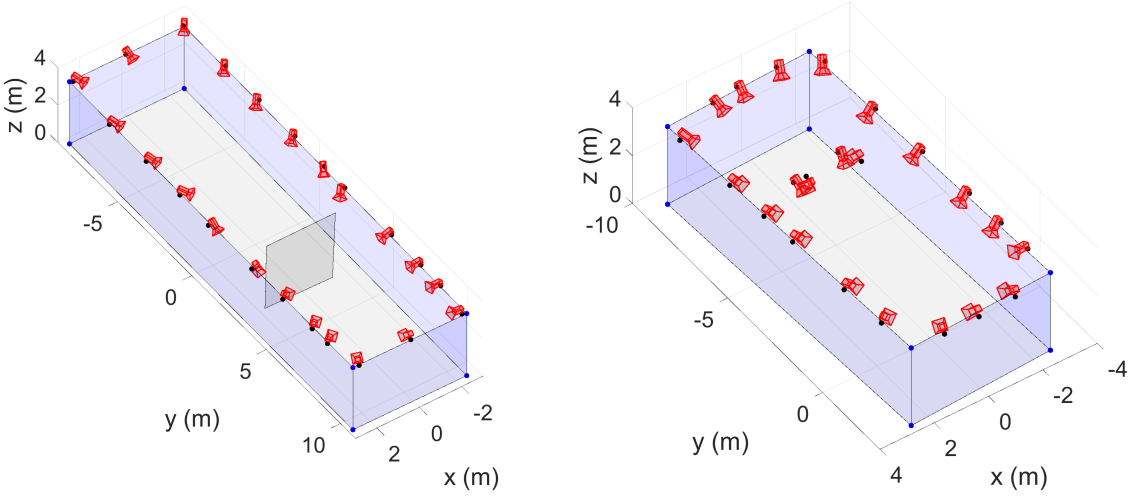
We computed the translation of the cylinder as the trajectory of the circumcentre of the triangle. We estimated it using a cubic smoothing spline, that takes into account the number of markers reconstructed.

We define the rotation of the cylinder's coordinate system as follows: its y-axis is always parallel to the origin's velocity vector $\vec{v}_{origin}$, and its x-axis is always parallel to the floor plane of the lab. This mimics the actual motion of the target, which was in tension between two lines, one pulled by the linear motor and one looped around the initial pulley.

B.4 Walls, floor and ceiling geometry

The motion capture cameras were mounted on a fixed scaffold from which camouflage netting was hung. We estimated its location using the position and orientation of the motion capture cameras. These were computed during the calibration. We used the cameras' positions and orientations to estimate points on the scaffolding, taking into account the dimensions of the cameras and their mounts. We then fitted the estimated scaffolding points to a line for each wall section. This provided an estimate of the scaffolding rungs holding the netting and the cameras. We defined each wall as the plane perpendicular to the floor that contained the corresponding scaffolding rung, from which the camouflage netting was hung.

Figure 17 shows the estimated walls and the motion cameras for the pursuit and the obstacle avoidance flights. The mean angle at the walls' corners was 90°, with standard deviation $\sigma = 0.2$ for the pursuit flight, and $\sigma = 0.5$ for the obstacle avoidance flights. We set the ceiling at the mean height of the estimated scaffolding points: 3.33 m for the pursuit flight and 3.25 m for the obstacle avoidance flights.



(a) Pursuit trial (b) Obstacle avoidance trial

Fig. 17 Estimated location of the walls. We estimated the location of the walls from the motion capture cameras' positions and orientations for the pursuit flight (a) and the obstacle avoidance flights (b). Cameras are shown in red. From the dimensions of the cameras and their mounts, we estimated points in the scaffolding (black markers). We then used the scaffolding points to the define the walls' planes (blue transparent planes). The corners of the volume are highlighted with blue markers. A plane representing the curtain's location is shown for reference in (a), computed from the markers attached to it

Appendix C Visual coordinate system calibration

To define the visual coordinate system relative to a headpack coordinate system, we made use of three assumptions (see Section 2.2). We identified specific periods in the collected data in which we expect these assumptions to hold most reliably. We used the data during these periods to estimate $\vec{v}_{gaze}$, the bird's gaze direction, and $\vec{n}_{sagittal}$, the normal to the bird's head symmetry plane (i.e., the sagittal plane). These vectors define the basis of the visual coordinate system: the x-axis is defined parallel to $\vec{n}_{sagittal}$ (pointing to the left side of the head) and the y-axis parallel to $-\vec{v}_{gaze}$. We defined the x and y axes in this way so that the visual coordinate system is close to parallel to the motion capture coordinate system at the start of the trial; this facilitates the interpretation of the computed rotations. The following sections describe the estimation of $\vec{v}_{gaze}$ and $\vec{n}_{sagittal}$ for the pursuit and the obstacle avoidance flights.

C.1 Estimate for the pursuit flight

For the pursuit flight, we estimated the visual coordinate system using data from the gaze calibration trial. During this trial, we recorded the bird's head movements while it was held on the falconer's fist and presented with a food reward that had a marker attached to it. In the corresponding reference video, we identified a 'fixation phase': a range of frames in which the bird is likely to be fixating on the food reward while holding its head level. We identified these frames by a characteristic behaviour of the bird, in which it lowers its head and prepares its wings for a downstroke; three samples of the typical posture of the bird in this phase are shown in Figure 18. We selected two sequences of 342 and 120 frames of motion capture data within the trial in which this behaviour was most apparent.

We estimated the bird's forward gaze direction $\vec{v}_{gaze}$ by fitting a 3D line to the food marker's trajectory during the fixation phase, in an auxiliary coordinate system linked to the headpack (RMSE = 47.5 mm for $N = 462$ samples); the resulting vector is shown in Table 3 and Figure 19.

For the same range of frames, we estimated the orientation of the imaginary line connecting the bird's eyes, assuming the animal keeps its head

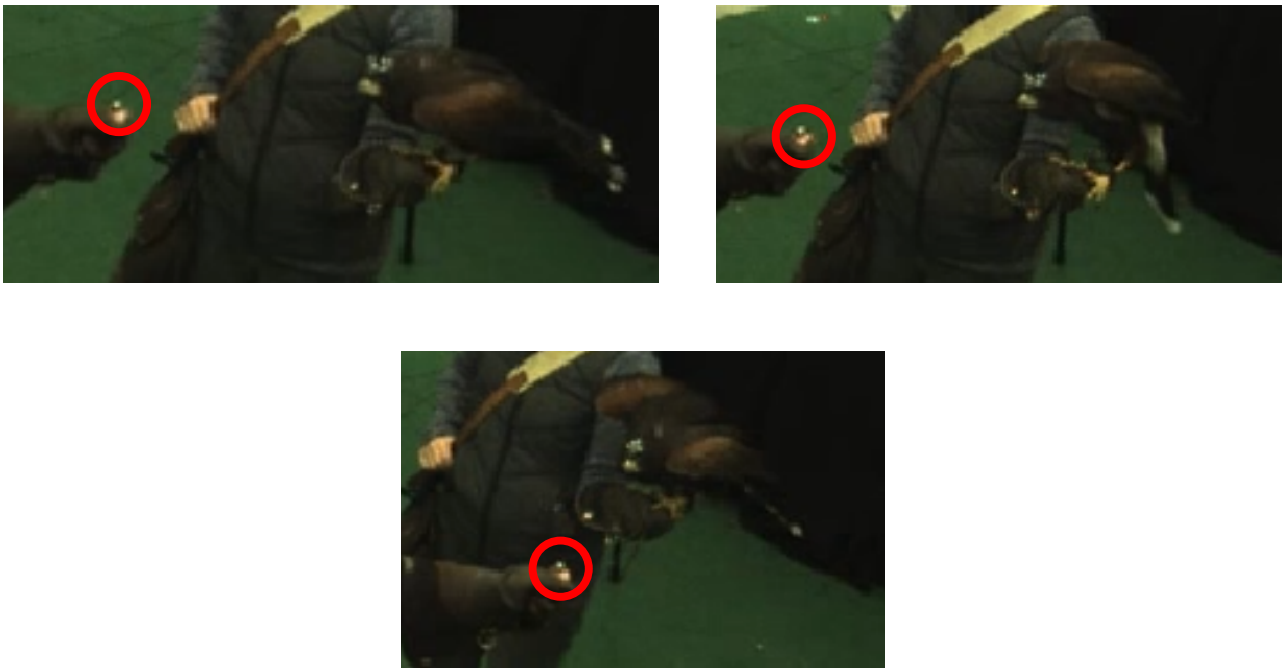


Fig. 18 Examples of the bird's behaviour in the fixation phase of the gaze calibration trial. For certain periods during the gaze calibration trial, the bird appeared to focus on the food reward presented to it, with a characteristic behaviour of lowering its head and preparing its wings for a downstroke. The marker on the food reward is highlighted with a red circle

approximately level during that period. We identify this line with $\vec{n}_{sagittal}$, the normal to the bird's head symmetry plane (i.e. the sagittal plane), which we define as positive when pointing to the left side of the bird's head. We compute $\vec{n}_{sagittal}$ as the unit vector perpendicular to $\vec{v}_{gaze}$ that best approximates the normal to the sagittal plane in the headpack coordinate system. We do this by solving the following least-squares problem:

