The application of wearable cameras, accelerometers and motion capture for the analysis of complex gait

Stefan Bostock

St Peter's College
University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Hilary Term 2022

Abstract

Gait analysis is an increasingly useful biomedical tool, with applications including diagnosis, treatment planning, and disease monitoring. Most gait analysis is conducted in a clinical setting using motion capture, however its greatest impact and most pertinent uses are outside the laboratory in a free-living environment. To enable the transition from clinic to real-world, wearable technologies have been developed to provide affordable, accessible, remote and unobtrusive assessment. The most popular by far are accelerometry-based devices, however their reliability as free-living tools is questionable due to: the limitations of complex gait definitions; lack of robust experimental validation; an absence of ground truth in free-living environments; and their sensitivity to a range of parameters relating to the devices themselves and how they are implemented.

It is the aim of this thesis to confront each of these issues, enabling a better understanding and measurement of complex gait. To achieve this, novel spatio-temporal gait definitions and visualisations are proposed to accommodate all forms of gait. It is also hypothesised that computer vision methods applied to video from a wearable camera can extract a range of gait features in all environments, providing less sensitivity in accuracy to complex movement than current accelerometry techniques.

Experimentally, a study was devised and conducted using novel protocols to better imitate free-living gait in a motion capture laboratory; as well as utilising a wearable camera and range of accelerometers. It is shown that simplistic straight protocols are not representative of complex gait, and that all wearables suffer substantial loss of accuracy when submitted to such protocols. Furthermore, given robust definitions and a simple set up, complex gait can be accurately measured by motion capture; however, gait features exist on a spectrum that is highly sensitive to definition. Both points elucidate the need for unification in the field, and the questionability of the intra-study comparability and validity of gait measuring methods.

Extracting gait features from wearable camera video yielded varying results. For example, turns were detected and measured with greater than 90% accuracy, step detection precision and stride length correlation with motion capture ground truth ranged from 0.78% to 0.98% and 0.55 to 0.80 respectively depending on protocol, whereas step width could not be accurately measured at all. However, these fared comparably to all research-grade and consumer accelerometers; especially when compared step-by-step. Of these, the foot-mounted devices were by far the most reliable, whereas smartwatches and smartphones showed expected high accuracy during standard validation protocols, but this in no way transposed to imitation free-living movements.

The application of wearable cameras, accelerometers and motion capture for the analysis of complex gait



Stefan Bostock

St Peter's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary Term 2022

Acknowledgements

Institutional

I would like to thank Professors Maarten de Vos, Amy Zavatsky, and Victor Prisacariu for all of their help and guidance over the years; even through complication, desperation and a plague or two. Whether it be the perspicacious appraisal of work, the attainment of data and resources, or the navigation of administration; you were always kind and ready to help. I would also like to thank Professor Laurent Servais for his passion, interest and time; as well as Dr Silvia Del Din and Professor Lynn Rochester for provision of their data sets all those years ago. I wish you all great future happiness and success!

Additionally, I would like to thank all those at the Oxford Gait Lab for their friendliness, help and patience - but particularly Julie and Bart. My apologies for getting in the way of the important work that you do, especially at the most difficult of times. Thank you for your help with everything experimental - I thought by doing computational work I had escaped the lab!

Finally, I would like to send my warmest regards over to Switzerland and the whole team at Roche who were so welcoming and friendly when occasion was found to visit; but especially Florian, Alan, Detlef and Michael. It was inspiring to see the clinical application of the academic work; and I will be proud if I ever manage to emulate your work to any degree.

Personal

Well, what a sojourn its been. Well over half a decade, and a most peculiar one at that - for which I have of course to thank the unrelenting passage of time. Unexpectedly, the real thesis was the friends that were made along the way - and I would like to quickly thank as many of them as possible.

I am truly grateful for all the amazing people I got to work alongside before the department slowly (then suddenly) emptied, and whom I am honoured to now call friends (of a particularly incredible type - those who can also help you out with really befuddling machine learning problems), especially: Kirubin, Navin, Andy, Eoin, Maddy, the two Ollies, Julius, Fernando, Huy, and many more besides. You all provided incandescent company in an otherwise windowless world!

Beyond the IBME, my time at Oxford will be mainly marked by those unique people I got the pleasure of experiencing it alongside. Oxford feels very much like a peculiar and entirely too-brief union of itinerants, all of whom will no doubt go on to do fascinating and wonderful things. Monty, Nacho, Rodrigo, Sofia, and the gang - you made Oxford the best place I have ever lived. Beyond Cowley and recent years; Ciara, Anka, Harriet, Mike, Nat, Dan, Kaya, Olivia, Phil and plenty more still - you are missed, and I hope you will always feel welcome and compelled to suddenly turn

up at my door. I hope to see you all soon to reminisce about sunsets in South Park, formals in hall, board games, and boxing, and races, and more.

Beyond the University there has been some truly crucial players that have helped get me to this stage. I wish to send my kindest thanks to the doctors and nurses at the Churchill and JR - where I probably spent more time than college. Especially Graham, Mel and Teena - you have genuinely changed my life in a way I did not think possible. On a very different note, over a year of this was undertaken away from Oxford while working to actually put the learning to some use. I wish to offer my huge gratitude to both Sam and Jess for providing motivation, distraction, inspiration and big beers - all while putting up with my interminable complaints and abandonment of activities.

Last and of course most importantly, I want to thank my family for their faith and determination to see me through not just this but the entire journey to date. I just about managed to finish my education before the next generation started it - so that's got to count for something, surely! It has been great being so close to you many of you these past years, I will miss being able to cycle or drive down of a Sunday afternoon. Grandma, thank you for always providing unfaltering inspiration for perseverance. Most of all, thank you Mum for all your love, and entirely unfounded yet unwavering belief in me!

To everyone else, and I am forgetting many, I should really get this finally handed in now - so my apologies, I will thank each of you in person (not least to collect on the various bets and bargains that were stricken)!

Abstract

Gait analysis is an increasingly useful biomedical tool, with applications including diagnosis, treatment planning, and disease monitoring. Most gait analysis is conducted in a clinical setting using motion capture, however its greatest impact and most pertinent uses are outside the laboratory in a free-living environment. To enable the transition from clinic to real-world, wearable technologies have been developed to provide affordable, accessible, remote and unobtrusive assessment. The most popular by far are accelerometry-based devices, however their reliability as free-living tools is questionable due to: the limitations of complex gait definitions; lack of robust experimental validation; an absence of ground truth in free-living environments; and their sensitivity to a range of parameters relating to the devices themselves and how they are implemented.

It is the aim of this thesis to confront each of these issues, enabling a better understanding and measurement of complex gait. To achieve this, novel spatio-temporal gait definitions and visualisations are proposed to accommodate all forms of gait. It is also hypothesised that computer vision methods applied to video from a wearable camera can extract a range of gait features in all environments, providing less sensitivity in accuracy to complex movement than current accelerometry techniques.

Experimentally, a study was devised and conducted using novel protocols to better imitate free-living gait in a motion capture laboratory; as well as utilising a wearable camera and range of accelerometers. It is shown that simplistic straight protocols are not representative of complex gait, and that all wearables suffer substantial loss of accuracy when submitted to such protocols. Furthermore, given robust definitions and a simple set up, complex gait can be accurately measured by motion capture; however, gait features exist on a spectrum that is highly sensitive to definition. Both points elucidate the need for unification in the field, and the questionability of the intra-study comparability and validity of gait measuring methods.

Extracting gait features from wearable camera video yielded varying results. For example, turns were detected and measured with greater than 90% accuracy, step detection precision and stride length correlation with motion capture ground truth ranged from 0.78% to 0.98% and 0.55 to 0.80 respectively depending on protocol, whereas step width could not be accurately measured at all. However, these fared comparably to all research-grade and consumer accelerometers; especially when compared step-by-step. Of these, the foot-mounted devices were by far the most reliable, whereas smartwatches and smartphones showed expected high accuracy during standard validation protocols, but this in no way transposed to imitation free-living movements.

Contents

List of Figures **xi**

List of Tables **xv**

List of Abbreviations **xvii**

1 Introduction **1**
1.1 Motivation . . . 1
1.2 Research aims . . . 4
1.3 Research approach & thesis structure . . . 4
1.4 Key contributions . . . 7

2 Background & literature review **9**
2.1 Introduction . . . 9
2.2 Gait and the clinic . . . 9
2.2.1 Gait assessment & digital biomarkers . . . 10
2.2.2 Wearables & digital biomarkers . . . 11
2.2.3 Gait cycle . . . 12
2.3 Gait features . . . 12
2.3.1 Walking bouts & speed . . . 13
2.3.2 Turns . . . 14
2.3.3 Steps . . . 14
2.4 Free-living gait . . . 15
2.4.1 The frequencies of human motion . . . 17
2.5 Gait assessment by camera . . . 17
2.5.1 Motion capture & static cameras . . . 17
2.5.2 Wearable cameras . . . 18
2.6 Gait assessment by inertial measurement unit . . . 19
2.6.1 Precision & personalisation . . . 21

3 Essential methods & technical background **23**
3.1 Introduction . . . 23
3.2 Gait & graphs . . . 23
3.2.1 Axes & planes . . . 24
3.2.2 Step maps . . . 24
3.2.3 Gait events . . . 25
3.3 Optics & multiview geometry . . . 30
3.3.1 Projections, planes & homographies . . . 31
3.3.2 Homography estimation . . . 38

3.3.3 Relative camera pose estimation . . . 43
3.3.4 Camera calibration . . . 48
3.4 Object detection . . . 49
3.4.1 Classification statistics . . . 49
3.4.2 Assessment metrics . . . 50
3.5 Conclusion . . . 52

4 Validation study & data sets 53
4.1 Introduction . . . 53
4.2 Subjects . . . 54
4.3 Hardware, software & set up . . . 55
4.3.1 Camera . . . 55
4.3.2 Accelerometers . . . 56
4.3.3 Motion capture . . . 57
4.4 Protocols . . . 59
4.4.1 Straight walking protocol . . . 59
4.4.2 Turning protocol . . . 61
4.4.3 Activities of daily living protocols . . . 61
4.5 Ethical considerations . . . 65
4.5.1 Data management & security . . . 65
4.5.2 COVID-19 considerations . . . 65
4.6 Discussion . . . 66
4.7 Newcastle data set . . . 68
4.8 Ad hoc data . . . 69
4.9 Conclusion . . . 70

5 Developing robust definitions for spatio-temporal gait features 71
5.1 Introduction . . . 71
5.2 Gait feature definitions & theory . . . 72
5.2.1 Defining direction of progression . . . 73
5.2.2 Defining turning . . . 75
5.2.3 Defining the stride . . . 78
5.2.4 Defining stride length, step length and step width . . . 90
5.2.5 Defining walking . . . 93
5.3 Methods . . . 99
5.3.1 Preprocessing . . . 99
5.3.2 Filtering . . . 100
5.3.3 Differentiation . . . 101
5.3.4 Orientation . . . 101
5.4 Results . . . 102
5.4.1 Motion capture reliability . . . 102

5.4.2 Parameter sensitivity . . . 103
5.5 Discussion . . . 104
5.5.1 Motion capture reliability . . . 104
5.5.2 Parameter sensitivity . . . 107
5.5.3 Marker set . . . 108
5.5.4 Gait feature definitions . . . 109
5.6 Conclusion . . . 110

6 Measuring human movement with a wearable camera 111
6.1 Introduction . . . 111
6.1.1 Technical background & theory . . . 113
6.1.2 Unique challenges & key considerations . . . 114
6.2 Methods . . . 117
6.2.1 Template matching . . . 117
6.2.2 Pose estimation . . . 120
6.2.3 Feature extraction . . . 122
6.3 Results . . . 122
6.3.1 Rotational speed & turns . . . 123
6.3.2 Translational speed . . . 126
6.4 Discussion . . . 128
6.4.1 Measuring rotational speed . . . 128
6.4.2 Turn detection & measurement . . . 130
6.4.3 Translational speed . . . 135
6.4.4 Subject dependency . . . 137
6.4.5 Translational speed during linear walking . . . 138
6.5 Conclusion . . . 139

7 Foot tracking & gait features from a wearable camera 141
7.1 Introduction . . . 141
7.1.1 Object tracking challenges . . . 143
7.2 Methods . . . 151
7.2.1 Model selection & training . . . 152
7.2.2 Foot tracking . . . 153
7.2.3 Foot trajectories to step maps . . . 158
7.3 Results . . . 159
7.3.1 Foot detection . . . 161
7.3.2 Step detection . . . 161
7.3.3 Step features . . . 162
7.4 Discussion . . . 163
7.4.1 Foot detection . . . 163
7.4.2 Step detection . . . 165
7.4.3 Step features . . . 166
7.5 Conclusion . . . 169

8 Comparing gait analysis from multiple accelerometry-based wearable devices 171
8.1 Introduction . . . 171
8.2 Methods . . . 173
8.2.1 Orientation . . . 174
8.2.2 Signal filtering . . . 176
8.2.3 Step detection . . . 176
8.3 Results . . . 184
8.4 Discussion . . . 184
8.4.1 Inter-device comparison . . . 184
8.4.2 Comparison to published results . . . 188
8.4.3 Algorithm selection . . . 191
8.4.4 Parameter selection . . . 192
8.4.5 The importance of the individual . . . 193
8.4.6 The importance of ADL protocols and free-living gait . . . 197
8.4.7 Sensor location & orientation . . . 200
8.4.8 Other methods, limitations and uncontrolled parameters . . . 201
8.5 Conclusion . . . 203

9 Conclusions & Future Work 205
9.1 Summary of results & contributions . . . 205
9.1.1 Defining and assessing complex gait . . . 205
9.1.2 Gait analysis with a wearable camera . . . 207
9.1.3 Accelerometry-based wearables during complex gait . . . 208
9.2 Future work . . . 208
9.2.1 Expansion and additional use of the study data . . . 208
9.2.2 Further development of the wearable-camera system . . . 209
9.2.3 Expanding on the multiple accelerometry device comparison . . . 210
9.2.4 General observations & trends . . . 211

A Appendix: Accelerometry Thresholds 213

References 217

List of Figures

1.1 Thesis structure 5

2.1 The gait cycle 12

3.1 Anatomical and camera axes 25
3.2 Step map example 26
3.3 GET graph examples 27
3.4 GET graph examples with event catagorisation 29
3.5 Linear assignment problem example 31
3.6 The pinhole camera model 32
3.7 Camera intrinsics 34
3.8 Back-projection of image points to rays 35
3.9 Planar induced homography 36
3.10 Multiple homography decomposition solutions 44
3.11 Triangulation 46
3.12 Precision-Recall curve example 52

4.1 Study participant statistics 55
4.2 Experimental device and equipment set up 58
4.3 Protocol schematics 60
4.4 Protocol durations 61
4.5 Motor-cognitive dual tasking exercise sheets 64
4.6 Newcastle data set activities and context 69

5.1 Spatial gait features during non-linear walking 74
5.2 Travelling through a door foot map 75
5.3 Direction of progression 76
5.4 Heel and toe movement during pivot turns. 80
5.5 Foot speed threshold 81
5.6 Gaussian foot velocity model example 83
5.7 Gaussian foot velocity model 84
5.8 Sensitivity of proportion of stride detected 90
5.9 Stride length, step length, and step width definitions 92
5.10 Gait with no conveyance 96
5.11 Camera orientation from motion capture 102
5.12 Motion capture marker location accuracy 104
5.13 Turn detection sensitivity to minimum total turn angle and threshold turn speed. 105
5.14 Stride detection sensitivity to foot speed and stride length 106

6.1 Feature-based visual odometry . . . 114
6.2 Template matching . . . 118
6.3 The limitations of image feature detection. . . . 119
6.4 Normalised cross-correlation . . . 120
6.5 Rotational speed results. . . . 124
6.6 Turn angle . . . 126
6.7 Mean and maximum turn speeds . . . 127
6.8 Translational speed results. . . . 129
6.9 Examples of measuring rotational speed . . . 131
6.10 Examples of turn detection . . . 133
6.11 Examples of measuring translational speed. . . . 134
6.12 Relative distribution of translational speed by protocol . . . 137
6.13 Rotational and translational speed accuracy by subject. . . . 138
6.14 Use of VO for steady-state gait around a running track. . . . 140

7.1 Object tracking attribute prevalence and importance. . . . 145
7.2 Analysis of the object tracking attributes in the data sets . . . 147
7.3 Object tracking attributes example frames . . . 148
7.4 Blurring and false positives . . . 150
7.5 Detection overlap . . . 155
7.6 Erroneous detection removal . . . 155
7.7 Tracklets example . . . 157
7.8 Left and right foot determination . . . 158
7.9 Foot movement tracking . . . 160
7.10 Foot map and step feature extraction . . . 160
7.11 PR and mAP sensitivity to IoU . . . 161
7.12 Step feature results . . . 164
7.13 False positive strides as a function of stride length . . . 165
7.14 Step count results . . . 167
7.15 Step feature variation with walking speed and acceleration . . . 168

8.1 Step count sensitivity to SSD and AYM for lumbar IMU . . . 178
8.2 Step count percentage error . . . 180
8.3 Lumbar-IMU and Smartphone step count sensitivity to AYM and SSD 181
8.4 Comparing lumbar-mounted IMU and hip-pocket Smartphone AP accelerometry signals . . . 182
8.5 Feet-IMU step count sensitivity to A3M and SSD . . . 183
8.6 Smartwatch step count sensitivity to λ and σ . . . 184
8.7 Comparison of step count accuracy between multiple devices . . . 185
8.8 Threshold sensitivity for foot-IMU algorithm . . . 194

8.9 Threshold sensitivity for all devices between the Straight and ADL3 protocols 195
8.10 Dependency of step counting accuracy on subject 196
8.11 Impact of subject characteristics on step counting accuracy 198

A.1 Threshold sensitivity for lumbar-IMU algorithm 214
A.2 Threshold sensitivity for smartphone-IMU algorithm 215
A.3 Threshold sensitivity for smartwatch-IMU algorithm 216

List of Tables

4.1 Distribution of walking locations in the Newcastle data set 70

5.1 Walking speeds from literature 88
5.2 Gait asymmetries from literature 89
5.3 Gait cadences from literature 91
5.4 Bout definitions from literature 95
5.5 Gap filling statistics for motion capture signals 103

6.1 Turn event detection results 125
6.2 Continuous turn detection results 125
6.3 Turn detection precision results 125

7.1 Object tracking attributes 144
7.2 Step detection results 162
7.3 Step features results 163

List of Abbreviations

2D . . . . . Two dimensional

2MWT . . . . Two minute walk test

3D . . . . . Three dimensional

6MWT . . . . Six minute walk test

A3M . . . . . The magnitude of total triaxial acceleration

AD . . . . . Alzheimer's disease

ADL . . . . . Activity of daily living

AP . . . . . Antero-posterior axis (Chapter 8)

AP . . . . . Average precision (Chapter 3)

API . . . . . Application programming interface

AR . . . . . Augmented reality

AYM . . . . . Modulus of vertical acceleration

BC . . . . . Background clutter

CGI . . . . . Computer generated image

CNN . . . . . Convolutional neural network

COCO . . . . . Common objects in context

CoM . . . . . Centre of mass

COPD . . . . . Chronic obstructive pulmonary disease

CSV . . . . . Comma separated value

CUREC . . . . Central University Research Ethics Committee

CWT . . . . . Continuous wavelet transform

DEF . . . . . Deformation

DiMe . . . . . Digital Medicine Society

DLB . . . . . Dementia with Lewy bodies

DLT . . . . . Direct linear transform

DMD . . . . . Duchennes muscular dystrophy

DMO . . . . . Digital mobility outcome

DNN . . . . . Deep neural network

DoP . . . . . Direction of progression

EDSS . . . . . Extended disability status scale

EMA . . . . . European Medicine's Agency

EMG . . . . . Electromyography

EPSRC . . . . Engineering and Physical Sciences Research Council

FES . . . . . Functional electronic stimulation

FFT . . . . . Fast Fourier transform

FM . . . . . Fast motion

FN . . . . . False negative

FOG . . . . . Freezing of gait

FOV . . . . . Field of view

FP . . . . . False positive

FPN . . . . . Feature pyramid network

FPS . . . . . Frames per second

GET . . . . . Gait event time series

GGD . . . . . Generalised Gaussian distribution

GPS . . . . . Global positioning satellite

GRF . . . . . Ground reaction force

HC . . . . . Healthy control

HMM . . . . . Hidden Markov model

IBME . . . . . Institute of Biomedical Engineering

iOS . . . . . iPhone operating system

IoU . . . . . Intersection over union

IMU . . . . . Inertial measurement unit

IPR . . . . . In-plane rotation

IQR . . . . . Interquartile range

IV . . . . . Illumination variation

KAZE . . . . . KAZE 2D feature detection method

LAP . . . . . Linear assignment problem

LED . . . . . Light-emitting diode

LR . . . . . Low resolution

LSTM . . . . . Long short term memory network

mAP . . . . . mean average precision

MAPE . . . . Mean average percentage error

MB . . . . . Motion blur

MEMS . . . . Micro electro-mechanical systems

ML . . . . . Medio-lateral axis

MSAC . . . . . Maximum likelihood estimation sample consensus.

MMSE . . . . Mini-Mental State Exam

MOLC . . . . . Minor overlap coefficient

MS . . . . . Multiple sclerosis

NAPE . . . . . Negative absolute percentage error

NCC . . . . . Normalised cross correlation

NPV . . . . . Negative predictive value

OCC . . . . . Occlusion

OF . . . . . Optical flow

OGL . . . . . Oxford Gait Laboratory

OLC . . . . . Overlap coefficient

OPR . . . . . Out-of-plane rotation

OUH . . . . . Oxford University Hospitals

OV . . . . . Out-of-view

PASCAL . . . Pattern analysis, statistical modelling, and computational learning

PCA . . . . . Principal component analysis

PD . . . . . Parkinson's disease

PDR . . . . . Pedestrian dead reckoning

PNS . . . . . Personal navigation systems

PPV . . . . . Positive predictive value

PR . . . . . Precision-recall

QOL . . . . . Quality of life

RANSAC . . . Random sample consensus

RGB-D . . . . Red, green and blue colour image with depth channel

ROI . . . . . Region of interest

SDC . . . . . Sørensen-Dice coefficient

SFM . . . . . Structure from motion

SI . . . . . Superior-inferior axis

SLAM . . . . . Simultaneous localisation and mapping

SSD . . . . . Sum of the standard deviations of triaxial acceleration

SV . . . . . Scale variation

SVD . . . . . Singular-value decomposition

T25FW . . . . Timed 25 foot walk

TM . . . . . Template matching

TN . . . . . True negative

TP . . . . . True positive

TUG . . . . . Timed up and go

UPDRS . . . . Unified Parkinson's Disease rating scale

VO . . . . . Visual odometry

VOC . . . . . Visual object classes

XML . . . . . Extensible markup language

1 Introduction

Contents

1.1 Motivation 1
1.2 Research aims 4
1.3 Research approach & thesis structure 4
1.4 Key contributions 7

1.1 Motivation

The assessment of gait is an important clinical tool. Understanding an individual's gait, and how it changes over time can provide valuable insight into what health conditions a person may have, how progressed they are, and what prognosis they might expect. Furthermore, it may illuminate the best mode of treatment, the efficacy of therapies as they are administered, and how a patient's quality of life is affected by treatment [1]. Despite this range of important applications, clinical evaluation of gait has traditionally been highly qualitative; with a clinician or trained assessor asking a patient to conduct various tasks, and scoring their capability accordingly [2]. This method suffers from a range of issues that limit its applicability, reliability and veracity for the aforementioned tasks.

To counteract the subjectivity and low resolution of the assessment, a plethora of tools and devices were developed. The leading technologies for accurate and high resolution gait assessment rapidly became motion capture and instrumented walkways. These facilitate high resolution analysis, and are now wide spread throughout healthcare systems. However, beyond other limitations, they are highly expensive and require great expertise to operate and interpret. Therefore, they are generally limited to

specific high value use cases; such as in the preoperative assessment of children with cerebral palsy undergoing orthopaedic surgery to improve function and symptoms [3, 4]. This high resource requirement means that the current standard methods of gait assessment are unsuitable for ongoing monitoring of patient symptoms, disease severity, and therapy response.

To combat this, a number of bespoke devices have been developed as well as consumer wearables utilised, in order to both increase quantification of current clinical assessments and replace more advanced systems. However, this does not counteract other key issues with in-clinic assessment: little diversity of motion and activity, limited clinic access and accessibility, change in gait due to the environment, and assessment infrequency.

All in-clinic assessments are conducted under controlled conditions, with prescribed protocols. This is both due to the limited capture capacity of the equipment (e.g. a three metre instrumented walkway), and to ensure inter-assessment repeatability and comparability. Often this means that the assessment involves short periods of linear walking. This however is not representative of real-life free-living gait. However, it is the impact of gait abnormality on free-living activities of daily living that has greatest impact on the patient's quality of life. Furthermore, it is during these activities and in these environments that gait is more variable and the most clinically relevant events occur; such as falls and freezing of gait in Parkinson's disease [5–7].

An additional issue with the current methods is that it requires the subject to travel to a clinic. This is both costly and inconvenient, especially as those most likely to undergo gait assessment have compromised mobility. Furthermore, infrequent assessment of gait fails to identify the longitudinal trends and variations that are observed in all individuals. This is especially pronounced in diseases that have highly erratic trajectories, such as relapsing remitting multiple sclerosis [8]. Therefore the desire for simple, affordable, and accessible devices that can assess gait continuously, passively and non-invasively outside the clinic in the free-living environment is clear.

With the advent of smartphone and smartwatch ubiquity many people now carry an array of sensors on their person at all times. These, and other commercial sensors, have been used to provide continuous gait assessment in free-living environments - and a wide range of algorithms and sensor set ups have been developed and tested

[9]. These developments however, have highlighted several key challenges and concerns that must be confronted; and it is these that this thesis investigates.

Firstly, gait assessment terminology and definitions are largely based around the premise of a series of similar successive steps in a straight line. This however, is not representative of all possible situations. As such free-living gait features cannot be directly compared between subjects and measurement sessions. For example, highly perturbed shuffling gait may not include the removal of contact of the feet with the ground, making it hard to define and extract key events in the gait cycle necessary for feature extraction [10, 11]. Similarly, current definitions lack the robustness to extract spatio-temporal gait features from complex movements associated with everyday living. Therefore, new robust definitions must be established, justified and tested.

Secondly, it has proven difficult to test the validity of these wearables systems in a free-living environment, as no validated gold standard system currently exists that can follow an individual around while they live their lives and extract accurate gait features. This is demonstrated clearly even by the inconsistency and inaccuracy of step counts between different commercial devices [12, 13]. To further understand this problem, several research groups began attaching wearable cameras to subjects and comparing device step count to those manually observed from the video [14–21]. This method is time consuming, somewhat subjective, and cannot be used to extract features beyond step count. It was therefore hypothesised that modern computer vision methods could be used to extract accurate spatio-temporal gait features from such videos.

Thirdly, reasoning and evidence suggests that the validity of the algorithms developed to extract gait features from wearables are dependent on device type, location, orientation, and a range of other parameters [22–24]. Therefore it is important to investigate the comparability between different devices and set ups, both in linear and free-living gait. To achieve this it is also necessary to develop protocols by which free-living gait can be imitated within a clinical environment where gold standard technologies can be used to precisely extract gait features.

1.2 Research aims

The general aim of this thesis is to understand the validity of gait analysis tools beyond highly restricted clinical gait assessment, and to suggest improvements to both protocols and methodologies while highlighting limitations and concerns. This work is multipartite, and thus can be broken down into the following four key aims:

1. To formulate and justify robust engineering definitions for common spatio-temporal gait features that can be used in all environments, for all individuals, and account for any limitations or pathologies.

2. To design motion capture laboratory protocols that enable the thorough testing of gait analysis tools in a manner representative of free-living gait; conferring reliability of validation studies to real life.

3. To develop and validate a computer vision-based gait analysis system for wearable cameras so that gait features can be measured outside of the lab, in a free-living environment.

4. To test the absolute and relative reliability of various wearables at performing gait analysis, compare them to the wearable camera-based system, and identify limitations in current validation studies.

1.3 Research approach & thesis structure

There is much interdependence within the thesis chapters, therefore the following section outlines the narrative of the work. Figure 1.1 schematically demonstrates how the outputs of each chapter feed into others.

- **Chapter 1** begins the work by outlining the motivations, aims and contributions of this thesis; alongside an overview of the thesis structure and key contributions.

- **Chapter 2** provides a background and literature review of each of the concepts core to the following chapters. Emphasis is given to the clinical importance of gait analysis, and the driving factors towards the rise of free-living gait and wearables.

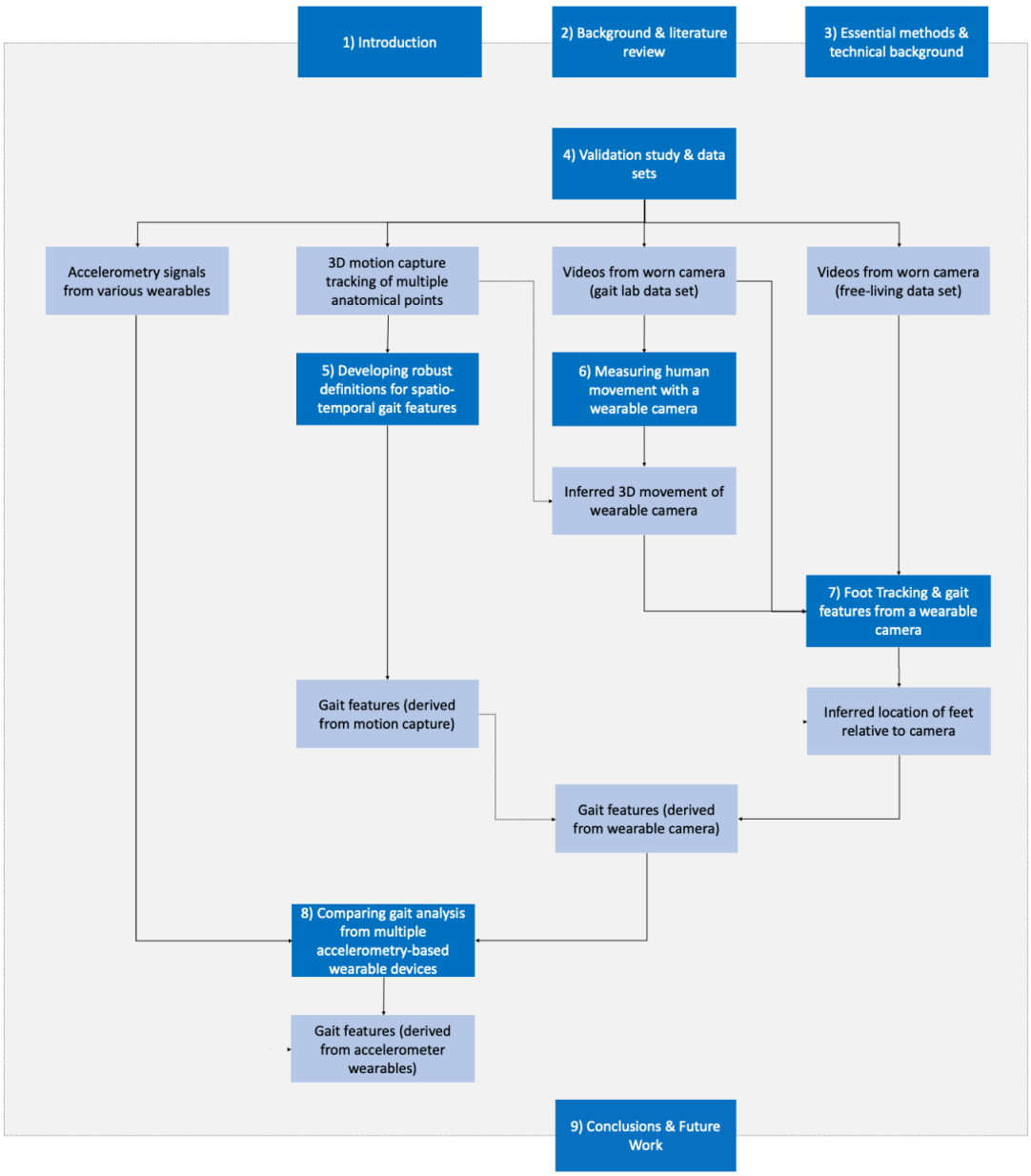


Figure 1.1: Schematic of the structure of thesis. Each dark blue node represents a chapter. Each light blue node represents the output of a chapter. Dashed lines indicate the use of an output as validation for the results of a different chapter.

- **Chapter 3** covers the required technical knowledge to follow the development of the computer vision, accelerometry, and motion capture methods. This allows the following chapters to focus on the unique challenges each method encounters with the data considered, and how they were adapted to account for them.

- **Chapter 4** explains the data gathered to test the developed gait analysis systems.

The primary source of data was a study set up as part of the research; referred to as the Oxford study throughout. This enabled the direct comparison of computer vision and accelerometry-based methods to the current validated ground truth; provided by motion capture. Additional free-living data was however obtained through collaboration - the Newcastle study.

- To understand gait in all circumstances, there must be robust definitions for each of the gait features. **Chapter 5** develops these from fundamental principles, while providing justification for the decisions made in the form of both developed models and the range of gait feature values in the literature. These definitions are used on the motion capture data from the Oxford study, and their sensitivity to threshold values is investigated. The accuracy of the motion capture system during complex movement is also scrutinized.

- A key component of the research was the development of a computer vision method for measuring gait from a wearable camera in a free-living environment. To achieve this, two general methods were utilised and combined: measuring the movement of the camera from the video footage, and tracking the subject's feet throughout the videos. **Chapter 6** explains the first method, from which several gait measurements such as walking speed and turn angle can be directly measured. The accuracy of these measurements is then assessed.

- **Chapter 7** then proceeds to cover the second method, which when combined with the former enables the estimation of gait features such as step count and stride length. The accuracy of these features is compared to the features extracted from the motion capture data from Chapter 4 and analysed by Chapter 5. Difficulties specific to this form of object tracking are also considered in length.

- **Chapter 8** explains the development of accelerometry methods to extract gait features from the wearables used in the Oxford study; a smartwatch, smartphone, lumbar worn intertial measurement unit (IMU) and foot mounted IMU. These results are then compared to the ground truth, computer vision method, and prior research in literature.

- **Chapter 9** explores the difference between free-living gait and laboratory gait; using both the Newcastle and Oxford data sets. It then uses these comparisons to establish a critique of the novel protocols developed for the Oxford study.
- The conclusions of the thesis are then given in **Chapter 10**, as well as notes on the potential directions and opportunities of future work.

1.4 Key contributions

The key contributions provided by this thesis are:

- Step maps, gait event time series (GET) graphs, and the application of methods for solving the linear assignment problem to enable analysis and visualisation of complex movements in free-living gait.
- A data set is gathered comprising seven subjects undertaking novel protocols to emulate free-living gait. This is conducted in a motion capture lab alongside the subjects wearing an action camera, smartphone, smartwatch, lumber-IMU and feet-IMUs.
- A new set of robust definitions is provided for several spatio-temporal gait features in any environment, for any individual. Development of this includes a step velocity model.
- The spectrum on which gait features exist, and thus the importance of strict definitions, is demonstrated.
- A system is devised to extract gait features from videos gathered by wearable-cameras. This includes the measurement of the wearer's egomotion and the detection in the videos of their feet.
- Insights are provided as to the unique limitations and problems faced when developing computer vision methods for gait analysis from a wearable camera.
- Several accelerometry-based gait analysis devices are simultaneously utilised and compared to one another across the range of protocols. The importance of step-by-step analysis, rather than comparison of gross statistics, is also demonstrated.

- It is shown that simple studies involving linear gait do not confer the validity of wearables in extracting gait features beyond these protocols, especially not in a free-living environment.

- The importance of a range of parameters when developing accelerometer-based gait analysis devices and algorithms, such as thresholds selection, the individual's physical characteristics, device location and orientation, and several more, are highlighted.

2

Background & literature review

Contents

2.1 Introduction 9
2.2 Gait and the clinic 9
2.3 Gait features 12
2.4 Free-living gait 15
2.5 Gait assessment by camera 17
2.6 Gait assessment by inertial measurement unit 19

2.1 Introduction

The motivations for this work are outlined in Section 1.1. The following chapter provides deeper insight, and justification for this motivation. First, the importance of gait in the clinic is explained, and how and why there is an increased push to perform such analysis outside of the clinic. How gait can be broken down into its component elements is then explained and the importance of the subsequent features that can be extracted covered. With the value of such gait analysis clear, how it is measured by both cameras and inertial measurement units (IMUs) is elucidated.

2.2 Gait and the clinic

Gait abnormality, or change, is a symptom of many pathologies and health conditions. This is particularly true for neuromuscular disorders, as small abnormalities in the complex neurological and muscular systems required to orchestrate posture and gait can result in noticeable movement aberrations [25]. Indeed, James Parkinson's original medical description of the 'shaking palsy' focuses heavily on the hallmark Parkinsonian

gait characteristics such as slow shuffling of the feet, festinations, and stooped posture [26]. Some gait characteristics, such as reduced gait speed and asymmetry, are relevant to a range of conditions and demographics; however, there are many that are indication specific. Examples of this include freezing of gait (FOG) in Parkinson's disease (PD) [27], hemiplegic gait resulting from stroke [28], and foot drop in Multiple Sclerosis (MS) [29]. It is therefore unsurprising that gait assessment is an important tool in the diagnosis, prognosis and monitoring of many diseases.

2.2.1 Gait assessment & digital biomarkers

Methodologies for gait assessment have evolved over the years to incorporate new technologies and objective quantification of gait features. However, many of the assessments are still the result of qualitative observations of a clinician or healthcare professional. The gait section of the Unified Parkinson Disease Rating Scale (UPDRS) simply distinguishes its grading of 0 to 5 as: no problem walking, independent walking with minor gait impairment, independent walking with substantial impairment, walking requires assistive device but no person, and unable to walk or only able to walk with assistance from another person [30]. Similarly, the 0 to 10 Extended Disability Status Scale (EDSS) for MS is largely based on walking and mobility, but distinguishes between scores 4.0 and 6.0 based on the ability of the subject to walk 500m and 100m metres respectively [31]. These scores show the importance of the clinical attribute, but also the very low resolution and partial subjectivity of its observation; demonstrated by high intra- and inter-rater variability [32, 33].

To account for this, more objective tests have been developed as part of standard gait assessment batteries. An example of this is the six minute walking test (6MWT), where the distance covered in the test's duration is recorded [34]. Similarly, timing of a specified task can be conducted, such as in the timed 25 foot walk (T25FW) test or the timed up and go (TUG) test [35]. In the latter, subjects are timed to ascertain how long it takes them to stand from a seated position, walk three metres, turn around, return to the chair and sit down [36]. These have been shown to be simple to conduct and well tolerated by patients, however they may be challenging for individuals with

very severe gait pathology. Most importantly, although they may increase objectivity they still suffer from low resolution or insight into the underlying medical issue.

To confront this challenge an ever increasing utilisation of technology has enabled further gait quantification and insight. Two key technologies are instrumented walkways and motion capture. The former utilises pressure sensors or force transducers to detect footfall over a short distance; and the latter uses the combination of an array of cameras, computer vision, and reflective markers attached to the subject, to triangulate the location and thus trajectories of various parts of the body. These techniques have enabled the highly precise and granular analysis of both kinetic and kinematic gait features [37].

There are however several key limitations to these methods. They are greatly expensive, generally immobile, require a high level of expertise to operate, and substantially limit the type of movement that can be analysed. To combat this, consumer electronics have been leveraged to provide gait insight beyond the clinic.

2.2.2 Wearables & digital biomarkers

Wearables are devices worn by an individual that enable the extraction of physiological or behavioural measurements of the wearer. Digital biomarkers are the result of the conversion of these signals into health or medical insights. Although the field began with the production of bespoke devices to measure gait, the trend of research has been to increasingly leverage ubiquitous consumer electronics to achieve the same goal. For example gait features can be extracted from the sensors present in everyday smartwatches and smartphones. Advances in this area are highly valuable as a growing body of evidence also suggests that subtle changes in gait can be used as a significant predictor of disease susceptibility prior to clinical onset [38–41]. However, it would be impractical to conduct regular detailed gait analysis on every individual using gait labs or instrumented walkways, and furthermore much of the valuable pathological gait signal may be most apparent when individuals are living their lives and not conducting constrained and specified gait tests.

2.2.3 Gait cycle

Normal non-pathological linear gait is a cyclical and repetitive motion. The process of a single period, from the heel strike of one leg in front of the individual, to the next heel strike of the same (ipsilateral) leg, is known as the gait cycle. Many gait features are defined in terms of the events and phases of the gait cycle, for example the initiation of a stride at toe off and termination at heel strike. Therefore, Figure 2.1 demonstrates an entire such cycle for future reference. In pathological gait, heel strike and toe off may not be representative of the aforementioned moments in the gait cycle; for example due to foot dragging. Therefore, in this circumstance, foot strike can be used for initial contact and foot off for final contact.

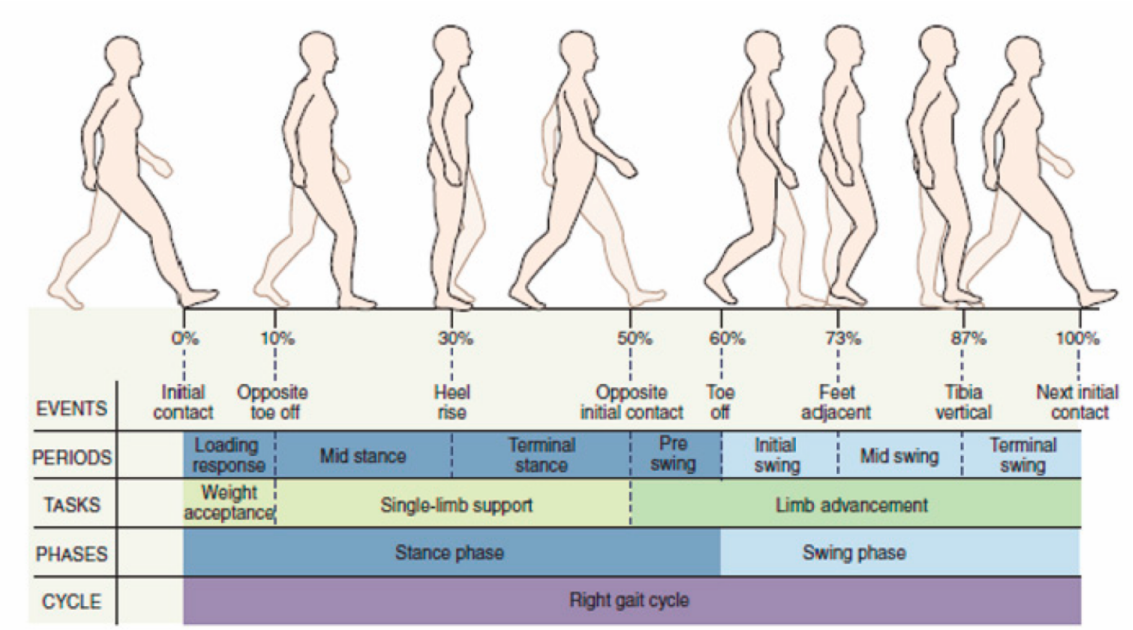


Figure 2.1: The gait cycle. The cycle is decomposed into the stance phase where the body weight is supported by the foot, and the swing phase where it travels forward. These in turn can be subdivided into more granular tasks, periods and events. The image is taken from *Di Gregorio and Vocenas's* paper on identification of gait cycle phases for prosthesis control [42].

2.3 Gait features

This work focuses on fundamental spatio-temporal gait features. This is partly due to constraint by the available research time, but also the limitations of the devices considered. Other features may include: statistical analysis of the clinical features, such

as coefficients of variation, asymmetries or cadence; kinetic features, for example ground reaction force; energetic features, including harmonic ratios or Lyapunov exponents; or more complex kinematic features, for instance knee flexion or ankle angle [43–45].

The features considered in this thesis include: walking bout count, walking speed, turn count, turn angle, turn speed, step count, step length, step width, and stride length. These are fundamental units of gait and their clinical importance is demonstrated below.

2.3.1 Walking bouts & speed

Walking speed, and its variation, is the most readily noticeable gait feature. A reduction in walking speed with age is a commonplace observation both within the clinic and beyond it [46]. Such observations were being made long before gait analysis was a formalised science [2]. More recent studies have shown that beyond being a symptom of aging, gait speed has substantial prognostic value in regards to longevity and survival [47]. Its importance is demonstrated by the emphasis of current clinical gait tests, such as the 6MWT, T25FW and TUG tests on the rapidity with which a subject can cover distance. Even beyond this, strong advocates of its clinical value released a white paper in 2009 extolling walking speed as *'the sixth vital sign'*; due to its validity and reliability in being a measure of balance, hospitalisation, function decline, future health status, mortality, and quality of life [48–50]. With the importance of gait speed being clear, many questions then arise regarding how to measure gait speed, under which conditions, over what distance, whether average speed is sufficient, if the variation in gait speed throughout the period of walking matters, and whether even the instructions for the clinical tests can affect measurement [51]. These parameters make the use of wearable-based gait analysis ideal for capturing this macro gait feature in all scenarios.

Another factor that is often casually observed is the reduction in the amount of walking that occurs in later life. A period of walking delineated by two periods of being stationary (whether sitting, standing or lying) is termed a 'walking bout'. Bout length and bout number have both been shown to be important biomarkers.

On one hand, during longer gait bouts, individuals may become more tired and erratic, on the other they may spend proportionally more time in cyclic linear

walking. The importance of this was shown by *McArdle et al.* when analysing free-living gait in healthy controls (HC), PD, and dementia with Lewy bodies (DLB). Although both pathological cohorts were clearly distinguishable from HC, gait features were only statistically differentiable between PD and DLB during gait bouts of less than ten seconds [52].

2.3.2 Turns

From a quality of life perspective, the ability to navigate, corner and avoid obstacles - and thus turn - during gait is crucial. In fact, *Glaister et al.* showed that ambulation during daily activities is comprised of between 8% and 50% turning steps depending on the activity and individual [53]. However, these proportions were only for movement in the office, store and parking lot; movement and activity around the home is likely even more dependent on turning. It is therefore fairly remarkable that the vast majority of research concerns linear walking.

Beyond the sheer volume of time spent turning, turns are of additional clinical value as it is during this movement that many clinically relevant events occur. For example, FOG in PD is substantially more common during turns than linear walking [54]. Furthermore, turning is associated with altered postural control in even the very early stages of PD, making it a key moment for powerful diagnostic analysis [55]. The correlation between clinically relevant events and turns is likely due to the increased cognitive requirement to undertake the movement [56].

Beyond simply the binary state of turning, the quality and speed of the turn is important. Indeed, turning quality has been shown to be the discriminatory factor between age matched individuals who fall and do not fall [56], with fallers also turning slower and less frequently [57]. Similarly, turn speed, magnitude and time are descriptive of MS patients [58, 59].

2.3.3 Steps

In recent years, step count has received broad acceptance and acknowledgement in the wider community as a metric of health, and a way to motivate improved health outcomes [60]. This has been largely due to the rapid rise in wearables and smartphone adoption.

Similarly, there has been substantial research on the prognostic value of the metric and its correlation with longevity and reduced mortality across a range of indications [61, 62].

Although the adoption of step count has brought the awareness of gait features to the public, the true power of such gait analysis is in the assessment of step-based spatio-temporal features, their variability, and their range. For example, reduced step width and step width variability is discriminatory of patients with chronic obstructive pulmonary disease (COPD) [63], increased stride speed asymmetry of PD [64], reduced stride length of Alzheimer's disease (AD) [65], and substantially increased first step length variability at gait initiation in elderly fallers [66].

As suggested by the previously acknowledged importance of feature variability, it is crucial to provide accurate step-by-step analysis of gait features, as statistics on the features are themselves important features. For example, step length variability has been shown to be a highly discriminatory clinical gait feature for PD [67]. Similarly, in 2019 the European Medicines Agency (EMA) approved one of the first digital secondary clinical endpoints in the form of the stride velocity 95th centile in Duchennes Muscular Dystrophy (DMD) patients using the ActiMyo device [68]. This move has demonstrated great potential in the field of digital biomarkers, while emphasizing the need for step-by-step measurement.

2.4 Free-living gait

Gait in the lab is generally highly constrained (undertaking a specific protocol in a limited space), purposeful and focused (subjects are cognitively engaged on the task), free of obstacles (everyday items, people and furniture), has reduced coincident cognitive requirement (unlikely to be solving problems or conducting complex tasks like operating a smartphone), is undertaken in a homogeneous environment (floor texture, lighting, et cetera), and is observed (potentially causing Hawthorne effect). Due to these factors, it is expected that gait in the lab is different from free-living gait; and likely less discriminatory of pathology (where free-living gait is considered that which occurs unsupervised, uncontrolled, non-standardised and in the real-world [69]). Several experiments have been conducted to validate this hypothesis. One such study recruited 47 age-matched PD patients and 50 HC, and subjects wore a

lumbar-mounted accelerometer during laboratory gait analysis and then for seven days of their normal lives. It was shown that not only did gait features differ outside of the lab, but four (rather than two) of the fourteen features monitored provided to be statistically discriminatory of PD in free-living gait [5]. Interestingly, and conversely to the aforementioned results of *McArdle et al.*, shorter gait bouts actually showed less prominence of pathology-distinguishing features.

Further evidence suggests that feature importance may actually change under these different conditions. When *Shah et al.* compared MS patients with healthy controls, not only did free-living gait provide substantially more discriminatory signals to those previously gathered in the lab, but the most salient feature changed from toe-off angle to gait speed [70]. Similar results were observed for PD patients, with the respective feature being lumbar coronal range of motion and foot-strike angle respectively [70].

One of the elements that likely affects the difference between free-living and laboratory gait is the need to navigate obstacles and variable walking surfaces. Certainly, when these conditions are imitated in a lab environment, changes in gait features are observed [71–73], and in fact episodes of FOG are observed more frequently [74]. As can be seen from the obstacle studies, acknowledgement of the importance of free-living gait has motivated increasing efforts to imitate and simulate the free-living environment in the lab; both for pathological gait and wearable-based gait feature validation. However even when attempted, differences to true free-living gait are acknowledged [75].

The Hawthorne effect is that by which a subject changes their behaviour due to being observed. Although there is some contention in this field of behavioural psychology [76], studies specifically in gait have suggested that it is a consideration. *Ardestani & Hornby* report variations in cadence and force symmetries in stroke patients under observation, although few other differences [77]. *Robles-Garcia et al.* however note significant change in a range of gait features including step length, cadence and gait speed [78].

Overall, it is therefore argued that free-living gait comprises more representative, more discriminatory, and more feature rich gait signals. Thus wearable technologies will be vital in the future of clinical gait analysis.

2.4.1 The frequencies of human motion

Throughout this work, and in the establishment of robust gait analysis methods generally, it is important to know the range of human movement and its frequency components. This is predominantly to aid filtering and anomaly detection.

Although it is widely quoted that 99% of the acceleration power in gait is below 15Hz [79], this was deduced from a population of 12 healthy subjects walking at normal speed over force plates in a lab setting, and as such has limited validity in a free-living data. However, everyday activities were found to be generally composed of frequency components less than 10Hz, with the majority of the power being found within 0.3Hz to 3.5Hz range [80], this was deduced using a a full room, 2.5m by 2.5m pressure sensor.

2.5 Gait assessment by camera

Cameras have an intimate and historic relation to gait analysis, dating back to the 19th century and the early days of photography. The famous images of *The Horse in Motion* taken by Eadweard Muybridge in 1878 were used to study and assess the phases of locomotion [81]. Similarly, *Braune & Fischer* analysed the three dimensional (3D) motion of human gait using photographic plates in the dark and Geissler tubes strapped to the limbs. As such, gait analysis has evolved alongside photography over the subsequent 140 years; with seminal papers describing pathological gait as emblematic of various diseases, such as PD and DMD [2].

2.5.1 Motion capture & static cameras

Most modern clinical gait analysis is conducted using motion capture. This involves an array of high resolution and high sampling frequency cameras in a room over-viewing a capture volume. Reflective markers are then attached to the subject and their location within the capture volume triangulated from all non-obstructed camera views. This technology is the same as that used for visual effects and computer generated images (CGI) in the film industry. The most common motion capture systems have been shown to have sub-millimetre positioning accuracy for static objects. This reduces to between millimeter and centimeter level for dynamic systems, depending on numerous factors.

These include the number and positioning of the cameras, occlusion of the markets by objects and subjects, the appearance of ghost markers due to reflective objects in the capture volume, the speed and complexity of movement, and marker size [82–84].

Recent advances have seen the increasing use of markerless motion capture systems. These systems utilise depth cameras and are preferable as they substantially reduce the time and resources required to set up the markers, as well as the limitations on clothing and activities. However, the accuracy of these systems is still not to the level of the marker-based motion capture systems [85]. As much of the clinical gait analysis work is to help determine the best surgical and therapeutic intervention to improve the patient's gait, based on individual neuromuscular systems, high accuracy is vital [86].

A modern motion capture system (for example, Vicon solutions) can cost well into hundreds of thousands of pounds, and requires substantial expertise and resources to maintain and operate [87]. It is therefore unsurprising that a spate of more affordable technologies have been developed to help make motion capture more affordable and accessible. These range from single camera augmented reality (AR) systems [88], to markerless multi-camera systems [89]. The most common affordable solution in the field of biomechanics and gait analysis is the use of a single RGB-D (Red Green Blue-Depth) camera, such as the widely commercially available Microsoft Kinect [90–93]. Use of such consumer electronics enables the distributed and regular use of motion capture and gait analysis, for example in the monitoring of the elderly in their homes or regular assessment of stroke rehabilitation [94, 95]. However, the accuracy of such systems limits their clinical use [96]. Additionally, almost all such solutions require the cameras used to remain static. This therefore greatly limits the capacity for the analysis of gait in a free-living environment whereby an individual will regularly leave a room, move behind an object obstructing the view of the lower limbs, or walk somewhere completely beyond the view of any static camera. Therefore, it would be greatly beneficial to free-living gait analysis for the the camera to be mounted on the individual themselves.

2.5.2 Wearable cameras

Due to the increased availability and affordability of high resolution wearable cameras and smartphones with sufficient memory to record long videos, there has been a natural

progression towards these devices for gait and activity analysis. The vast predominance of this work is in the use of such cameras to provide a contextualised ground truth for activity recognition [97–100]. However, some work has been conducted where the video footage was used to provide a ground truth for other gait analysis methods [14, 101]. Most notably, *Hickey et al.* used a wearable action camera to provide a ground truth for step count and walking bout count in free-living environment against which to compare their accelerometer algorithm [14]. To extract the ground truth, gait specialists manually watched the videos and counted step and walking events. It was in fact this study that motivated the original conception of automatic gait feature extraction from a wearable camera (Chapters 6 and 7), and the free-living data mentioned in Chapter 4.

There is also some limited precedence of attempts to extract gait features from videos of gait outside of the lab. Both *Aubeck, Isert & Gusenbauer* and *Ozcan & Velipasalar* extracted step count from videos collected by smartphones attached to- or carried by the subject [102, 103]. However, similar to previous works, although the purpose of the work was to enable the collection of data outside the lab, the validation studies consisted of very short bouts of linear gait, in a constrained environment and with no robustness for different clothes, background, obstacles, lighting, subjects, pathological gaits, or activity type.

Finally, *Young, Simic & Simic* attached a camera to an automatic guided vehicle which travels in front of the subject. Computer vision methods were then used to extract step features such as step length. This was however only tested with small stints of linear gait. Furthermore, the need for the guided vehicle to travel in front of the subject massively limits the method's practicality in real environments with stairs, clutter, mixed terrain, and unpredictable movements.

2.6 Gait assessment by inertial measurement unit

The aforementioned motivations for affordable and wearable gait analysis technology have lead to vast growth and development of systems based on inertial measurement units (IMUs). These micro-electromechanical systems (MEMS) are small devices that can detect some combination of linear acceleration (accelerometer), rotational acceleration (gyroscope), and magnetic field (magnetometer). These devices are now

sufficiently small and affordable that high performance IMUs are found in a wide range of wearables and consumer electronics; for example in smartwatches to measure activity and smartphones to detect orientation. Due to their ubiquity, these devices can then be used to conduct gait analysis passively, remotely, non-invasively and continuously; making them ideal for disease monitoring.

Beyond a single device, multiple IMUs may be used simultaneously, at different locations on the subject, to provide more in-depth kinematic insight [104–106], or infer joint angles and kinetics [107, 108]. Similarly, accelerometers have been used in conjunction with an array of other sensors to gather additional kinetic or neurophysiological information. These include but are not limited to: electromyographs (EMG) to measure muscle activity [109], electrogoniometers to measure joint moment [110, 111], instrumented shoes with force or pressure sensors to measure ground reaction force (GRF) [112, 113], and ultrasonic sensors to measure the distance between feet for step length and width measurements [9].

How gait features are extracted from IMU signals depends hugely on: the number of IMUs used, the types of IMUs used, the sensors within the IMUs used, the location and orientation of the IMUs, the algorithms used for extraction, the type of activity conducted, any gait abnormality, and subject specific considerations. The high interest in the use of IMUs for gait analysis, activity recognition and fall detection has resulted in a vast array of research studies being conducted with various combinations of the above parameters, however the comparability of the methods for each combination is limited [114]. For example, use of gait feature extraction algorithms such as the inverted pendulum model would not apply to IMUs located on the extremities; and even then there are variations on the model itself [115, 116]. Even the orientation and specific location of a sensor on the same body segment can render different results; whether it be trunk, wrist, leg, or other body segment [117–119]. Some of these issues can be ameliorated by using sensor fusion, orientation invariant representations of the signals, or deep networks; and even these are comprised of a catalogue of options [22, 120, 121]. However, many methods depend on the identification of specific gait events within the signal, such as heel strike. This may be complicated or obfuscated by low sensor sampling rate, compound or complex movements during non-linear gait,

negligible heel strike during pathological shuffling gait, or the sensor not being rigidly attached to the subject (for example a smartphone in a loose fitting pocket) [122–124].

This multiple-parameter dependency is part of the motivation for developing a robust computer vision system that requires simple set up and can be used in any environment (Chapters 4 to 6).

An individual's gait is increasingly regarded as unique [125], which has inevitable consequences on how effective any particular algorithm is at accurately extracting features. Therefore, a range of subject specific considerations also affect the extraction of gait features from IMU signals. These include but are not limited to: age, height, weight, fitness, physiology, pathology, mood, fatigue, clothing [126–134].

2.6.1 Precision & personalisation

The importance of accuracy and consistency in measurement, and thus the previously mentioned parameters, is exemplified by the richness and depth of information contained within an individual's gait. This depth is demonstrated, not only by the fact that presence and severity of disease can be detected and measured using accelerometry [135, 136], but by an increasing body of evidence that each individual's gait is unique [137]. Therefore, there is evermore research on how to utilise gait as a biometric measure for security, identity management, military, safety, healthcare and many other applications [138].

This chapter can thus finish in a similar place to where it began. Ambulation is so fundamental to the human experience, and so constantly observed, that much of what the scientific community is proving now was acknowledged long before the technologies existed for its formal measurement. Much as we can now identify neurological diseases in accordance with gait disturbance, as James Parkinson (and many before him) observed centuries ago, we can now prove what Shakespeare acknowledged through Cassius; *'...I do know him by his gait'* [139].

3

Essential methods & technical background

Contents

3.1 Introduction 23
3.2 Gait & graphs 23
3.3 Optics & multiview geometry 30
3.4 Object detection 49
3.5 Conclusion 52

3.1 Introduction

This chapter contains technical background as well as various mathematical and computational methods that are used in the following chapters. Although these methods may not be chapter specific, where they apply is generally referenced. Three key areas are explored.

First, some basic concepts with regards to anatomical axes, sensor axes, and gait diagrams are covered. This will enable all future chapters to discuss what is being measured, in what way, and how measurements from different devices can be compared. Secondly, the fundamental principals of optics and multiview geometry are discussed; as pertaining to the particular case of visual odometry utilised in Chapter 6. Finally, there an explanation on how the veracity of the object detection method deployed in Chapter 7 is assessed.

3.2 Gait & graphs

In order to discuss gait, its measurement, and the devices that are conducting the measurement, it is useful to utilise various diagrams and graphs. This is necessary when

comparing multiple measurements of the same ambulatory period, from different devices; especially when considering complex, free-living gait. Those used here are: diagrams to establish the relative coordinate systems of wearables and the individuals wearing them, symbolic step maps of the chronology of a gait bout, and gait event time series graphs that enable the comparison of multiple measurements of the same gait bout.

3.2.1 Axes & planes

This work deals predominately with the development and assessment of gait analysis from wearable technologies. Therefore, it is useful to understand the basic anatomical planes, and the relative coordinate systems in which the devices work.

Figure 3.1 demonstrates the anatomical axes and planes which will be useful throughout the thesis; it is worth noting the axes in relation to specific limbs and physiological features, as complex movement results in much deviation from this idealised pose. The wearable camera discussed in Chapters 6 and 7 is also shown as well as its coordinate system.

3.2.2 Step maps

As this thesis discusses complex movement during free-living gait, it is useful to discuss non-linear walking patterns. To achieve this and best convey the movement, 'step maps' are proposed and will be used. Figure 3.2 gives an example of what these look like. A movement is demonstrated as a series of numbered foot images. The numbers represent the chronology of the events and the feet are coloured by side. Foot images can overlay each other if the same location is stepped upon multiple times. An arrow demonstrates the approximate movement of the individual's centre of mass; enabling some information to be conveyed about step symmetry or pose adjustment. Arrows are split into walking bouts, so as to enable to delineation of bouts separated by a pause.

These step maps are a minor expansion on already fairly standard diagrams - which generally have been used to demonstrate simple, predominantly linear, gait [141–143]. With this extension they can help in both the description of complex gait and the establishment of more robust methods for measuring spatio-temporal gait features (both in Chapter 5). If in doubt, trying acting out the step maps.

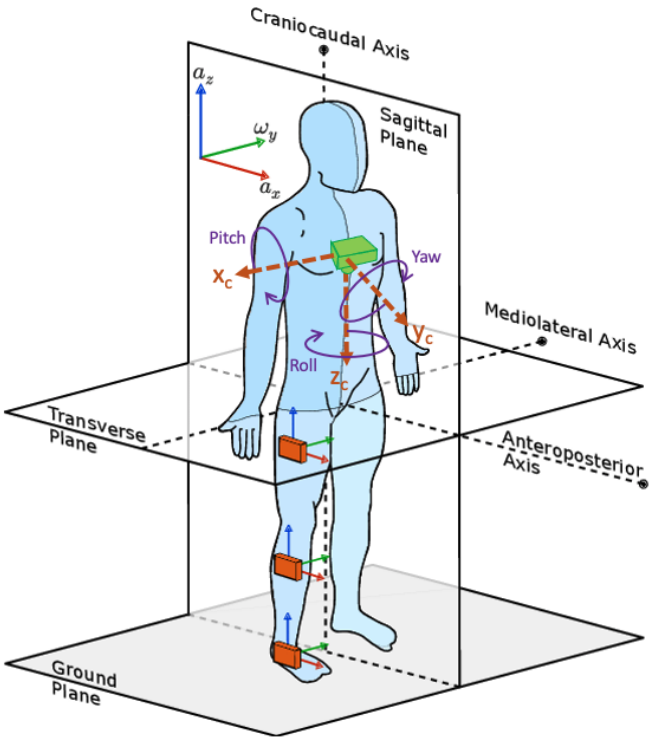


Figure 3.1: Anatomical and camera axes. The figure has been adapted from the work by *Laufer et al.* on extracting gait features from IMUs attached to the lower limbs [140]. The IMUs used in the experiments discussed in this thesis are attached in different locations. A wearable camera is attached to the chest (green), and has its own coordinate system (orange). Rotation in the camera's planes can be described as pitch, roll and yaw (purple). Generally, the camera is pointing towards the ground and slightly forward, so the planes of movement almost align. If a subject turns while walking, the camera undergoes predominantly roll. However this is dependent on stance, movement and camera setup. The craniocaudal axis can also be referred to as the superior-inferior axis with superior being the direction of up the body against the gravity vector.

3.2.3 Gait events

Gait events are any delineated actions or events that occur over time during ambulation. Those considered in this thesis include steps, walking bouts and turns. However the following section on how to visualise, compare, and assign correspondences across multiple signals, could be used for any temporal event.

3.2.3.1 Gait event time series graphs

Given that a period of free-living movement is comprised of numerous and complex walking bouts, and walking bouts themselves are comprised of steps and turns of varying lengths, it would be useful to quickly observe and assess a gait analysis

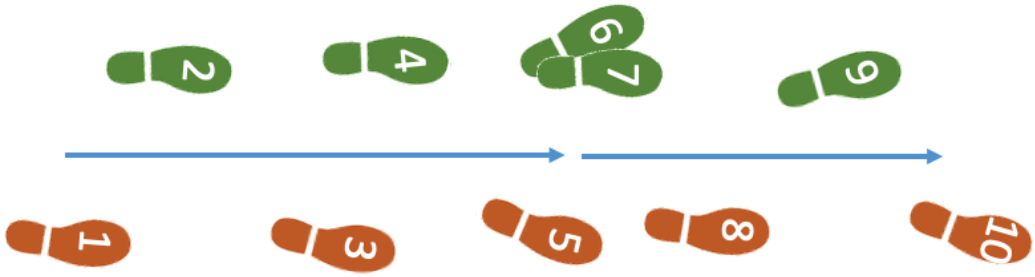
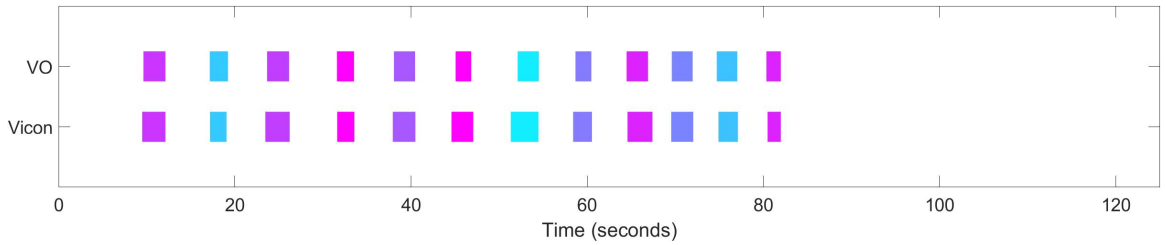


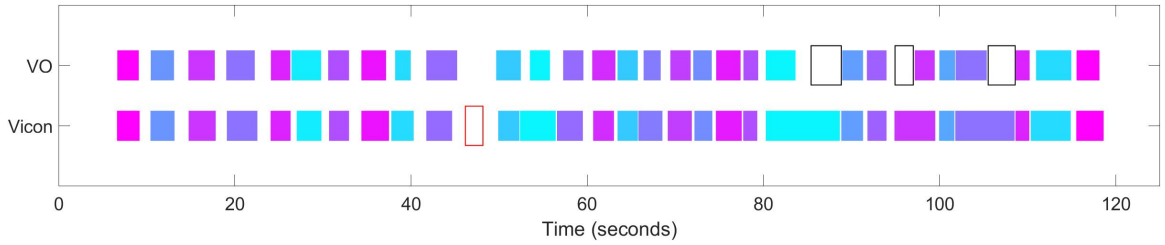
Figure 3.2: Step map example. Left steps and right steps are numbered by the order they are taken, and coloured by side. The approximate movement of the centre of mass of the individual is represented by arrows, with each arrow representing an individual walking bout. In this example, an individual has walked in a straight line, stopped walking, adjusted their stance, and then continued walking again.

methodology's accuracy. To do this one needs to compare both the presence and temporal location of an event in one signal, when compared to another source of information. For example, when assessing the validity of a gait analysis wearable, it might be necessary to investigate the success of the device versus a motion capture gold standard, or against other competitor devices. Generally this is done by comparing gross statistics such as total step count, however this lacks the nuance of insight provided by assessing the series event-by-event.

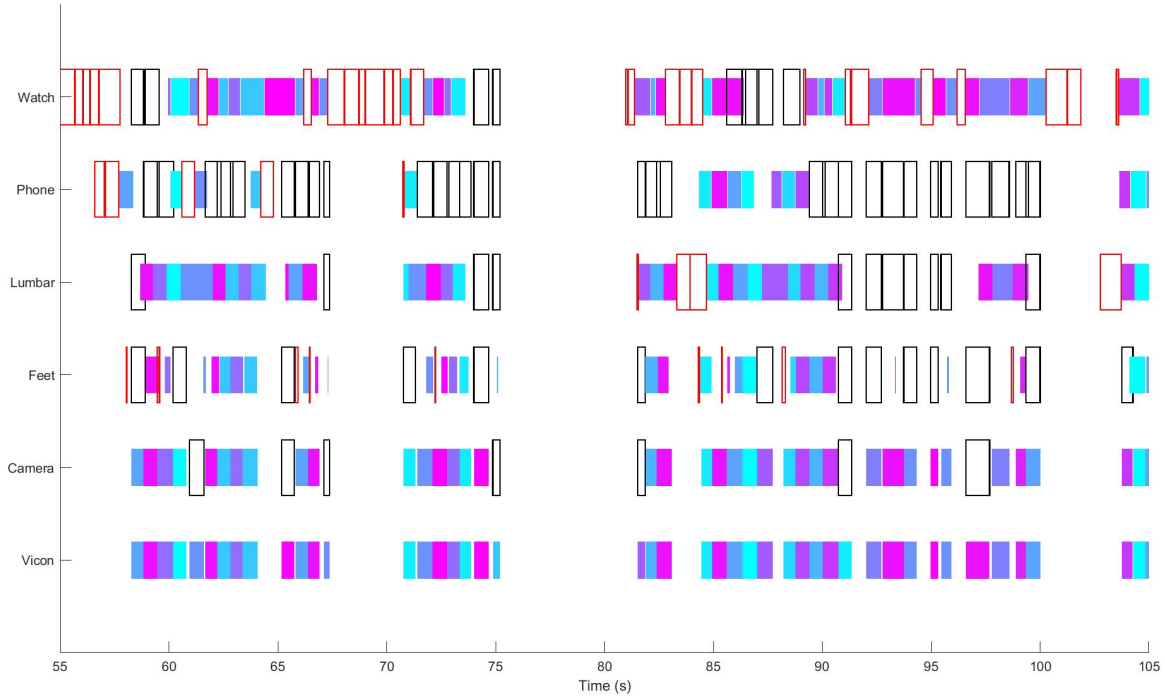
No suitable method was found in order to quickly and qualitatively make this assessment, so here, Gait Event Time series (GET) graphs are proposed. Fundamentally, a GET graph demonstrates the periods of temporal events graphically. They are most useful when comparing two series of contemporaneous events. In this scenario, concurrent events in both series that can be matched to one another are similarly coloured (Figure 3.3a). Subsequent events are always coloured differently to previous events to ensure that if events happen consecutively, they are still discernibly distinct (Figure 3.3b). Events that are unmatched in either signal are deemed either false negatives if part of the ground truth series, or false positives if part of the comparator series. These are also shown in their own series as empty black (for false positives) or red (for false negatives) outlined events (Figure 3.3b). These erroneous events are made more distinct by elongating them beyond the y-axis limits of the matched events.



(a) Turn events detected using two methods while walking backwards and forwards over 5 metres.



(b) Turn events detected using two methods while walking a path containing a range of turn angles.



(c) Stride events detected using six methods during a period of imitation free-living activities.

Figure 3.3: GET graph examples. (a) shows a simple scenario whereby turn detection is compared between two methods of measurments; Vicon motion capture and visual odometry (VO). Similar colours correspond to turn events that are assigned to one another. (b) demonstrates that subsequent events are coloured differently to enable visual delineation and false positives (black rectangle) and false negatives (red rectangle) can also be shown on the graph. (c) enables direct comparison across multiple devices and methodologies; whereby the bottom motion capture row is considered the ground truth against which all others are compared. These graphs are generated using a bespoke script in Matlab R2020a [144].

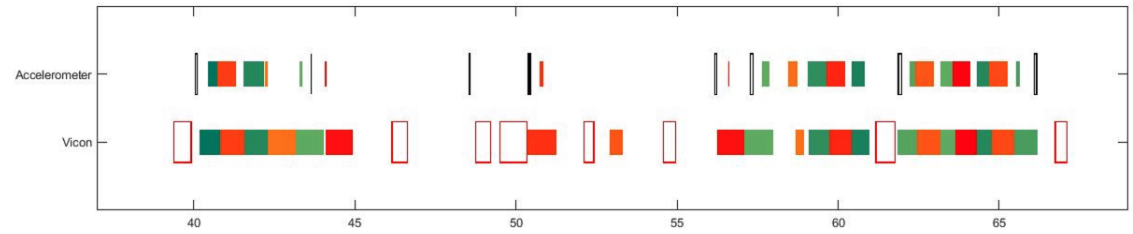
The GET graph can be extended to simultaneously compare any number of synchronised devices. For example, Figure 3.3c shows the comparison of stride detection accuracy for multiple devices of different types during activities in imitation of daily living. In this case, the only difference to previous GET graphs is that each series being compared against the ground truth hosts both the false positive and false negative events relevant to its series.

Additionally, the GET graph can be adapted to elucidate additional information about the events based on a category description. In the examples in Figure 3.4, events are coloured various shades of green for left foot strides and various shades of red for right foot strides. This enables us to ascertain more in-depth information about the movement, for example if an individual tends to lead short bouts in free-living gait with one foot or the other, or during what kinds of bouts are strides missed (Figure 3.4a).

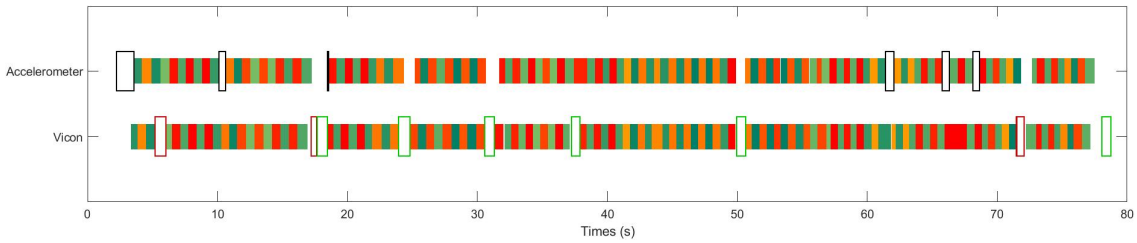
The GET graphs have further extensibility in the colouring of the false negatives. By following the same colour scheme, Figure 3.4b shows that the majority of missed strides appear to be left foot strides during turns - likely the result of small adjustment steps during 180° about turns. This figure also indicates the relative temporal symmetry of the gait during linear gait, and long strides during about turning at the end of each straight. This is because the lumbar-IMU only provides a single signal - and the methodology used to extract stride events only allows for non-overlapping consecutive strides (Chapter 8). If however events can overlap, for example when using two separate foot-mounted IMUs (Figure 3.4c), then due to the automated way the graphs are constructed, it seems that there is great asymmetry in stride times with much longer left footed strides and much shorter right footed strides. In order to prevent this obscuring of overlapping events, the GET graph can be separated out into its component categories (Figure 3.4d).

3.2.3.2 The linear assignment problem

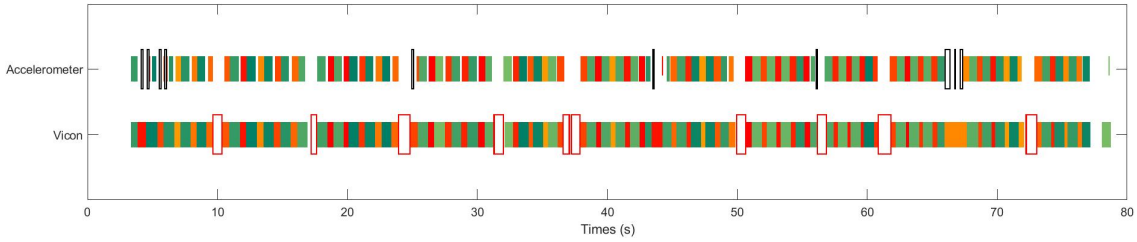
The GET graphs in Section 3.2.3.1 enable direct inspection and comparison of a single event, such as a stride, from multiple sources. The problem that arises while building such graphs or investigating event-by-event features is how to ascertain which events in each series correspond to one another. In the simplest possible form, event



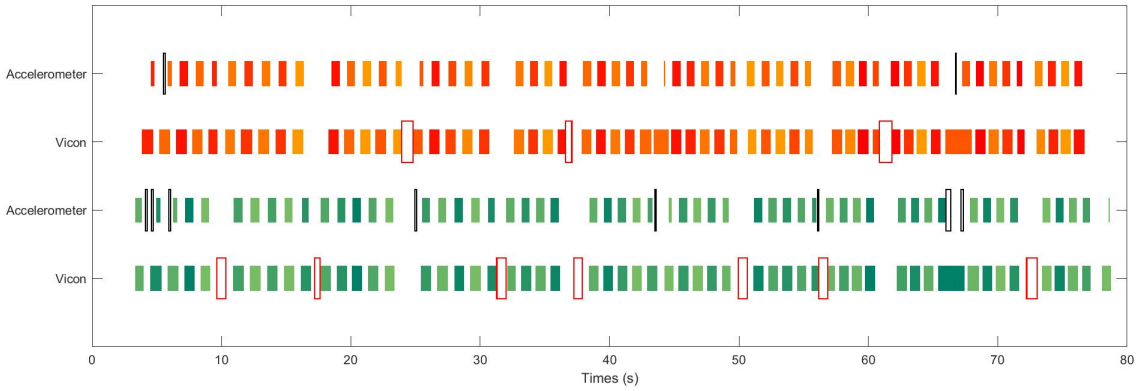
(a) Left and right stride events for a lumbar-IMU during imitation free-living activities



(b) Left and right stride events for lumbar-IMU while walking up and down a 5m path



(c) Left and right stride events for feet-IMUs during while walking up and down a 5m path



(d) Left and right stride events for feet-IMUs, walking up and down a 5m path, split by foot

Figure 3.4: GET graph examples with event catagorisation. In all figures, green denotes left foot strides and red demonstrates right foot strides. In (a), nuanced information about which foot is moving during short bouts can be seen for short erratic movements during imitation free-living activities. False positive and false negatives can also be assigned to a category with appropriate colouring, as can be seen in (b). Here we can also see that seemingly errors arise periodically - in this case in accordance with the about turn at the end of each straight stretch while walking up and down a 5m path. (c) shows that there is some limitation and information loss in the use of the graphs if events are not strictly subsequent and can in fact overlap. Therefore, in this scenario, it is best to prise apart the categories for greater clarity; as shown in (d).

correspondences are obvious (for example in Figure 3.3a), however in more complex circumstances, how to automatically and optimally assign correspondences is not necessarily clear. Figure 3.5a demonstrates the problem.