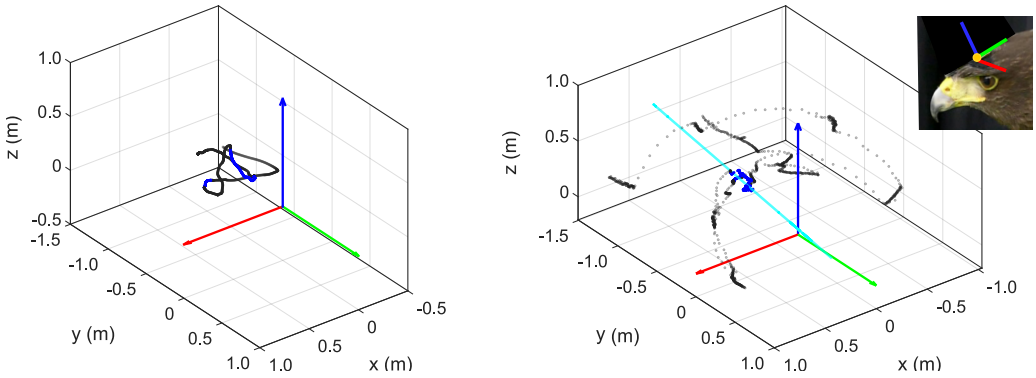
$$\begin{aligned} \theta^* &= \arg\min_{\theta} \sum_{i=1}^{N} (\vec{z}_{world,i} \cdot \vec{n}_{sagittal}(\theta))^2, \\ \vec{n}_{sagittal}(\theta) &= (cos\theta)\vec{a} + (sin\theta)\vec{b}, \\ \vec{n}_{sagittal} &= \vec{n}_{sagittal}(\theta^*), \end{aligned} \tag{3}$$

where $\vec{z}_{world,i}$ is the world's z-axis in the headpack coordinate system (positive opposite to gravity), $\vec{a}$ and $\vec{b}$ are an arbitrary orthonormal basis of the plane perpendicular to $\vec{v}_{gaze}$, and θ is the angle between $\vec{n}_{sagittal}$ and $\vec{a}$. The fit over the $N = 462$ fixation frames yields a root-mean square residual of 0.04; the resulting vector is shown in Table 4 and Figure 20.

The origin of the visual coordinate system in the pursuit flight is defined as the midpoint between the two lateral markers of the headpack (see Figure 23a). We estimated this point to be close to the midpoint between the eyes, the ideal origin of the visual coordinate system (Potier et al, 2016). From reference images of the headpack on the bird's head, we estimate that the origin is < 10 mm from the midpoint between the bird's eyes in the direction perpendicular to the sagittal plane, and < 20.1 mm in the direction perpendicular to the headpack's baseplate (see Appendix C.3).

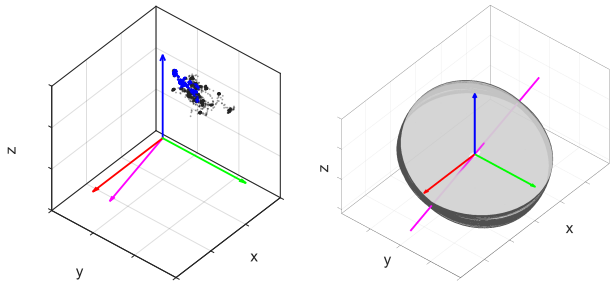
C.2 Estimate for the obstacle avoidance flights

For the obstacle avoidance flights, we estimated the visual coordinate system using data from 15 trials with the same bird and headpack placement. Note that each trial consists of two flights, which correspond to the two legs of the trial (see Figure 3). From those 15 trials, 8 were recorded with obstacles in place, and 7 without them. We identified two phases in all the flights within this set of trials: a final approach phase, in which we expect the bird to fixate on the centre of the perch;



(a) Motion capture coordinate system (b) Headpack coordinate system

Fig. 19 Estimation of gaze direction for the pursuit flight. The trajectory of the marker attached to the food reward in the gaze calibration trial is represented in two coordinate systems: one that translates with the headpack and is parallel to the motion capture coordinate system (a), and one fixed to the headpack (b). The estimated gaze direction $\vec{v}_{gaze}$ is the orthogonal regression line (cyan) to the samples in the fixation phase (blue) in the headpack coordinate system (b). The samples outside this phase are shown for reference (black, semi-transparent). The food reward is presented to the bird at a distance between 0.5 and 1.5 m. Note how the position of the food marker in the fixation phase in (b) falls within a small region, in contrast to the curved trajectory it follows in (a), suggesting the bird is actively positioning the food reward on a small area of its visual field during that phase. Inset in (b) shows the approximate orientation of the headpack coordinate system relative to the bird's head



(a) Estimation of sagittal plane (b) Local horizon

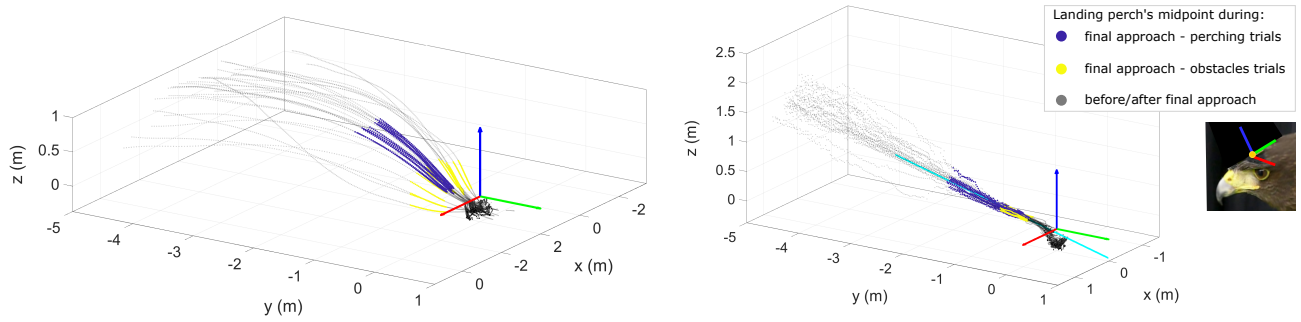
Fig. 20 Estimation of the normal to the sagittal plane for the pursuit flight We estimated the normal to the sagittal plane $\vec{n}_{sagittal}$ (magenta), assuming the bird keeps its eyes level during the fixation phase of the gaze calibration trial. In (a), dots represent the direction of the world's z-axis in the headpack coordinate system, during the fixation phase (blue markers) and outside the fixation phase (black, semi-transparent). Samples outside the fixation phase show larger variation in roll (i.e. rotation around the y-axis). In (b), the local horizon (i.e. the plane perpendicular to the world's z-axis in the headpack coordinate system) is represented for every frame in the fixation phase (gray circular planes). The direction of the $\vec{n}_{sagittal}$ vector (magenta) is the best-fit line that is perpendicular to the previously estimated $\vec{v}_{gaze}$ direction and is contained in all these planes, in least-squares sense

and a mid-flight phase, in which we expect the bird to keep its head level. We used these two phases to estimate $\vec{v}_{gaze}$ and $\vec{n}_{sagittal}$, respectively.

We estimated the bird's forward gaze direction $\vec{v}_{gaze}$ using data from the final approach phase. This phase was defined based on the bird's distance to the landing perch (0.5 to 1.0 m for the obstacle trials, 0.5 to 2.0 m for the no obstacle trials; note that the obstacles were placed 1.5 m ahead of the end perch, as per Figure 3b). We fitted a 3D line to the trajectory of the midpoint of the landing perch during this phase, in a headpack coordinate system (see Figure 21). The root-mean square error of the fit was RMSE = 65.6 mm, for N = 1274 samples; the resulting vector is shown in Table 3. Results were similar to those obtained by fitting each trial individually (see Appendix C.4, Table 5).

We computed the normal to the sagittal plane $\vec{n}_{sagittal}$ following the same approach as in the pursuit flight, using the data of the mid-flight phase. We defined this phase to be when the bird was > 2 m away from either perch and flying at a speed > 2.5 m/s; for all flights these conditions defined a continuous range of frames. We solved for $\vec{n}_{sagittal}$ using the set of equations in 3, and obtained a root-mean square residual of 0.08 for $N = 5056$ samples; the resulting vector is shown in Table 4 and Figure 22.