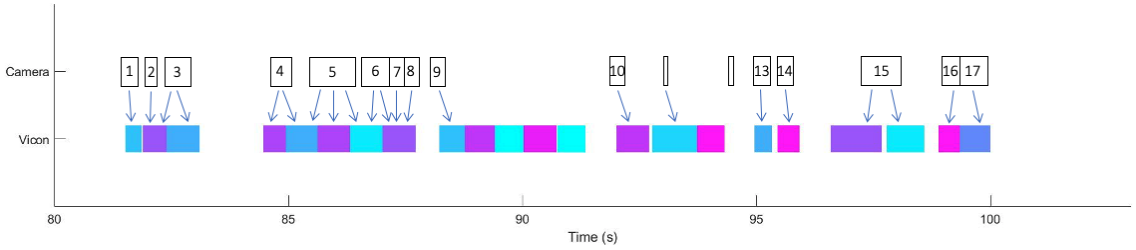
With appropriately synchronised signals, this can be perceived as a version of the linear assignment problem (LAP). Whereby a matrix is formed of the cost of assigning each event in one series to each of those in the other. Each row event is then assigned to a column element event in order to minimise the total cost. In this circumstance, the cost can be regarded as the lack of overlap in time between the two events. Figure 3.5b, shows how this may look for the example in Figure 3.5a. Here, hotter colours represent greater overlap and thus a lower value. White represents no overlap at all and thus a highly punitive cost.

It is clear that this will generate a large sparse matrix with values approximately on the diagonal; even within the relatively short protocols of the Oxford study, each sample may contain hundreds of stride events. In order to solve the LAP, the Matlab implementation of *Duff & Koster's* algorithm for the permutation of large entries to the diagonal of a sparse matrix is used [145].

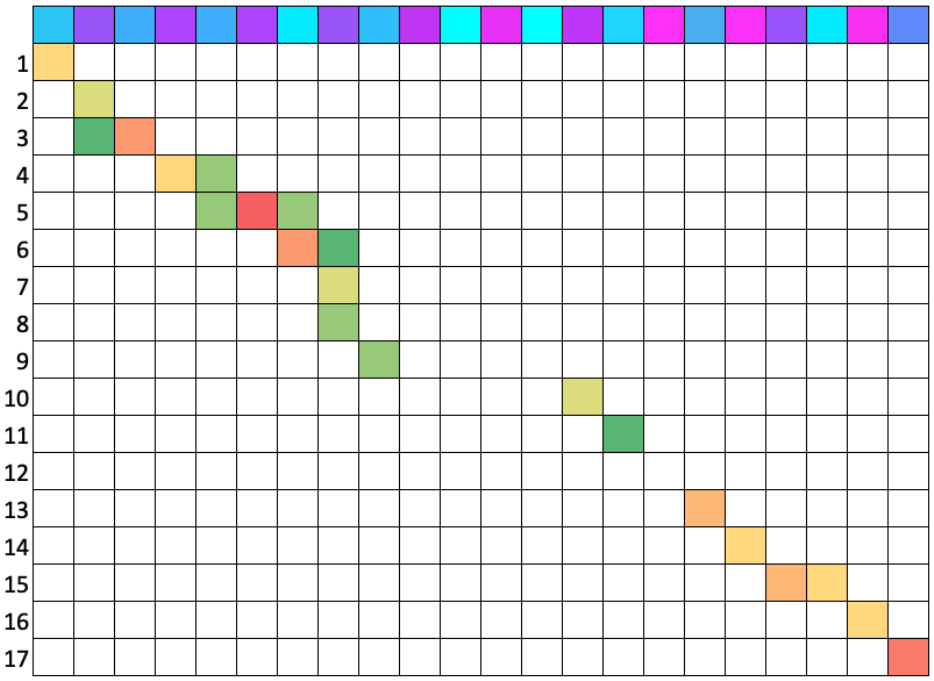
3.3 Optics & multiview geometry

Although the physics and mathematics behind cameras and optics generally is a vast field with many approaches and nuances, this section covers exclusively the mathematical and algorithmic background and methods that are utilised in the processing of video data relevant to this thesis. These methods apply to chapters 6 and 7.

Firstly, an introduction to cameras and how they capture a 3D world in a two dimensional (2D) image is explained. The relationship between world planes and the camera's image plane, and how it can be estimated, is then explored. With this relationship established, it can then be used to estimate the relative locations of the real-world planes and the camera.



(a)



(b)

Figure 3.5: Linear assignment problem example. The two series of events are compared in (a) and it is not immediately clear how events should be paired. The suitability of the pairing is measured by the amount of time the events overlap by; and a cost matrix can be built. (b) demonstrates the according matrix on which the bipartite matching algorithm operates.

3.3.1 Projections, planes & homographies

To understand how information about the real world can be deduced from multiple images, it is first important to understand how a camera works, and how the real world relates to the images taken by a camera.

3.3.1.1 The pinhole camera model

A camera is an optical device that captures an image of the world, generally converting 3D information into a 2D projection. The pinhole camera, or camera obscura, exploits a simple model for the projection of 3D points within the homogeneous world coordinates

$\boldsymbol{X}$ =(X,Y,Z,1)$^\top$ through a point aperture and onto a homogeneous location on the image plane, $\boldsymbol{x} = (x, y, 1)^\top$ (Figure 3.6). The aperture is also known as the centre of projection and occupies the origin of the camera coordinate system. $\boldsymbol{Z_c}$ therefore becomes the optical axis and is defined as perpendicular to the image plane. The intersection of the optical axis and the image plane is the principal point.

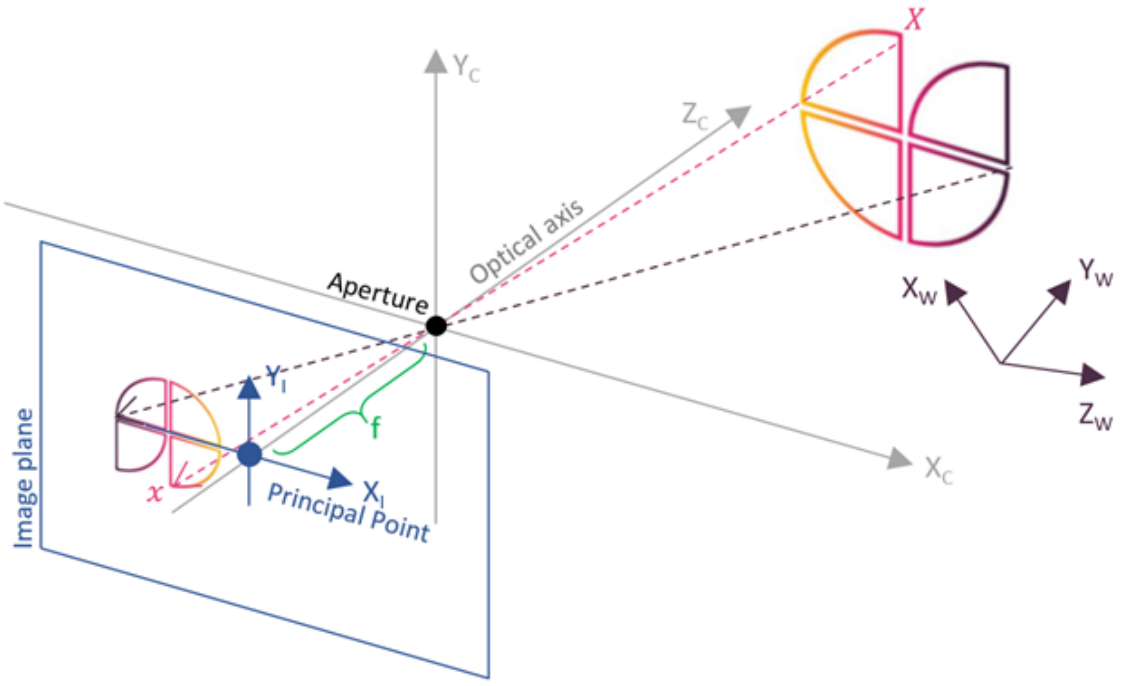


Figure 3.6: The pinhole camera model. The larger real world image on which X lies, is projected onto the camera's image plane via the aperture. The camera, world, and image plane all have their own coordinate systems.

Assuming the coordinates of the point $\boldsymbol{X}$ are known in the camera coordinate system, the homogeneous 3D point $\boldsymbol{X}$ is therefore related to the homogeneous image point $\boldsymbol{x}$ by the projectivity:

$$\boldsymbol{x}_{3\times1} = \boldsymbol{P}_{3\times4}\boldsymbol{X}_{4\times1} \tag{3.1}$$

If we consider this model (Figure 3.6) as viewed from the axis Y_c (Figure 3.7a), and then did the same from the axis X_c, we can use the properties of similar triangles to deduce that:

$$x = -f\frac{X}{Z}, \tag{3.2}$$

$$y = -f\frac{Y}{Z}, \tag{3.3}$$

where f is the focal length. An additional consideration is that the image is actually composed of pixels (Figure 3.7b). Therefore the the focal length is actually expressed in terms of pixel height and pixel width (f_x and f_y). The first pixel in both image dimensions is also defined as being in the top left corner of the image. Therefore the image point must undergo a transformation (measured in pixels) of $[u_o, v_o]^T$.

Finally, it is also possible that the pixels are not perfectly square, which induces a skew to the projection (Figure 3.7c). It is acknowledged that this effect is negligible in modern cameras [146, 147], however it is included here alongside the calculation for accounting for it for completeness (Figure 3.7d).

It can be shown by superimposing the rectangular pixel and the skewed pixel, that:

$$x_{skew} = x_{rect} - y_{rect}\cot\Theta \tag{3.4}$$

$$y_{skew} = \frac{y_{rect}}{\sin\Theta} \tag{3.5}$$

$$\begin{bmatrix} x_{skew} \\ y_{skew} \end{bmatrix} = \begin{bmatrix} 1 & -\cot\Theta \\ 0 & \frac{1}{\sin\Theta} \end{bmatrix} \begin{bmatrix} x_{rect} \\ y_{rect} \end{bmatrix} \tag{3.6}$$

This leaves Equation 3.1 as:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & s & u_o & 0 \\ 0 & f_y & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3.7}$$

which can be expressed as,

$$\boldsymbol{P} = \begin{bmatrix} f_x & s & u_o \\ 0 & f_y & v_o \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \boldsymbol{K}[\boldsymbol{I}|\boldsymbol{0}], \tag{3.8}$$

where s is the skew coefficient and is equivalent to $f_x \cot\Theta$, and $\boldsymbol{K}$ is known as the camera's intrinsics matrix. This however assumes that the world coordinate system is the same as the camera coordinate system. To transform a point in world

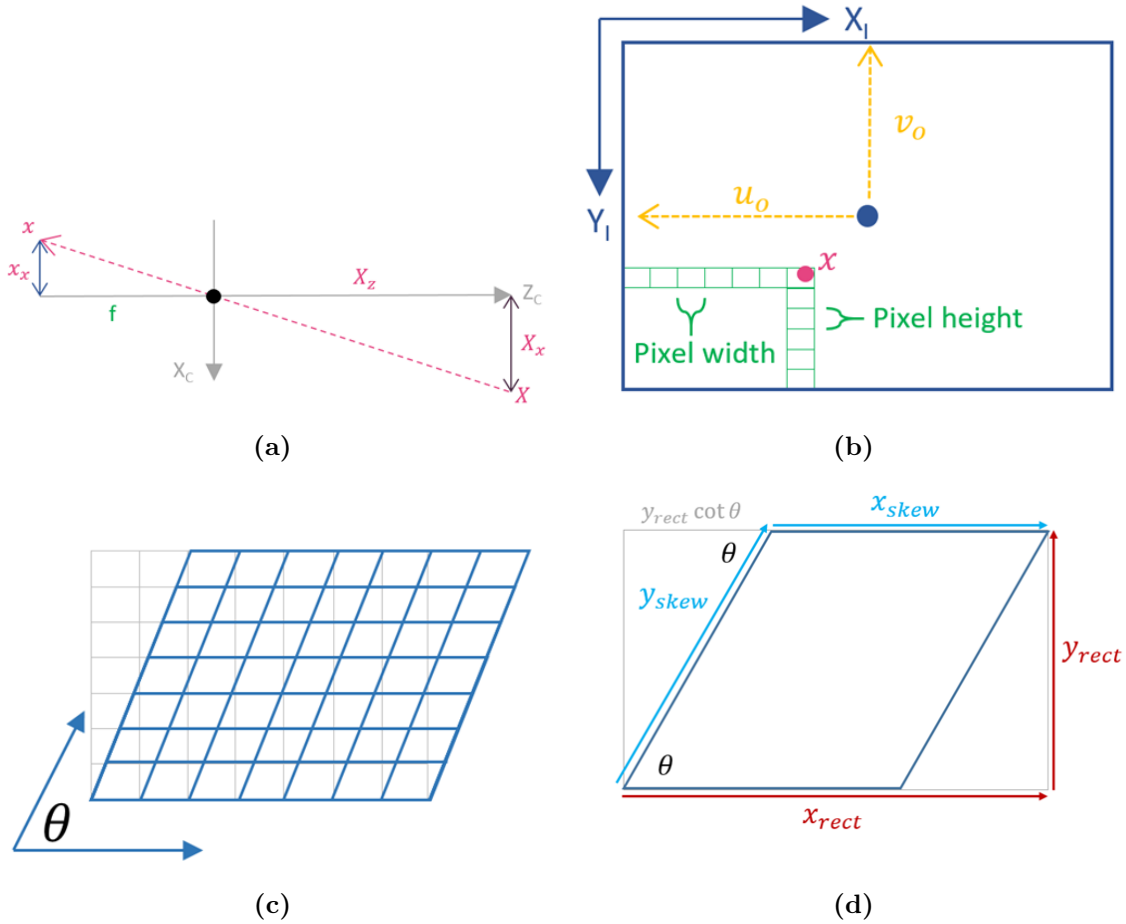


Figure 3.7: Camera intrinsics. To help explain how point $\boldsymbol{x}$ in the image plane relates to point $\boldsymbol{X}$ in the real world, the similar triangles formed by the light beam connecting the two can be observed (a). (b) shows that the image plane is actually composed of pixels, and that the origin for pixel number is in the top left corner of the plane. (c) demonstrates that the pixels may not be precisely perpendicular and may in fact undergo some skew, Σ. Which can be related to the rectilinear image plane by trigonometry (d).

coordinates X_w to camera coordinates, the vector representing the point must be translated and rotated. This gives,

$$\boldsymbol{X} = \boldsymbol{R}(\boldsymbol{X_w} - \boldsymbol{T}), \tag{3.9}$$

where $\boldsymbol{T}$ and $\boldsymbol{R}$ are respectively the translations and rotations between the camera and world coordinate systems. This results in the general formulation of equation 3.1 as,

$$\boldsymbol{x} = \boldsymbol{K}\boldsymbol{R}[\boldsymbol{I}|-\boldsymbol{C}]\boldsymbol{X}, \tag{3.10}$$

with $\boldsymbol{C}$ being the camera aperture's location in the world coordinate system.

3.3.1.2 Back-projection

As the direct converse of Equation 3.1, an image point corresponds to a set of points in the 3D world that project into that image point. This set of points forms a ray that travels through the image point and the camera centre (Figure 3.8). One way of expressing this ray is as a line that connects these points.

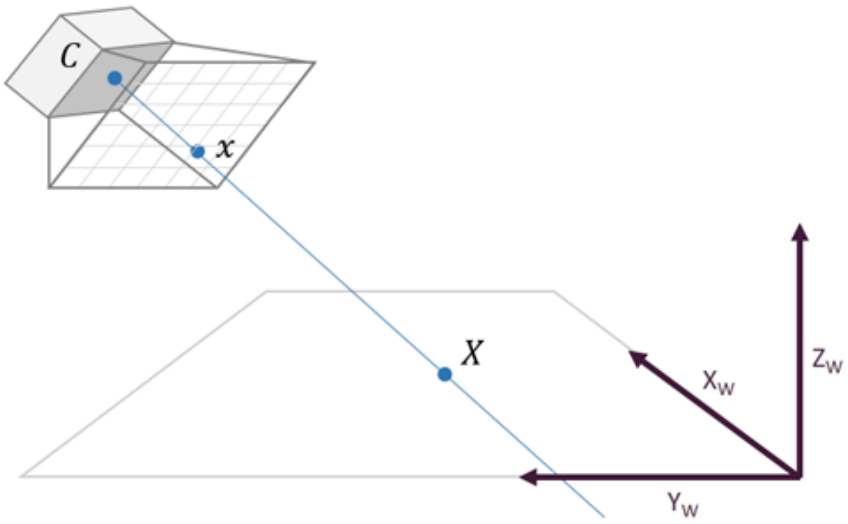


Figure 3.8: Back-projection of image points to rays. The image point $\boldsymbol{x}$ is back-projected to $\boldsymbol{X}$ in the $\boldsymbol{X_W Y_W}$ world plane by connecting the camera centre, $\boldsymbol{C}$, (equivalent to the aperture or centre of projection) with the image point and drawing a beam until it intersects the plane.

To obtain the 3D point $\boldsymbol{X}$ from Equation 3.1, the pseudo-inverse of $\boldsymbol{P}$ must be found as $\boldsymbol{P}$ is not square. The pseudo-inverse of $\boldsymbol{P}$ is defined as:

$$\boldsymbol{P}^{+} = \boldsymbol{P}^{\top}(\boldsymbol{P}\boldsymbol{P}^{\top})^{-1} \tag{3.11}$$

for the case where $\boldsymbol{P}\boldsymbol{P}^{+} = \boldsymbol{I}$. Point $P^{+}x$ is on the ray as it projects to $\boldsymbol{x}$, which can be shown as $\boldsymbol{P}(\boldsymbol{P}^{+}\boldsymbol{x}) = \boldsymbol{I}\boldsymbol{x} = \boldsymbol{x}$. The ray can therefore be expressed as the connection of the two points:

$$\boldsymbol{X}(\lambda) = \boldsymbol{P}^{+}\boldsymbol{x} + \lambda\boldsymbol{C} \tag{3.12}$$

If it is desired to project the image point $\boldsymbol{x}$ into the ground plane (arbitrarily designated as $z = 0$ for purpose), then λ can be found from the third row of the components of Equation 3.12:

$$\lambda = -\frac{[\boldsymbol{P}^{+}\boldsymbol{x}]_3}{C_3}. \tag{3.13}$$

Using this in Equation 3.12, and then converting from homogeneous to inhomogeneous coordinates gives the euclidean location of $\boldsymbol{X}$ on the ground plane.

3.3.1.3 Planar induced homography

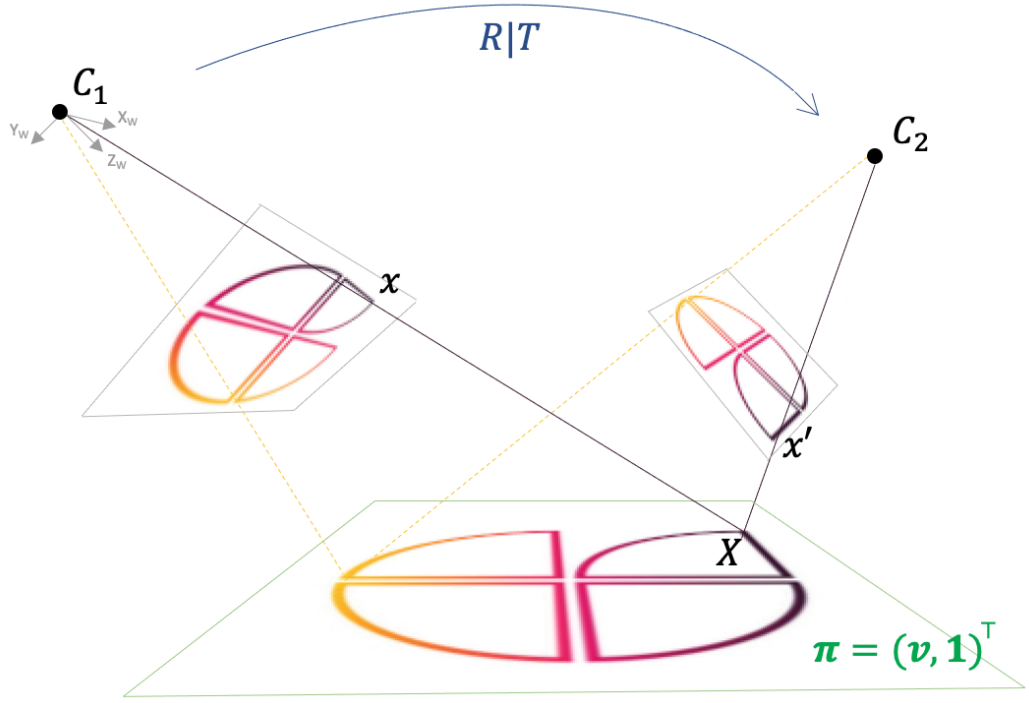


Figure 3.9: Planar induced homography. Two cameras may observe the same point, $\boldsymbol{X}$ in the plane $\boldsymbol{\pi}$, defined by the vector $\boldsymbol{v}$. This point will be projected into the two camera image planes as $\boldsymbol{x}$ and $\boldsymbol{x'}$ respectively. The two cameras are connected by the transformation $\boldsymbol{R|T}$, and the world coordinate system can be centred and aligned with the first camera's coordinate system for the sake of convenience.

If a plane is being observed by two cameras, a point in the plane $\boldsymbol{X}$ can be projected to image points $\boldsymbol{x}$ and $\boldsymbol{x'}$ (Figure 3.9). These projections can be calculated by Equation 3.1. The first camera, C_1 can then be aligned with world coordinate system, so that:

$$\boldsymbol{x} \simeq \boldsymbol{K}[\boldsymbol{I}|\boldsymbol{0}]\boldsymbol{X}, \tag{3.14}$$

$$\boldsymbol{x'} \simeq \boldsymbol{K'}[\boldsymbol{R}|\boldsymbol{T}]\boldsymbol{X}, \tag{3.15}$$

where $\boldsymbol{R}$ and $\boldsymbol{T}$ is the relative rotation and translation between camera coordinate for camera $\boldsymbol{C_1}$ and $\boldsymbol{C_2}$. A ray, γ, can be back-projected from camera centre $\boldsymbol{C_1}$, through $\boldsymbol{x}$ so that from Equation 3.14, any 3D point on the ray can be defined as,

$$\boldsymbol{X} = ((\boldsymbol{K^{-1}x}^\top, \rho)^\top. \tag{3.16}$$

It is assumed that $\boldsymbol{X}$ lies on the plane defined by

$$\pi = (\boldsymbol{v}, 1)^\top, \tag{3.17}$$

where $\boldsymbol{v}$ is a vector defining the plane. As $\boldsymbol{X}$ lies in the plane, $\pi^\top \boldsymbol{X} = 0$, and therefore

$$\boldsymbol{X} = ((\boldsymbol{K^{-1}x}^\top, -\boldsymbol{v}^\top \boldsymbol{K}^{-1}\boldsymbol{x})^\top. \tag{3.18}$$

The point $\boldsymbol{X}$ as defined by Equation 3.18 is then projected into the second image plane with Equation 3.15, giving:

$$\begin{aligned} \boldsymbol{x'} &\simeq \boldsymbol{K'}[\boldsymbol{R}|\boldsymbol{T}]((\boldsymbol{K^{-1}x}^\top, -\boldsymbol{v}^\top \boldsymbol{K}^{-1}\boldsymbol{x})^\top, \\ &\simeq \boldsymbol{K'RK}^{-1}\boldsymbol{x} - \boldsymbol{K'Tv}^\top \boldsymbol{K}^{-1}\boldsymbol{x}, \\ &\simeq (\boldsymbol{K'R} - \boldsymbol{K'Tv}^\top)\boldsymbol{K}^{-1}\boldsymbol{x}. \end{aligned} \tag{3.19}$$

If the cameras have the same intrinsic matrix, $\boldsymbol{K} = \boldsymbol{K'}$, as is the case when the two cameras are in fact the same camera at different points in time, then

$$\boldsymbol{x'} \simeq (\boldsymbol{R} - \boldsymbol{Tv}^\top)\boldsymbol{x}. \tag{3.20}$$

It can thus be stated that the points $\boldsymbol{x}$ and $\boldsymbol{x'}$ are related by a homography, up to scale α, induced by the plane of the form,

$$\boldsymbol{x'} = \alpha \boldsymbol{H}\boldsymbol{x}, \tag{3.21}$$

where the homography $\boldsymbol{H} = \boldsymbol{R} - \boldsymbol{Tv}^\top$ [148].

3.3.2 Homography estimation

To understand the relationship between images and the world, it is useful to estimate the homography. $\boldsymbol{H}$ in Equation 3.21 has nine elements. However, as it is only defined up to scale, this leaves eight degrees of freedom to a projectivity. Each world point viewed by both cameras causes two constraints on $\boldsymbol{H}$, and therefore four image point correspondances are sufficient to calculate $\boldsymbol{H}$ [149]. This can be done by utilising the direct linear transform (DLT).

3.3.2.1 Direct linear transform

The DLT algorithm is the process of reforming equations into homogeneous linear equations and finding a solution by utilising conventional methods [150]. By expressing Equation 3.21 fully, in the form:

$$\begin{bmatrix} x'_{hom} \\ y'_{hom} \\ w'_{hom} \end{bmatrix} = \alpha \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \begin{bmatrix} x_{hom} \\ y_{hom} \\ w_{hom} \end{bmatrix}, \tag{3.22}$$

and by acknowledging that $x' = x'_{inhom} = x'_{hom}/w'_{hom}$ and similarly $y' = y'_{inhom} = y'_{hom}/w'_{hom}$, it can be deduced that,

$$x' = \frac{H_{11}x_{hom} + H_{12}y_{hom} + H_{13}w_{hom}}{H_{31}x_{hom} + H_{32}y_{hom} + H_{33}w_{hom}}, \tag{3.23}$$

$$y' = \frac{H_{21}x_{hom} + H_{22}y_{hom} + H_{23}w_{hom}}{H_{31}x_{hom} + H_{32}y_{hom} + H_{33}w_{hom}}. \tag{3.24}$$

Without loss of generality, it can be assumed that $x = x_{inhom} = x_{hom}/w_{hom} \simeq x_{hom}/1$ and similarly $y = y_{inhom} = y_{hom}/w_{hom} \simeq y_{hom}/1$, converting Equation 3.23 and Equation 3.24 to a pair of equations linear in the components of $\boldsymbol{H}$ [151]:

$$x'(H_{31}x + H_{32}y + H_{33}) - H_{11}x - H_{12}y - H_{13} = 0 \tag{3.25}$$

$$y'(H_{31}x + H_{32}y + H_{33}) - H_{11}x - H_{12}y - H_{13} = 0. \tag{3.26}$$

These can be rewritten in the form,

$$\boldsymbol{a}_{\boldsymbol{x}}^{\top}\boldsymbol{h} = \mathbf{0} \tag{3.27}$$

$$\boldsymbol{a}_{\boldsymbol{y}}^{\top}\boldsymbol{h} = \mathbf{0} \tag{3.28}$$

where,

$$\begin{aligned} \boldsymbol{h} &= (H_{11}, H_{12}, H_{13}, H_{21}, H_{22}, H_{23}, H_{31}, H_{32}, H_{33})^{\top} \\ \boldsymbol{a_x} &= (-x, -y, -1, 0, 0, 0, x'x, x'y, x')^{\top} \\ \boldsymbol{a_y} &= (0, 0, 0, -x, -y, -1, y'x, y'y, y')^{\top} \end{aligned} \tag{3.29}$$

3.3.2.2 Homogeneous linear least squares problem

The system of equations derived from Equations 3.27, 3.28, and 3.29, and the four sets of corresponding image points; $(x_n, y_n) \rightarrow (x'_n, y'_n)$, for $n = 1, 2, 3, 4$, can be written as:

$$\begin{bmatrix} -x_1 & -y_1 & -1 & 0 & 0 & 0 & x_1x'_1 & y_1x'_1 & x'_1 \\ 0 & 0 & 0 & -x_1 & -y_1 & -1 & x_1y'_1 & y_1y'_1 & y'_1 \\ -x_2 & -y_2 & -1 & 0 & 0 & 0 & x_2x'_2 & y_2x'_2 & x'_2 \\ 0 & 0 & 0 & -x_2 & -y_2 & -1 & x_2y'_2 & y_2y'_2 & y'_2 \\ -x_3 & -y_3 & -1 & 0 & 0 & 0 & x_3x'_3 & y_3x'_3 & x'_3 \\ 0 & 0 & 0 & -x_3 & -y_3 & -1 & x_3y'_3 & y_3y'_3 & y'_3 \\ -x_4 & -y_4 & -1 & 0 & 0 & 0 & x_4x'_4 & y_4x'_4 & x'_4 \\ 0 & 0 & 0 & -x_4 & -y_4 & -1 & x_4y'_4 & y_4y'_4 & y'_4 \end{bmatrix} \begin{bmatrix} H_{11} \\ H_{12} \\ H_{13} \\ H_{21} \\ H_{22} \\ H_{23} \\ H_{31} \\ H_{32} \\ H_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{3.30}$$

or more succinctly,

$$\boldsymbol{A}\boldsymbol{h} = \mathbf{0} \tag{3.31}$$

which is the formulation of a homogeneous linear least squares problem [152]. To solve this problem, singular-value decomposition (SVD) can be used.

3.3.2.3 Singular-value decomposition

The SVD theorem states that any matrix $\boldsymbol{A_{n\times m}}$ can be expressed as:

$$\boldsymbol{A_{n\times m}} = \boldsymbol{U_{n\times n}}\boldsymbol{\Sigma_{n\times m}}\boldsymbol{V}_{\boldsymbol{m\times m}}^{\top} = \sum_{i=1}^{9} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\top}, \tag{3.32}$$

where $\boldsymbol{\Sigma}$ is diagonal and contains the singular values, and the left and right singular vectors, $\boldsymbol{U}$ and $\boldsymbol{V}$, are orthonormal [151, 153]. The columns of $\boldsymbol{V}$ correspond to the eigenvectors or $\boldsymbol{A}^\top\boldsymbol{A}$. SVD is a common technique for a range of scientific and engineering problems and as such there are software implementations for solving it. SVD is solved in this work using Matlab [144, 154]. When solved, the last column, or right singular vector, of $\boldsymbol{V}$ gives the coefficients of $\boldsymbol{h}$.

3.3.2.4 Overdetermined solutions & robustness

Although $\boldsymbol{H}$ can be theoretically solved by four corresponding point pairs, in practical application, the image points projected into the camera are subject to both noise and errata. Therefore a large number of features are utilised to increase the robustness of the solution [146]. In the case of n image point pairs, $\boldsymbol{A}$ will be a $2n \times 9$ matrix. Therefore the homography will be overdetermined, and the value σ_9 from Eq. 3.32 will be non zero and will represent the residual, or goodness of fit, of the solution (when the homography is exactly determined, $\sigma_9 = 0$) [151].

To remove errata from the point correspondences, and provide the most likely correct solution to the projectivity, a random sample consensus (RANSAC) method can be utilised.

3.3.2.5 Random sample consensus

The RANSAC algorithm is an iterative sampling methodology, first proposed by *Fischler and Bolles*, that removes erroneous data from that used in the estimation of a model's parameters [155]. Data, in this case image point correspondences, that are representative of the projectivity are considered to be 'inliers', and those that are errata are considered to be 'outliers'. Many sampling techniques use the maximum amount of data possible, and then aim to remove outliers, RANSAC however uses the minimal required amount of data and then expands the selection to include inliers. In the case of planar homography, the minimum required amount of data is four point correspondences. Therefore, four points are randomly selected, and $\boldsymbol{H}$ is solved for using the aforementioned methods. All of the image points in the data set $\boldsymbol{x}$ are then projected into the second image plane by $\boldsymbol{Hx}$ and the error, e, between the projected

points, $\boldsymbol{x'_{projected}}$ and the observed points, $\boldsymbol{x'}$ is measured as the Euclidean distance between them. The error for point i is therefore,

$$e_i = \sqrt{(x'_{i,projected} - x'_i)^2 + (y'_{i,projected} - y'_i)^2} \tag{3.33}$$

If the proportion, λ of the data set that falls within a threshold distance, T, of the observed points is above a pre-specified threshold, γ, then the algorithm stops looping, and recalculates $\boldsymbol{H}$ with all of these inliers. Otherwise, the algorithm continues to resample and estimate $\boldsymbol{H}$ until a predefined limit of iterations, N, is reached. If this limited is reached, the best prior solution is utilised (Algorithm 1).

Algorithm 1 RANSAC for solving overdetermined planar homography

Input:

$\boldsymbol{x}, \boldsymbol{x'}$ - image point correspondences in image 1 and 2
N - Maximum number of estimation attempts
T - Maximum Euclidean pixel distance
γ - Minimum required ratio of inliers within $\boldsymbol{x}$

Output:

$\boldsymbol{H}$ - Estimated homography
$\boldsymbol{x_{inliers}}, \boldsymbol{x'_{inliers}}$ - Image point correspondences that fit $\boldsymbol{H}$

procedure RANSAC($\boldsymbol{x}, \boldsymbol{x'}, N, T, \gamma$)
  for Iteration number 1 : N **do**
    Randomly select 4 point correspondences from $\boldsymbol{x} \rightarrow \boldsymbol{x'}$
    Calculate $\boldsymbol{H}$ using DLT and SVD
    Project all points $\boldsymbol{x'_{projected}} = \boldsymbol{H}\boldsymbol{x}$
    Calculate $\boldsymbol{e} = |\boldsymbol{x'} - \boldsymbol{x'_{projected}}|$
    Calculate and store, λ, as proportion $\boldsymbol{e} < T$
    if $\lambda > \gamma$ **then**
      Designate $\boldsymbol{x_{inliers}}$ as those with $e < T$
      break for loop
    else if Iteration number == N **then**
      Find from stored λ, $max(\lambda)$
      Designate $\boldsymbol{x_{inliers}}$ as those with $e < T$ for $max(\lambda)$
    else
  Recalculate $\boldsymbol{H}$ using all $\boldsymbol{x_{inliers}}$
  Reproject and designate $\boldsymbol{x_{inliers}}$ based on final $\boldsymbol{H}$

To estimate a suitable number for maximum number of iterations, N, it can be deduced that the probability, p, of selecting a representative set of image point correspondences is [156]:

$$p = 1 - (1 - (1 - \epsilon)^s)^N, \tag{3.34}$$

where ϵ is the expected proportion of errata within the point correspondences, and s is the number of image points in the selection. Therefore, if a bad scenario is envisaged, where only 25% of the data are correct, and the desired probability of 99% certainty that the correct homography be found, the number of required iterations would be:

$$N = \frac{ln(1-p)}{ln(1-(1-\epsilon)^s)} = \frac{ln(1-0.99)}{ln(1-(1-0.75)^4)} \approx 1177. \tag{3.35}$$

It is also important to note that if there are errata, that are themselves consistent to an incorrect homography, and these are of greater number than the data that represent the true homography; then it is more likely that the incorrect homography will be returned.

3.3.2.6 M-estimator sample consensus

A limitation of the RANSAC method is that if the required distance threshold, T, is set too high, then the homography estimation becomes quite poor. This is because points that fit the model well are considered equally to those that fit less well. As such RANSAC can be considered the minimisation of the cost function, C:

$$C = \sum_i \rho(e_i) \tag{3.36}$$

where,

$$\rho(e) = \begin{cases} 0 & e < T \\ constant & e \geq T. \end{cases} \tag{3.37}$$

It is therefore perhaps beneficial to utilise a different cost function, C_2, whereby points beyond the threshold still incur a constant penality, however points within the threshold are weighted based on their fit:

$$C = \sum_i \rho_2(e_i), \tag{3.38}$$

where,

$$\rho_2(e) = \begin{cases} e & e < T \\ T & e \geq T. \end{cases} \tag{3.39}$$

This is an example of a redescending M-estimator, and its implementation is termed the M-estimator Sample Consensus (MSAC) [156, 157].

3.3.3 Relative camera pose estimation

Knowing the homography that links $\boldsymbol{x} \rightarrow \boldsymbol{x'}$, it is often desirable to calculate the 3D rotational and translational transformation between camera 1 and camera 2 (Figure 3.9).