We defined the origin of the visual coordinate system as the centroid of the headpack markers projected onto the headpack's baseplate plane. This is different to the definition used in the pursuit flight because we used a different headpack



(a) Motion capture coordinate system

(b) Headpack coordinate system

Fig. 21 Estimation of the gaze direction for the obstacle avoidance flights. The gaze direction $\vec{v}_{gaze}$ is estimated over 8 trials with obstacles in place (obstacle trials) and 7 trials without them (perching trials). Note that each trial consists of two flights, which correspond to the two legs of the trial (see Figure 3). The landing perch's midpoint (grey and coloured dots) is represented from 5 m away until landing, in two coordinate systems: one that translates with the headpack and is parallel to the motion capture coordinate system (a), and one fixed to the headpack (b). The estimated gaze direction $\vec{v}_{gaze}$ is the orthogonal regression line (cyan) to the samples in the final approach phase (plotted blue for perching trials, and yellow for obstacle trials) in the headpack coordinate system (b). Note the straightness of the perch's midpoint trajectories in the headpack coordinate system in (b), and compare this to the curvature of the lines in (a), which confirms unequivocally that the head pose was being stabilised in relation to the perch. Inset in (b) shows the approximate orientation of the headpack coordinate system relative to the bird's head

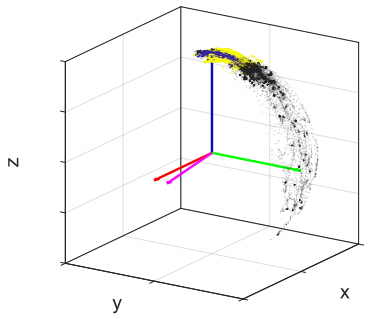


Fig. 22 Estimation of the normal to the sagittal plane for the obstacle avoidance flights. We estimated the normal to the sagittal plane $\vec{n}_{sagittal}$ (magenta) assuming the bird keeps its eyes level during the mid-flight phase. The dots represent the direction of the world's z-axis in the headpack coordinate system during the mid-flight phase, for perching (blue) and obstacle trials (yellow). Note that each trial consists of two flights, which correspond to the two legs of the trial (see Figure 3). Samples outside of the mid-flight phase (black, transparent), show qualitatively larger variation in pitch (i.e., rotation around the x-axis) and roll (i.e., rotation around the y-axis)

design. From reference images, we estimate this point is within 12 mm of the midpoint between the bird's eyes in the direction perpendicular to the sagittal plane, and < 20 mm in the direction perpendicular to the headpack's baseplate (see Appendix C.3).

C.3 Distance between the origin and the midpoint between the eyes

In the pursuit flight, the origin of the visual coordinate system is defined as the midpoint between the two lateral markers of the headpack (see Figure 23a). In the obstacle avoidance flights, it is defined as the centroid of the headpack markers projected onto the plane defined by the headpack baseplate (see Figure 25a; note that the headpack designs are not the same in the pursuit flight and the obstacle avoidance flights). We selected these points aiming to be close to the midpoint between the eyes. We estimated how much these points deviate from the midpoint between the eyes using snapshots of the videos of the birds wearing the headpacks, recorded before and after the trials.

The reference images used to estimate this offset for the pursuit flight are shown in Figures 23 and 24. Figure 23a shows the selected origin lies between markers 2 and 5, and Figure 23b shows it is closer to the sagittal plane (red dashed line) than to marker 5. We therefore estimate that the offset in the direction perpendicular to the sagittal plane is, at most, the distance between the origin and marker 5, which is 10 mm. To estimate the offset in the direction perpendicular to the headpack's baseplate, we selected three frames in which the baseplate plane was almost perpendicular to

Table 3 Results of the gaze direction fitting, for the pursuit and the obstacle avoidance flights. The estimated gaze direction $\vec{v}_{gaze}$ and the point in the fitted line A_{gaze} are both expressed in the headpack coordinate system. The distance between the fitted line and the origin of the headpack coordinate system is d. RMSE stands for root mean square error, where the error is the distance between the sample points and the fitted line. The number of samples per fit is denoted by n. For the pursuit flight, it corresponds to the number of frames in the fixation phase in the gaze calibration trial. For the obstacle avoidance flights, it corresponds to the number of frames in the final approach phase, across all the flights recorded with that bird and headpack placement

Trial	Bird	$\vec{v}_{gaze}$	A_{gaze} (mm)	d (mm)	RMSE (mm)	n
Pursuit	Toothless	$[0.16,\ -0.94,\ 0.32]$	$[103.7,\ -572.3,\ 251.5]$	54.3	47.5	462
Obstacles	Drogon	$[0.13,\ -0.94,\ 0.31]$	$[131.4,\ -969.9,\ 247.0]$	68.4	65.6	1274

Table 4 Results of the normal to the sagittal plane fitting, for the pursuit and the obstacle avoidance trial. The estimated normal $\vec{n}_{sagittal}$ is expressed in the headpack coordinate system. RMSE stands for root mean square error, where the error is the residual of the least-squares problem solved. The number of samples per fit is denoted by n. For the pursuit trial, it corresponds to the number of frames in the fixation phase in the gaze calibration trial. For the obstacle avoidance trial, it corresponds to the number of frames in the final approach phase, across all the trials recorded with that bird and headpack placement

Trial	Bird	$\vec{n}_{sagittal}$	RMSE	n
Pursuit	Toothless	$[0.98,\ 0.11,\ -0.16]$	0.04	462
Obstacles	Drogon	$[0.99,\ 0.14,\ -0.01]$	0.08	5056

the camera plane. We estimated for each of them the real length of the yellow segment, based on the known real length of the green segment, and obtained a mean value of 20.1 mm.

We followed the same approach for the obstacle avoidance flights; the reference images used are shown in Figures 25 and 26. From these we estimate that the midpoint between the eyes is within 12 mm from the selected origin in the transverse direction (perpendicular to the sagittal plane), and 17.3 mm in the direction perpendicular to the headpack baseplate.

C.4 Gaze direction fit

To estimate $\vec{v}_{gaze}$ in both the pursuit and the obstacle avoidance trials, we computed the orthogonal regression line to the assumed points of fixation in the headpack coordinate system. We used the singular-value decomposition approach described in Söderkvist (2021):

$$USV^T = svd(X - \overline{X}), \tag{4}$$

$$\vec{v}_{gaze} = U(:, 1), \tag{5}$$

$$A_{gaze} = \overline{X} \tag{6}$$

where X denotes the array of sample points (size $3 \times N$, with N the number of sample points) and $\overline{X}$ is the mean of the sample points (size 3×1). The singular-value decomposition is represented by the operator $svd()$. The matrices U, S, V correspond to the matrix containing the left singular vectors in columns, the diagonal matrix containing the singular values, and the matrix containing the right singular vectors in columns, respectively. The estimated gaze direction unit vector $\vec{v}_{gaze}$ is computed as the first column of U. A_{gaze} is the centroid of the sample points, and a point in the fitted line. Table 3 shows the complete fitting results for the flights considered in the main text.

The root mean square error (RMSE) in Table 3 is computed with the error being the distance between the samples and the fitted line. We additionally computed the distance d, between the origin of the headpack coordinate system and the fitted line (Table 3).

For both the pursuit and the obstacle avoidance trial, the distance d is larger than the distance we estimated between the origin of the headpack coordinate system and the midpoint between the eyes in the previous section, Appendix C.3. Therefore, we don't expect either of these lines to closely represent a line that goes through the midpoint between the eyes. This could be due to the bird fixating on a different points than the ones we assume, due to eye movements that we are not

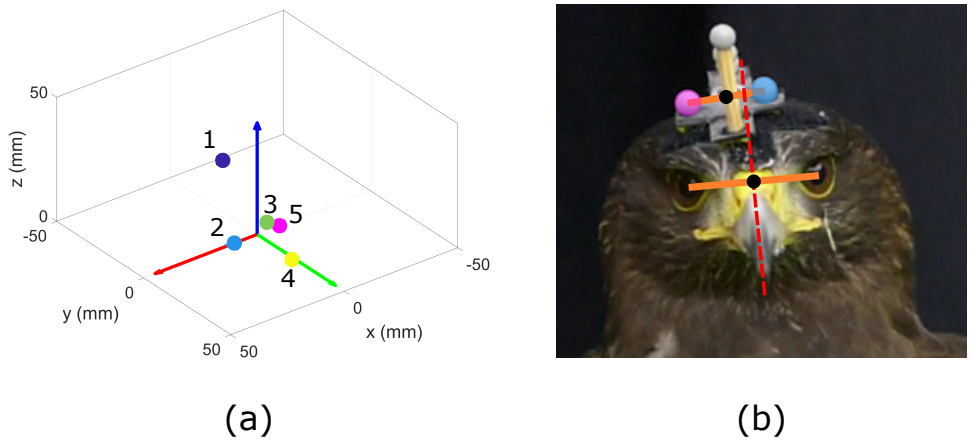


Fig. 23 Distance between selected origin and midpoint between the eyes for the pursuit flight, in the direction perpendicular to the sagittal plane. Figure (a) shows the disposition of the markers in the headpack coordinate system. The origin (which is the same for the headpack, visual and trajectory coordinate systems) lies at the midpoint between markers 2 (blue) and 5 (magenta). A front view of the bird wearing the headpack is shown in (b), with the orange segments highlighting the line between markers 2 (blue) and 5 (magenta), and the approximate line between the eyes. The midpoints of these segments are show as black dots. The trace of the sagittal plane (red dashed line) contains the midpoint between the eyes, and we estimate it based on the symmetry of the bill. We can see that the distance between the sagittal plane and the selected origin is smaller than the distance between the origin and marker 5 (10 mm)