3.3.3.1 Euclidean homography decomposition

In Equation 3.20 and 3.21, $\boldsymbol{H}$ was defined as,

$$\boldsymbol{H} = \boldsymbol{R} - \boldsymbol{T}\boldsymbol{n}^\top, \tag{3.40}$$

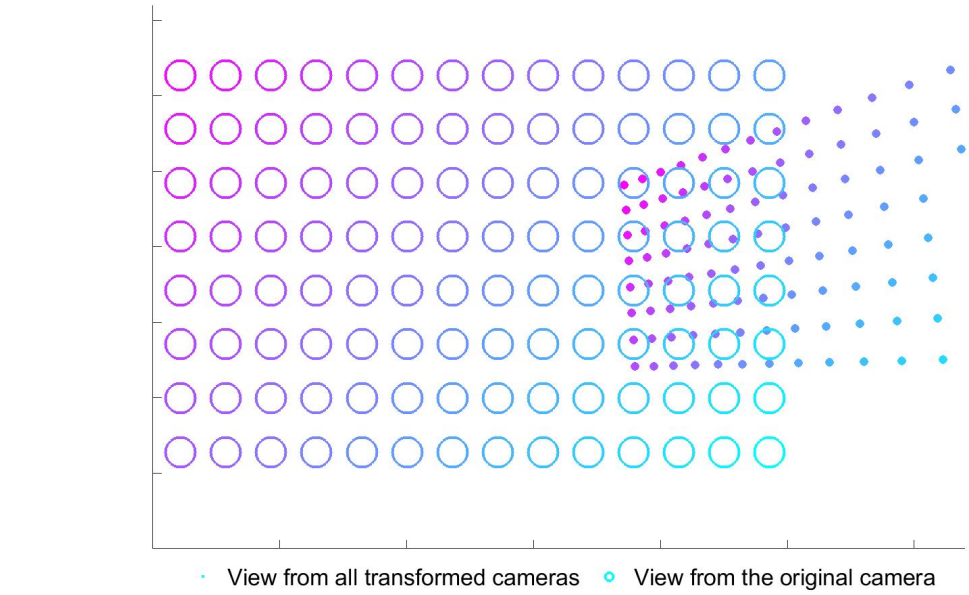
and the extraction of relative camera rotation, translation and normals to the plane, $\boldsymbol{n}$ that induces the homography, from $\boldsymbol{H}$ itself, is known as Euclidean homography decomposition:

$$\boldsymbol{H} \Rightarrow \{\boldsymbol{R}, \boldsymbol{T}, \boldsymbol{n}\}. \tag{3.41}$$

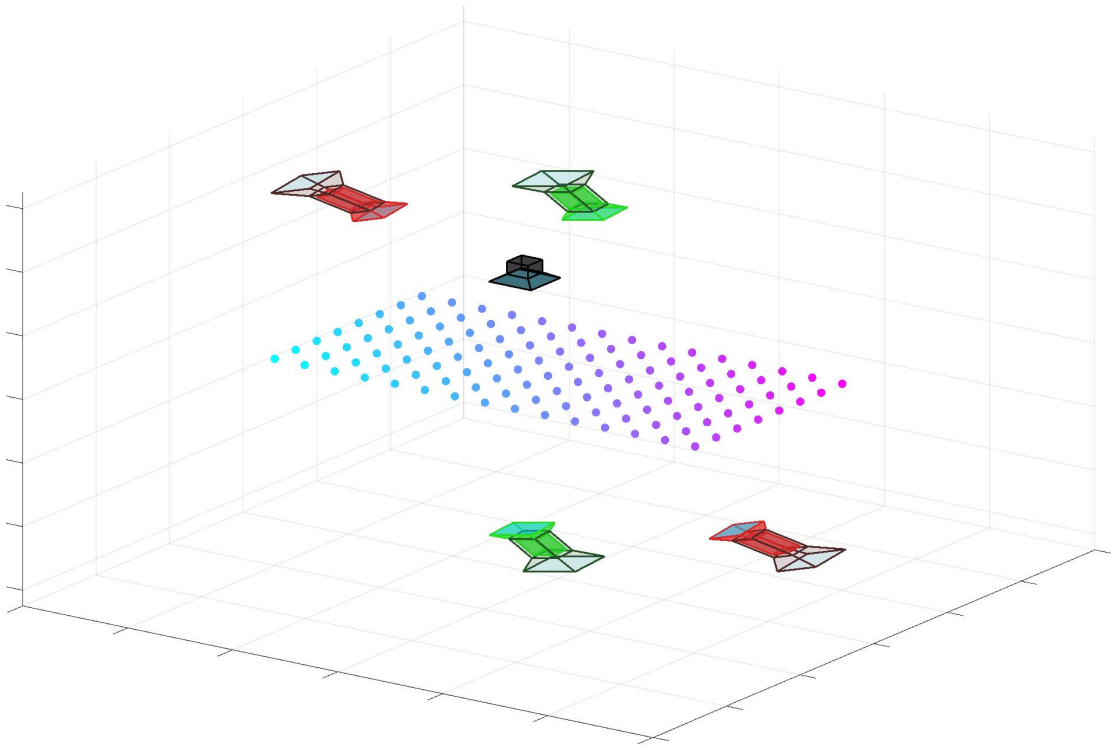
Various methods have been developed to achieve this, including numerical SVD-based methods by *Faugeras & Lustman* [158], and *Zhang & Hanson* [159], as well as analytical methods by *Malis & Varags* [160].

These methods are complex and have standard implementations in openCV for use by the computer vision community [161]. There are two key results of these implementations. Firstly, the translation vector $\boldsymbol{T}$ is normalised with respect to the plane depth, d (distance to the plane). Secondly, that $\boldsymbol{H}$ has up to eight mathematical solutions (Figure 3.10b) [162].

The two valid, realisable physical solutions are the red and green cameras on the same side of the plane as the first camera (black). To remove those on the other side of the plane, a reference-plane non-crossing constraint is applied (i.e. both frames must be the same side of the plane). Plane crossing occurs when the relative translation



(a)



(b)

Figure 3.10: Multiple homography decomposition solutions. (a) Shows a grid of planar points and the relative position of an original camera (black) and the cameras resulting from all eight possible transformations that yield the same perspective change. This demonstrates that there are two distinct sets of solutions (red and green). Each transformation has a valid counterpart on the other side of the plane, and each of these has a mathematically valid inverted solution (transparent framework cameras). (b) Shows the superimposition of the two perspectives of the planar points; that from the original camera and all possible solutions.

that occurs in the direction of the plane normal is greater than the distance from the plane to the first camera, d. This makes the constraint [160]:

$$\boldsymbol{T}^\top \boldsymbol{n} < d. \tag{3.42}$$

As translation is normalised with respect to the plane depth, this enables the constraint to be represented as,

$$1 + \boldsymbol{n}^\top \boldsymbol{R}^\top \boldsymbol{T} > 0. \tag{3.43}$$

Applying this constraint leaves the four solutions on the same side of the plane as the first camera. To reduce these down to the two physical solutions, a reference point visibility constraint can be applied. This asserts that point, $\boldsymbol{X}$, that corresponds to the image points $\boldsymbol{x}$ and $\boldsymbol{x'}$ must be in front of both cameras. To achieve this however, the world coordinates of this point must be calculated; this is done via triangulation.

3.3.3.2 Triangulation

Triangulation is the localisation of a point in space given that it is viewed in at least two images. Given the camera matrices of the two cameras, finding the point at which two rays intersect should be trivial. However, if there is noise in the image, or uncertainty in the the relative camera poses or intrinsics, then it is not guaranteed, and in fact unlikely that the rays intersect (Figure 3.11). There are a number of methods for finding the most likely solution but the 'midpoint method' is that which is outlined here. It is acknowledged that this method is considered the least accurate or robust, but it is highly computationally simplistic. As the purpose of triangulation in this case is to ascertain which of the camera pose solutions provide negative or positive real world points, accuracy is of little concern but computational burden is.

Let the camera matrix of the first camera be decomposed as,

$$\boldsymbol{P} = (\boldsymbol{M} | -\boldsymbol{MC}), \tag{3.44}$$

then, similar to back-projection (Equation 3.12), any point on the ray, $\boldsymbol{X}$ can be expressed as,

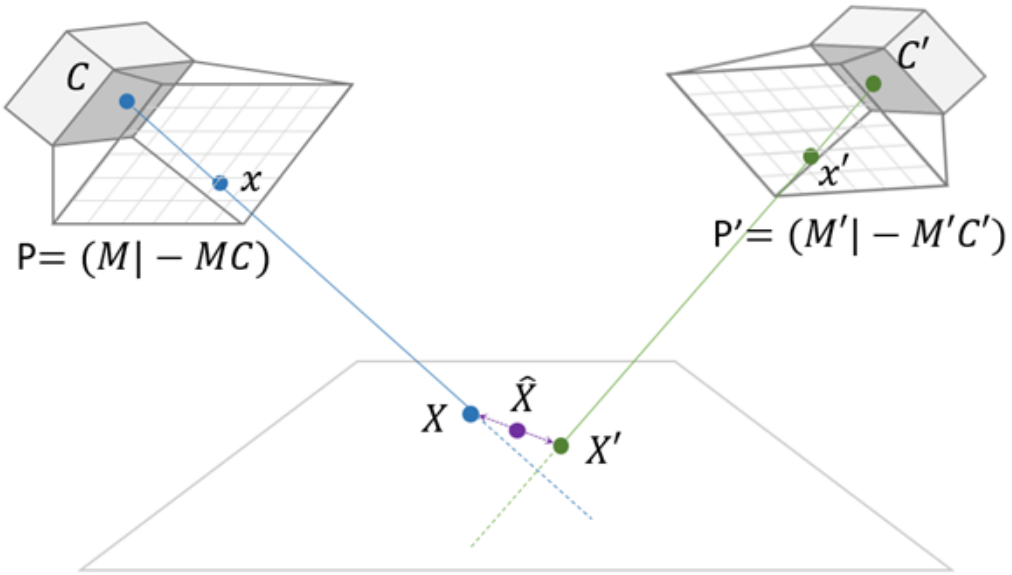


Figure 3.11: Triangulation. If two image points do not project to intersecting rays, a midpoint between the rays $\hat{\boldsymbol{X}}$ can be found using the midpoint method.

$$\boldsymbol{X} = \alpha \boldsymbol{M}^{-1}\boldsymbol{x} + \boldsymbol{C}, \tag{3.45}$$

for some α. The same can be applied to the other camera's ray, and acknowledging that the rays should intersect, it results that:

$$\alpha \boldsymbol{M^{-1}x} - \alpha' \boldsymbol{M'}^{-1}\boldsymbol{x'} = \boldsymbol{C} - \boldsymbol{C'} \tag{3.46}$$

In homogeneous coordinates, this is an over-determined linear equality and linear least-squares methods may be used to solve for α and α'. With these unknowns found, the midpoint between the two rays is clearly [163]:

$$\hat{\boldsymbol{X}} = \frac{\boldsymbol{X} + \boldsymbol{X'}}{2} = \frac{(\alpha \boldsymbol{M}^{-1}\boldsymbol{x} + \boldsymbol{C}) + (\alpha' \boldsymbol{M'}^{-1}\boldsymbol{x'} + \boldsymbol{C'})}{2}. \tag{3.47}$$

With the locations of the points determined, they can be rotated and translated into the second camera's coordinate system, and then those with negative z components are deemed behind the camera. In application, erroneous image points will likely be included. Therefore it is possible that all solutions have a combination of negative and positive z distances. Therefore the valid solutions are deemed those with the highest ratio of positive distances.

Application of these constraints leaves only the two physically valid solutions. However, assuming that $\boldsymbol{H}$ has been calculated from a real transformation, only one of them is the actual 'truth'. There are a number of ways of determining which is the correct solution, including the use of a second plane and known geometric relationships within it, or utilising geometric relationships or features within the plane [158]. However, here the additional constraint of a third view of the plane shall be used.

3.3.3.3 True solution determination

To determine which of the two remaining solutions is the true solution, a third view of the plane can be used. The decomposition of the first-to-second view homography $\boldsymbol{H}_{12}$ renders two physically valid estimations for the normal to the plane $\{\boldsymbol{n}^{a}_{12}, \boldsymbol{n}^{a}_{12}\}$. If the entire process outlined in Sections 3.3.1, 3.3.2 and 3.3.3 is repeated for a third view of the plane, then a first-to-third view homography is generated, $\boldsymbol{H}_{13}$, and offers two additional solutions to the plane $\{\boldsymbol{n}^{a}_{13}, \boldsymbol{n}^{a}_{13}\}$.

A compatible pair of solutions that represents a common plane therefore determines the correct solution for both homography decompositions. Due to possible errors in all previous calculations, the two correct planes may not be perfectly coherent. Therefore, to determine the most compatible pair, the plane normals with the smallest angular distance, Θ between them is selected [158, 164]:

$$\Theta = \cos^{-1}\left(\frac{\boldsymbol{n}^{a/b}_{12} \cdot \boldsymbol{n}^{a/b}_{13}}{\left\|\boldsymbol{n}^{a/b}_{12}\right\| \left\|\boldsymbol{n}^{a/b}_{12}\right\|}\right). \tag{3.48}$$

3.3.3.4 Scale determination

Alongside the solution ambiguity problem discussed above, the scale (or similarity) ambiguity problem is one of the most fundamental and persistent problems in homography decomposition, and its fields of application [146, 165, 166]. Also similarly, there are approaches that utilise planar structures within the views [165, 167]. As covered, the homography is only determined up to scale (Equation 3.21). As such, the recovered estimation of translation $\boldsymbol{T}$ is a normalised vector. To determine the absolute metric scale of translation, additional information is required. A range of methods have been developed based on use case and data set, including the provision

of a stereo camera [168], prior knowledge of the 3D structure of the environment [169], integration with other sensors such as IMUs or global positioning system (GPS) [170, 171], and wheel odometry [172].

Many applications use an identified plane in the scene, and a known 'distance to the plane', h_K, to solve this problem. This is often a desirable solution as many of the applications are related to autonomous vehicles which have a readily observable planar environment, the road, and a known distance to the plane, the height of the camera mount on the vehicle . Although a range of embellishments and case-dependent improvements are available, the fundamental premise revolves around the fact that Equation 3.20 can be represented as [160, 173]:

$$\boldsymbol{x'} = \boldsymbol{K}(\boldsymbol{R} + \frac{\boldsymbol{T}\boldsymbol{n}^\top}{h_E})K^{-1}\boldsymbol{x} \tag{3.49}$$

Where h_E is the estimated height of the camera from the plane. From this, a scaling factor, s, can be established, where;

$$s = \frac{h_K}{h_E}. \tag{3.50}$$

3.3.4 Camera calibration

In order to determine the parameters that define the camera's intrinsics matrix (Section 3.3.1.1); camera calibration must be conducted. To achieve this, all the aforementioned methods are combined for multiple views of a known calibration pattern. Practically, the camera is moved around with a checkerboard pattern always in the field of view (FOV). The non-symmetric checkerboard is on a flat surface with known dimensions. Therefore, the corners of each square can be identified in every image and the relative camera poses and 3D locations of the checkerboard can be estimated in 3D. Given these estimations, the projection error of the corners in the checkerboard can be estimated and minimised by altering the internal camera parameters. The parameters that give the reprojections with the lowest error are taken as the camera intrinsics. The computational process behind this is conducted by the Matlab Camera Calibrator App [144].

3.4 Object detection

Chapter 7 utilises deep learning to detect feet in video frames. This is the field of object detection. Although the models and methods used are explained in the Chapter; here some fundamental metrics are discussed that pertain to the assessment of whether an object detection method has correctly identified an object. This involves: simple classification statistics (which are also used in Chapters 6 and 8 to determine the accuracy of the detection of gait events), how to judge the spatial similarity of two identified objects, and how to combine classification and spatial similarity into a tangible metric of success.

3.4.1 Classification statistics

In a classification problem, classifications can be categorised whereby true positives, TP, are those correctly identified, false positives, FP, those incorrectly identified, true negatives, TN, those that are correctly ignored, and false negatives, FN, those that are erroneously neglected. Several derivative statistics can be calculated from these classifications: Sensitivity, specificity, recall, negative predictive value (NPV), positive predictive value (PPV), and $F - Score$. These may have slightly different interpretations depending on whether the classification task is binary or multi-class.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{3.51}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{3.52}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3.53}$$

$$\text{Precision} = \text{Positive predictive value} = PPV = \frac{TP}{TP + FP} \tag{3.54}$$

$$\text{Negative predictive value} = NPV = \frac{TN}{TN + FN} \tag{3.55}$$

$$\text{F-Score} = \frac{2}{\text{recall}^{-1} + \text{precision}^{-1}} = \frac{2.\text{precision.recall}}{\text{precision} + \text{recall}} = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \tag{3.56}$$

3.4.2 Assessment metrics

To evaluate an object detection model, two separate components must be assessed. One is the classification aspect; if an object exists within the image, and if so, what type, or class of object is present. The other is the localisation component; that assesses the similarity in locality of the object within the image to the ground truth. The location of objects within an image are delineated by bounding boxes.

3.4.2.1 Metrics of spatial similarity

Given two detections A and B, delineated by a rectangular bounding boxes, several metrics of similarity can be computed. The Jaccard similarity, or in object detection, known as the Intersection over Union (IoU) is the ratio between the intersection of A and B, over the union of A and B,

$$IoU(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}. \tag{3.57}$$

Similarly, the Szymkiewicz-Simpson coefficient, or Overlap Coefficient (OLC) is the union over the smaller of the two sets,

$$OLC(A, B) = \frac{|A \cap B|}{min(|A|, |B|)}. \tag{3.58}$$

.

The minor overlap coefficient is similar but accounts for the larger of the sets this has no formal name but its calculations will be important in algorithms,

$$MOLC(A,B) = \frac{|A \cap B|}{max(|A|,|B|)}. \tag{3.59}$$

.

And finally, the Sørensen-Dice coefficient (SDC):

$$SDC(A,B) = \frac{2|A \cap B|}{|A| + |B|}. \tag{3.60}$$

.

3.4.2.2 Precision-recall curves & mean average precision

By setting a threshold value for spatial similarity, for example an IoU of 0.5, an object detection instance can be determined as being a true positive, false positive or false negative when compared to the ground truth bounding box. With this established, the threshold confidence score can be varied, the relationship between precision and recall can be established, and a Precision Recall (PR) graph can be constructed (Figure 3.12). In accordance with standard practices, the PR curve is represented by interpolated precision where;

$$p_{interp} = \max_{r' \geq r}\{p(r')\} \tag{3.61}$$

and for precision, p, and recall, r [174]. Average precision (AP) can then defined as the area under the curve,

$$AP = \sum_{i=1}^{n-1}(r_{i+1} - r_i)p_{interp}(r_{i+1}). \tag{3.62}$$

This can be generalised across K object classes, as mean average precision (mAP), which is the most commonly reported figure in object detection evaluation;

$$mAP = \frac{\sum_{i=1}^{K} AP_i}{K}. \tag{3.63}$$

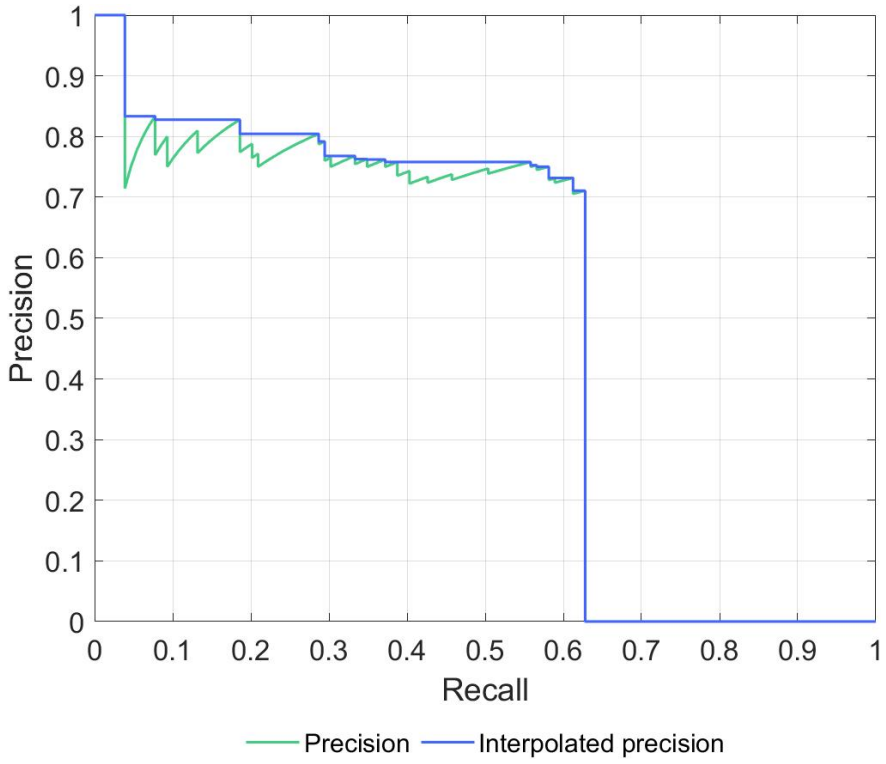


Figure 3.12: Precision-Recall curve example.

3.5 Conclusion

This chapter provides a basic grounding in the underlying methods used in subsequent chapters. The bulk of the mathematics pertains to the multiview geometry relevant to the visual odometry in Chapter 6. The appraisal of object detection methods (for Chapter 7) are also covered - as well as classification statistics that are used throughout to assess the accuracy of event or object detection.

Some useful novel concepts are also introduced in this chapter. These are step maps and GET graphs, which will be used throughout this work to explain and analyse complex gait and its measurement by various wearables.

4

Validation study & data sets

Contents

4.1 Introduction **53**
4.2 Subjects **54**
4.3 Hardware, software & set up **55**
4.4 Protocols **59**
4.5 Ethical considerations **65**
4.6 Discussion **66**
4.7 Newcastle data set **68**
4.8 Ad hoc data **69**
4.9 Conclusion **70**

4.1 Introduction

The following chapters of this thesis will investigate the capacity for motion capture, wearable cameras, and wearable accelerometers to assess macro gait features. In order to both validate and directly compare each method, a study was designed and conducted that enabled the concurrent testing of each of the technologies.

An additional aim for this study, beyond the direct comparison of the capacity for wearables to extract a wide range of gait features, is the exploration of complex free-living gait. Therefore, this chapter outlines a battery of novel protocols, and offers explanation and justification for their development. Beyond the protocols, other key considerations are outlined and discussed, including: participants, hardware, software, and ethics.

In order to further explore the application of wearables in assessing free-living gait and inform the development of robust computer vision methods applicable to

the real world, a data set was obtained from Newcastle University. This data set is also described as well as ad hoc data collected to demonstrate and test the wearables beyond the remit of the official study.

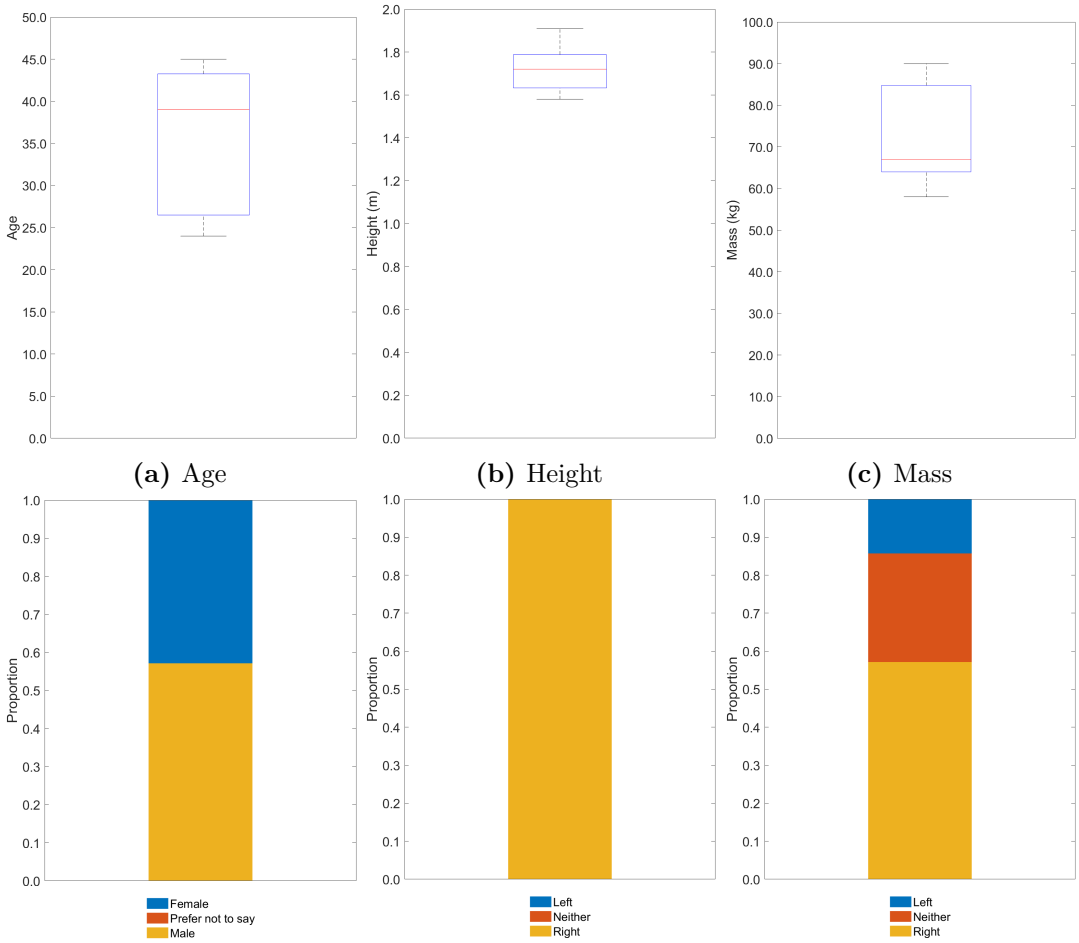
4.2 Subjects

Seven healthy adult subjects were recruited into the study. Figure 4.1 shows the distributions of the key attributes of the subjects. Leg and arm dominance were recorded as the wearable device set up was anatomically asymmetrical (Figure 4.2a). Inclusion and exclusion criteria were as follows:

- Inclusion criteria
  - Participants must be able to walk unaided.
  - Participants must be between the ages of 18 and 65 inclusive.
  - Participants must be capable of providing free and informed consent.

- Exclusion criteria
  - Health problems, recent surgery, injuries or medical problems that may impede their capacity to walk and safely carry out the protocols.
  - Abdominal proportions that would obstruct the view between the chest-mounted camera and the participants' feet; for example pregnancy or obesity.

In addition to these strictly enforced criteria, it was also requested of participants that they wear generic, plain coloured, close-fitting shoes and garments. This was to prevent the obstruction either of the field of view of the camera (Figure 4.2e), or of the motion capture cameras from the worn markers.



(a) Age **(b)** Height **(c)** Mass

(d) Gender **(e)** Dominant arm **(f)** Dominant leg

Figure 4.1: Group statistics of study participants.

4.3 Hardware, software & set up

This study incorporated the simultaneous use of motion capture, a wearable camera and wearable IMUs to track movement of various parts of the body. The recorded videos, accelerometry signals and motion capture trajectories were synchronised by informing each subject before each protocol to jump up and down twice, ensuring that their feet left the floor. This ensured that a sharp vertical movement and acceleration could be identified in each device's signal.

4.3.1 Camera

Subjects wore a Garmin VIRB Ultra 30 action camera (88g) in a waterproof case, attached to the body by a generic action camera chest harness (Fig. 4.2c). The camera was set to point downwards towards the feet (Fig. 4.2e). Precise positioning was

unnecessary as part of the purpose of the study was to prove the system's robustness for use, however it was advantageous to provide the most consistent view of the feet. The camera settings were: 720p wide field of view, 30 frames per second (FPS), standard orientation, lens correction off, stabilizer off. Recordings were stored on a SanDisk Ultra 64Gb memory card.

4.3.2 Accelerometers

4.3.2.1 Smartphone

An iPhone 6S was placed in the front right pocket of the participant. The orientation of the smartphone was left to the participant's discretion. The IMU data was logged at 30Hz using the SensorLog iPhone application. This choice of location was based on: the fact that it is a common place to keep a smartphone in everyday life, precedent for using the right side specifically [175], and that there is strong evidence that the asymmetrical movement observed in the upper limbs is substantially less pronounced in the lower [176]. This allowed for the focus on consistency, and subject leg dominance was recorded (Fig. 4.1f).

4.3.2.2 Smartwatch

An Apple Watch Series 5 44mm was worn on the left wrist. The internal IMU was logged using the SensorLog Apple Watch application at a sampling rate of 30Hz [177]. The app was set to directly log recording to the iPhone via Bluetooth. The watch location was decided upon as previous studies have shown that devices on the left wrist provide higher accuracy activity detection [178]. Similarly, better step detection was observed in smartwatches worn on the non-dominant wrist [179]. To ensure consistency and real-life applicability (the majority of individuals wear smartwatches on their left wrist [180]), the left was chosen and the subjects' dominant hand recorded (Fig. 4.1e).

4.3.2.3 Foot and Lumbar IMUs

Foot and lumbar movement was recorded using stand alone Blue Trident IMUs (Vicon. Oxford, UK). The IMUs were set to record using the Low-G setting (range of $\pm 16g$) at a sampling frequency of 1125Hz. Device initialisation and data management was conducted using the Capture.U iOS (iPhone operating system) application connected

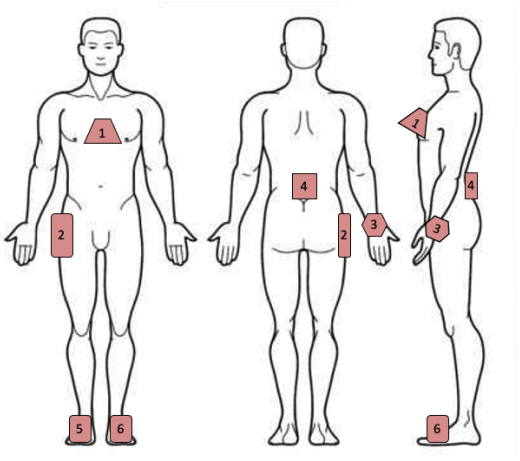
to the iPhone via Bluetooth. The data was uploaded onto a local machine using the Capture.U Windows Desktop software.

4.3.3 Motion capture

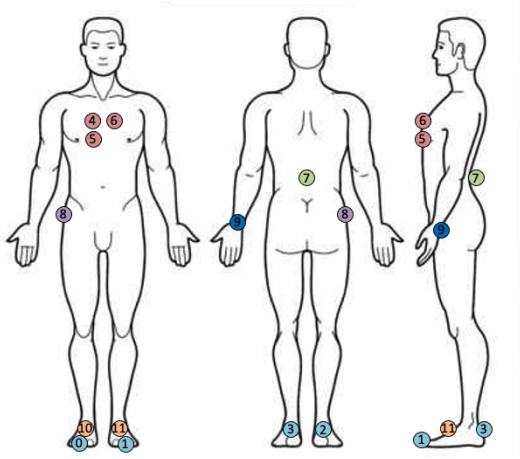
Motion capture ground truth was provided by the Oxford Gait Lab's 16 camera Vicon MX T40S motion capture system (Vicon. Oxford, UK), sampling at 100 Hz using the Vicon Nexus 1.8 software. The raw data was then processed to provide an output CSV (comma-separated values) of the 3D coordinates of each marker's location in the lab's coordinate system (more detail of this process can be found in Chapter 5).

Vicon is a thoroughly validated motion capture system that can track reflective markers attached to the participant to sub-millimetre accuracy [181]. The marker arrangement that was used in this study is shown in Figures 4.2b and 4.2d. They were attached to the front and back of the subjects' feet to enable tracking, as well as three co-planar markers on the camera case to determine orientation as well as location of the camera (Fig. 4.2d). A marker was placed on the smartwatch strap, close to the watch face, and over the participant's right pocket approximately above the location of the smartphone when standing upright and still. Additional markers were placed directly on top of the lumbar and feet IMUs. For the sake of brevity, Vicon will be used as a synonym for motion capture from this point onward. All markers were 14mm in diameter.

Before each capture session, the Vicon system was set up according to standard practice. The system was turned on and unwanted reflection were removed by applying camera masks in the software. Calibration was then conducted by waving the calibration wand about the desired capture volume until all camera-calibration indicators stopped flashing and turned green. The volume origin was then set with the same wand affixed to the central force plate. For quality assurance, the wand position is rotated by 180 degrees and the measured distances between the lights on the wand assessed. Maximum error in distance measurement was deemed to be 2mm.



(a) Device set up



(b) Marker placement



(c) Camera set up



(d) Camera markers



(e) Camera FOV

Figure 4.2: Experimental device and equipment set up. (a) Orientation and location of devices. (b) Markers are located on the toes and heels of the feet, on the pocket over the smartphone, on the smartwatch strap, on the lumbar IMU and on the camera. (c) The camera is worn using a chest strap approximately over the sternum. (d) the markers on the camera are three vertices of the camera screen to ensure they approximately form perpendicular vectors for orientation estimation. (e) The camera is pointed downwards with a clear view of the feet within the FOV.

4.4 Protocols

To validate the camera, accelerometry and motion capture systems a battery of protocols was developed. These protocols enabled the investigation of linear walking, turning while walking, and natural free-living movement. As none of these protocols have been used before specifically, justification for them is provided alongside the protocol descriptions. Instructions for each protocol were given in the order provided in the below sections. Throughout this thesis, the protocols are referred to in short as Straight, Turns, ADL1, ADL2, and ADL3 (activities of daily living).

4.4.1 Straight walking protocol

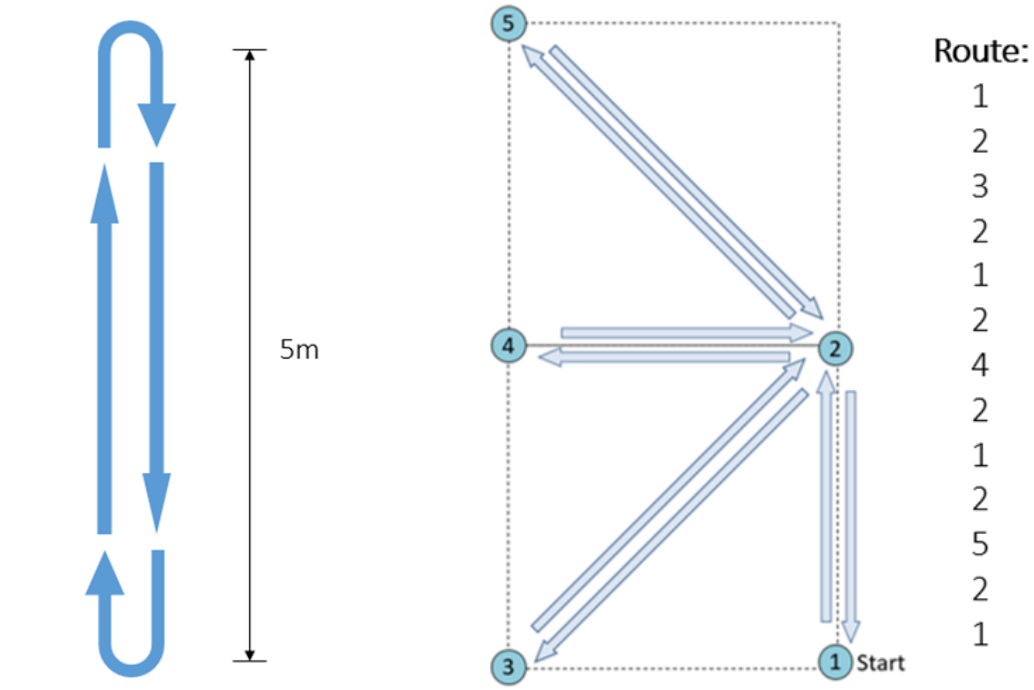
The participants were asked to walk between two markers spaced five metres apart, and upon reaching the end, turn around and return (Fig. 4.3a). This was then to be performed twice at slow, normal and fast speeds consecutively, in this order, and without stopping. Specific walking speed was determined by the participant as considered comfortable and natural.

This simplistic protocol was developed to enable direct comparison between this study and others. The most common protocols for the validation of gait features as extracted from wearables involve subjects walking short, linear distances down a corridor, on a treadmill, over an instrumented walkway, or across a motion capture laboratory [182, 183]. These protocols rarely involve turning or a large number of steps.

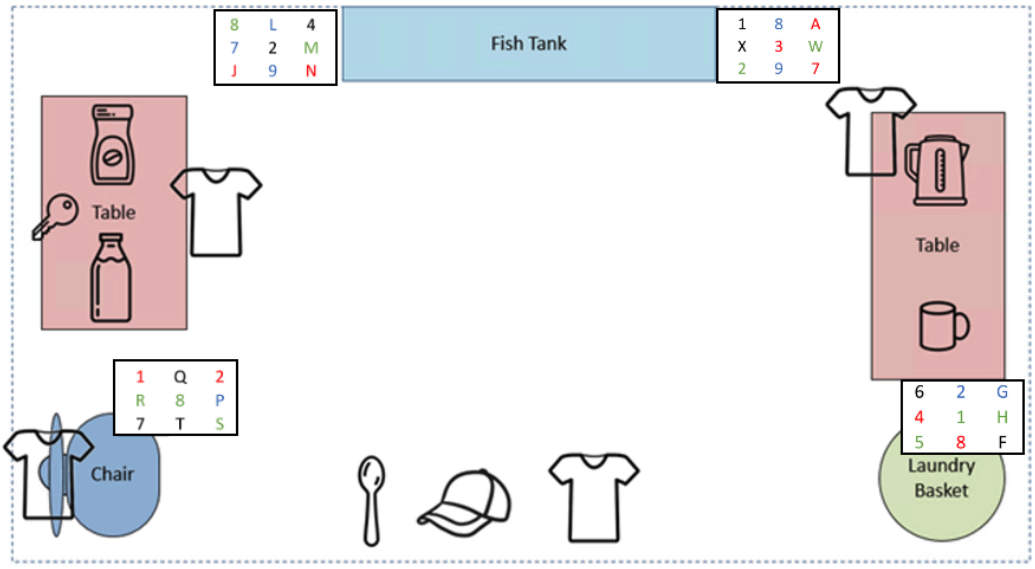
In addition, some research is conducted using protocols akin to the 2-minute walk test (2MWT) or 6MWT. To ensure constant walking over this period, subjects generally walk short linear distances including turns [184–186]. In fact the protocols used in this study often ran to about two minutes (Figure 4.4).

This protocol therefore exploits the larger number of steps and the inclusion of turns from the latter studies and the higher accuracy and resolution validation from the motion capture more comparable to the former type.

The importance of encouraging gait of different speeds, is that walking speed has been shown to affect the accuracy of gait features extracted from wearables. This is particularly true at slower and faster speeds and is relevant to step counts [187], walking speed [188], step length [189], turn detection and measurement [190].



(a) Straight protocol **(b)** Turns protocol



(c) ADL protocol

Figure 4.3: Schematics of the protocols. (a) The Straight protocol involves the subject walking up and down twice over a distance of five metres, and turning about at the end of each straight. (b) The Turns protocol incorporates 45°, 90°, 135° and 180° turns by getting the participant to follow the route outlined in the figure. The edges of the boxes are 2 metres. (c) An example of the arrangement of the space used for the ADL protocols, all the items are present throughout each of the three protocols. The protocol space is 2.5 by 4 metres.

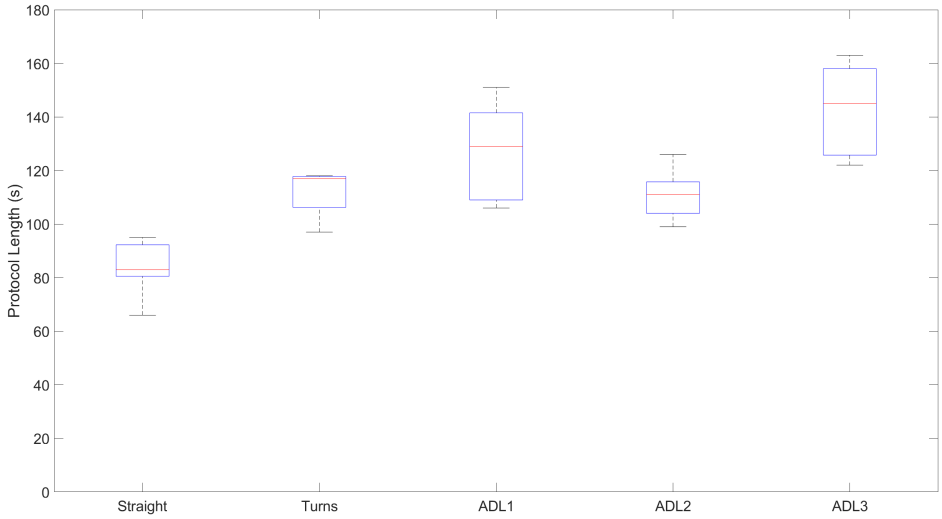


Figure 4.4: Protocol durations. Distribution of the times taken to complete each protocol.

4.4.2 Turning protocol

The participants are asked to walk the approximate route outlined by a series of markers on the floor (Fig. 4.3b). This route includes multiple turns of 45°, 90°, 135° and 180° to both sides. No prescription is given as to how to turn or move between the markers. A specific route is deliberately not marked out to encourage participants to act as naturally as possible. The balance is therefore struck between encouraging natural gait while also ensuring a range of turn angles and methods. The whole route is walked at slow, normal and fast speeds for the same reasons as explained in Section 4.4.1.

The importance of being able to detect and measure turning in gait is outlined in Section 2.3.2. However, there are remarkably few studies that actually investigate this specifically. Those that do, tend to limit the subjects to 90° turns. Furthermore, those that conduct high resolution investigation using motion capture, often only include a single highly modulated turn of 90° [191, 192]. Although some studies attempt to vary the turning angle, often this includes those that would not be highly representative of everyday gait, such as a 360° turn [193].

4.4.3 Activities of daily living protocols

As discussed in Section 2.4, a substantial proportion of gait is conducted in very short bouts around the home. The is especially true for the demographics for whom gait analysis offers the most potential useful outcomes: the elderly, partially mobile, and

patient populations. Despite this, there is a paucity of rigorously tested protocols that try and imitate these free-living conditions in the gait lab. Numerous studies conduct protocols outside of a motion capture lab, however in such a scenario there is no real provision of a high accuracy and high resolution ground truth.

This leaves one of two options: attempting to install motion capture in the home and other everyday locations, or trying to imitate everyday activities in the motion capture lab. The former options are covered in Section 2.4, but are intrinsically limited in quality. There are very few examples of the latter option. In one study, less prescribed gait was encouraged by getting subjects to walk continuously and intermittently in a room with no defined path [75]. Other studies tend to equate ADLs with categories common in activity recognition, such as running, walking, standing, chair-to-standing transition, stair climbing, etc. However, these are of course only macro events within complex free-living movement.

Therefore there was little precedent for attempting to study free-living gait in a motion capture lab. As a result, three protocols were developed that could cover the greatest range of ADLs that are both most common [194], and also the most relevant to quality of life (QOL) [195, 196]. An additional key consideration in protocol development was the best utilisation of the limited capture volume afforded by the Vicon lab, and the limited items and objects that would be practical within a motion capture lab. For example, most furniture would obscure the Vicon camera's view of the markers.

4.4.3.1 ADL1 protocol: Dressing & tidying

The participant is asked to sit down in the chair, then stand up. They are informed that they are to imitate getting ready to leave their home. They are then encouraged to move around and choose items to take with them (e.g. gloves, keys, mask). Sometimes objects are hidden to encourage searching.

Next they are informed that they are to imitate returning home. Items are removed from their person and placed around the room wherever they choose. They are then informed that they should tidy the room and are left to interpret this as they see fit.

4.4.3.2 ADL2 protocol: Preparing & eating food

The participant is asked to sit down in the chair, then stand up. They are informed that they are to imitate preparing breakfast. They are then asked to tidy away afterwards. Empty cereal boxes, milk cartons, juice cartons, coffee containers, crockery and cutlery are present for their manipulation. They are encouraged to act out use of a fridge, sink or cupboard if they are so inclined. Throughout the task, the researcher engages them in casual conversation to help distract them from the contrivance of the activities and imitate a more natural food preparation routine.

4.4.3.3 ADL3 protocol: Motor-cognitive dual tasking

The participant is asked to sit down in the chair, then stand up. They are informed that there are four sheets of A4 paper in the room (Fig. 4.5) which have been placed so that they cannot all be seen from a single location. They are then asked the following questions which they should answer without touching any of the sheets, or obstructions:

- How many red numbers are there?
- What is the sum of the green numbers?
- Which colour has the most numbers?
- Which numbers or letters appear on every sheet?

Part of the aim of the questions and mental tasks in this protocol is to distract from the contrivance of the experiment and the unnaturalness of the equipment (it was qualitatively noticed that the subjects moved more naturally during this protocol). However, beyond this, motor-cognitive dual task tests are commonly used to investigate the relationship between gait and cognition - especially in the elderly and individuals with neurodegenerative diseases [197]. This is partly because the combination is considered to be a method of assessing a subject's executive function [198], but also pragmatically because divided attention and reduced dual tasking function is associated with falling prevalence [199, 200]. Furthermore, dual tasking has a direct effect on gait in healthy adults, such as causing a decrease in walking speed [201].

Therefore its inclusion increases the range of gait types, and thus applicability of the study to true free-living gait.



Figure 4.5: Motor-cognitive dual tasking exercise sheets used in the ADL3 protocol.

4.4.3.4 ADL protocol preparation & additional considerations

A region well within the Vicon capture volume (approximately $5 \times 3 \times 2$ metres, and thus comparable to a typical room [202]) was prepared by the addition of chairs, several surfaces of varying height, and an array of objects. The objects for all of the protocols were present throughout, as this increases clutter and realistic interactions. The exact set up of the ADL study environment was not the same for every subject and session, this was to increase the range of movements and actions that were recorded, and thus increase the study's applicability to true free-living gait.

Additional considerations went into the design of these ADL protocols. Many of these derived from qualitative observations from the Newcastle data set (Section 4.7).

- People move differently around objects and obstructions. Studies observe this and it is important in fall avoidance and imitating real-world gait [203–205].
- Furniture and clutter may cause problems such as view occlusion and false positive object detections [206]; but are heavily present in an everyday environment.
- Having tasks to do removes the contrivance of the experiment, hopefully reducing the Hawthorne effect [76].
- Accelerometry in smartwatches is likely heavily influenced by the movements of the arms during gait and activities [207].

4.5 Ethical considerations

Ethical approval to conduct the study was applied for and obtained prior to commencement. The approval was granted by the Medical Sciences Interdivisional Research Ethics Committee of Research Services at the University of Oxford (Ref. R70970/RE001, 7th October 2020). Additionally, Good Research Practice for clinical researchers training was undertaken by the researcher.

All participants provided informed written consent at the beginning of each data collection session. Participants were informed that they could withdraw from the study at any time and request their data destroyed.

4.5.1 Data management & security

All data collected during study sessions was immediately transported from the Oxford Gait Lab (OGL) to the Institute of Biomedical Engineering (IBME) whereby it was uploaded from the devices onto a password-protected department computer that is on an independent intranet within the institute. All data was then deleted from the devices. All subject measurement information was digitised and stored similarly. All personally identifiable subject data including physiological measurements was pseudo-anonymised and stored in accordance with the Central University Research Ethics Committee (CUREC) Management and Protection of Data Collected for Research Purposed Best Practice Guidance 09 Version 3.0. All data and consent forms will be kept for ten years from the date of data collection completion in accordance with Engineering and Physical Sciences Research Council (EPSRC) funding policy.

4.5.2 COVID-19 considerations

As the study was conducted during the COVID-19 pandemic, additional University, Oxford University Hospitals (OUH) and OGL COVID-19 risk assessments were conducted and the following measures were put in place in accordance:

- Researchers, staff and participants were required not to attend the OGL if they or a member of their household was diagnosed with, or had any symptoms of, COVID-19.

- All devices and equipment were sterilised before and after sessions.
- Participants entered the OGL directly through the rear door.
- Personal protective equipment including a disposable apron, masks and gloves were worn by the researcher. Participants were required to wear masks.
- Social distancing was maintained in the OGL whenever possible. Participants were instructed on how to put on the devices and markers - help was only provided if requested.
- Participants were asked not to attend if they belonged to any high-risk demographics. However, when the study was conducted, none of the participants were high-risk.

4.6 Discussion

A noteworthy development in the field of gait assessment is in the increased efforts by consortia, such as Mobilise-d, to create validated mobility and gait outcomes that will enable better cross-study comparability, as well as greater clinical applicability. As part of this effort, and alongside the acknowledgement of the importance of free-living gait, a battery of protocols has been concluded upon. Several of these protocols bare great resemblance to all those in this study, validating the greater consensus that such protocols are both relevant and necessary [208].

It would have been preferable to have a greater range in subject characteristics. The clearest example of this is in the dominance of right handedness (Figure 4.1e), however it would have also been preferable to have a much greater range of ages and physiologies. The initial proposed and approved number of study participants was twenty; unfortunately however, the study ran concurrent with multiple full national lockdowns in response to the COVID-19 pandemic. Therefore, due to the OGL being a functional clinical gait laboratory, only a limited number of sessions could be undertaken. The figure of twenty was initially arrived at, partly because it would theoretically enable a relatively broad range of ages, handedness, heights and physiologies. It was also deemed a manageable number given the time and

resources. Furthermore, comparable validation studies commonly have a similar number of participants including the Newcastle study that formed the driving motivation behind the research initially [14, 209, 210].

The most important parameter for the wearable camera is subject height, as the worn height of the camera is the only variable that is used in the computer vision system. Of all the parameters, this one has the greatest representative range of the population despite the small sample size [211]. The dominant right handedness will need to be considered when assessing the results from the smartwatch.

It is noted that the clothing limitations required by both the wearable camera and the Vicon markers are not necessarily representative of real everyday life. However, the actual effect of this on gait is negligible, and controlling for clothing is standard practice in gait analysis [212].

It was qualitatively observed throughout all of the sessions, that engaging with the subjects in conversation during the ADL protocols increased their natural movement about the study space and reduced their perception of the wearables. Simply requesting the imitation of an everyday activity, followed by silence, sometimes led to unnatural accentuated motions and acting. A similar effect was observed during ADL3 when the subjects were challenged with a problem.

Assessment of the Vicon system's accuracy and reliability during these protocols is performed in Chapter 5. This is important as although Vicon is considered highly accurate and thus the gold standard; these protocols are longer, more complex, more diverse, and involve more objects and furniture than standard motion capture procedures.

A number of activities were conducted to ensure consistency between protocols, subjects and sessions. The location of markers were measured or defined by physical or physiological locations. The Straight and Turns protocols were defined by markers on the floor which were measured out each time. The same instructions. were given each time. Despite this, some inter-subject variation is likely, for example due to the variability in how the devices sat upon the subject's clothes, or moved throughout the protocol. On the other hand however, some inter-subject variability was desired in order to gather as great of a variation in movements as possible - to test the edge

cases of the devices. Therefore, the ADL protocol space was set up slightly different each time, and conversations varied during them.

It is worth considering that the wearing of masks by participants may have affected the data collected. Such personal protective equipment has been show to affect gait and activity by reducing the field of view [213, 214]. This would have the greatest effect during the ADL activities where many items below eye line needed to be interacted with. It is also worth noting, and considering in respect to the the ecological validity of the protocols, the impact that the provision of instructions, and how they are given, has on both gait style and speed [51].

4.7 Newcastle data set

As explained in Section 2.5, the motivation behind automated extraction of gait features from a wearable camera derived in part from a study conducted at Newcastle University which collected free-living video data very similar to that gathered by the study outlined above [14]. The study aimed to validate a step counting and walking detection algorithm developed for a lumbar-worn IMU. However, the key motivation behind the study design was the acknowledged lack of free-living ground truth data. This data was procured and used as an additional step to help devise, test and validate the wearable camera method for extracting gait features in a real-world environment. Therefore the protocol, equipment and some results are explained below.

This study was conducted in 2016 with ten healthy adults aged 20 to 33 years; six men and four women. The age, height and weight distributions were 27.5 ± 4.7 years, 1.74 ± 0.07 metres, and 70.4 ± 8.8kg respectively. Each subject wore an AX3 (Axivity, York, UK) accelerometer positioned over the fifth lumbar vertebra, which logged motion at 100 Hz. They also wore a GoPro HERO (GoPro Inc., CA, USA) action camera, attached to the chest by a chest strap and pointing towards their feet, recording at 50Hz and 720p (progressive scanning).

The study protocol involved each participant wearing the devices for two separate one hour sessions. The participants were to continue their normal lives and the capture sessions were initiated at various times of day so that activities were undertaken beyond the work environment. An overview of the range of activities undertaken and their

contextual information is shown in Figure 4.6. Approximately 30% of the data includes some form of gait for a total of 350.6 minutes. How this is distributed amongst the environments and contexts is shown in Table 4.1. As demonstrated, the data set was very useful as representative of everyday life and has a wide range of activities.

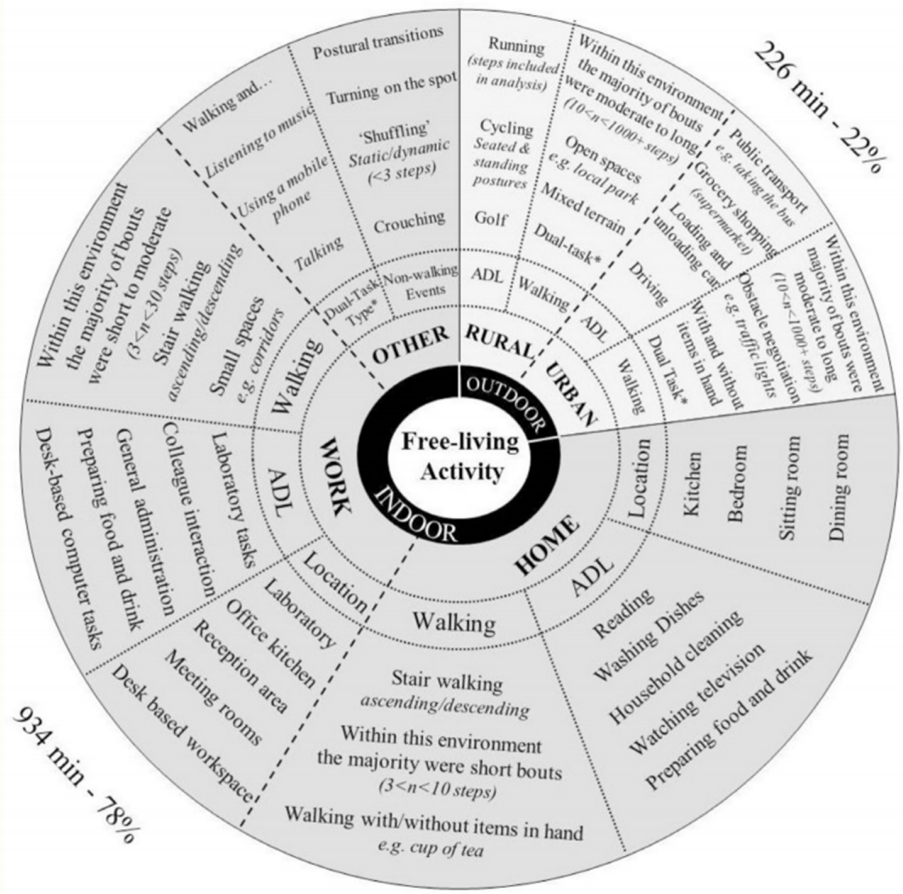


Figure 4.6: Newcastle data set activities and context. A wide range of locations, activities, scenarios and situations is represented by the data set. Figure taken from original publication [14].

4.8 Ad hoc data

In order to test and demonstrate the capabilities of the wearable camera system beyond the confines of the OGL several ad hoc data collection sessions were undertaken. These involved the researcher wearing the Garmin camera from the study while walking in various locales about Oxford. These data sets have no motion capture ground truth against which to compare them, however the camera provides a ground truth for step count and turn detection, and satellite imagery provided by Google Maps

Work		Home		Other	
Location	Mins	Location	Mins	Location	Mins
Office workspace*	51.8	Sitting-room	4.9	Urban outside	98.6
Office kitchen	10.6	Kitchen	3.7	Rural outside	71.9
Laboratory	6.8	Bedroom	2.7	Golf	48.4
Stair-walking	4.1	Dining-room	1.7	Running	21.6
Corridor & other	3.2	Corridor & other	1.1	Cycling	18.2
Reception	0.6	Stair-walking	0.7		
Total	77.1		14.8		258.7

Table 4.1: Distribution of walking locations in the Newcastle data set. The time spent walking in each type of environment during the approximately 20 hours of free-living data is shown; 350.6 minutes in total. The table is adapted from the original paper [14]. (*predominantly desk-based workspace in a cluttered office environment).

(Google, Mountainview, California, United States of America) can inform the accuracy of the path taken, and as such the movement measurements. These results are used to illustrate points and test the system during longer linear gait bouts that are not possible within the OGL, but are common in the validation of similar wearables [10, 209, 215]. Ad hoc data gathering session descriptions:

- AH01: Constant pace rapid walking around a single lap of the Roger Bannister 400m running track.

4.9 Conclusion

This chapter outlines the various sources of data for the following chapters. This includes the development and execution of standard and novel gait protocols within a motion capture lab. These protocols have been developed to better explore complex gait from the perspective of various analysis techniques.

Particular focus is on the inclusion of ADL protocols that are truly representative of free-living gait. Chapter 5 develops and explores robust definitions for extracting gait features from such movement; both conceptually, as well as practically using motion capture. Chapters 6 and 7 then cover the development of computer vision methods to extract such gait features from wearable camera videos. Finally, Chapter 8 compares the capacity for each of the wearable accelerometers to accurately detect steps throughout the protocols.

5

Developing robust definitions for spatio-temporal gait features

Contents

5.1 Introduction 71
5.2 Gait feature definitions & theory 72
5.3 Methods 99
5.4 Results 102
5.5 Discussion 104
5.6 Conclusion 110

5.1 Introduction

The following chapters of this thesis develop and compare methods of extracting gait features from various wearable devices in a range of scenarios. To validate these methods, they must be compared to a ground truth or gold standard. In gait research, motion capture is often considered to be this gold standard due to its capacity to track markers attached to an individual to within sub millimeter accuracy. However, it is often then taken for granted that the 3D trajectories of markers can be converted into accurate gait features.

Some literature outlines methods that can achieve this automatically (without an examiner having to manually label gait events in the signals), however, almost all studies universally apply to short periods of highly prescribed, linear walking. There is therefore no strict set of engineered rules that can be applied to the trajectory data to extract gait features in any scenario, especially for pathological, disturbed,

erratic, complex or free-living movement. The aim of this chapter is to develop and test such rules that could be applied universally.

When one expands beyond linear gait, the agreed definitions for macro gait events begin to be challenged. Therefore, in the following theory section, robust definitions are proposed, explained and justified. It is acknowledged that a truly robust set of rules would consider every possible kind of gait and type of person. Access to such a cohort of individuals would be difficult and beyond the scope of this work. Therefore, throughout the definitions, thought experiments are used to consider as many scenarios as possible - and account for the least typical form of gait. One of the difficulties in the consideration of free-living gait and complex movements is how to convey them on paper. To help with this, step maps are provided to demonstrate movements (Section 3.2.2).

In the following methods section, how these definitions are applied to Vicon motion capture data specifically are explained. The results from doing so within the study are given and then explored, especially in regard to the limitations of the techniques during activities of daily living (ADLs).

5.2 Gait feature definitions & theory

Gait, and walking, are so fundamental to the human experience that it seems that we all have an intrinsic understanding and opinion of what it means 'to walk', 'to take a step', and 'to turn' without necessarily consulting a strict definition. However, these interpretations are highly subjective and fall upon uncertainty when considering extreme cases of pathological gait, complex movements, assisted gait, and the increasing abstraction of locomotion in AI and robotics [216, 217].

This lack of consensus for concepts as fundamental as 'what is a step' is increasingly problematic when trying to compare ever more sensitive methods of gait measurement; especially in a free-living environment. Indeed, it is acknowledged by *Hickey et al.* in their validation of an algorithm to detect macro gait features in a free-living environment, that there exists no definitive classification for steps or walking bouts [14]. Therefore, the following section aims to remedy this by providing and justifying robust definitions of the following macro gait features:

- Direction of progression
- Turns
- Strides & stride length
- Steps, step length & step width
- Walking bouts

To ensure robustness, many complex or pathological gait movements were considered. Therefore, examples are provided to help elucidate why and when a definition may seem convoluted. In the context of individuals without two lower limbs, the definitions apply equally readily when feet are allocated to prostheses or assistive devices. Additionally, each definition has been developed within the context of precedent - which is provided. Finally, the definitions are predominantly provided as functions of velocity. The motivation for this is that measurement of gait is often either by direct measurement of location (motion capture or GPS), or by direct measure of accelerations (accelerometry and gyroscopy). Therefore, providing values as velocities reduces the compounded error associated with integrating or differentiating a signal multiple times.

For the sake of clarity, the definitions are provided as a series of logical steps for each feature, and then each step is explained in a separate accordingly numbered subsection.

5.2.1 Defining direction of progression

1. **The direction of progression at any moment is the projection of the mean normalised velocity vector of the centre of mass over a window period of 2 seconds onto the horizontal plane (that which is perpendicular to gravity).**

The direction of progression (DoP) is a common concept within the field of gait analysis, and is utilised in the determination of several features such as distance walked, step length and gait events [218, 219]. It is colloquially considered to be: 'the direction the individual is walking in'. Although this is clear for linear walking, it becomes less so for free-living gait. This limitation of definition is acknowledged by *Huxham et al*

in their paper *'Defining spatial parameters for non-linear walking'* [141]. Here they define the DoP as: *along the line from a point on one foot at the first foot contact to the same point on the same foot at the next foot contact* (Fig. 5.1). Similarly, *Tranjaniello, Cereatti and Croce* define it as the projection onto the horizontal plane of the mean velocity vector over a full gait cycle (as integrated from the accelerometry signal of an IMU attached to the shank) [220].

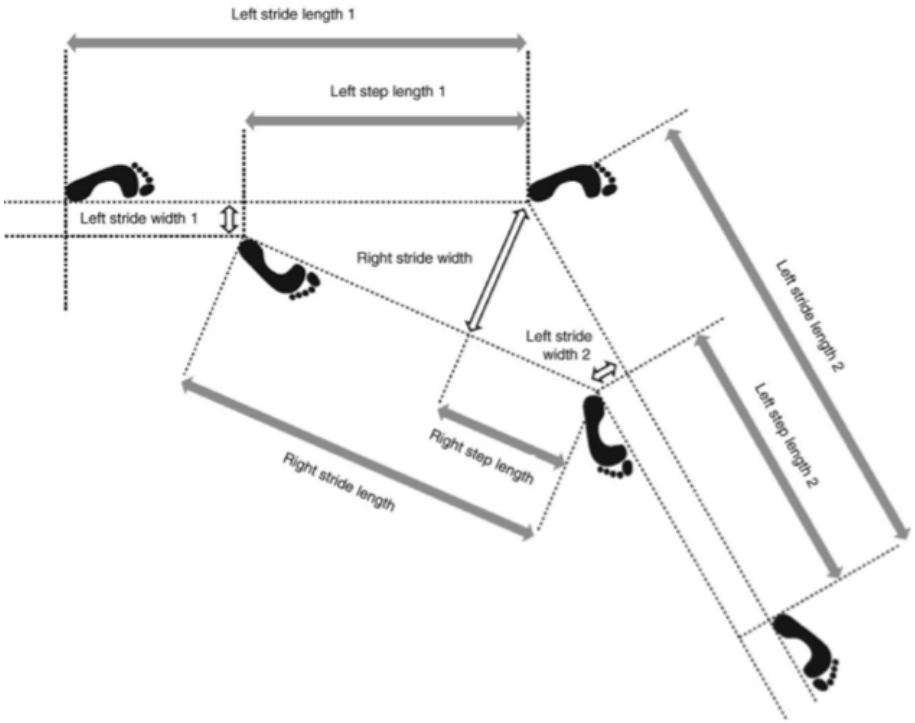


Figure 5.1: Spatial gait features during non-linear walking. Method for extracting features from non-linear gait as suggested by *Huzham et al.* Figure taken from original paper [141].

Although these definitions offer a more applicable definition for discrete turns within linear gait, they do not account for all free-living and pathological gait eventualities. Therefore the definition that is proposed here is that the DoP at any instance is the projection of mean velocity vector over a window of $\pm 1s$. An example of where this may provide improvement is shown in Figure 5.2.

The choice of a two second window is derivative of the fact that this is considered to be the threshold period for defining whether walking has stopped (see Section 5.2.5). Furthermore, it is preferable for the DoP to account for multiple steps regardless of the walker. Cadence values provided in Section 5.2.3.4 shows that even in highly

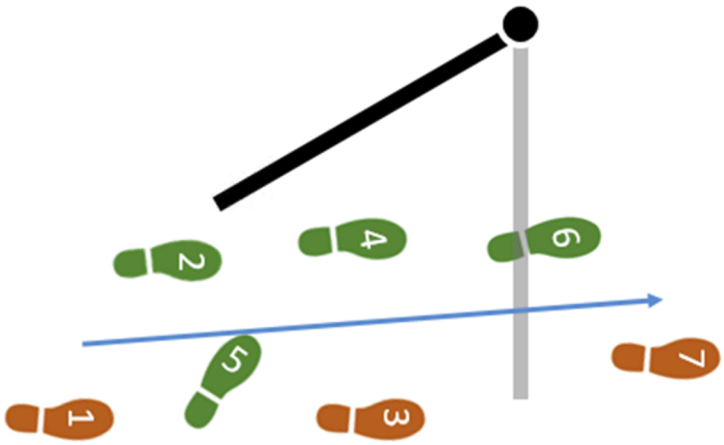


Figure 5.2: This foot map shows that while opening a door to pass through it, the left foot steps back. Using alternative definitions, this would result in a positive step length and an about turn in DoP. Using the definition proposed here (and assuming that pause at the door is less than two seconds), the DoP continues forward and the step length is therefore negative.

disturbed gait, this occurs within 2 seconds. The importance of this window length is demonstrated in Figure 5.3; whereby a shorter window length results in substantial continual changing of the DoP over linear walking.

The decision to use the projection of the mean velocity vector on the plane perpendicular to gravity (as opposed to the ground plane which may be incline or decline) is both because the human body must always adhere to gravity, and to aid universal application (for example, IMUs can detect the gravity vector).

5.2.2 Defining turning

1. **A turn is the rotation of the subject by a minimum of 30 degrees.**

2. **A turn begins when the speed of rotation exceeds 15 degrees/sec and ends when it falls below this threshold for more than 0.5 seconds, or if the direction of turn changes.**

3. **Turns must occur during a walking bout.**

A turn is generally understood as a change in the direction of walking and is vital in the navigation of free-living environments, and implicated as important in the consideration of life-impacting falls in the elderly and disease diagnosis (Section

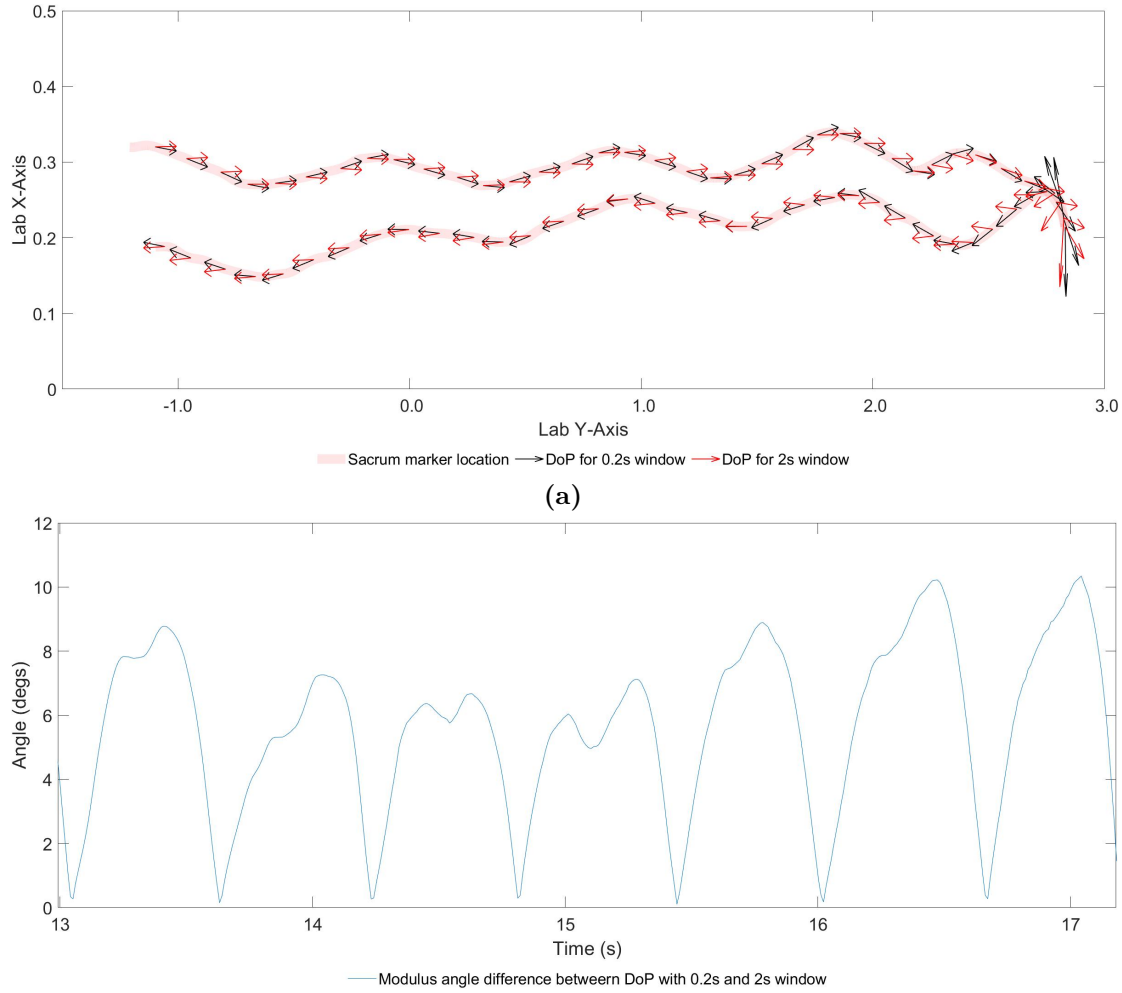


Figure 5.3: Direction of progression. The importance of temporal considerations in determining DoP. (a) Shows a short sample of the path of the Straight Walking protocol. Oscillation in the XY-plane is observed due to the medio-lateral oscillations of the pelvis and centre of mass (CoM) during walking. The DoP at various points is demonstrated by vectors for two different window lengths. (b) Shows the absolute difference in DoP angle over the period of one of the Straight protocol lengths.

2.3.2). There are therefore numerous studies on turns without a unified definition of what they are constituted by [56].

5.2.2.1 Minimum turn angle

Although there is no clear consensus on the minimum angle subtended to constitute a turn, many turning studies utilise an algorithm developed by *El-Gohary et al* [56]. Here they used a lumbar worn IMU to discern turns defined as 'trunk rotation about the transverse plane with a minimum of 45° and accompanied with at least one right and one left foot stepping' [221]. Further qualification involves the limitation of turns to those

between 0.5 and 10 seconds, and with vertical rotational rate exceeding a threshold of $15°s^{-1}$. Finally, turns separated by pauses of less than 50 milliseconds are combined.

The use of 45° is not universal however, others have used similar definitions but with thresholds of 40° [222], 50° [57], and 90° [190, 223]. The case for using a lower threshold of 30° is due to the distribution of the detected turns in free-living studies clearly implying that the majority of turns exist below 45 [57]. This is supported by the results of the ADL protocols within the Oxford study (Section 6.4.2). Additionally, there is a little precedent for researchers to attempt to capture 30° turns in free-living environments [224].

5.2.2.2 Minimum turn speed

This rule has plenty of precedent, and is in accordance with the majority of previous work and definitions mentioned in Section 5.2.2.1. The greatest concern however, would be that similarly to turn angle, the threshold of $15°s^{-1}$ would not capture the majority of turns. However, it was shown that even the slowest turners in a cohort of 7 recurrent fallers (who have been shown to turn more slowly than normal) turned at a mean speed of $> 55°s^{-1}$ [56]. Similarly, in a study of 133 subjects with idiopathic PD, even those 60 who exhibited FOG (mean UPDRS 35.2 ± 14.1), had mean turn speeds over a full 180° turn that exceeded $34°s^{-1}$ [225].

It is noteworthy that no limit has been placed on the length of time a turn can take. This is because there is no theoretical maximum turn angle or time. One could walk in a circle indefinitely. In fact, even within the confines of the Oxford study protocols, several long turns greater than even 360° were detected (Section 6.4.2). These occurred predominantly in the Turns and ADL3 protocols. In the former it is caused by an individual, through choice of path or style of turning, combine multiple turns into one continuous turn according to this definition. In the latter, a person may spin or circle looking for items in the experimental space. The distribution of turn angles can be found in Figure 6.6.

The cessation of a turn when the direction changes is a novel addition that is important in the consideration of free-living gait. Turns are not intrinsically

discrete events, and this rule ensures that subsequent turns in separate directions are independently captured and do not cancel each other out.

5.2.2.3 Turns only occur while walking

Turns that occur while standing stationary, or at least, not during a walking bout are excluded from consideration. Although much turning does occur during this period, its relevance to gait is limited. Furthermore, its clinical application to quality of life and prognosis (for example its relevance to chances of falling), are less distinct than actually during ambulation. This decision has precedent and is common in the field [57]. Any turning that occurs outside of a walking bout can be considered as stance rotation or adjustment.

5.2.3 Defining the stride

1. **All spatio-temporal measurements of the foot are done from the centre of the foot in the x-y plane**

2. **A foot event is defined as a period of movement at a velocity consistently greater than** $0.1ms^{-1}$**, and as such does not drop below the threshold for more than** $0.1s$

3. **For a foot event to be considered a stride event it must exceed a lateral displacement of** $0.1m$

4. **For a stride event to be a stride it must be part of a walking bout.**

The most fundamental unit of ambulation is 'the stride'. However, despite being referred to in a great many studies, the stride lacks a formal and recognised definition. It seems that the majority of studies consider the unit of the stride to be self evident however this is anything but true when it comes to free-living or pathological gait.

5.2.3.1 Measuring foot location

Traditionally, a stride is considered the entire gait cycle between two consecutive foot strike events of the same foot. The step is then the period between the foot strikes of opposing feet. However, from the reference frame of the foot, the progression only occurs during the swing phase; between foot-off and foot-strike. This definition holds in stereotypical linear walking, however free-living and pathological gait require an alteration in definition.

Hypokinetic-rigid, or shuffling, gait is a common phenotype observed in both the elderly and a range of common diseases such as Parkinson's and Alzheimer's [226]. Similarly foot-dragging is a common element of spastic gait, itself a symptom of numerous indications including Pyramidal syndrome and Cerebral Palsy [227]. As such it is feasible for gait to occur with negligible foot clearance, obscured heel-off events, and at the least with the traditional gait cycle model disrupted.

In free-living non-pathological gait, foot movements may also occur that do not fit the model. For example if we consider solely the toes or the heel, then toe tapping, heel shaking and more importantly pivot turns could potentially be missed as the linear velocity at the point of rotation is negligible (Fig 5.4) [228]. Rotation about the foot's centre is the least likely to occur naturally and therefore the best to consider to provide the most robust single-location reference point.

5.2.3.2 Minimum velocity for a foot event

In the idealised model of the gait cycle, the foot is stationary with respect to the ground during the stance phase. Therefore the swing phase of the stride occurs between two stationary periods. However all systems of measurement incorporate some degree of noise. Furthermore, during the stance phase it is possible for there to be some shaking of the foot due to pathology or instability, as well as slippage of the foot against the walking surface. Therefore a non-zero minimum threshold for movement is required.