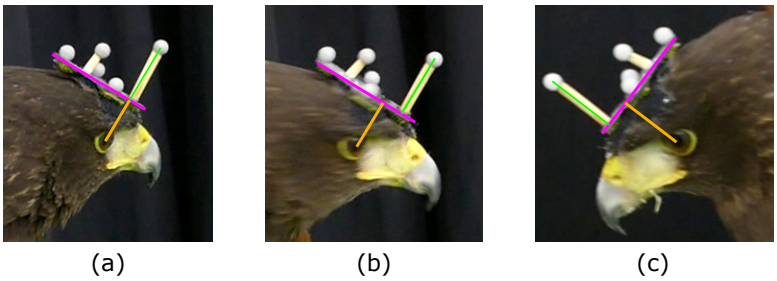


Fig. 24 Distance between selected origin and midpoint between the eyes for the pursuit flight, in the direction perpendicular to the headpack's plate. We selected three frames in which the headpack baseplate was almost perpendicular to the camera plane. We then estimated on each of them the real length of the yellow segment (the approximate distance from the headpack's baseplate to the bird's eye), based on the known length of the green segment (the height of marker 1 above the baseplate). We obtained a mean value of 20.1 mm

accounting for, or due to existing movement of the headpack relative to the head. We expect this last effect to be minimal, as we didn't find any evidence of it in the reference videos. In any case, we expect the estimated visual coordinate system to be more representative of the forward direction of the head and of the sagittal plane orientation than a coordinate system simply aligned to the headpack. The fact that our preliminary results seem in agreement with observations previously reported in the literature would support this, although it should be confirmed with more flights.

For the obstacle avoidance trial, we checked for every trial recorded with the same bird and headpack placement, whether the slope of the fitted line per trial was similar across trials. The results are shown in Table 5). We computed a deviation metric $\sigma_{\vec{v}_{gaze}}$ for the estimated unit gaze vector per trial, as the square-root of the trace of the covariance matrix:

$$\sigma_{\vec{v}_{gaze}} = \sqrt{\sigma^2_{v_{x,gaze}} + \sigma^2_{v_{y,gaze}} + \sigma^2_{v_{z,gaze}}}, \quad (7)$$

obtaining $\sigma_{\vec{v}_{gaze}} = 0.0463$. The mean gaze vector was $\overline{\vec{v}_{gaze}} = [0.11, -0.95, 0.29]$, very similar to the gaze direction estimate considering all trials aggregated (second row in Table 3). We also computed the distance from the origin of the headpack coordinate system to the fitted line per trial, which yielded a mean value of $\overline{d} = 65.56$ mm and

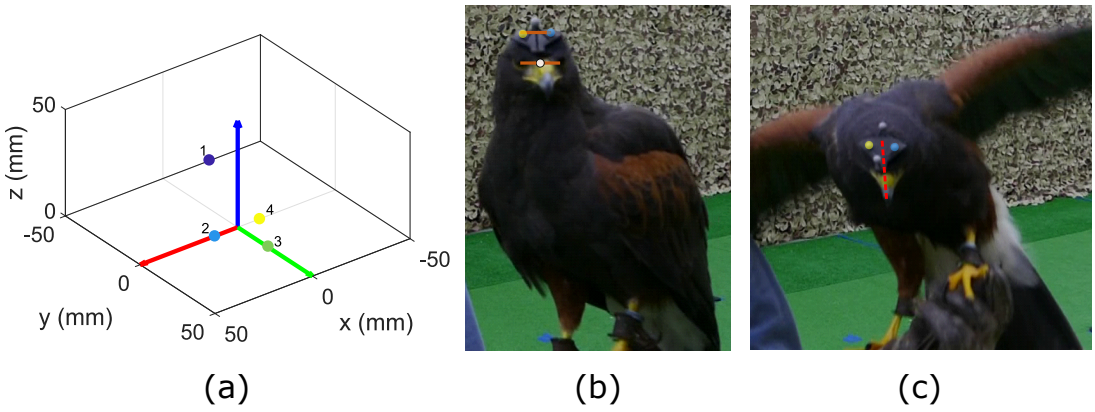


Fig. 25 Distance between selected origin and midpoint between the eyes for the obstacle avoidance flights, in the direction perpendicular to the sagittal plane. Figure (a) shows the disposition of the markers in the headpack coordinate system. The origin (which is the same for the headpack, visual and trajectory coordinate systems) lies approximately between markers 2 (blue) and 4 (yellow). A front view of the bird wearing the headpack is shown in (b), with the line between markers 2 (blue) and 4 (yellow) and the approximate line between the eyes. The midpoint of the line between the eyes (white dot) falls within the segment connecting markers 2 and 4. We therefore estimate that the distance between the midpoint between the eyes and the origin is at most, the maximum distance between the origin and one of the highlighted markers (12 mm). In (c), the red dashed line drawn following the bill's symmetry shows the approximate location of the sagittal plane, which contains the midpoint between the eyes



Fig. 26 Distance between selected origin and midpoint between the eyes for the obstacle avoidance flights, in the direction perpendicular to the headpack's plate. We selected three frames in which the bird's sagittal plane was almost parallel to the camera plane. We then estimated on each of them the real length of the yellow segment (the approximate distance from the headpack's baseplate to the bird's eye), based on the known length of the green segment (the height of marker 1 above the baseplate). We obtained a mean value of 17.3 mm

a standard deviation of $\sigma_d = 22.88$ mm, showing the variability in the landing perch midpoint trajectories per trial.

C.4.1 Retinal margins

The retinal margins define the set of directions in the visual field that project on the right or left retina. We represented the retinal margins of Harris' hawks' eyes in the estimated visual coordinate system, using data from Potier et al (2016) (Figure 27).

The authors in Potier et al (2016) measured the visual field experimentally, aligning the bird's sagittal plane with a visual perimeter and using an ophthalmoscopic reflex technique. They determined the degree of overlap between the retinal

Table 5 Results of the gaze direction fitting per trial, for all trials with the same bird and headpack placement as the rendered obstacle avoidance trial. The estimated gaze direction $\vec{v}_{gaze}$ and the point in the fitted line A_{gaze} are both expressed in the headpack coordinate system. The distance between the fitted line and the origin of the headpack coordinate system is d. RMSE stands for root mean square error, where the error is the distance between the sample points and the fitted line. The number of samples n corresponds to the number of frames in the final approach phase, for each trial

Trial	$\vec{v}_{gaze}$	A_{gaze} (mm)	d (mm)	RMSE	n
Drogon perching 1	[0.10 −0.94 0.33]	[136.50 −1108.30 306.00]	84.78	21.94	122
Drogon perching 2	[0.13 −0.92 0.37]	[112.38 −1086.61 317.02]	113.15	48.073	123
Drogon perching 3	[0.17 −0.93 0.33]	[212.42 −1094.46 310.33]	74.83	37.44	122
Drogon perching 4	[0.12 −0.95 0.28]	[122.18 −1138.82 285.44]	46.50	44.70	117
Drogon perching 5	[0.13 −0.96 0.25]	[196.22 −1115.80 263.51]	50.17	29.78	117
Drogon perching 6	[0.12 −0.93 0.36]	[149.82 −1095.98 315.56]	99.115	54.63	124
Drogon perching 7	[0.09 −0.96 0.27]	[172.20 −1126.35 248.55]	91.52	43.80	112
Drogon obstacles 1	[0.12 −0.96 0.26]	[141.08 −686.26 172.10]	59.95	62.09	57
Drogon obstacles 2	[0.07 −0.95 0.29]	[65.88 −707.29 154.66]	62.08	50.43	54
Drogon obstacles 3	[0.09 −0.96 0.28]	[19.18 −711.67 158.51]	65.75	48.99	56
Drogon obstacles 4	[0.15 −0.96 0.26]	[117.33 −699.39 160.86]	29.74	61.56	58
Drogon obstacles 5	[0.11 −0.96 0.27]	[85.59 −709.25 155.57]	44.30	15.51	53
Drogon obstacles 6	[0.09 −0.96 0.27]	[32.18 −710.18 156.70]	57.27	24.38	55
Drogon obstacles 7	[0.12 −0.96 0.27]	[109.94 −697.48 153.10]	50.87	73.59	55
Drogon obstacles 8	[0.12 −0.95 0.28]	[83.00 −707.49 154.79]	53.41	20.28	49

margins of the bird's eyes $\Delta\theta$, at several angles measured from the top of its head ϕ. We digitized the data from the paper (figures 5A, 5C and 6 from Potier et al (2016)) and fitted a smoothing spline periodic boundary conditions (8 polynomial pieces of order 2) to the relation $\Delta\theta$ vs ϕ (RMSE = 1.6°, Figure 28). From the interpolated overlap $\Delta\theta$ we derived the retinal margins for each eye, θ_{left} and θ_{right}, assuming symmetry with respect to the sagittal plane. We also assumed our estimated gaze direction corresponds to the ϕ angle where the binocular overlap is widest ($\phi = 90°$) (in line with the point raised in the supplementary figure S1 from Potier et al (2016)).