To demonstrate this, examples gathered from the motion capture lab can be considered. Figure 5.5 shows a proposed threshold of $0.1ms^{-1}$ in both straight walking and imitation ADLs. Given the arbitrary cut off of $0.1ms^{-1}$ it is clear that some very minor foot movements will be missed entirely, but also that if the stride time is

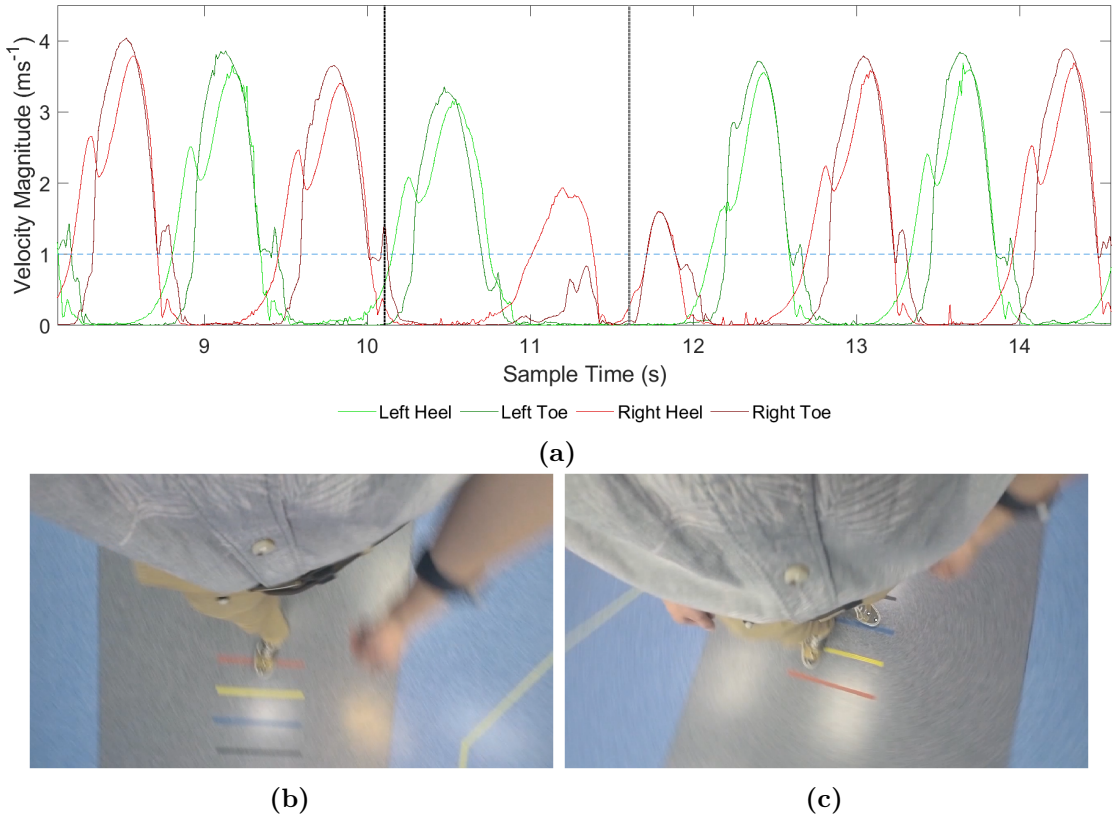
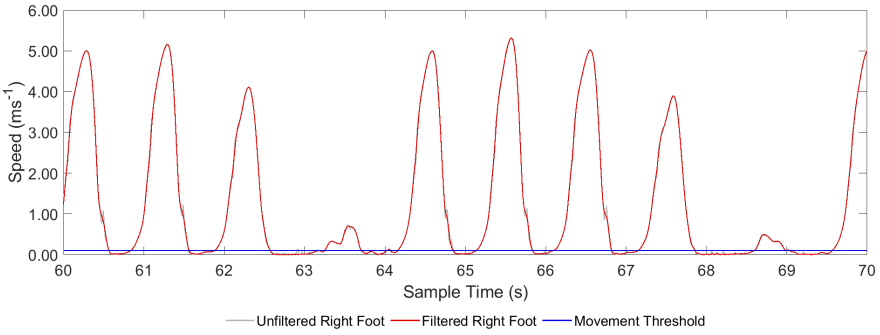
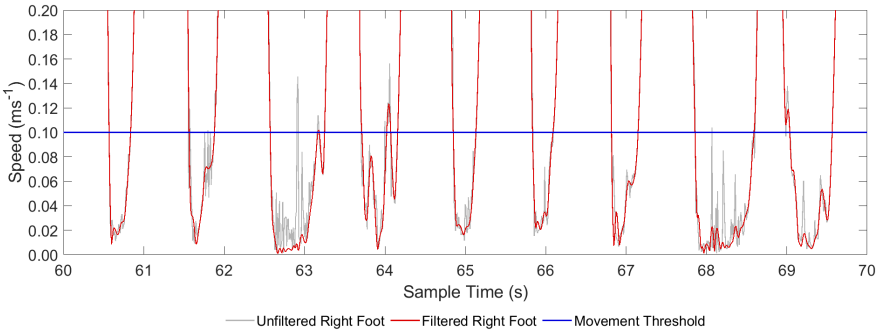


Figure 5.4: Heel and toe movement comparison during pivot turn. (a) Shows the velocity magnitude of the toes and heels of both feet during the last two and first two gait cycles of a subject walking up and down in a straight line at a slow pace. In between the linear walking is a pivot turn; the heel of the right foot is swung around 180°, also slightly moving the anterior foot. The difference in heel and toe velocity magnitudes can be clearly seen. (b) Shows the foot immediately before the pivot turn corresponding to the black dotted line in (a). (c) Shows the foot immediately after the pivot turn; corresponding to the grey dotted line in (a). Comparison of the coloured striped markings between (b) and (c) clarify the change in orientation and location of the right foot. It is therefore demonstrated that if a threshold of $1ms^{-1}$ is used as the definition of movement, then the algorithm would miss the pivot turn if the reference point was on the toe. Although in this example the toe does move, it will not always do so. Therefore the centre of the foot (average velocity magnitude of the toe and heel) is chosen as the single location reference point for the foot in further calculations.

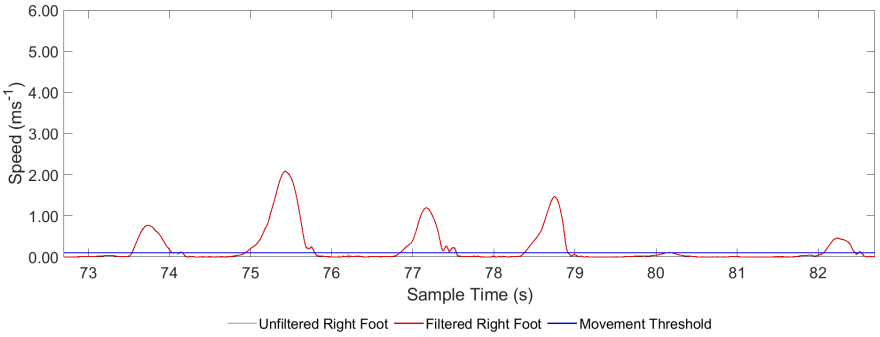
considered from the point of exceeding the threshold to the point of dropping below it; that some portion of the stride's distance and time will be lost. This is only a single example and does not represent the range of free-living movements, subjects and pathologies. Therefore it would be beneficial to have a model that can provide an approximation of the proportion of stride lost when compared to the threshold chosen.



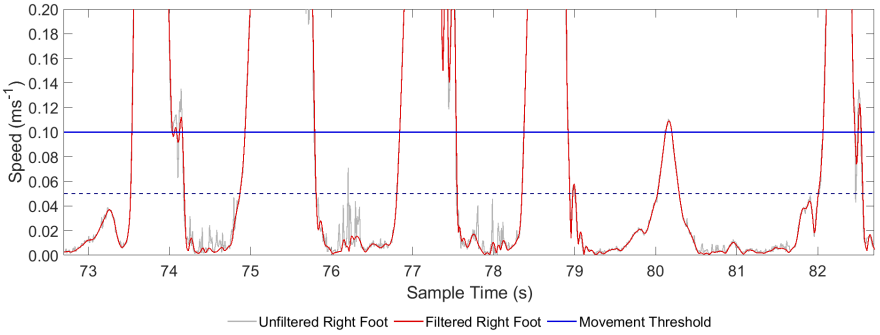
(a)



(b)



(c)



(d)

Figure 5.5: Foot speed threshold. (a) and (c) show the speed of the right foot over periods of approximately 10 seconds during both straight walking (with 180^o turns) and imitation activities of daily living. The arbitrary threshold of $0.1ms^{-1}$ is shown, and seems to readily discern individual strides. However, on closer inspection of the foot movement at lower speeds, (b) and (d) show that even a low threshold of $0.1ms^{-1}$ will miss some foot movements, severely foreshorten others, and pick up signal fluctuations.

5.2.3.3 Foot velocity model

To estimate the proportion of stride time and stride length that is detected when using a specific threshold foot velocity, $V_{Threshold}$, the foot velocity, v, must be expressed as a function of time, t. From qualitative observation, it seems that the Gaussian distribution would be a suitable estimation when scaled by a constant, A, limited between -2σ and 2σ, and with a mean μ equivalent to the peak in foot velocity (Figure 5.6a). Therefore foot velocity during a stride can be expressed as:

$$f(v) = A\mathcal{N}(\mu, \sigma^2) = \frac{A}{\sigma\sqrt{2\pi}} e^{\frac{-(t-\mu)^2}{2\sigma^2}} \text{for} - 2\sigma \leq t \leq 2\sigma. \tag{5.1}$$

Although this model will not be perfect, especially in pathological gait, it is always necessary for the foot to accelerate and decelerate through the swing phase and this model looks to be a sufficient approximation. Next, the aim is to express the proportion of detected stride time, $P_{StrideTime}$, and the proportion of detected stride length, $P_{StrideLength}$, as functions of the chosen minimum velocity threshold, $V_{Threshold}$. From comparison of the stride model function (Fig. 5.7a) and the Generalised Gaussian Distribution (Fig. 5.7b), or GGD, which expresses the probability density, Φ, as a function of the Z value, it can be stated that:

$$P_{StrideLength} = \frac{\int_{-Z_{Threshold}}^{+Z_{Threshold}} \mathcal{N}(\mu, \sigma^2) dz}{\int_{-Z_{Start}}^{+Z_{End}} \mathcal{N}(\mu, \sigma^2) dz}. \tag{5.2}$$

As the temporal limits of the stride are defined as $\pm 2\sigma$,

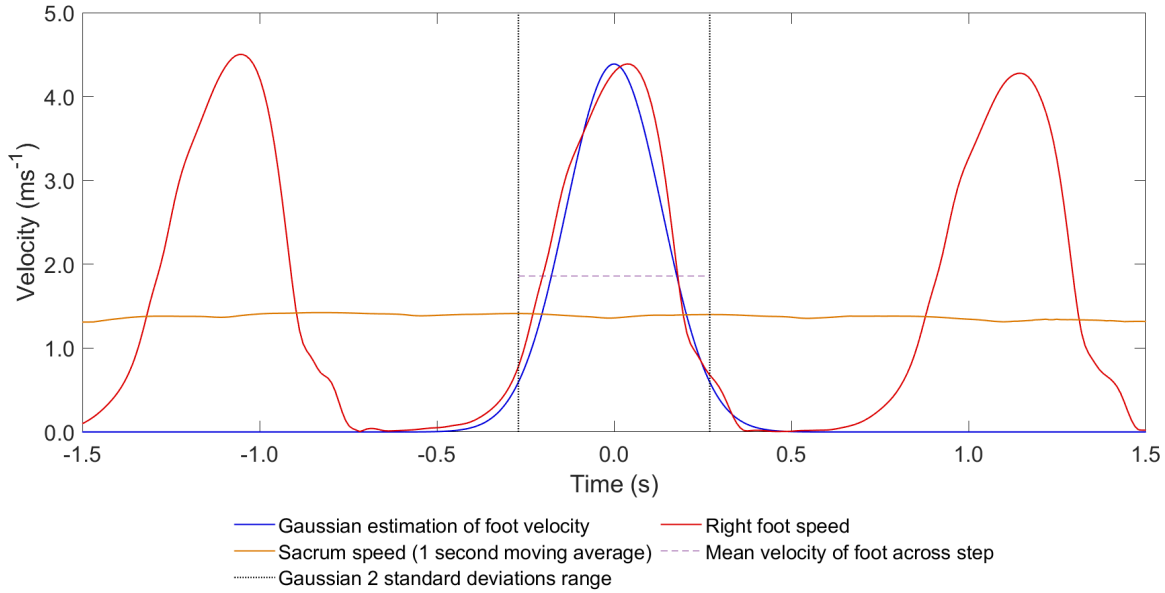
$$\int_{-Z_{Start}}^{+Z_{End}} \mathcal{N}(\mu, \sigma^2) = 0.95 \approx 1. \tag{5.3}$$

Therefore Equation 5.3 can be represented as,

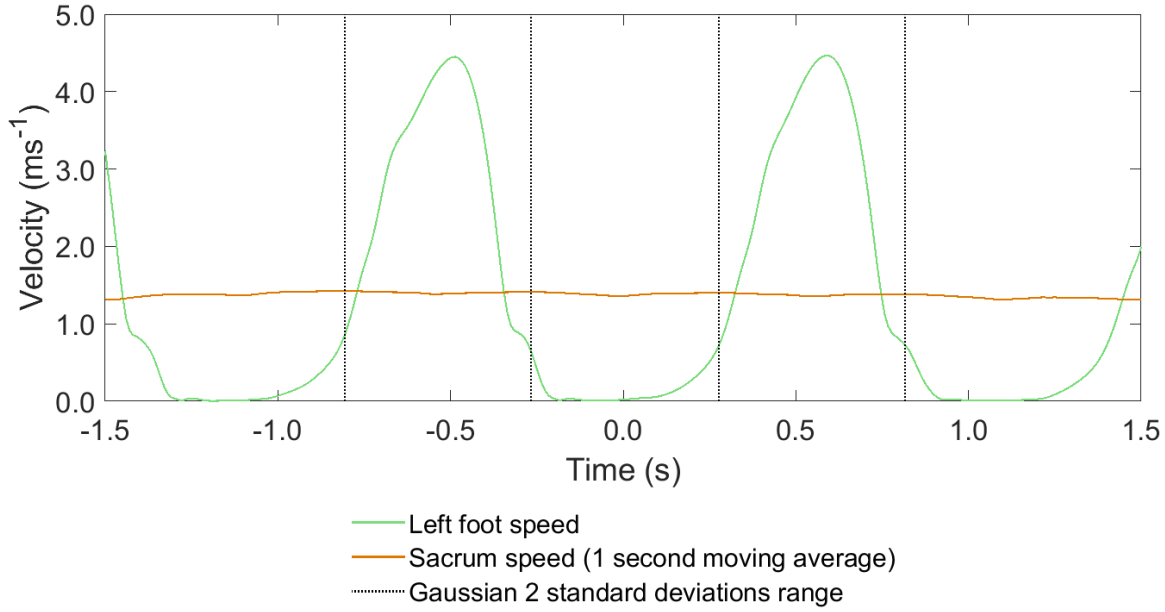
$$P_{StrideLength} \approx 1 - 2\int_{-\infty}^{-Z_{Threshold}} \mathcal{N}(\mu, \sigma^2) dz. \tag{5.4}$$

Similarly we can deduce,

$$P_{StrideTime} = \frac{2Z_{Threshold}}{Z_{End} - Z_{Start}} = \frac{Z_{Threshold}}{Z_{End}}. \tag{5.5}$$



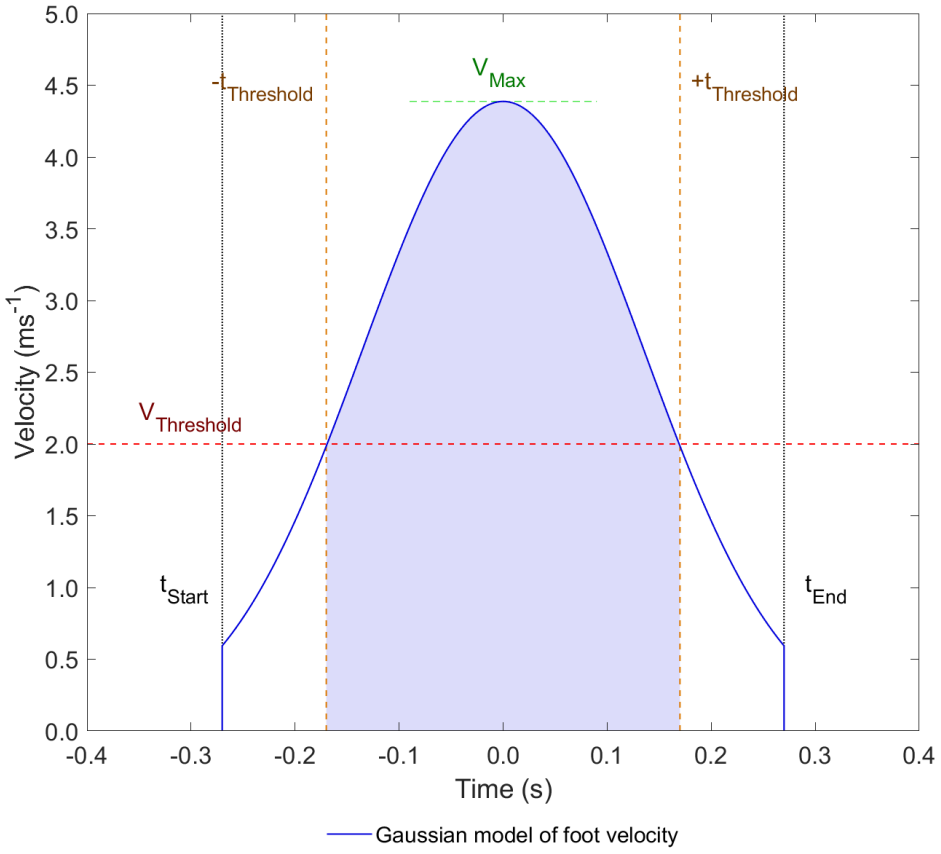
(a)



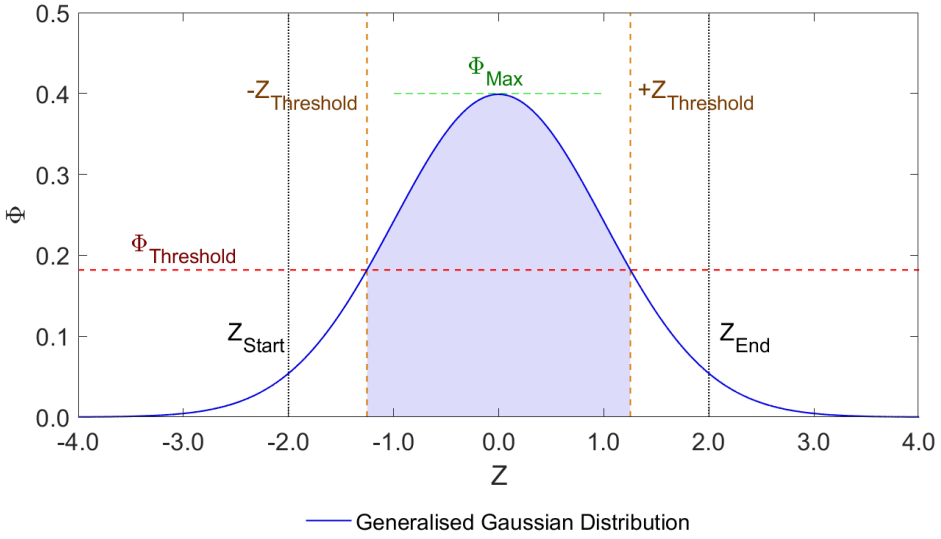
(b)

Figure 5.6: Gaussian foot velocity model example. (a) Shows the comparison of a scaled Gaussian distribution when compared to the speed of the right foot during walking. By comparison to the velocity of the left foot during the same period in (b) it can be seen that limiting the distribution to between -2σ and 2σ ensures that the majority ($\sim 95\%$) of the stride is detected while left and right strides are considered basically temporally distinct from each other.

Knowing that in the GGD, $\sigma = 1$, and $\mu = 0$ we can utilise Equation 5.1 to establish that:



(a)



(b)

Figure 5.7: Gaussian foot velocity model. (a) shows the foot speed function, and in shaded blue, the proportion of stride that is detected by a certain threshold velocity. This can then be compared to the GGD in (b); where the maximum foot speed, V_{Max}, corresponds to the maximum probability density Φ_{Max}, the true start and finish times of the stride (t_{Start} and t_{End} respectively), correspond to the Z_{Start} and Z_{End}, and the detected start and finish times of the stride ($-t_{Threshold}$ and $+t_{Threshold}$ respectively) correspond to $-Z_{Threshold}$ and $+Z_{Threshold}$.

$$Z_{Threshold} = \sqrt{\ln\left(\frac{1}{2\pi.\Phi_{Threshold}^2}\right)}. \tag{5.6}$$

By comparing Figures 5.7a and 5.7b it is also clear that,

$$\frac{\Phi_{Threshold}}{\Phi_{Max}} = \frac{V_{Threshold}}{V_{Max}}. \tag{5.7}$$

As $Z = 0$ at Φ_{Max}, Equation 5.1 can be substitute into Equation 5.7 to give,

$$\Phi_{Threshold} = \sqrt{2\pi}\frac{V_{Threshold}}{V_{Max}}. \tag{5.8}$$

We therefore need an estimation for V_{Max}. However, there is limited literature on the peak velocity of feet during strides, let alone sufficient literature to establish a suitable range to ensure that a robust stride definition that includes all possible cases. Although still limited, there is much more information on walking velocities across demographics and indications. Therefore, an estimation of V_{Max} based on walking velocity, V_{Trunk}, would be preferable. An additional consideration is that, especially in pathological gait, considerable gait asymmetry may occur. In this situation one foot may travel at a velocity, V_{Slow}, substantially slower than the other foot at V_{Fast}. The relationship between the feet is known as the asymmetry ratio, R_{Asym}.

$$R_{Asym} = \frac{V_{Slow}}{V_{Fast}} = \frac{d_{Slow}/t_{Slow}}{d_{Fast}/t_{Fast}} \tag{5.9}$$

Where d and t are distances and times of each stride respectively. Since, for net translation of the individual to occur, both feet must travel approximately just as far as the trunk,

$$d_{Trunk} \approx d_{Slow} \approx d_{Fast}. \tag{5.10}$$

Additionally, we can assume that when one foot is moving the other is not (see Fig. 5.6). Although this might not be exactly true, any extended periods of double support would result in a required increase in V_{Slow} in order to achieve the same V_{Trunk}, and as we are concerned with finding a threshold for minimum velocity, the assumption is reasonable for purpose. Therefore,

$$t_{Trunk} = t_{Slow} + t_{Fast}. \tag{5.11}$$

By combining Equations 5.9, 5.11 and 5.10 we can express V_{Trunk} as,

$$V_{Trunk} = \frac{d_{Trunk}}{t_{Trunk}} = \frac{d_{Slow}}{t_{Slow} + t_{Fast}} = \frac{d_{Slow}/t_{Slow}}{1 + t_{Fast}/t_{Slow}} = V_{Slow}\left(\frac{R_{Asym}}{R_{Asym}+1}\right), \tag{5.12}$$

$$V_{Slow} = V_{Trunk}(R_{Asym} + 1). \tag{5.13}$$

As we are looking to estimate the slowest possible foot at a certain walking speed the mean velocity in the model of Equation 5.1, $\overline{v}$, is the same as V_{Slow} and can be expressed as:

$$V_{Slow} = \overline{v} = \frac{d_{Slow}}{t_{Slow}} = \frac{\int_{-2\sigma}^{+2\sigma} f(v)dv}{4\sigma} \approx \frac{A}{4\sigma}. \tag{5.14}$$

Similarly, as $f(v) = V_{Max}$ at $v = 0$,

$$V_{Max} = \frac{A}{\sigma\sqrt{2\pi}} \tag{5.15}$$

Combining equations 5.13, 5.14 and 5.15 obtains:

$$V_{Max} = \frac{4}{\sqrt{2\pi}} V_{Trunk}(R_{Asym} + 1). \tag{5.16}$$

Therefore, we can combine the equations 5.4, 5.5, 5.6, 5.8, and 5.16 to get the detected proportion of stride length and time as:

$$P_{StrideLength} \approx 1 - 2\int_{-\infty}^{-\sqrt{\ln\left(\frac{V_{Trunk}^2(R_{Asym}+1)^2}{2\pi^3 V_{Threshold}^2}\right)}} \mathcal{N}(\mu, \sigma^2)dz. \tag{5.17}$$

$$P_{StrideTime} = \frac{1}{2}\sqrt{\ln\left(\frac{V_{Trunk}^2(R_{Asym}+1)^2}{2\pi^3 V_{Threshold}^2}\right)}. \tag{5.18}$$

With these relationships established, we can model the impact of a chosen $V_{Threshold}$ on stride detection and measurement, to ensure that the definition of a stride is robust and will work for the slowest possible gait. Table 5.1 gives examples of walking speeds

in the literature from a range of demographics, gait pathologies and comorbidities including Alzheimer's, stroke, Multiple Sclerosis and spinal cord injury. A representative range of values from this table are used to demonstrate the modelled impact of $V_{Threshold}$ on stride detection across the range of possible gait speeds (Fig. 5.8a). The minimum feasible gait speed is therefore taken to be $0.15ms^{-1}$; justification for this is discussed in more depth in Section 5.2.5.3.

With minimum gait speed identified, gait asymmetry must also be considered. Table 5.2 demonstrates the range of measured asymmetries in a range of pathologies and demographics. Using examples from this table, the sensitivity of the proportion of stride detected at different gait asymmetries can also demonstrated for the slowest considered speed (Fig. 5.8b). It can be observed from this that in the extreme scenario of an individual having a mean gait speed of $0.15ms^{-1}$ and a gait asymmetry ratio of 0.25, with a threshold velocity of $0.1ms^{-1}$, the expectation is detection of $\approx$ 75% of stride time and $\approx$ 85% of stride length. Once substantially more likely gait speeds are considered this proportion approaches approaches 100%. Therefore, this is the threshold that shall be used.

Due to signal noise it is possible for the foot velocity to drop below the threshold prematurely during the stride (as can be seen in Figure 5.5b). Therefore, to ensure that the the velocity has indeed dropped below the threshold, it must remain so for $0.1s$. This value is chosen as even at the very highest cadences, it will still be true. Furthermore, it is a convenient figure for an integer number of samples for common sampling rates such as 30Hz and 100Hz, while still be detectable as low as 10Hz.

5.2.3.4 Minimum displacement for a stride event

As the velocity thresholding is at $0.1ms^{-1}$, the stride detection is sensitive to sudden jerks of the foot, vertical movement, foot adjustment and signal noise, a minimum stride distance ensures that these are filtered out. There is limited literature on the shortest reasonable stride length, however, it can be deduced that the shortest stride length, L_{Min}, will be twice the result of the slowest walking speed, V_{Min} (as both feet must cover the same distance) at the fastest cadence C_{Max};

Table 5.1: Walking speeds from literature. Speeds and ages are provided alongside their standard deviations where relevant and stated, and the sample size is shown alongside the type of measurement. *1minimum speed threshold for inclusion. *2mini-mental state examination. *3functional electrical stimulation.

Paper	Speed	Type	Age	Population
Bohannon 1997 [229]	1.27 (0.21)	Mean (20)	73.1 (3.1)	Comfortable walking speed of healthy women in their 70s
Bohannon 1997 [229]	1.46 (0.16)	Mean (22)	44.9 (2.6)	Comfortable walking speed of healthy men in their 40s
Bohannon 1997 [229]	2.53 (0.29)	Mean (15)	23.9 (3.2)	Maximum walking speed of healthy men in their 20s
Bohannon 1997 [229]	1.75 (0.28)	Mean (20)	73.1 (3.1)	Maximum walking speed of healthy women in their 70s
Rantanen & Avela 1997 [230]	1.22 (0.29)	Mean (24)	85	Maximum walking speed of women aged 85 with a range of chronic conditions and walking assistance
Purser et al 2005 [231]	0.53 (0.24)	Mean (1388)	74.2 (5.7)	Walking speed of elderly male veterans on discharged from hospital with a range of indications including stroke, dementia, and ambulatory difficulty
Hardy et al 2007 [232]	0.20	Min*1 (439)	73.9 (5.6)	Walking speed of individuals aged 65 and over including assistive devices, comorbidities, minimum MMSE*2 score of 16
Hardy et al 2007 [232]	0.88 (0.24)	Mean (439)	73.9 (5.6)	Walking speed of individuals aged 65 and over including assistive devices, comorbidities, minimum MMSE*2 score of 16
Langlois et al 1997 [233]	0.38 (0.18)	Mean (1231)	79.2	Walking speed of pedestrians (minimum age 72, maximum age 105) who self-report to have difficulty with crossing the street, with a range of comorbities including fracture, heart disease, cancer, diabetes and stroke.
Langlois et al 1997 [233]	0.59 (0.22)	Mean (1231)	79.2	Walking speed of pedestrians (minimum age 72, maximum age 105) who self-report to not have difficulty with crossing the street, with a range of comorbities including fracture, heart disease, cancer, diabetes and stroke.
Callisaya et al 2017 [234]	0.70 (0.21)	Mean (26)	82.9 (5.9)	Preferred walking speed of female Alzheimer's patients with moderate dementia, elderly with comorbities and medications
Callisaya et al 2017 [234]	0.98 (0.22)	Mean (94)	73.1 (5.1)	Preferred walking speed of female Alzheimer's patients with mild cognitive impairment, elderly with comorbities and medications
Miller et al 2015 [235]	0.97 (0.15)	Mean (9)	46.6 (4.5)	Self-selected walking speed of Multiple Sclerosis patients who are faster walkers ($\geq 0.8ms^{-1}$) without FES*3, EDSS (4.0 ± 1.5)
Miller et al 2015 [235]	0.46 (0.18)	Mean (11)	54.2 (7.1)	Self-selected walking speed of Multiple Sclerosis patients who are slower walkers ($< 0.8ms^{-1}$) without FES*3, EDSS (6.0 ± 0.0)
Miller et al 2015 [235]	0.53 (0.17)	Mean (11)	54.2 (7.1)	Self-selected walking speed of Multiple Sclerosis patients who are slower walkers ($< 0.8ms^{-1}$) with FES*3
Yogev et al 2006 [64]	1.00 (0.25)	Mean (21)	71.9 (7.3)	Normal pace walking speed of PD patients ranking 2-3 on the Hohn and Yahr scale
Yogev et al 2006 [64]	0.97 (0.25)	Mean (15)	76.3 (4.9)	Normal pace walking speed of fallers
Yang et al 2007 [236]	0.58	Minimum (13)	59.5 (11.8)	Walking speed of stroke patients with least limited community ambulation, (1.1 - 9.5 years post stroke)
daCuhna-Filho et al 2003 [237]	0.23	Minimum (10)	38.1 (12.9)	Walking speed of patients with spinal cord injury (50:50 split between cervical and thoracic)
daCuhna-Filho et al 2003 [237]	0.55 (0.31)	Mean (10)	38.1 (12.9)	Walking speed of patients with spinal cord injury (50:50 split between cervical and thoracic)
Soaz & Diepold 2015 [238]	0.12	Minimum (4)	87.0 (2.0)	Walking speed of severely frail seniors
Soaz & Diepold 2015 [238]	0.34 (0.15)	Mean (4)	87.0 (2.0)	Walking speed of severely frail seniors

$$L_{Min} = 2V_{Min} . \frac{60}{C_{Max}} \tag{5.19}$$

It is established in Sections 5.2.5.3 and 5.2.5.4 that the minimum walking speed is $0.15ms^{-1}$. Table 5.3 shows a distribution of cadences, and although it is clear that cadence can easily exceed 100, it is also widely acknowledged that substantial reduction

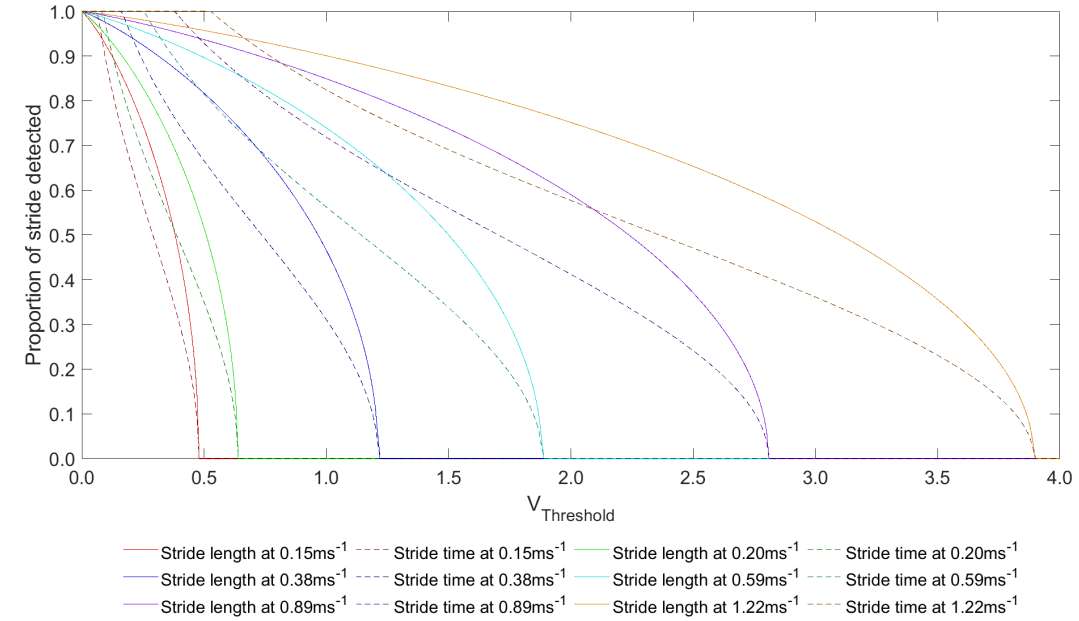
Table 5.2: Gait asymmetries from literature. Gait asymmetry can be reported in a number of ways. For the sake of comparability, the figures below have been adjusted to simply show the ratio of the slowest foot to the fastest foot. Sample size is shown alongside statistic type.

Paper	Ratio	Type	Age	Population
Yogev et al 2006 [64]	0.987 (0.002)	Mean (11)	67.5 (3.5)	Healthy controls
Yogev et al 2006 [64]	0.963 (0.008)	Mean (15)	76.3 (4.9)	Fallers
Yogev et al 2006 [64]	0.959 (0.008)	Mean (21)	71.9 (7.3)	PD patients scoring 2-3 on the Hahn Yahr scale
Kramers-de Quervain et al 2004 [239]	0.950	Minimum (10)	16.8 (7.4)	Female idiopathic scoliosis patients
Plotnik et al 2020 [240]	0.960 (0.011)	Mean (60)	43.2 (10.4)	Multiple Sclerosis patients with mild disability, EDSS 0.0 - 2.5
Plotnik et al 2020 [240]	0.931 (0.015)	Mean (26)	53.0 (8.1)	Multiple sclerosis patients with moderate disability, EDSS 3.0-4.0
Plotnik et al 2020 [240]	0.798 (0.027)	Mean (6)	54.1 (11.2)	Multiple Sclerosis patients with severe disability, EDSS 5.0-6.5
Patterson et al 2010 [241]	0.880 (0.162)	Mean (161)	62.7 (13.6)	Community-dwelling stroke patients
Patterson et al 2010 [241]	0.430	Minimum (161)	62.7 (13.6)	Community-dwelling stroke patients
Balasubramanian 2007 [242]	0.444	Minimum (49)	62.7 (10.2)	Chronic hemiparesis patients

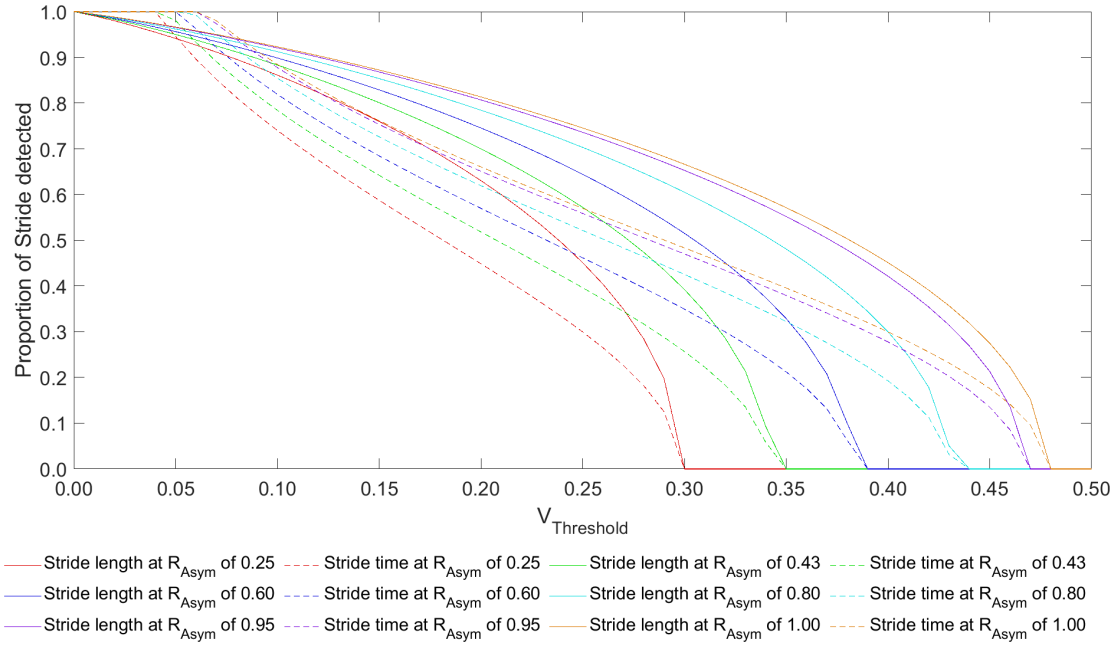
is observed in relation to increased age and disease severity [243]. Therefore cadence of any population walking at $0.10ms^{-1}$ will be substantially lower than 100, and it can be deemed a reasonable maximum cadence. With these values, Equation 5.19 gives us a minimum stride length of $0.18m$. For the sake of clarity and to ensure the inclusion of the very extreme ends of the gait spectrum, a minimum stride length of $0.1m$ is chosen.

5.2.3.5 Strides must be part of a walking bout

It can be argued that the movement of a single foot does not represent gait. In fact as will be evidenced later in Table 5.4, walking bouts are often defined by the numbers of steps of which they are composed. Therefore, there are a range of foot movements that although they might be stride-like, must be ignored. Therefore, for a stride-event to qualify as a stride, it must be part of a walking bout.



(a)



(b)

Figure 5.8: Sensitivity of proportion of Stride detected. (a) Shows the proportion of the modelled stride that is detected both in length and time for walking speeds from the lowest recognised to an average pace. (b) Demonstrates the sensitivity of the proportion of stride length and time detected on walking asymmetry.

5.2.4 Defining stride length, step length and step width

With strides detected and defined as the duration of foot movement between stationary periods, subsequent gait features can be defined. The following definitions are a generalisation of the methods proposed by *Huxham et al* [141] based on only the

Table 5.3: Gait cadences from literature. Sample size is shown alongside statistic type. A thorough overview of typical cadences in both active tests and free-living environments can be found in the review conducted by *Tudor-Locke et al* [243].

Paper	Cadence	Type	Age	Population
Monticone et al 2014 [244]	88.29 (16.71)	Mean (25)	49.2 (11.5)	MS patients conducting motor-motor dual task walking
Monticone et al 2014 [244]	76.20 (14.23)	Mean (25)	49.2 (11.5)	MS patients conducting motor-cognitive dual task walking
Elble et al 1992 [245]	106 (10)	Mean (19)	76 (6)	Older adults with treated comorbidities including hypertension, COPD, and mild diabetes mellitus
Elble et al 1992 [245]	80 (18)	Mean (10)	79 (4)	Patient with a range of indications associated with 'senile gait', including normal pressure hydrocephalus, Alzheimer and vascular dementia and peripheral neuropathy
Gardner el al 2010 [246]	90.8 (14.0)	Mean (98)	66 (12)	Peak 1 min cadence over 1 week of subjects with intermittent claudication
Gardner el al 2010 [246]	52.4 (17.0)	Mean (98)	66 (12)	Peak 30 min cadence over 1 week of subjects with intermittent claudication
Mudge et al 2010 [247]	111.8 (9.3)	Mean (30)	27.7 (8.9)	Peak 30 min cadence over two 3-day periods in healthy adults
Mudge et al 2008 [248]	58 (22)	Mean (40)	69.2 (12.6)	Peak 30 min cadence over two 3-day periods in participants more than six months post-stroke
Judge et al 1996 [249]	116 (7)	Mean (26)	79 (6)	Mean cadence over 5 metres in elderly participants from care homes.

knowledge of the locations of the stationary feet; and that will apply to all free-living scenarios. A diagram of the definitions is provided in Figure 5.9.