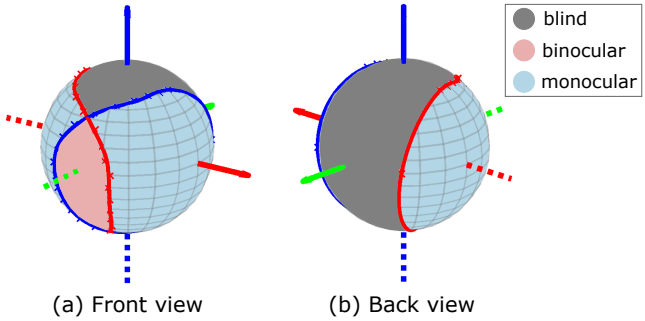


Fig. 27 Retinal margins of Harris' hawks eyes. The retinal margins for the left (blue) and right (red) eyes are shown on the unit sphere in the visual coordinate system, from the front (a) and the back (b). The data points for the retinal margins from Potier et al (2016) are represented by crosses. The blind (grey), binocular (pink) and monocular (blue) areas are shown. The inset shows the approximate position of the visual coordinate system relative to the bird's head. The parallels on the sphere are plotted every 9° in latitude and the meridians every 18° in longitude

C.5 Trajectory coordinate system

We define the trajectory coordinate system as a coordinate system with its y-axis tangent to the forward direction of the trajectory and its x-axis parallel to the horizontal:

$$\vec{y}_T = -\frac{\vec{v}}{\vec{v}}, \tag{8}$$

$$\vec{x}_T = \frac{\vec{y}_T \times \vec{z}_{world}}{\vec{y}_T \times \vec{z}_{world}}, \tag{9}$$

$$\vec{z}_T = \vec{x}_T \times \vec{y}_T, \tag{10}$$

where $\vec{x}_T, \vec{y}_T, \vec{z}_T$ represent the versors of the trajectory coordinate system, $\vec{v}$ represents the velocity vector of the head, and $\vec{z}_{world}$ represents the world z-axis. We computed the velocity vector of the head $\vec{v}$ with a central differences scheme on the head interpolated trajectory using the *gradient* function in MATLAB. In the pursuit trial, we filtered the rotations of the trajectory coordinate

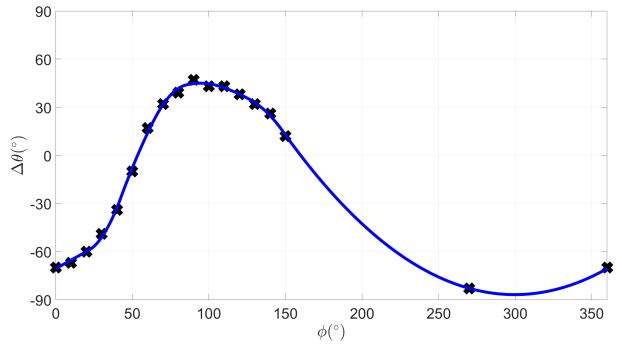


Fig. 28 Overlap between retinal margins of Harris' hawks' eyes. We digitized the overlap between the retinal margins of the bird's eyes ($\Delta\theta$) at several angles measured from the top of its head (ϕ). We fitted the data (black crosses) to a smoothing spline (blue dots) with periodic boundary conditions (RMSE = 1.6°)

system using the same approach applied to the visual coordinate system.

Appendix D Model of the lab environment

D.1 Geometry of pursuit flight

The pulleys used to guide the target's trajectory were diabolos of 6.5 cm radius and 15 cm length, glued to a metallic base of 1 cm width. We reduced their geometry in Blender to cones of 6.5 cm radius at the base, and height determined by the corresponding reference markers placed at the diabolos' axes. Some pulleys had additional markers placed on the contour but these were not used to model their geometry.

We modelled the starting boxes as cuboids of 1 m length, with their height and orientation defined by the reference markers placed at their top front vertices.

We modelled the target as a cylinder of 15 cm length and 2.54 cm diameter, based on reference images of the markers' location on the target and on the average distances measured between the markers over the whole trial (see Table 6). In determining the target's width, we took into account the diameter of the markers attached to it (6.4 mm).

We rendered one trial in which we added a texture to the wall of the motion capture room, that mimicked the texture of the camouflage netting. We did this by tiling a picture of the netting with appropriate scaling. The results are presented in the Online Resources 1 and 2.

Table 6 Mean distances between the markers fixed to the target. Computed from applying *k-means* clustering to the distances between markers labelled as belonging to the target. Length, width and diagonal refer to the dimensions of the cylinder with which we approximate the target's shape

	mean value (mm)
length	91.2
width	31.8
diagonal	97.4

D.2 Geometry of obstacle avoidance flights

We defined the geometry and locations of the perches and obstacles based on the positions of their corresponding markers (placed at the perches' edges and at the centre of the obstacles' tops).

We reduced the perches to their top rungs and modelled them as cylinders of radius 4 cm (based on measurements of the actual perches). From reference images we estimated that a line between the centres of the markers on the perch's edges would be approximately tangent to the top rung cylinder.

We modelled the obstacles as vertical cylinders of 0.3 m diameter. We estimated the offset between the larger markers' centre and the obstacles' tops by computing the mean deviation from 2 m height for these markers. We took this offset into account (14 mm, including the markers' base) when defining the obstacles' dimensions in Blender.

D.3 Dense 3D map of the curtain

The 3D mapping was carried out after collecting all bird trials of the day. To evaluate the error in the captured mesh, we recorded a motion capture trial just before carrying out the 3D mapping, that registered the static motion capture markers in the scene and on the ArUco calibration plate. We call this the mapping trial. We noticed the markers fixed to the curtain weren't clearly identifiable in the captured meshes, so we placed red tape around

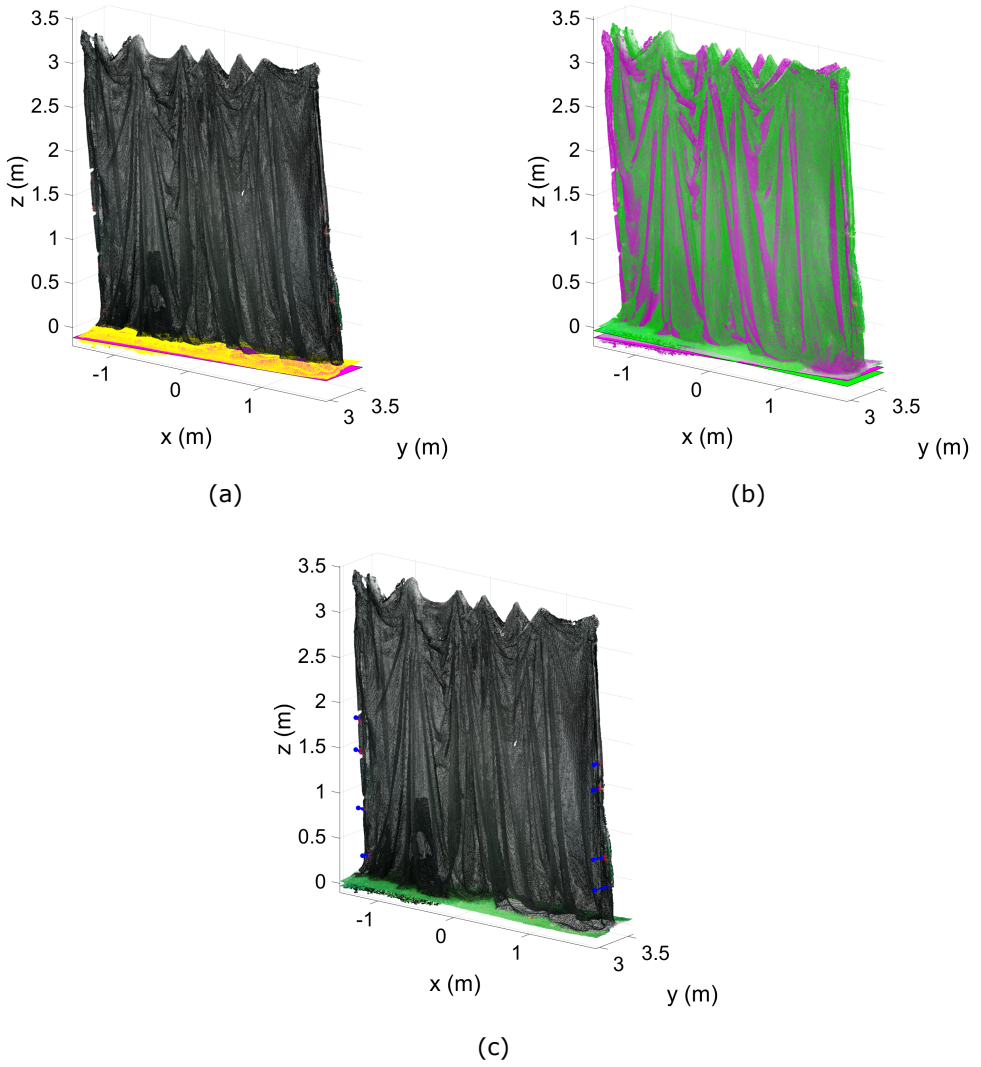


Fig. 29 Correction of the captured curtain mesh. We fitted a plane to the floor of the mesh; figure (a) shows the fitted plane (magenta) and the inlier points (yellow). We rotated the mesh so that its fitted floor plane was parallel to the floor plane of the motion capture coordinate system; figure (b) shows the original mesh and the fitted floor plane (magenta), and the rotated mesh and floor plane (green). We then translated the rotated mesh so that its floor plane contained the origin of the motion capture system (not shown). Figure (c) shows the median position of the markers fixed to the curtain during the pursuit trial (blue dots) and the segments (blue lines) connecting them to the nearest vertices from the red subset of the mesh (highlighted in red)

them to signal their approximate region prior to mapping. These pieces of tape were also in place during the bird trials.

We started the mapping procedure by setting up a local wireless network between the motion capture computer, the augmented reality smartphone and the laptop running *SemanticPaint*, making use of its streaming functionality. We then started the *SemanticPaint* application on the laptop and connected the smartphone's augmented reality app (*Tango Streamer*) to it. Once the connection was verified, the ArUco calibration plate was brought into camera view to compute the required coordinate system transformation. After visually confirming the computed coordinate system for the mesh was correct, the mapping was carried out.

To obtain the 3D map for the curtain, we cropped it from a partial mesh of the lab using MeshLab (Cignoni et al, 2008). We cleaned the resulting mesh by removing duplicate vertices and isolated pieces. We noticed that the floor plane in the captured mesh was slightly deviated from the motion capture coordinate system's floor so we corrected the mesh in MATLAB to match them.

Table 7 Distances between the motion capture markers fixed to the curtain and the subset of red vertices in the curtain's mesh. The values shown represent the maximum, minimum and mean distance (in cm) between the median position of the curtain markers and the 20 nearest neighbours in the red vertices subset of the curtain's mesh

	Pursuit trial		Mapping trial	
	Original mesh	Transformed mesh	Original mesh	Transformed mesh
Mean (cm)	9.7	9.3	9.1	8.9
Minimum (cm)	5.2	5.7	5.7	5.9
Maximum (cm)	16.4	16.6	13.0	13.3

To do this we first fitted a plane to the floor of the mesh with a reference normal equal to $[0, 0, 1]$, using *pcfitplane* in MATLAB. This is an implementation of the M-estimator sample consensus algorithm (MSAC), a variant of RANSAC. The plane was computed with a 99% confidence of finding the maximum number of inliers, with a maximum distance between an inlier point and the plane of 10 cm, and a maximum angular distance between the normal vector of the fitted plane and the reference orientation $[0, 0, 1]$ of 2.5°. The resulting angle between the fitted plane and the reference orientation was 2.4° and the mean distance between inlier points and the fitted plane was 3.8 cm.

We rotated and translated the input mesh so that the fitted floor plane contained the origin of the motion capture system, and its normal was parallel to $[0, 0, 1]$. This implied the following transforms:

$$R = \begin{pmatrix} 1.00 & 0.00 & -0.04 \\ 0.00 & 1.00 & 0.005 \\ 0.04 & -0.005 & 1.00 \end{pmatrix}, \tag{11}$$

$$\vec{t} = \begin{bmatrix} 0.00, & 0.00, & 3.53 \end{bmatrix} \tag{12}$$

where R denotes the rotation matrix applied and $\vec{t}$ the translation vector, in centimetres.

We computed the deviation from the reference motion capture markers as follows. For the markers placed on the curtain edges, we computed their reference position as the median of its coordinates across all frames, for the pursuit trial and for the mapping trial. From all the points in the mesh, we selected a subset of 400 vertices, whose colour was closest to red (i.e., the vertices whose RGB vector had lowest sum-of-square error from the red vector, $[255, 0, 0]$). For each reference marker, we computed the mean distance to the 20 nearest vertices from the red subset (see Figure 29). The mean, maximum and minimum distances for the original and the transformed mesh are shown in Table 7.

Appendix E Supplementary videos

Figure 30 shows a representation of the orthographic projection used in the rendered videos. With an orthographic projection the distortion is less than if we use an equirectangular projection, but the most peripheral parts of the bird's field of view are not included. Both projections are used to represent the RGB data in the supplementary videos. A description of the supplementary videos is summarised in Table 1.

For the pursuit trial, we exported the RGB channel in PNG format, and the rest of channels as part of a multilayer OpenEXR file. In the frame numbering shown in the video, the split between the two phases of the trial (before and after the curtain) is at frame 1132, and the interception frame corresponds to 1374.

For the obstacle avoidance trial, we exported all rendered passes as a mulitlayer OpenEXR file. For the RGB video, we used the *tonemap* function in MATLAB to read the HDR images. The analysis presented in the paper excludes the first and last 20 frames of the videos for each leg of the trial. In the frame numbering shown in the video, the landing perch becomes fully visible at frame 958 in leg 1, and frame 1997 in leg 2 of the trial. The observed saccades in the second leg of the trial occur at approximately frame 2029 and frame 2184. These are more evident in the orthographic projection.

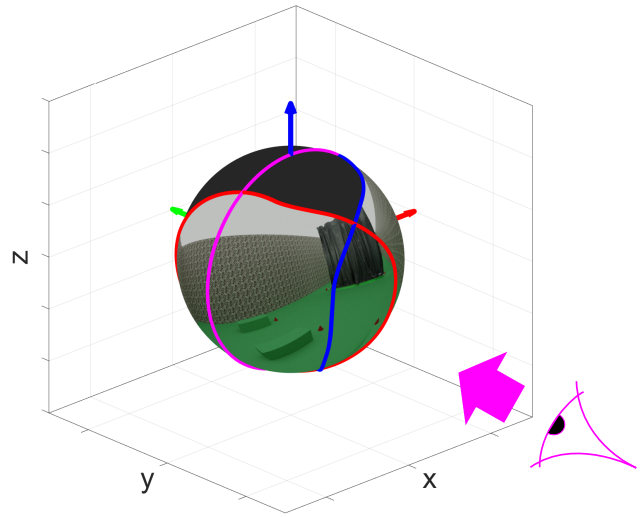


Fig. 30 Orthographic projection of the rendered output. The point of view of the rendered videos in orthographic projection is represented schematically. The rendered output (here, the RGB data) is represented in the visual field unit sphere of the bird. The retinal margins for the left (blue) and right (red) eyes are represented, as well as the blind area (black). The direction of the orthographic projection is represented with a magenta arrow and an eye icon. Note that when projecting the data as indicated by the arrow, the most peripheral parts of the field of view of the bird (beyond the magenta contour) are not included

We used the OpenEXR bindings available at https://github.com/skycaptain/openexr-matlab to read the OpenEXR files in MATLAB.

References

Aloimonos J, Weiss I, Bandyopadhyay A (1988) Active vision. Int J Comput Vis 1(4):333–356. https://doi.org/10.1007/BF00133571

Altshuler DL, Srinivasan MV (2018) Comparison of visually guided flight in insects and birds. Front Neurosci 12:157. https://doi.org/10.3389/fnins.2018.00157

Ardin P, Mangan M, Wystrach A, et al (2015) How variation in head pitch could affect image matching algorithms for ant navigation. J Comp Physiol A 201(6):585–597. https://doi.org/10.1007/s00359-015-1005-8

Ardin P, Peng F, Mangan M, et al (2016) Using an insect mushroom body circuit to encode route memory in complex natural environments. PLoS Comput Biol 12(2). https://doi.org/10.1371/journal.pcbi.1004683

Baird E, Boeddeker N, Srinivasan MV (2021) The effect of optic flow cues on honeybee flight control in wind. Proc R Soc B Biol Sci 288(1943). https://doi.org/10.1098/rspb.2020.3051

Bhagavatula PS, Claudianos C, Ibbotson MR, et al (2011) Optic flow cues guide flight in birds. Curr Biol 21(21):1794–1799. https://doi.org/10.1016/j.cub.2011.09.009

Bian X, Chandler T, Laird W, et al (2018) Integrating evolutionary biology with digital arts to quantify ecological constraints on vision-based behaviour. Methods Ecol Evol 9(3):544–559. https://doi.org/10.1111/2041-210X.12912