The stride vector, $\boldsymbol{S}_i^{L|R}$ is that which connects the last stationary location of a point of a foot of one side $\boldsymbol{X}_{i-1}^{L|R}$, to subsequent stationary location of the same point on the same foot $\boldsymbol{X}_i^{L|R}$ (Fig. 5.9). Therefore, stride length can be defined as,

$$StrideLength_i^{L|R} = |\boldsymbol{S}_i^{L|R}| = |\boldsymbol{X}_i^{L|R} - \boldsymbol{X}_{i-1}^{L|R}| \tag{5.20}$$

In order to calculate the step length and width, the location of the other foot must be known. Whereby $\boldsymbol{X}_{\boldsymbol{j-1}}^{\boldsymbol{L|R}}$ and $\boldsymbol{X}_{\boldsymbol{j}}^{\boldsymbol{L|R}}$ reflect the current (or most recent) stationary location, and the next stationary location, of the opposing foot.

The stride vector $\boldsymbol{S}_j^R$ can be expressed in the form,

$$\boldsymbol{S}_j^R = \boldsymbol{X}_{j-1}^R + \lambda \boldsymbol{n}, \tag{5.21}$$

where n is the normalised unit vector in the direction of $\boldsymbol{S}_j^R$, defined by,

$$\boldsymbol{n} = \frac{\boldsymbol{X}_j^R - \boldsymbol{X}_{j-1}^R}{|\boldsymbol{X}_j^R - \boldsymbol{X}_{j-1}^R|} \tag{5.22}$$

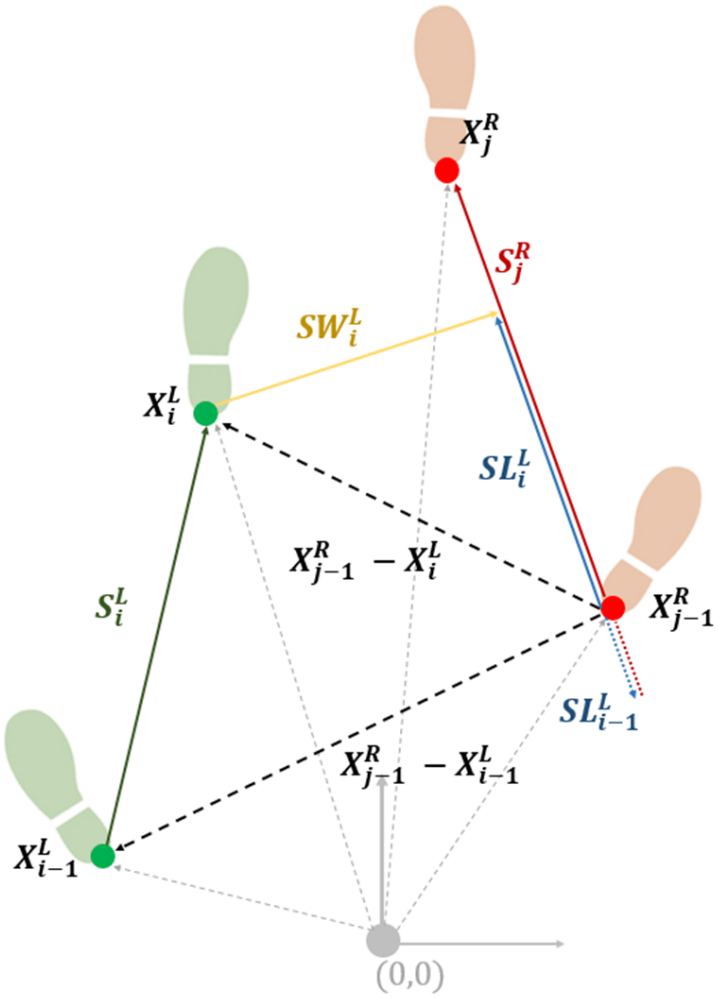


Figure 5.9: Stride length, step length, and step width definitions. Diagram explaining the novel method for defining step and stride features and how they can be extracted from a step map of stationary foot locations during stance phase. All vectors are (solid and dashed arrows) are in the world coordinate system (grey axes). Step width (yellow) and stride lengths for the left (green) and right (red) feet are shown as spatial vectors. The step length of the left foot (blue) is derived as a projection of one stride vector on the other. Dashed arrows are not gait features, but vectors needed in the calculation of the features.

The step length of stride $\boldsymbol{S_i^L}$ is thus defined as the length of the projection of the vector connecting $\boldsymbol{X}_i^L$ and $\boldsymbol{X}_{j-1}^R$ on $\boldsymbol{S}_j^R$. However, it is possible that the starting point of this stride is ahead of $\boldsymbol{X}_{j-1}^R$, therefore the projection of the vector connecting $\boldsymbol{X}_{i-1}^L$ and $\boldsymbol{X}_{j-1}^R$ on $\boldsymbol{S}_j^R$ must be subtracted from the step length when $\boldsymbol{X}_{i-1}^L$ is on the same side of $\boldsymbol{X}_{j-1}^R$ as $\boldsymbol{X}_i^L$.

$$StepLength_i^L = \begin{cases} |\boldsymbol{SL_i^L}| & if \boldsymbol{SL_i^L}.\boldsymbol{SL_{i-1}^L} <= 0 \\ |\boldsymbol{SL_i^L}| - |\boldsymbol{SL_{i-1}^L}|, & \text{otherwise} \end{cases} \tag{5.23}$$

where,

$$|\boldsymbol{SL}_i^L| = |((\boldsymbol{X}_{j-1}^R - \boldsymbol{X}_i^L) \cdot \boldsymbol{n})\boldsymbol{n}| \tag{5.24}$$

$$|\boldsymbol{SL}_{i-1}^L| = |((\boldsymbol{X}_{j-1}^R - \boldsymbol{X}_{i-1}^L) \cdot \boldsymbol{n})\boldsymbol{n}| \tag{5.25}$$

Similarly, step width is the length of the vector perpendicularly connecting $\boldsymbol{X}_i^L$ with the stride vector $\boldsymbol{S}_j^R$,

$$StepWidth_i^L = |\boldsymbol{SW}_i^L| = |(\boldsymbol{X}_{j-1}^R - \boldsymbol{X}_i^L) - ((\boldsymbol{X}_{j-1}^R - \boldsymbol{X}_i^L) \cdot \boldsymbol{n})\boldsymbol{n}| \tag{5.26}$$

In the case whereby there is not a 'next' or 'previous' step by the opposing foot (as will occur at the beginning and the end of each walking bout), the previous two foot locations of the opposite foot are used to form the vector. By defining these gait features in vector form, this ensures that a backwards stride (such as in Figure 5.2) will have a negative step length, and crossing of the feet will have a negative step width.

5.2.5 Defining walking

1. **Walking is a period of movement of at least three strides.**
2. **Walking must within the bout include the net lateral conveyance of the individual beyond their immediate area of influence,** $1m$**.**
3. **The velocity of the trunk in the direction of progression must not drop below** $0.1ms^{-1}$ **for more than 2 seconds**
4. **Walking begins at the start of the first stride and ends on the completion of the final stride of those strides considered part of the bout.**

Walking is defined in Whittle's Gait Analysis as: *'...a method of locomotion involving the use of the two legs, alternately, to provide both support and propulsion'.* However this is immediately qualified by the acknowledgement that such a definition may not apply to several pathological gaits [250]. The subject is thus concluded with the aphorism: *'it is probably both unreasonable and pointless to attempt a definition of*

walking which will apply to all cases...' [250]. Although this may be true from a traditional clinical perspective (where treatment is more dependent on observation, patient interaction, and context, than on strict definitions), it does not hold when applied to engineered solutions that try to delineate periods of walking from non-walking. This becomes even more problematic when applied to free-living gait. Accurate walking detection is evermore important as it is a fundamental step in an increasing number of clinically relevant technological systems; these include fall detection [251], rehabilitation [252–256], and disease management [257]. As such, its objective accuracy of detection has important real-life applications.

Often the definition of walking is dependent on the method of measurement. For example, walking detection via GPS may define walking as the movement of the individual between two threshold speeds likely to be discriminant of walking, flanked by two resting periods [258, 259]. Whereas, accelerometry researchers are more likely to define it in terms of numbers of steps detected. The method outlined here, however, ties all gait features to the spatio-temporal parameters of the feet.

5.2.5.1 Minimum number of strides in a walking bout

It is common to define a walking bout by its minimum number of steps. Table 5.4 lists numerous variations of how this may be applied. It is noted that a substantial proportion of walking occurs in short bouts with brief pauses. *Orendurff et al* saw that 17.0% of all walking bouts were 4 ± 1 steps in duration [260]. The decision therefore, is that 3 strides should be the cut-off, this is partly with the aim of providing consistency with the majority of previous work, as well as acknowledgement that generally two strides qualitatively appears to be more of an adjustment than an actual period of locomotion.

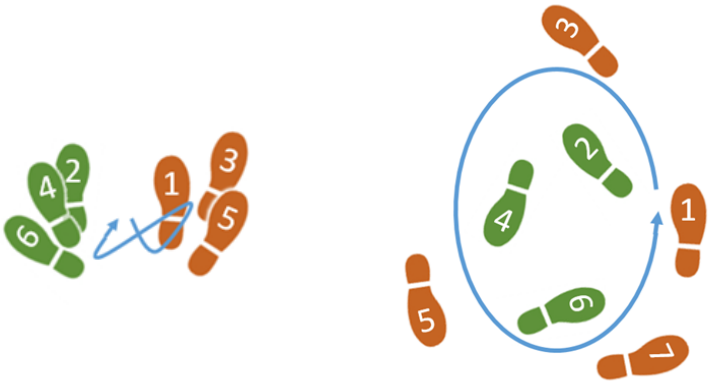
5.2.5.2 Minimum conveyance in a walking bout

It is insufficient to define a walking bout simply in terms of a number of consecutive strides and velocity as this does not necessarily ensure the 'propulsion' element of Whittle's definition is respected. It is clearly possible for numerous strides to be taken without the individual actually moving anywhere (Figure 5.10), but this movement, or 'conveyance' is fundamental to the principle of walking.

Table 5.4: Bout definitions from literature. How walking is defined in studies for a range of sensors and device-setups.

Paper	Device	Definition
Najafi et al 2003 [261]	Chest worn gyroscopes and accelerometers	At least 3 consecutive walking steps which must be between 0.25 and 2.25 seconds apart
Godfrey et al 2015 [262]	Lumbar worn accelerometer	At least 3 consecutive walking steps which must be between 0.25 and 2.25 seconds apart
Hickey et al 2017 [14]	Lumbar worn accelerometer and video annotation of chest mounted camera video	At least 3 consecutive walking steps which must be between 0.25 and 2.25 seconds apart, bouts must be more than 0.5 seconds
Fortune 2014 [263]	Activity Monitoring System (Thigh, waist and ankle worn accelerometers) and handheld camera video manual annotation	At least 4 consecutive steps separated by a pause of at least 2 seconds
Brodie et al 2016 [264]	Manual video annotation from handheld camera	A minimum of 3 consecutive steps. Steps must be no more than 3 seconds apart and did not include 'shuffling movements'.
Dijkstra et al 2010 [265]	Lumbar worn accelerometer	A minimum of 2 steps, where a step was defined as a forward displacement of the foot together with the forward displacement of the trunk.
Orendurf et al 2008 [260]	Ankle worn StepWatch Activity Monitor (accelerometer [Schmidt et al 2011])	A 10 second interval in which a minimum of 2 right footed steps occur. A bout ceases after the first 10 second interval without steps.
Culhane et al 2004 [266]	Thigh and chest worn accelerometers	Any 1 second moving window that was deemed to be either 'static' or 'dynamic'
Cavanaugh et al 2007 [267]	Ankle worn StepWatch Activity Monitor (accelerometer [Schmidt et al 2011])	Consecutive 1 minute intervals with a stride count greater than 0.
Le Faucher et al 2007 [258]	Backpack carried Garmin GPS 60	Movement between threshold velocities whereby walking bouts must be longer than 4 seconds, and pauses greater than 2 seconds.
Bussmann et al 1998 [268]	Accelerometers attached to both thighs and chest	Bouts of any activity must have a duration of more than 5 seconds
Storm et al 2018 [75]	Waist worn IMU	A minimum of 5 consecutive steps with no more than 1 second between steps
Soaz & Diepold 2016 [238]	Waist worn accelerometer	A step lasts no more than 2 seconds

Although many papers consider a short walking bout to be still of a sizeable length, for example, *Roos et al.* define it as anything less than 40 steps [269], it is noteworthy that clinically relevant gait tests such as the TUG test and 3-meter walk test [270, 271] are only across very short distances. Similarly, the standard instrumented walkways and motion capture technologies are generally designed to capture walking data over 2 to 6 metres (e.g. GAITRite [37, 272, 273] and Vicon respectively [274]). It is clearly important to consider very short distances as *Orendurff et al.* has shown that a substantial proportion of bouts are of less than four steps [260].



(a) **(b)**

Figure 5.10: Gait with no conveyance. Step maps show how stepping may occur with negligible conveyance, such as by: (a) constant readjustment of the feet during conversation, and (b) turning around on the spot while looking for an item in a room. Both these examples have been qualitatively observed in the Newcastle and Oxford data sets.

The distance that has therefore been chosen for minimum conveyance is $1m$ in any direction. This figure has been chosen by balancing three factors:

1. **Practicality:** From a purely theoretical perspective, the main purpose of locomotion is to move the individual so they may interact with things outside of their immediate sphere of physical influence. The area that one could hypothetically interact with is a function of physiology, flexibility, skill, and task. These themselves are complex functions of age, gender, and a vast range of other considerations. However general ergonomics guidelines suggest that immediate work spaces are to be approximately considered in the range $0.51 - 1.07m$ [275], and this seems sensible.

2. **Proxemics:** What an individual considers within their sphere of influence is also dependent on psychological factors. Although this varies by culture and age, the region defined by *Hall* as 'Personal Distance' is generally divided into the 'Close Phase' $(0.46 - 0.76m)$ and the 'Far Phase' $(0.76 - 1.22m)$ [276]. Similarly, there is evidence that the peripersonal space, colloquially referred to as the area 'within arms length' has a distinct neurophysiological basis and can extend to the limit of reach [277, 278].

3. **Simplicity:** Taking into account the variability of the other factors, it is therefore considered prudent to make the figure both reasonable and simple for the sake of recollection and dissemination.

It would of course be possible to bind the strides themselves with a minimum requirement for displacement in the direction of movement, indeed *Dijkstra et al* use this in their walking definition [265]. However this has two key limitations. Firstly, it prevents us from analysing strides as independent events. Secondly, not every stride will necessarily contribute to conveyance (Figure 5.10).

5.2.5.3 Minimum speed of walking

Progressing from the more philosophical, and similar to the stride definition, it is also necessary to define a minimum threshold walking velocity. Otherwise, continuous sporadic foot movements that result in a potentially infinitely slow drift in a single direction would be considered walking. Therefore, walking velocity is measured in terms of the direction of progression, which itself is defined in Section 5.2.1. To ensure that the walking bout definition is as robust as possible, the slowest feasible gait speed must be considered.

Unfortunately, minimum walking speeds are rarely reported, and it is clear from Table 5.1 that some individuals have very low gait speeds. Indeed *Graham et al* noted in their study of 173 ambulatory adults, aged 65 years and older, who had been admitted to an Acute Care unit in Texas had a mean walking speed of $0.43 \pm 0.23 ms^{-1}$ [279], with 75% being slower than the recommended lowest community-based risk threshold of $0.60 ms^{-1}$ [47]. Furthermore, from cut-point analysis they recommend a value somewhere in the range of $0.30 - 0.35 ms^{-1}$ as a meaningful threshold for maintaining in-hospital walking independence [47]. Pushing the considered threshold ambulatory speed lower, *Langois et al* report $0.38 \pm 0.18 ms_{-1}$ as the mean walking speed of elderly individuals who self reported difficulty crossing the street [233], and *Hardy et al* considered a gait speed of $0.2 ms^{-1}$ as the minimum threshold for inclusion for their study into the prediction power of a number of metrics in the survival of older adults (of which gait speed was one metric) [232]. Therefore there are clearly individuals who are community-living, and thus may undergo gait analysis with exceptionally low gait speeds. In an attempt

to establish gait speed as a predictive metric of rehabilitation in the elderly, *Friedman et al* concluded from their study of 125 patients that a gait speed threshold of $0.15ms^{-1}$ was discerning of 54.2% sensitivity and 100% specificity of predicting a patients long-term placing into hospital care versus home alone or rest home. However even they acknowledge a prospective threshold of 'immobility' as $0.23ms^{-1}$ when conducting the research. As individuals confined to hospitalisation are unlikely to be analysed, we shall therefore select out slowest gait speed as $0.10ms$ to ensure a margin of error.

In order to ascertain that walking has stopped, a period of time spent below the threshold walking velocity is required. Similar to with stride velocity, this ensures that the period of conveyance has indeed ended. For example, an individual's velocity may drop below the threshold briefly when they stop to open a door but this perhaps should not be considered a cessation of walking. The importance of defining a minimum resting time is emphasized by the finding by *Orendurf et al* that 33.1% of rests in free-living gait are $\leq 10s$. Table 5.4 shows a range of definitions that are used to delineate the end of walking. If focusing on free-living gait, which would likely have the most brief pauses, *Hickey et al* defined the minimum resting period between bouts as 2.5 seconds [14]. However, this was through manual labelling and was based on the assumption by *Aminian et al* that a gait cycle is always less than two seconds [280]. Once again, for the sake of both consistency with prior work and simplicity, a minimum resting period of $2s$ is decided upon.

5.2.5.4 Walking commencement and cessation

With a walking bout defined in both terms of strides taken and walking speed, a decision must be made on which takes priority in the delineation of walking commencement and cessation. It is decided that walking begins with initiation of the first stride that is included in the walking bout and ends at cessation of the last stride that is included. The prioritisation of strides over movement is necessary as, for example, an individual may stand up from a chair before commencing walking. During this period, their trunk velocity will likely be above the threshold walking velocity, however their feet will not have yet moved.

5.3 Methods

In order to extract gait features from Vicon trajectories using the rules developed in Section 5.2, the location, orientation, and velocity of various anatomical regions must be known. How these signals are produced is covered below. First however, the location signals themselves must be processed and filtered.

5.3.1 Preprocessing

The raw marker trajectories are imported into Vicon Nexus 2.11 on a local machine and processed. With the markers labelled in accordance with the marker set shown in Figure 4.2b, the software tracks that marker throughout the recording. Although the Vicon system is comprised of an array of cameras, marker trajectories can be lost during the recording. This can occur for various reasons such as: obscuring the marker with furniture, a marker falling between a fold of clothing, and a marker trajectory being obfuscated by an erroneous marker (which can be any reflective surface or object).

Loss of marker tracking results in a gap within its trajectory. These gaps must be filled and the software provides several standard methods including spline filling, cyclic filling, or pattern filling. The particularities of the algorithms are not revealed but recommendations for their use ensure the optimum choice in each scenario [281].

For short gaps (less than or equal to 0.5 seconds), or those during a period where there is little movement, gap filling will likely cause minimal problems. However, if the subject steps outside of the capture volume, or markers trajectories are lost during activities, the Vicon ground truth results will be erroneous. This is especially likely during the protocols developed for the Oxford study due to the types of activities which are unusual for clinical gait motion capture. Similarly, the filling algorithms are predominantly designed for linear, cyclic gait, and therefore often offer poor gap filling guesses during ADLs and turns. In these circumstances, the most appropriate gap filling algorithm was chosen on a case-by-case basis; predominantly the best was the pattern filling algorithm. The only occurrence where this was not the case was if there was no appropriate other marker to match the trajectory of, and the marker had clearly barely moved during the gap; in which case the spline filling algorithm was used.

For large gaps (more than 0.5 seconds), pattern filling was used. If another marker on the same foot was present during this gap, the trajectories were reasonable, if not, they were very poor. Either way, gait features were not extracted from periods that contained large gaps to ensure the reliability of the results. A large gap is defined as one comprising more than 0.5 seconds (50 samples) to ensure that no strides can be completely missed, given the maximum cadence decided upon in Section 5.2.3.4.

5.3.2 Filtering

Due to the noise in the Vicon marker location signals, it is beneficial to filter them before further processing. Each signal was lowpass filtered with cutoff at 10Hz. In deciding the filter cut-off frequency, three key considerations were made. Firstly regarding the key frequency components of gait, secondly the range of frequency components of everyday activities, and thirdly how these are affected by anatomical points. For example, it seems likely that the extremities of articulated limbs may have higher frequency components than the trunk; and this must be accounted for.

The frequency components of gait are heavily researched. *Sharma et al* found the key frequency ranges throughout periods of gait to be 0.5Hz while stationary, 1.5 to 2.5Hz for normal walking, and 2.5 to 3.0Hz while running [282]. Although specific values vary somewhat, the general range reported by other studies agrees with this. For example, 1.8 to 2.0Hz found by *Pachi & Ji* [283], and 1.3 - 2.8Hz by *Osaki et al* [284].

Similarly, *Al-Naime, Al-Anbuky & Mawston* found walking activities to be predominantly between 0.75 and 1Hz. However they expanded on this and investigated the frequency components of both the wrist and ankle during nine different sporting activities. The wrist accelerometer consistently recorded higher frequencies with the upper limit of dominant frequency reaching 4.0 to 4.5Hz for intense cycling [285]. Furthermore, it is worth noting that outlier frequency components of up to 6Hz have been noted in limited subject cases, while for example, stair climbing [286]

The majority of everyday life is neither walking nor sports. Many papers suggest that the all body movement falls below 20Hz, with 99% of acceleration power below 15Hz however this seems to be predominately based on work focused on both gait specifically and on force plates (and thus not necessarily representative of the whole

anatomy) [79, 287, 288]. This is indirectly supported by research that sampled everyday activities at 10Hz increment sampling frequencies between 10Hz and 60Hz. Very little change in activity classification was seen across the range of sampling frequencies for almost all classification techniques [289], a validation of similar previous work at 20Hz suggested that the vast majority of activity occurs below 10Hz [290]. A more specifically useful study utilised a gonionmeter on the wrist which recorded that the mean dominant component was approximately 1Hz with 75% of the energy falling below 5Hz in 24 different ADLs [291]. This is supported by the development of a whole-room force plate and calorimeter that suggested the majority of the energy of everyday living being in the 0.3 and 3.5Hz range [80].

5.3.3 Differentiation

The velocity, $\boldsymbol{v}$, at sample number, n, is extracted from the Vicon 3D location measurement, $\boldsymbol{s}$ by numerical differentiation:

$$\boldsymbol{v_n} = F_s(\boldsymbol{s_n} - \boldsymbol{s_{n-1}}), \tag{5.27}$$

where F_s is the sampling frequency.

5.3.4 Orientation

The orientation of the camera, and thus the torso, is derived from the three camera markers (Figure 4.2d). The location of these markers in the OGL coordinate system are M_{Cam}^{TL}, M_{Cam}^{TR}, and M_{Cam}^{BL} respectively (Figure 5.11). To convert these locations into an orientation, orthogonal vectors for the camera coordinate basis must be calculated. Directly, $\boldsymbol{x}$, $\tilde{\boldsymbol{y}}$, and $\boldsymbol{z}$ can be calculated as:

$$\boldsymbol{x} = \frac{M_{Cam}^{TR} - M_{Cam}^{TL}}{||M_{Cam}^{TR} - M_{Cam}^{TL}||} \tag{5.28}$$

$$\tilde{\boldsymbol{y}} = \frac{M_{Cam}^{BL} - M_{Cam}^{TL}}{||M_{Cam}^{BL} - M_{Cam}^{TL}||} \tag{5.29}$$

$$\boldsymbol{z} = \frac{\boldsymbol{x} \times \tilde{\boldsymbol{y}}}{||\boldsymbol{x} \times \tilde{\boldsymbol{y}}||} \tag{5.30}$$

Although the markers are set up so that they form perpendicular $\boldsymbol{x}$ and $\boldsymbol{y}$, some uncertainly is inevitable in marker placement (this is deliberately exaggerated in Figure 5.11). Therefore to ensure the orthogonality of the axes, $\boldsymbol{y}$ is defined as:

$$\boldsymbol{y} = -\frac{\boldsymbol{x} \times \boldsymbol{z}}{||\boldsymbol{x} \times \boldsymbol{z}||} \tag{5.31}$$

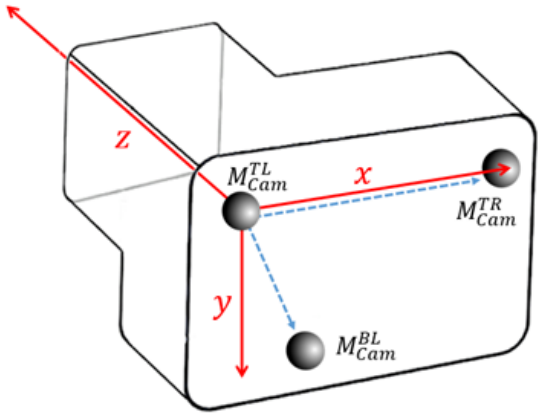


Figure 5.11: Camera orientation from motion capture. The camera basis is calculated from the 3D positions of the three coplanar camera markers to ensure orthogonality.

5.4 Results

The results in this chapter cover the ability of the methodology as applied to the proposed gait definitions to extract gait features from Vicon motion capture signals. Firstly, the reliability of the Vicon signals is explored via the proportion of the signal that requires gap filling, and the consistency of known distance measurement. Secondly, the sensitivity of gait feature results to the parameters selected in the definitions is shown through sensitivity analysis.

5.4.1 Motion capture reliability

As outlined in Section 5.3.1, the marker trajectories may be lost during the protocols, and need estimation. Table 5.5 gives the statistics of both how often and for how long marker trajectories are lost for each marker. The camera and feet markers were those lost most regularly and for the longest periods of time. The most complex movements and activities actually represent less marker trajectory ambiguity.

Table 5.5: Gap filling statistics for motion capture signals. The mean and standard deviation of the gap count across all subjects for each protocol and marker is shown, as well as the overall proportion of the recordings that these gaps represent. ADL values are averaged across the three protocols, and camera values are averaged across the three markers.

	Straight		Turns		ADLs	
	Gaps	Prop	Gaps	Prop	Gaps	Prop
Camera	16.0 (3.6)	9.0 (11.5)%	22.5 (22.2)	1.8 (2.1)%	1.6 (2.2)	0.1 (0.2)%
Sacrum	5.5 (3.3)	2.1 (2.4)%	2.3 (1.5)	1.2 (0.9)%	1.0 (2.0)	0.4 (0.6)%
Watch	4.0 (4.7)	1.8 (3.2)%	2.8 (3.1)	0.8 (1.0)%	1.3 (2.3)	0.6 (2.0)%
Pocket	2.8 (3.1)	1.1 (2.2)%	5.8 (0.3)	0.3 (0.5)%	0.1 (0.3)	0.1 (0.1)%
Left Foot	16.3 (13.3)	2.9 (2.7)%	16.5 (18.5)	4.2 (3.6)%	0.7 (1.4)	0.1 (0.2)%
Right Foot	11.3 (9.8)	4.0 (4.7)%	15.5 (15.5)	4.7 (2.3)%	0.2 (0.4)	0.2 (0.7)%

To investigate the marker tracking accuracy, the distance between camera markers M_{Cam}^{TL} and M_{Cam}^{TR} were compared through the trials. The markers are rigidly attached to the camera screen and thus should not move relative to one another. Figures 5.12a and 5.12b exhibit these measurements over 20s segments of a Straight and ADL1 protocol respectively. It is clear that there is considerable noise in the marker location measurement during both, and that gap filling has resulted in severe anomaly. The variance in distance between markers across all subjects was $2.03 \times 10^{-5} m^2$. Figure 5.12c shows that this uncertainty has an approximately normal distribution.

5.4.2 Parameter sensitivity

The choice of parameters and thresholds outlined in Section 5.2 are justified on theory and literature. It is therefore important to appraise these choices in the context of actual data. Figure 5.13 shows that the definition of a turn is robust; resulting in limited variation in the number of turns detected. Additionally, Figure 5.13c demonstrates that a substantial proportion of turns, or at least changes in direction, are below 45°.

Similar analysis is done on the stride definition. Figure 5.14 shows how step count is related to the foot speed threshold and minimum stride length parameters. Step count values are equivalent to stride counts, but step count is a more common term and so used here for clarity and comparability. The overall step count proves to be fairly robust and insensitive to the parameters of foot speed threshold and minimum stride length. However, it is noted that variation does occur; with a substantial number of foot movements at low speeds across distances less than 0.1m.

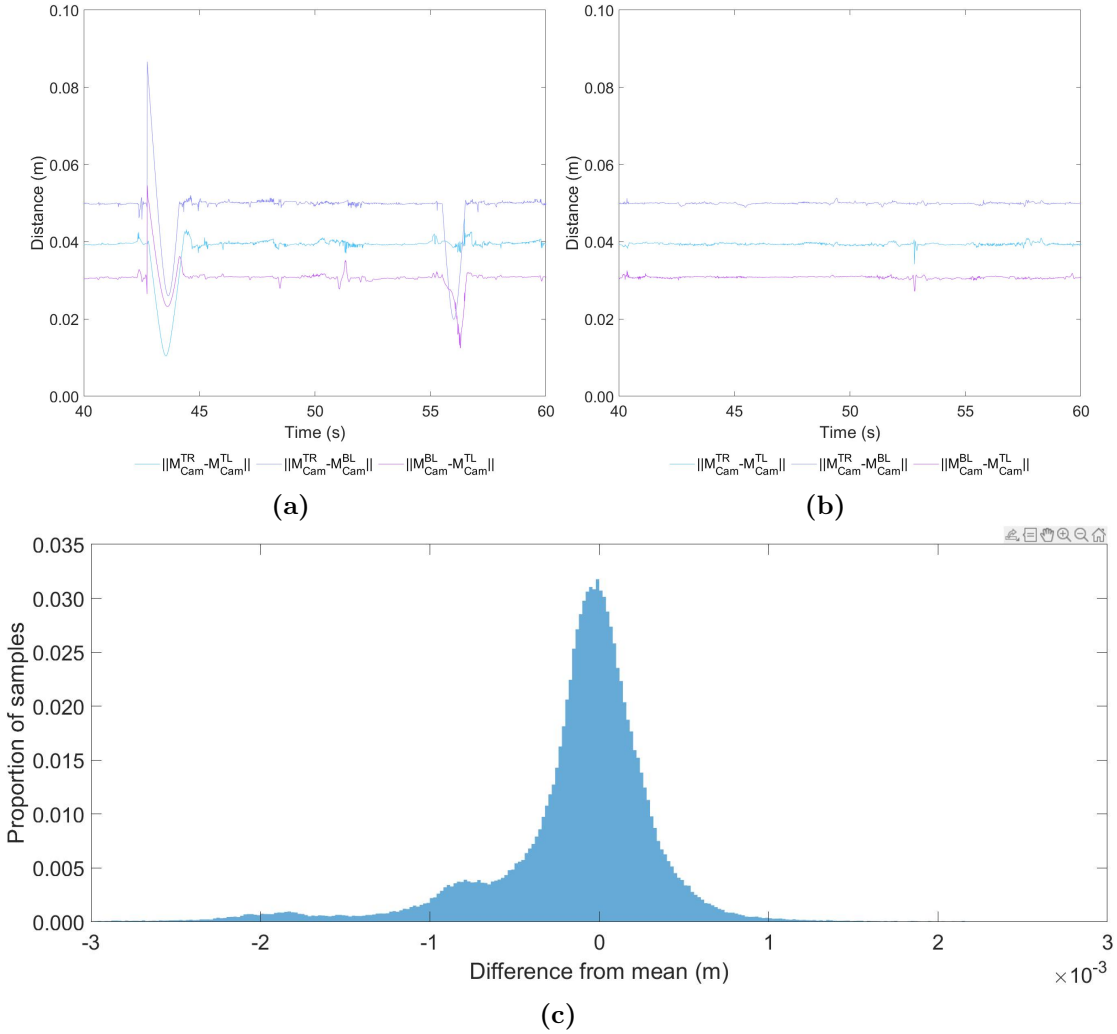


Figure 5.12: Motion capture marker location accuracy. (a) and (b) show the variations in the distances between the three camera markers across a short segment of a Straight and ADL1 protocol respectively. Erroneous gap filling and signal noise is clear. (c) shows the distribution of the difference between distance measurements in each sample, and the mean measurement across each trial - bin size 0.02mm.

5.5 Discussion

The data provided in exploration of both the capacity of the Vicon system to reliably monitor complex movement, and on the sensitivity of gait features to parameter choice is discussed below. Additional consideration is given to the marker set used.

5.5.1 Motion capture reliability

The amount of gap filling required in the Oxford study is actually remarkably high in the relatively simplistic Straight and Turns protocols. The fact that there was very little gap filling in the ADL protocols implies that this is not the result of marker occlusion

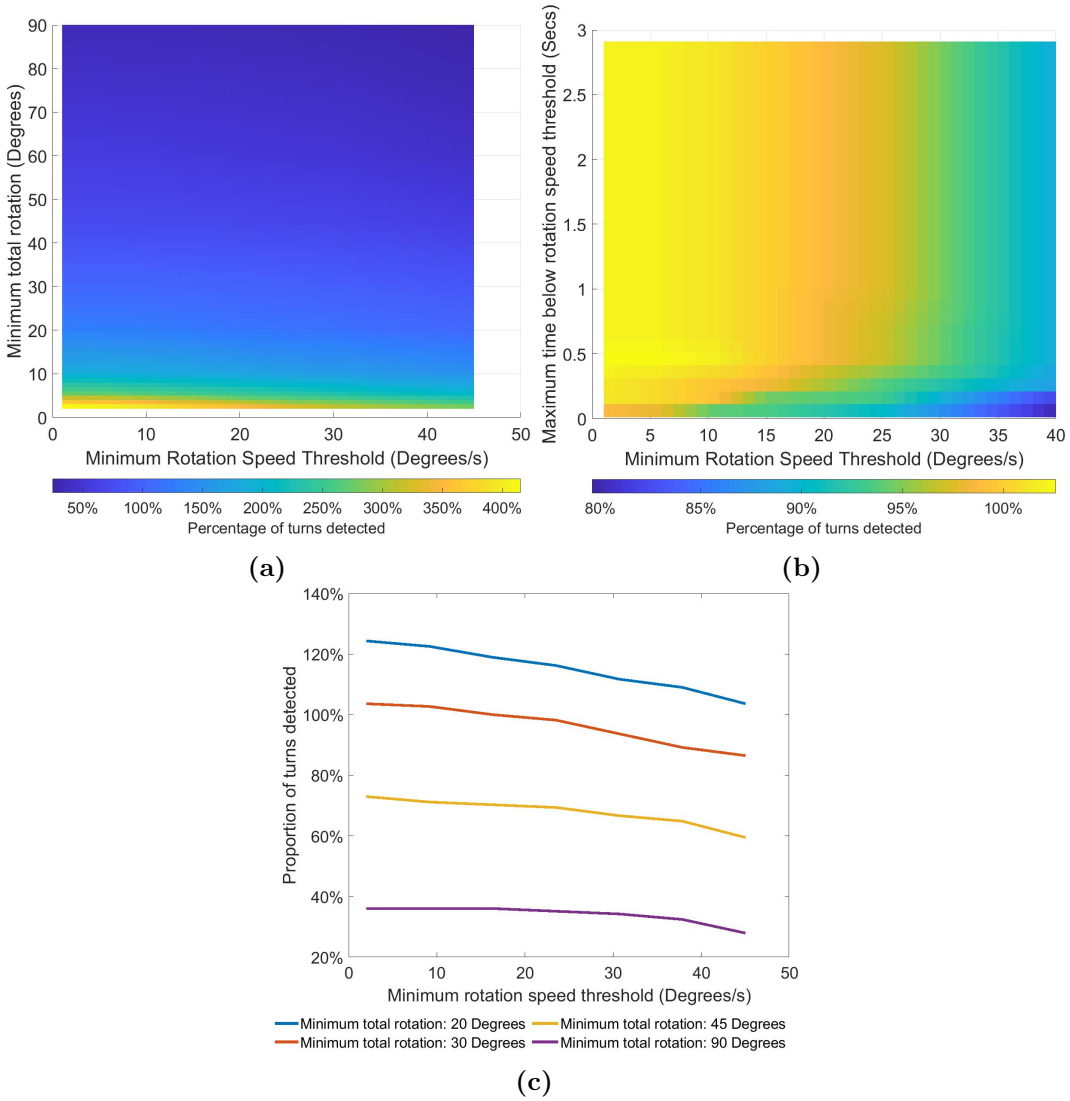


Figure 5.13: Turn detection sensitivity to minimum total turn angle and threshold turn speed. Turn counts are shown as the proportion of turns detected normalised by the total turn count at 15 degree threshold and minimum total turn angle of 30 degrees. The values are for all ADL protocols (c) Shows the differences in count for four common definitions of minimum turn angle.

by objects, furniture or body parts. It was qualitatively observed during processing that the vast majority of gaps, and especially long gaps, occurred due to the subject leaving the capture volume. This happened a lot as subjects walked rapidly backwards and forwards or through turns. It was qualitatively noted in provisional exploratory sessions that the larger (14mm rather than 9mm, and theoretically outdated) markers are tracked more easily by the system during complex movements. This amount of gap filling causes a slight limitation to how representative the Straight and Turns protocol step features are of those specific protocols. Although the filling of most gaps would not

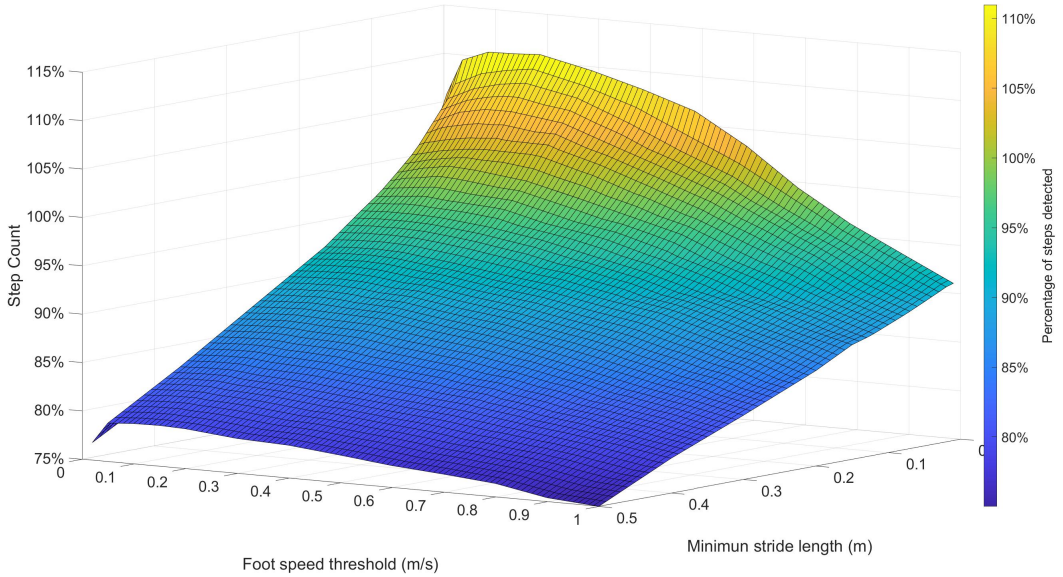


Figure 5.14: Stride detection sensitivity to foot speed and stride length. The step counts are normalised about the step value obtained by foot speed threshold $0.1ms^{-1}$, and a minimum stride length of $0.1m$.

affect the step features extracted from the trajectories (providing at least some period of the stationary foot is detected during the stance phases both before and after a stride is taken), this is not true of longer gaps. Though rarer these occurred predominantly when the subject accidentally left the capture volume during about turns and at the start and end of each protocol. If the gaps were over 0.5s, these would be removed from the data set, and so a disproportionate number of short and turning strides would be excluded.