Bian X, Chandler T, Pinilla A, et al (2019) Now you see me, now you don't: Environmental conditions, signaler behavior, and receiver response thresholds interact to determine the efficacy of a movement-based animal signal. Front Ecol Evol 7(APR):1–16. https://doi.org/10.3389/fevo.2019.00130

Bian X, Pinilla A, Chandler T, et al (2021) Simulations with Australian dragon lizards suggest movement-based signal effectiveness is dependent on display structure and environmental conditions. Sci Rep 11(1):1–11. https://doi.org/10.1038/s41598-021-85793-3

Blender Online Community (2021) Blender - a 3D modelling and rendering package. Stichting Blender Foundation, Amsterdam, URL http://www.blender.org

Brighton CH, Taylor GK (2019) Hawks steer attacks using a guidance system tuned for close pursuit of erratically manoeuvring targets. Nat Commun 10(1):1–28. https://doi.org/10.1038/s41467-019-10454-z

Brighton CH, Thomas AL, Taylor GK (2017) Terminal attack trajectories of peregrine falcons are described by the proportional navigation guidance law of missiles. Proc Natl Acad Sci 114(51):201714,532. https://doi.org/10.1073/pnas.1714532114

Cignoni P, Callieri M, Corsini M, et al (2008) MeshLab: an Open-Source Mesh Processing Tool. In: Scarano V, Chiara RD, Erra U (eds) Eurographics Ital. Chapter Conf. The Eurographics Association, https://doi.org/10.2312/LocalChapterEvents/ItalChap/ItalianChapConf2008/129-136

Dakin R, Fellows TK, Altshuler DL (2016) Visual guidance of forward flight in hummingbirds reveals control based on image features instead of pattern velocity. Proc Natl Acad Sci 113(31):8849–8854. https://doi.org/10.1073/pnas.1603221113

Eckmeier D, Geurten BR, Kress D, et al (2008) Gaze strategy in the free flying zebra finch (Taeniopygia guttata). PLoS One 3(12). https://doi.org/10.1371/journal.pone.0003956

Fair J, Paul E, Jones J (2010) Guidelines to the use of wild birds in research. Tech. Rep. August, URL https://birdnet.org/wp-content/uploads/2017/07/guidelines{_}august2010.pdf

Frost BJ, Wise LZ, Morgan B, et al (1990) Retinotopic representation of the bifoveate eye of the kestrel (Falco sparverius) on the optic tectum. Vis Neurosci 5(3):231–239. https://doi.org/10.1017/S0952523800000304

Garrido-Jurado S, Muñoz-Salinas R, Madrid-Cuevas FJ, et al (2014) Automatic generation and detection of highly reliable fiducial markers under occlusion. Pattern Recognit 47(6):2280–2292. https://doi.org/10.1016/j.patcog.2014.01.005

Gollob C, Ritter T, Kraßnitzer R, et al (2021) Measurement of forest inventory parameters with apple ipad pro and integrated lidar technology. Remote Sens 13(16):1–35. https://doi.org/10.3390/rs13163129

Golodetz S, Sapienza M, Valentin JPC, et al (2015) SemanticPaint: a framework for the interactive segmentation of 3D scenes. arXiv Prepr pp 1–33. https://doi.org/10.1145/2751556, URL http://arxiv.org/abs/1510.03727{%}0Ahttp://dx.doi.org/10.1145/2751556https://arxiv.org/abs/1510.03727, https://arxiv.org/abs/arXiv:1510.03727

Golodetz S, Cavallari T, Lord NA, et al (2018) Collaborative Large-Scale Dense 3D Reconstruction with Online Inter-Agent Pose Optimisation. IEEE Trans Vis Comput Graph 24(11):2895–2905. https://doi.org/10.1109/TVCG.2018.2868533

Haalck L, Mangan M, Webb B, et al (2020) Towards image-based animal tracking in natural environments using a freely moving camera. J Neurosci Methods 330(108455). https://doi.org/10.1016/j.jneumeth.2019.108455

Hazan Y, Kra Y, Yarin I, et al (2015) Visual-auditory integration for visual search: a behavioral study in barn owls. Front Integr Neurosci 9:1–12. https://doi.org/10.3389/fnint.2015.00011

Holmgren CD, Stahr P, Wallace DJ, et al (2021) Visual pursuit behavior in mice maintains the pursued prey on the retinal region with least optic flow. Elife 10:e70,838. https://doi.org/10.7554/eLife.70838

Kane SA, Zamani M (2014) Falcons pursue prey using visual motion cues: new perspectives from animal-borne cameras. J Exp Biol 217(2):225–234. https://doi.org/10.1242/jeb.092403

Kane SA, Fulton AH, Rosenthal LJ (2015) When hawks attack: animal-borne video studies of goshawk pursuit and prey-evasion strategies. J Exp Biol 218(2):212–222. https://doi.org/10.1242/jeb.108597

Kano F, Walker J, Sasaki T, et al (2018) Head-mounted sensors reveal visual attention of free-flying homing pigeons. J Exp Biol 221(17):1–13. https://doi.org/10.1242/jeb.183475

Kern R, Van Hateren JH, Michaelis C, et al (2005) Function of a fly motion-sensitive neuron matches eye movements during free flight. PLoS Biol 3(6):1130–1138. https://doi.org/10.1371/journal.pbio.0030171

Kress D, Van Bokhorst E, Lentink D (2015) How lovebirds maneuver rapidly using super-fast head saccades and image feature stabilization. PLoS One 10(6):1–24. https://doi.org/10.1371/journal.pone.0129287

Land MF (1999) Motion and vision: why animals move their eyes. J Comp Physiol - A Sensory, Neural, Behav Physiol 185(4):341–352. https://doi.org/10.1007/s003590050393

Lev-Ari T, Gutfreund Y (2018) Interactions between top-down and bottom-up attention in barn owls (Tyto alba). Anim Cogn 21(2):197–205. https://doi.org/10.1007/s10071-017-1150-2

Lin HT, Ros IG, Biewener AA (2014) Through the eyes of a bird: modelling visually guided obstacle flight. J R Soc Interface 11(96):1–12. https://doi.org/10.1098/rsif.2014.0239

McClune DW (2018) Joining the dots: reconstructing 3D environments and movement paths using animal-borne devices. Anim Biotelemetry 6:5. https://doi.org/10.1186/s40317-018-0150-6

Merel J, Tunyasuvunakool S, Ahuja A, et al (2020) Catch & carry: reusable neural controllers for vision-guided whole-body tasks. ACM Trans Graph 39(4):1–14. https://doi.org/10.1145/3386569.3392474, https://arxiv.org/abs/arXiv:1911.06636

Mildenhall B, Srinivasan PP, Tancik M, et al (2020) NeRF: representing scenes as neural radiance fields for view synthesis. In: ECCV, pp 405–421, https://doi.org/10.1007/978-3-030-58452-8_24, 2003.08934

Minano S, Taylor GK (2021) Through hawks' eyes: reconstructing a bird's visual field in flight to study gaze strategy and attention during perching and obstacle avoidance. bioRxiv https://doi.org/10.1101/2021.06.16.446415

Mitkus M, Potier S, Martin GR, et al (2018) Raptor vision. In: Oxford Res. Encycl. Neurosci. March, Oxford University Press, p 1–38, https://doi.org/10.1093/acrefore/9780190264086.013.232

Motion Lab Systems (2021) The C3D file format: a technical user guide. Tech. rep., Motion Labs Systems, Baton Rouge, Louisiana, URL https://www.c3d.org/docs/C3D{_}User{_}Guide.pdf

Naik H (2021) XR for all: closed-loop visual stimulation techniques for human and non-human animals. PhD thesis, Technische Universität München, Munich, URL http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:91-diss-20210308-1554403-1-6

Naik H, Bastien R, Navab N, et al (2020) Animals in Virtual Environments. IEEE Trans Vis Comput Graph 26(5):2073–2083. https://doi.org/10.1109/TVCG.2020.2973063, https://arxiv.org/abs/arXiv:1912.12763

Neumann TR (2002) Modeling insect compound eyes: space-variant spherical vision. Proc 2nd Int Work Biol Motiv Comput Vis (BMCV 2002) LNCS 2525:360–367. https://doi.org/10.1007/3-540-36181-2_36

Ochs MF, Zamani M, Gomes GMR, et al (2016) Sneak peek: Raptors search for prey using stochastic head turns. Auk 134(1):104–115. https://doi.org/10.1642/auk-15-230.1

Ozawa Y (2010) Vision and movement in birds. PhD thesis, University of Oxford, Oxford, URL https://isni.org/isni/0000000427104029

Potier S, Bonadonna F, Kelber A, et al (2016) Visual abilities in two raptors with different ecology. J Exp Biol 219(17):2639–2649. https://doi.org/10.1242/jeb.142083

Potier S, Lieuvin M, Pfaff M, et al (2020) How fast can raptors see? J Exp Biol 223(1). https://doi.org/10.1242/jeb.209031

Prisacariu VA, Kähler O, Golodetz S, et al (2017) InfiniTAM v3: a framework for large-scale 3D reconstruction with loop closure. arXiv Prepr https://arxiv.org/abs/arXiv:1708.00783

Raudies F, Mingolla E, Neumann H (2012) Active gaze control improves optic flow-based segmentation and steering. PLoS One 7(6). https://doi.org/10.1371/journal.pone.0038446

Ravi S, Bertrand O, Siesenop T, et al (2019) Gap perception in bumblebees. J Exp Biol 222(2):1–10. https://doi.org/10.1242/jeb.184135

Ravi S, Siesenop T, Bertrand OJ, et al (2022) Bumblebees display characteristics of active vision during robust obstacle avoidance flight. J Exp Biol 225(4). https://doi.org/10.1242/jeb.243021

Risse B, Mangan M, Stürzl W, et al (2018) Software to convert terrestrial LiDAR scans of natural environments into photorealistic meshes. Environ Model Softw 99:88–100. https://doi.org/10.1016/j.envsoft.2017.09.018

Romero-Ramirez FJ, Muñoz-Salinas R, Medina-Carnicer R (2018) Speeded up detection of squared fiducial markers. Image Vis Comput 76:38–47. https://doi.org/10.1016/j.imavis.2018.05.004

Ros IG, Biewener AA (2016) Optic flow stabilizes flight in ruby-throated hummingbirds. J Exp Biol 219(16):2443–2448. https://doi.org/10.1242/jeb.128488

Ros IG, Biewener AA (2017) Pigeons (C. livia) follow their head during turning flight: Head stabilization underlies the visual control of flight. Front Neurosci 11(DEC):1–12. https://doi.org/10.3389/fnins.2017.00655

Rothkopf CA, Ballard DH (2009) Image statistics at the point of gaze during human navigation. Vis Neurosci 26(1):81–92. https://doi.org/10.1017/S0952523808080978

Ruppertsberg AI, Wuerger SM, Bertamini M (2003) The chromatic input to global motion perception. Vis Neurosci 20(4):421–428. https://doi.org/10.1017/S0952523803204077

Schiffner I, Srinivasan MV (2015) Direct evidence for vision-based control of flight speed in budgerigars. Sci Rep 5:1–7. https://doi.org/10.1038/srep10992

Schonberger JL, Frahm JM (2016) Structure-from-Motion revisited. In: 2016 IEEE Conf. Comput. Vis. Pattern Recognit., pp 4104–4113, https://doi.org/10.1109/CVPR.2016.445

Schulte P, Zeil J, Stürzl W (2019) An insect-inspired model for acquiring views for homing. Biol Cybern 113(4):439–451. https://doi.org/10.1007/s00422-019-00800-1

Seara JF, Schmidt G (2004) Intelligent gaze control for vision-guided humanoid walking: Methodological aspects. Rob Auton Syst 48(4 SPEC. ISS.):231–248. https://doi.org/10.1016/j.robot.2004.07.003

Seara JF, Lorch O, Schmidt. G (2001) Gaze Control for Goal-Oriented Humanoid Walking. Proc IEEE/RAS Int Conf Humanoid Robot (Humanoids), S 187–195, Tokio, Japan

Seara JF, Strobl KH, Schmidt G (2002) Information management for gaze control in vision guided biped walking. IEEE Int Conf Intell Robot Syst 1(October):31–36. https://doi.org/10.1109/IRDS.2002.1041357

Shoemake K (1985) Animating rotation with quaternion curves. In: Proc. 12th Annu. Conf. Comput. Graph. Interact. Tech., pp 245–254, https://doi.org/10.1145/325165.325242

Söderkvist I (2021) Using SVD for some fitting problems. Lecture notes for data analysis PhD course from Lulea University of Technology. Retrieved on 30 April 2021, URL https://www.ltu.se/cms{_}fs/1.51590!/svd-fitting.pdf

Stowers JR, Hofbauer M, Bastien R, et al (2017) Virtual reality for freely moving animals. Nat Methods 14(10):995–1002. https://doi.org/10.1038/nmeth.4399

Stuerzl W, Boeddeker N, Dittmar L, et al (2010) Mimicking honeybee eyes with a 280 deg field of view catadioptric imaging system. Bioinspiration and Biomimetics 5(3). https://doi.org/10.1088/1748-3182/5/3/036002

Stuerzl W, Grixa I, Mair E, et al (2015) Three-dimensional models of natural environments and the mapping of navigational information. J Comp Physiol A Neuroethol Sensory, Neural, Behav Physiol 201(6):563–584. https://doi.org/10.1007/s00359-015-1002-y

Stuerzl W, Zeil J, Boeddeker N, et al (2016) How wasps acquire and use views for homing. Curr Biol 26:470–482. https://doi.org/10.1016/j.cub.2015.12.052

Tammero LF, Dickinson MH (2002a) Collision-avoidance and landing responses are mediated by separate pathways in the fruit fly, Drosophila melanogaster. J Exp Biol 205(18):2785–2798. https://doi.org/10.1242/jeb.205.18.2785

Tammero LF, Dickinson MH (2002b) The influence of visual landscape on the free flight behavior of the fruit fly Drosophila melanogaster. J Exp Biol 205:327–343. https://doi.org/10.1007/bf00355543

Tancik M, Casser V, Yan X, et al (2022) Block-NeRF: Scalable Large Scene Neural View Synthesis. arXiv URL http://arxiv.org/abs/2202.05263, https://arxiv.org/abs/arXiv:2202.05263

Tatsumi S, Yamaguchi K, Furuya N (2022) ForestScanner: A mobile application for measuring and mapping trees with LiDAR-equipped iPhone and iPad. bioRxiv pp 95–100. https://doi.org/10.1101/2021.12.11.472207

Taylor GK, Bacic M, Bomphrey RJ, et al (2008) New experimental approaches to the biology of flight control systems. J Exp Biol 211(2):258–266. https://doi.org/10.1242/jeb.012625

Taylor LA, Taylor GK, Lambert B, et al (2019) Birds invest wingbeats to keep a steady head and reap the ultimate benefits of flying together. PLoS Biol 17(6):1–20. https://doi.org/10.1371/journal.pbio.3000299

Tucker VA (2000) The deep fovea, sideways vision and spiral flight paths in raptors. J Exp Biol 203(Pt 24):3745–3754. https://doi.org/10.1016/S0169-5347(01)02393-X, URL http://www.ncbi.nlm.nih.gov/pubmed/11076738

Tuia D, Kellenberger B, Beery S, et al (2022) Perspectives in machine learning for wildlife conservation. Nat Commun 13(1):1–15. https://doi.org/10.1038/s41467-022-27980-y

Tyrrell LP, Goller B, Moore BA, et al (2018) The orientation of visual space from the perspective of hummingbirds. Front Neurosci 12(16):1–9. https://doi.org/10.3389/fnins.2018.00016

Warrick ADR, Bundle MW, Dial KP, et al (2002) Bird maneuvering flight: blurred bodies, clear heads. Integr Comp Biol 42(1):141–148. https://doi.org/10.1093/icb/42.1.141

Windsor SP, Taylor GK (2017) Head movements quadruple the range of speeds encoded by the insect motion vision system in hawkmoths. Proc R Soc B Biol Sci 284(1864). https://doi.org/10.1098/rspb.2017.1622

Yorzinski JL (2021) Great-tailed grackles can independently direct their eyes toward different targets. Exp Brain Res 239(7):2119–2126. https://doi.org/10.1007/s00221-021-06122-8

Yorzinski JL, Platt ML (2014) Selective attention in peacocks during predator detection. Anim Cogn 17(3):767–777. https://doi.org/10.1007/s10071-013-0708-x

Yorzinski JL, Patricelli GL, Babcock JS, et al (2013) Through their eyes: selective attention in peahens during courtship. J Exp Biol 216(16):3035–3046. https://doi.org/10.1242/jeb.087338

Yorzinski JL, Patricelli GL, Platt ML, et al (2015) Eye and head movements shape gaze shifts in Indian peafowl. J Exp Biol 218(23):3771–3776. https://doi.org/10.1242/jeb.129544

Zeil J, Narendra A, Stürzl W (2014) Looking and homing: how displaced ants decide where to go. Philos Trans R Soc B Biol Sci 369(1636). https://doi.org/10.1098/rstb.2013.0034

Zhang H, Starke S, Komura T, et al (2018) Mode-adaptive neural networks for quadruped motion control. ACM Trans Graph 37(4):1–11. https://doi.org/10.1145/3197517.3201366

Zhu L, Mangan M, Webb B (2020) Spatio-temporal memory for navigation in a mushroom body model. In: Vouloutsi V, Mura A, Tauber F, et al (eds) Biomim. Biohybrid Syst. 9th Int. Conf. Living Mach. 2020, Freiburg, Ger. July 28–30, 2020, Proc., vol 12413 LNAI. Springer-Verlag, Berlin, Heidelberg, pp 415–426, https://doi.org/10.1007/978-3-030-64313-3_39