It is also noted that the markers with the most required gap filling were those attached to the camera. This is likely due to the coplanarity of the markers, which would also often be obscured from view by the head and neck. In future work it would perhaps be beneficial to have a 3D axes of markers that project out from the camera in order to ascertain its location and orientation.

The accuracy of marker tracking is also a concern, as small variation in location accuracy can have a substantial compounding impact on velocity and acceleration signals. This is further compounded when considering signals derived from the measurements of multiple markers relative to one another, such as orientation. The distribution in variation is approximately normal; this is expected as the system is calibrated to within a tolerance of 2mm using a static LED (light-emitting diode) wand each session. The deviations from normal distribution are likely due to the marker

being contacted and moved during a trial, and the slight variation in calibration between subjects.

Overall this shows that using motion capture for ADL-like protocols is both possible and successful. Importantly, it is also shown that the validity of motion capture as ground truth, that is usually taken as unequivocal, should be considered during motion - especially in derivative signals. To reduce the clearly substantial effects of erroneous gap filling, where possible throughout the study and further chapters, gait features extracted from periods with large gaps were excluded from further analysis.

5.5.2 Parameter sensitivity

The turning definition provided applies thresholds for turning speed, total rotation and minimum time spent below turning threshold to stop the turn. The minimum turn angle is easily justified, as without it almost all movement becomes some kind of turn. However, the decision to reduce it from 45° to 30° is contrary to the literature. Certainly in many real-world applications, the changing in direction of 30° would be considered a turn. Figure 5.13 shows the sensitivity of turn detection to the choice of minimum turn angle and threshold turn speed. It is clear that a substantial increase in detected turns occurs if the turn threshold drops below 5°. This is most likely the result of trunk sway being detected as turning. Figure 5.13c shows that selection of a minimum rotation angle of 30° results in an increase of approximately a quarter in detected turns. This corresponds well with the acknowledged greater prevalence of lower angle turns [57]. Additionally, relative insensitivity to rotation threshold is demonstrated. Figure 5.13b shows that turn detection is almost independent of the minimum amount of time beneath the threshold before a turn is stopped. This implies that nearly all turns are in fact ended by almost immediate rotation in the opposite direction. However, a slight limitation of using velocity thresholds for turn start and completion is that some very small but not negligible proportion of the turn will be lost, and this proportion will increase the slower the turn.

It has been noted that previous studies have included a minimum and maximum turn time (for example $0.5 < t_s < 10.0$) [221]. The choice to not include this in the definition is due to a desire for full free-living robustness. On the upper end of the

spectrum, there is no theoretical limit to how long someone may turn in one direction. On the lower end, it is desirable to keep very sudden and rapid turns as these occur in real life, however the minimum required turn angle of 30° provides an approximate limit to this under normal conditions (see Fig. 5.13).

Figure 5.14 shows moderate robustness of step count to parameter values, however clear variation is seen. The increase in steps at very short stride lengths and speeds indicates the importance of selecting a definitive point from which to measure foot movement, as this is likely caused by heel movement, toe tapping, foot twisting or position adjustment. This could cause substantial step count errors if defined by heel strike or toe off [292]. These sort of movements are clearly common, but only really seen in free-living movement, demonstrating the importance of ADL-style protocols.

It is noteworthy that there is some precedent for using the centre of the foot as a measurement point [293, 294]. Furthermore, the spectrum suggests that manual and qualitative assessment of gait events as is often performed in studies, is insufficient [37].

Overall, the sensitivity analysis demonstrates that gait definitions are indeed worth exploring and unifying, as all gait movements, even in a semi-orchestrated environment, exist on a spectrum. This is important, for example, when claiming that one population group demonstrates a discriminatory smaller number of turns than another. For example, if it is claimed that MS patients are discriminated from HC by turn quantity [222] or step count [295] then this could actually be a reduction in rotation speed and step length. This furthermore backs the importance of event-by-event comparison of well-defined gait features that is a theme throughout this work.

5.5.3 Marker set

It is acknowledged that the marker set used in the Oxford study is not one of the standard configurations, such as the Human Body Model or Conventional Gait Model. However, as it is much more sparse and reduced than these, each marker will be represented within these [296]. The features extracted have been designed to utilise the easiest possible set up, and be relatively unaffected by slight variations in marker position (as is demonstrated to some extent by the parameter sensitivity analysis).

This is important to note as variations in gait analysis due to marker set is a concern; although predominantly for more complex kinematic features [297].

To apply the same gait feature definitions to marker sets with no clear corollary, substitutions and adaptations could be made. These may include for example, using a single foot marker for stride extraction and lumbar, or head markers for walking speed and orientation. Additionally, if no coplanar or orthogonal markers are available for orientation, it would be possible to derive an approximate change in orientation from the DoP vectors in successive samples. Similarly, the definitions will be robust to technology as well. For example if non-depth computer vision methods are used to extract gait, then the centroid of the foot image can be used regardless of view angle (where toe off or heel strike might normally be obscured) [298].

5.5.4 Gait feature definitions

A key purpose of Section 5.2 is to try and theoretically build definitions of key gait features accounting for as robust a set of circumstances and subjects as possible. The aim is that through such considerations, unified definitions can be decided upon that will improve inter-study comparability. The importance of this has been acknowledged by recent similar work undertaken to come to such a consensus.

An example of such work is that of the efforts of *Kluge et al* in 2021 to establish a consensus based framework for gait terminology [69]. In this work, a consensus on the meaning of key terms such as walking bout and turning was obtained through questioning over 50 clinical and technical gait experts via an adapted Delphi method [299]. Although the resulting definitions are in many ways similar to those proposed here - many key differences are clear. One such example is that turns are deemed more prescriptive, including deceleration in the forward direction and a turn angle greater than 45°. Other forms of changes in heading while walking are considered to be curvilinear walking; and it is acknowledged that this form of walking needs clearer definition. Additionally, it seems that limits of these definitions may include the requirement for consecutive left and right stepping to define a walking bout - which may not account for more complex free-living activity. Similarly, walking bouts are defined as ended by turning regardless of turn style, strategy or speed.

Although the definitions differ from those proposed here, the mere existence of such work validates the necessity of such unified robust definitions. Combining these consensuses with edge-case considerations and sensitivity analysis of the parameters (e.g. the length of a resting period that ensure the division of a walking bout into two), is likely the best way to conclude on unified and robust definitions. A final consideration is that the field will likely always suffer to some extent from the difference between the colloquial understanding of a term - and its formal definition; so perhaps more specific and unique technical terms are required to avoid all ambiguity.

5.6 Conclusion

Overall, a framework for applying stride-level gait analysis in any environment with any individual is proposed, justified, developed and critiqued. The definitions of such fundamental concepts as 'the stride' and 'walking' are provided and exhaustively explained while giving consideration to the range of all possible normal, pathological and aberrant gait. The definitions are proposed using the least specific measurement requirements possible, however, they are tested using a Vicon motion capture system. The accuracy of the Vicon system is considered due to the unusual protocols devised in Chapter 4 to test the definitions' relevance to true free-living gait.

The actual gait features extracted from the study using the definitions, are utilised in Chapters 6, 7 and 8 as a ground truth against which to compare computer vision and accelerometry-based gait analysis. For completeness, the variability in the gait features as dependent on the definition parameters is investigated. Although the definitions are shown to be relatively robust, a key inference is that gait features exist on a spectrum and that small changes in parameter values can substantively affect the results. This is a widely under-acknowledged consideration in gait analysis and could have significant implications for study conclusions and inter-study comparability. Fundamentally, a discussion needs to be undertaken to account for extremes on both the engineering-clinical and theoretical-practical axes of consideration. A key element of future work would be in the discussion of these definitions with clinical gait experts.

6

Measuring human movement with a wearable camera

Contents

6.1 Introduction 111
6.2 Methods 117
6.3 Results 122
6.4 Discussion 128
6.5 Conclusion 139

6.1 Introduction

One of the key aims of wearable technologies in the biomedical field is the tracking of movement, gait speed, and activity. Wearables are highly congruous with this purpose as they allow monitoring of the subject outside of the controlled research environment and throughout their normal everyday activities. Although a vast array of devices have been developed for this purpose, both research and commercial, the field is dominated by inertial sensors [300].

As the size of cameras has decreased, and the affordability and on-board memory increased, wearable cameras have seen substantial adoption for a range of applications. In the field of clinical research, video cameras are beneficial as they offer rich contextual information. This has been noted in the field of activity recognition, where wearable cameras have been widely adopted both in their own right [97, 98, 301–304], and as the providers of ground truth against which to validate other technologies [305, 306], or in combination with them [307].

Visual odometry (VO) is the estimation of the egomotion of an entity solely by the use of cameras attached to it. Often, this is via the use of the first-person, or 'egocentric' view of a single camera; otherwise known as monocular VO [308]. Established as a term by Nister in 2004 [309], the methodology is fundamentally composed of inferring the frame-by-frame rotation and translation of a camera and combining these estimations to establish the 3D trajectory of the entity. To achieve this, a range of additional techniques and methods are used to improve and augment the results, including: keyframing, bundle adjustment, and Kalman filtering.

The computational deduction of an object's movement and location within an environment has a wide range of uses and is a fundamental component to a range of contemporaneous cutting-edge technologies. The application to autonomous vehicles is clear, and so the leading edge of visual odometry is forged by researchers and engineers in robotics [310], self-driving cars [311], unmanned aerial vehicles [312], autonomous weaponry [313], and guided ordnance [314]. Indeed, the pivotal public application of VO was in the guidance of the Mars rovers; where direct control is prohibitively delayed by distance, and GPS does not exist [315].

It is from this use-case, of navigation and movement monitoring in GPS-denied environments, that wearable cameras and VO have integrated. Pedestrian dead reckoning (PDR) and personal navigation systems (PNS) are important technologies for human guidance, monitoring and safety with examples being the navigation of the visually impaired [316, 317], and of workers in mines, enclosed spaces and urban canyons [318, 319].

The broader application of these techniques to the biomedical sciences and gait analysis specifically has been very limited. There has recently been some experimentation in the area; *Gomez et al* developed a bespoke spectacle-mounted camera rig to estimate a wearers walking speed [320], and *Schneider et al* achieved similar results with the use of smart glasses in structured and semi-structured environments [321]. However, these papers demonstrate numerous limitations with such methods.

This chapter will first outline the technical background of VO, then the unique challenges and considerations that are confronted in applying VO to the type of data collected in the Oxford and Newcastle studies (Chapter 4). This will then inform the

development of methods by which macro gait features and signals can be extracted. It will then use these methods to extract turn detection, turn speed and walking speed from such egomotion videos. The results will be compared to the ground truth, as established from Vicon motion capture (Chapter 5), and discussed in the context of gait assessment and monitoring.

6.1.1 Technical background & theory

Visual odometry is a specific case within the field of visual structure from motion (SFM), which attempts to deduce the 3D structure of an environment, and a camera's motion within it, from the series of frames captured. The field is multipartite with a range of parameters influencing both the methodology and the result.

Generally speaking there are two types of method: feature-based and appearance-based. The feature-based method extracts image features (abstractions of image components, such as corners, lines or shapes) in consecutive frames. Based on the feature parameters, matching occurs between the extracted features so that the same point can be located in consecutive frames [322, 323]. These points are then used to estimate the camera's transformation (translation and rotation) between the two frames (Figure 6.1). The appearance-based, or 'direct' method, utilises either the variation in pixel intensity and thus optical flow (OF), or template matching (TM) to directly infer the movement between frames [310].

One of the key concerns with either method is that small errors in each frame-to-frame estimation cause rotational and translational drift. This is especially disruptive if the VO is being used in simultaneous localisation and mapping (SLAM) of an area. To counteract this, keyframing and bundle adjustment can be conducted. The premise involves identifying a keyframe which includes features that can be observed from multiple frames and then estimating all their relative transformations from this keyframe [324]. Bundle adjustment then takes these groups of frames that are all related to one another by a keyframe and reduces the error in reprojection when considering all of the views as a system [325].

An additional method of reducing error and drift, especially over long distances or periods of time, is loop closure. This is achieved by identifying when the camera

has returned to a location previously visited. With this knowledge the intervening trajectories can be adapted to ensure that the path returns to the same place [326].

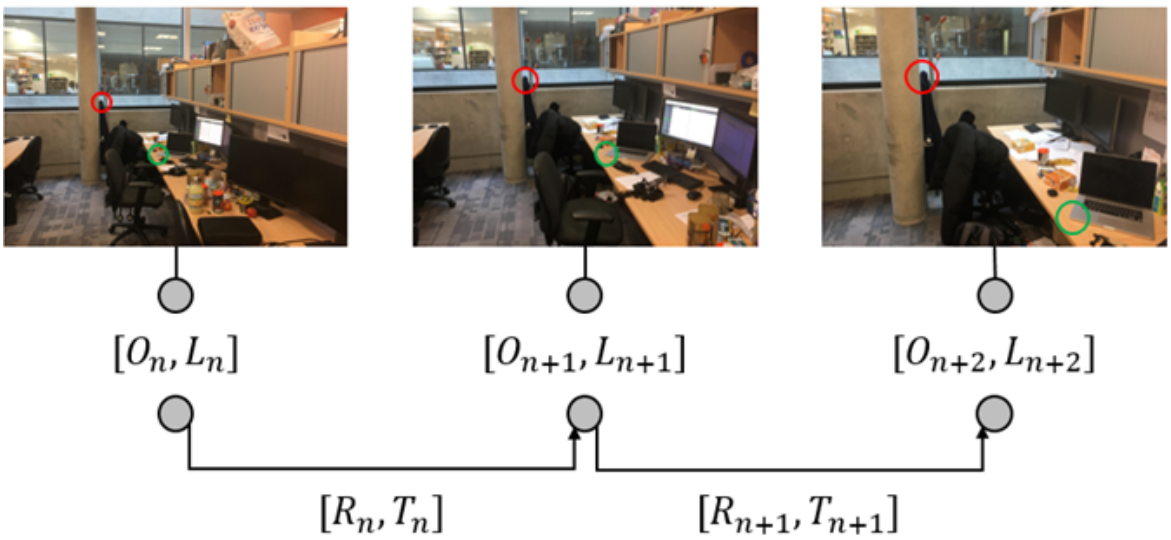


Figure 6.1: Feature-based visual odometry. Distinct features such as the shape of the top of the coat on the coat rack (red), or the corner of the laptop (green), can be identified in consecutive frames. This enables the estimation of the rotations and translations, $[R, T]$, between consecutive camera orientations and locations in the world coordinates, $[O, L]$. As the encircled features are visible in multiple frames, one of the frames could be selected and used as a keyframe.

6.1.2 Unique challenges & key considerations

The overall aim of the wearable camera system is to extract a range of gait features from egomotion videos in a free-living environment. Therefore, it is necessary for the camera to point towards the feet so that they can be tracked (Chapter 7), and to be able to track movement regardless of location or activity. Combined, these two requirements result in a range of unique challenges that must be considered when developing the VO methods.

The first key problem is the complexity and diversity of the FOV. Throughout the majority of videos collected in this way, the view is dominated by the lower body and the ground (as exemplified by Figures 6.2 and 6.3). The lower body is of no use in tracking the individual's movement within the environment and the ground is relatively featureless. Due to the homogeneity and low textural richness of flooring, any features that can be extracted from the ground plane are weak, generic and easily lost due to

blurring or illumination effects. Therefore, finding points of reference from which to infer the transformation is difficult. The homogeneity of the floor also prevents the use of many of the OF methods, as well as techniques for solving scale ambiguity.

The second key challenge is the speed of the camera combined with the lack of depth of scene. Although the camera is only travelling at the relatively low speed of human walking, with this FOV, this causes a complete change of scene within a sub-second time frame. Therefore, any distinct and salient image features actually found will rapidly disappear from view. This prevents the use of the valuable steps of keyframing and bundle adjustment. An additional complication that derives from this challenge is frequent and extensive blurring.

The third problem pertains to the prerequisite that the system work in a free-living environment. In this situation, there is no guarantee that the individual would return to the same locale twice, enabling loop closure. This problem is compounded by the first, in that with such a homogeneous view and weak features, how could it be established with any certainty that the same point had been returned to. One area of floor looks very much like every other area of floor.

Freedom of movement is also responsible for the fourth problem; the complete unpredictability of the camera's motion. Although the camera will move in a repetitive pattern during long bouts of linear gait, during ADLs the view and direction of movement may change suddenly and erratically. The camera can be perceived to be attached to the environment by a series of deformable links connected by articulated joints with many degrees of freedom. This is very much unlike a camera rigidly attached to an automobile, or smoothly moving through the air on an unmanned aerial vehicle. This prevents to use of methods that utilise motion prediction based on previous movement or data, such as Kalman filters and neural networks [327, 328]. This also prevents the use of scale estimation from known object motion [329].

The fifth concern is the somewhat ambiguous orientation and location of the camera on the body. Although it is generally known to be worn on the chest and pointing towards the feet, it is possible that: the chest strap may slip during movement, the camera might be knocked by the subject's hands or environment during activities, or simply the camera may be set up differently each time it is used. The first two of

these are unavoidable, and although the last may be countered by ensuring a highly prescribed set up, this is against the aims of the wearable camera system. One of the key issues in gait analysis by accelerometry is the dependence of the algorithms accuracy and reproducibility on IMU location and orientation. This is one of the issues that prevents the applicability of research to general application, is to be avoided if at all possible, and is explored in Chapter 8. The method should work regardless of precise camera location and orientation; this will also ensure robustness to subject physical requirements and anatomical dimensions.

View occlusion is the considerable, sixth challenge. Not only can a vast array of non-scene objects obscure the view of the rigid scene, they do so regularly, and can fill sizeable proportions of the frame. In this circumstance it is highly possible that the camera pose estimation will not occur relative to the static environment, but relative to the moving object. The most consistent of these is the subject's own lower body, legs and feet. These fill much of the top half of the frame: some parts constantly move, others such as the lower abdomen may remain static within the FOV. The arms, hands, and any held objects are also a substantial and consistent problem; especially in free-living data. Beyond this, non-subject objects may cause occlusion of the static scene, such as other people, pets, vehicles and furniture. Many examples of scene occlusion by moving objects can be found in Figure 7.3 and 7.4. This problem is quite a departure from the majority of VO data sets.

Problem seven is somewhat nuanced in that it is highly possible that the motion of the camera will be more-or-less planar and parallel to the ground plane, solely rotational, solely translational, or negligible altogether. Each of these can cause problems in at least one step of the transformation estimation process; for example it may result in ambiguous degenerate solutions [146].

Illumination in a free-living egomotion data set varies considerably, and thus is the eighth and final concern. This is explored and analysed in greater depth in Section 7.1.1.1, however it has two key impacts on VO. Firstly, very low illumination or deep shadow can obscure the scenery reducing the clarity of objects, texture richness, and feature density even further. Secondly, shine and shadow may in fact form more distinct

features than the floor itself. This could cause the motion inferred to be that relative to the illumination effect itself rather than the floor.

This is a considerable and perhaps daunting list of problems, however, several of these issues will be leveraged into advantages or confronted in a manner that helps solve a different challenge. These include the relative consistency in camera height being used to solve scale ambiguity, the relative certainty that the FOV will be dominated by a ground plane (or one parallel with it such as a desk or work surface) to solve camera pose estimation, and the requirements of dealing with low texture environments in the consideration of illumination effects.

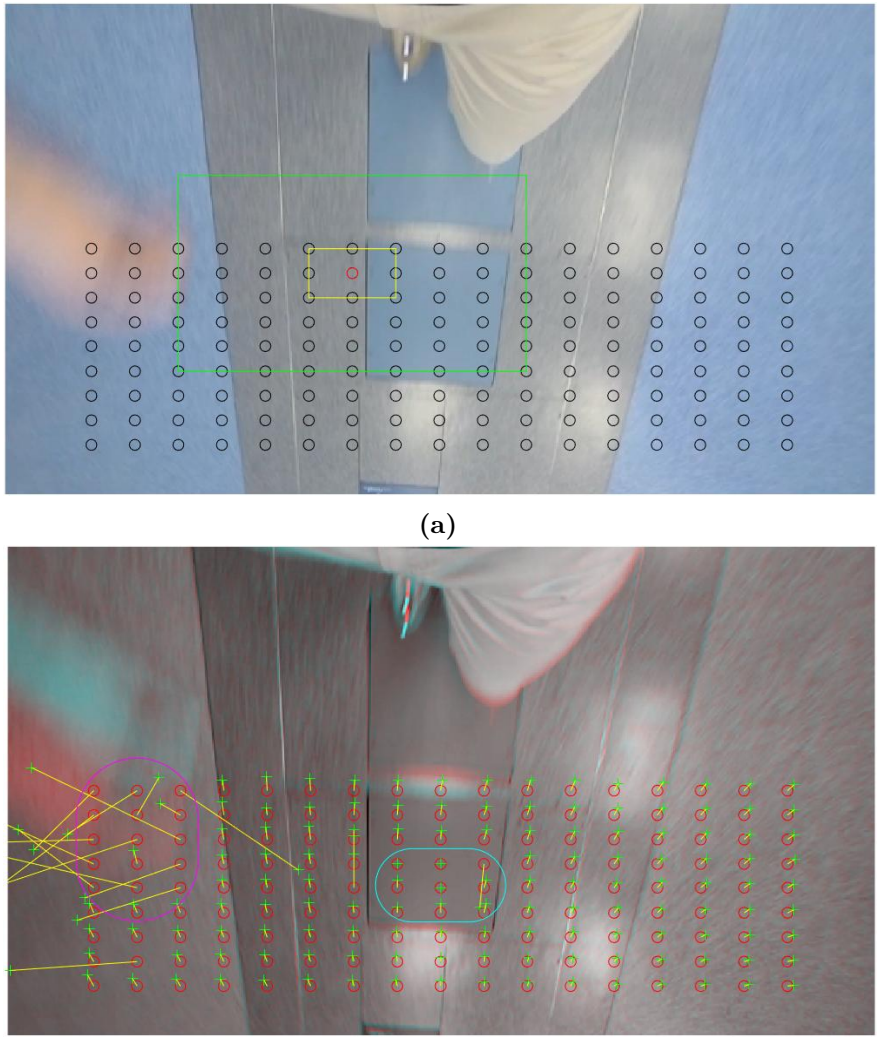
6.2 Methods

The following section covers the methods developed to best attempt VO given the unique challenges of the data set. The majority of the supporting mathematics was covered in Section 3.3, and referred to when appropriate. In implementation, quaternion representations of rotations and orientations are used throughout to reduce the effect of gimbal lock [330, 331].

6.2.1 Template matching

Due to the unique challenges involved in this problem, the frame-to-frame movement is inferred by a hybrid approach between the feature and appearance methods. Rather than extracting features from each frame and matching them across frames, a small region or sub-image of each frame is considered and searched for in an appropriate search area in the next frame; this is then repeated uniformly across a region of interest (ROI) within the frame (see Figure 6.2). This is substantially more effective than using traditional feature methods for a variety of reasons (see Figure 6.3). To reduce the effect of the lower part of the body on the movement estimation, the ROI is limited the to bottom half of the frame.

The identification of a template within an image is known as template matching (TM). This is achieved by taking the 2D fast Fourier transform (FFT) of the target region and the search area, and performing normalised cross-correlation (NCC) on the result [333] (Figure 6.4). The NCC of these two frequency domain arrays can be defined as:



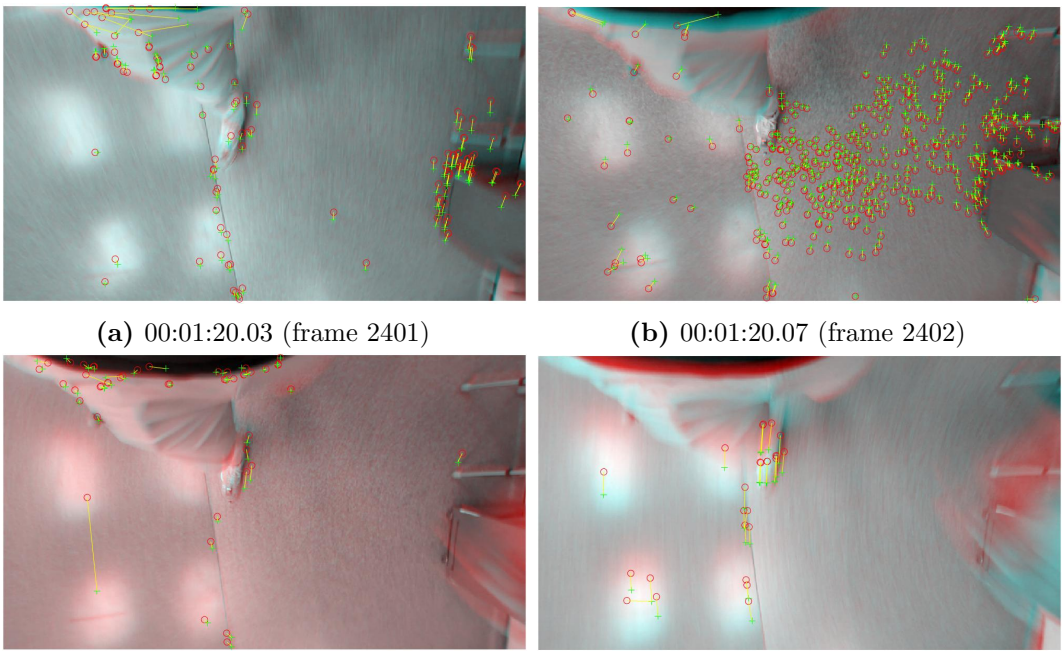
(a)

(b)

Figure 6.2: Template matching. (a) A location is selected within the image (red point). A region as a proportion of the image dimensions around this point is selected as the reference template (yellow bounding box). This template is searched for in the next frame within a greater but limited search area (green bounding box). This can then be done numerous times for a predetermined set of locations that can ensure both uniform distribution and a region less likely to have an obstructed view of the ground (black points). (b) The templates and their matched locations (red points and green points respectively) indicate the vector of movement of an image point between frames (yellow lines). Some templates are incorrectly matched due to obstruction by a moving object, such as the hand (magenta region), others result in incorrect vectors due to the complete homogeneity of the template texture (cyan region). This image is a false-color overlay stereogram showing the current, and next, frame. Red shading indicates higher pixel intensity in the original frame, blue, the latter.

$$NCC = \frac{\sum_{i=0}^{N-1}(\boldsymbol{T}_i - \overline{\boldsymbol{T}}) \times (\boldsymbol{I}_i - \overline{\boldsymbol{I}})}{\sqrt{\sum_{i=0}^{N-1}(\boldsymbol{T}_i - \overline{\boldsymbol{T}})^2 \times \sum_{i=0}^{N-1}(\boldsymbol{I}_i - \overline{\boldsymbol{I}})^2}}, \tag{6.1}$$

where $\boldsymbol{T}$ and $\boldsymbol{I}$ are the 2D-FFTs of the template and search areas respectively,



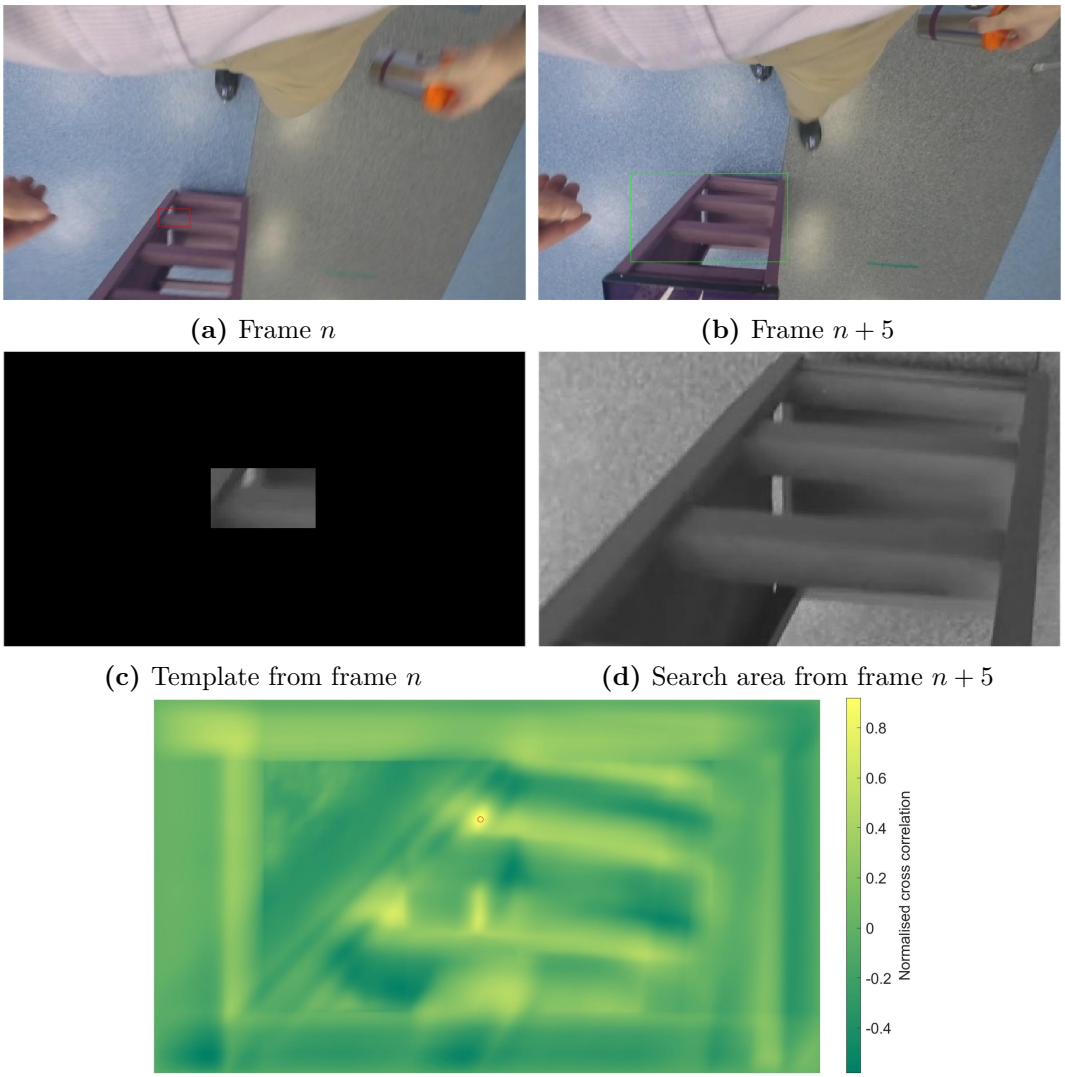
(a) 00:01:20.03 (frame 2401) **(b)** 00:01:20.07 (frame 2402)

(c) 00:01:20.10 (frame 2403) **(d)** 00:01:20.13 (frame 2404)

Figure 6.3: The limitations of image feature detection. Features were extracted, and matched for consecutive frames in a sample from the Oxford data set. KAZE features were used as they had been qualitatively assessed to be the most effective for this type of data; and the threshold parameter had been set an order of magnitude beneath normal operating levels [332]. Regardless, even within four consecutive frames (a period of just over a tenth of a second at 30Hz), key limitations of this method are demonstrated. (a) shows that the features are very unevenly distributed about the frame, and in fact are stronger on the subject and furniture, than on the ground plane which is what is being aimed for (see Section 3.3.1.3). (b) shows that although features in a relatively low texture surface can be extracted with a low enough threshold, this only occurs when the surface is in focus and is very inconsistent. (c) and (d) show that features cannot be found when there is blurring, and incorrect features are matched due to illumination and reflection.

and N is the number of pixels in the template [334, 335]. The NCC is used due to its higher accuracy when compared with other methods [336], and the 2D-FFT is utilised as NCC has been shown to be more efficient in the frequency domain [337]. Erroneous TMs are filtered out by the RANSAC process (Section 3.3.2.5) as part of the relative pose estimation.

The chief drawback of this method is that it is relatively slow, and increases linearly in time with the number of TM instances. However, as the aim is to attain high accuracy offline, this is preferential to faster, less robust methods. An additional benefit to this method over the feature-based method, is that no ambiguity is incurred by the point matching step of the process. Furthermore, this method is substantially



(a) Frame n

(b) Frame $n + 5$

(c) Template from frame n

(d) Search area from frame $n + 5$

(e) Normalised cross-correlation of template and search area

Figure 6.4: Normalised cross-correlation. The difference between the view in two frames is shown in (a) and (b). A frame difference of five is used to emphasize the movement. (a) contains a template (red bounding box) which is searched for, (c), in the green bounding box search area of (b), shown as (d). (e) shows the NCC and its maximum value (red point), is where the template and search area are most similar. This is clear to the eye despite the change in orientation and location of the camera, and thus perspective of the object.

less dependent on the resolution of the frame. A somewhat similar methodology was recently successfully used to estimate the relative motion of autonomous vehicles [335].

6.2.2 Pose estimation

In conventional VO, the epipolar geometry of points in the scene and of their corresponding homogeneous image points $\boldsymbol{x}_n$ and $\boldsymbol{x}_{n+1}$ in frames I_n and I_{n+1} respectively, can be determined uniquely given a minimum of five unique points. From this,

the fundamental matrix, $\boldsymbol{F}$, and thus relative pose of the camera at t_n and t_{n+1} can be determined up to scale and expressed as the combination of a rotation and translation, $R_n T_n$, where $\|\boldsymbol{T}_n\| = 1$ [308].

In the circumstance that all the points of the rigid scene lie on a plane, then the two camera centres and the world points lie on a critical surface, which is a degenerate quadratic surface [338], and as such the set of image points is geometrically degenerate with respect to $\boldsymbol{F}$. This is a key issue with this data set, as the vast majority of the points in any frame are likely on the floor, and thus coplanar.

Although problematic for 3D reconstruction; this coplanarity can be exploited to enable the extraction of camera movement. The entire mathematical process is outlined in Sections 3.3. In brief overview, the coplanarity of the world points induces a homography between the image points in each frame (Section 3.3.1.3). Once calculated (Section 3.3.2), the homography can then be decomposed and constrained to provide two physically realisable solution to the relative camera pose between the two frames and the plane that defines them (Section 3.3.3.1-3.3.3.2). To identify which of these two solutions is correct, the coplanarity of the points can be exploited again. A third view of the same points results in two more relative camera pose and plane estimations. The solutions that correspond to the most congruous pair of planes is deemed the correct solution (Section 3.3.3.3). To obtain this third view, the entire TM process in Section 6.2.1 is repeated for the frame combination: $\{I_n, I_{n+2}\}$.

The calculated homographies, and thus relative translations, are only valid up to scale. To extract the correct scale of translation, the scaling factor is calculated as a ratio of camera height from the ground. This is deemed to be approximately constant in this data set. Therefore, by providing the additional metric information of the camera's height from the ground while being worn, the scale can be calculated. This height is measured as part of the study preparation (Section 4.2).

The result of this process is that for each frame pair in the video, a relative rotation, translation, plane normal, and scaling factor are estimated from the perspective of the first camera's reference frame;

$$\{I_n, I_{n+1}\} \Rightarrow (\boldsymbol{R}_n, \boldsymbol{T}_n, \boldsymbol{N}_n, s_n). \tag{6.2}$$

By extracting the plane normal, the effects of rotational drift can be reduced both by directly converting the camera rotation into world coordinates, and by constraining the camera's rotation relative to the ground (i.e. predominantly camera yaw and pitch in this data set).

6.2.3 Feature extraction

The gait features to be extracted from the visual odometry are rotational speed, walking speed, turn detection, turn angle, and mean and max turn speeds. Both turns and walking are defined in terms of their linear, or angular distance in the transverse plane. Therefore measured transformations must be converted to world coordinates.

Firstly, each of the signals is Hampel filtered to remove extreme outliers with a window size of 1 second and an outlier threshold value of three standard deviations [339]. The signals are then lowpass filtered to remove frequency elements above 4Hz. This value will still include the vast majority of both gait and ADLs (Section 5.3.2). Then, as the scaling factor for each frame is a function of a range of variables, each with noise, it is calculated according to Equation 3.49 using the filtered signals.

To convert the camera transformations from the camera coordinates to the world coordinates, the rotation between the estimated normal of the plane, $\boldsymbol{n}$, and the world plane normal, $[0, 0, 1]^\top$ is calculated and the relative camera pose signals transformed accordingly. The normalised and world-oriented translation can then be multiplied by the scaling factor and measured camera height to provide the translation signal in the transverse plane. The magnitude of the velocity of movement in this plane is considered walking speed.

Similarly, the roll component of the rotation represents turning speed in the transverse plane. These signals are then submitted to the same feature extraction algorithms as the Vicon signals to extract and measure turns and movement bouts. For direct comparison, detected turns are matched to the ground truth (Section 3.2.3.2).

6.3 Results

The movement of subjects as measured by the VO system can be broken down into rotational speed and translational speed. Rotational speed can then be processed to

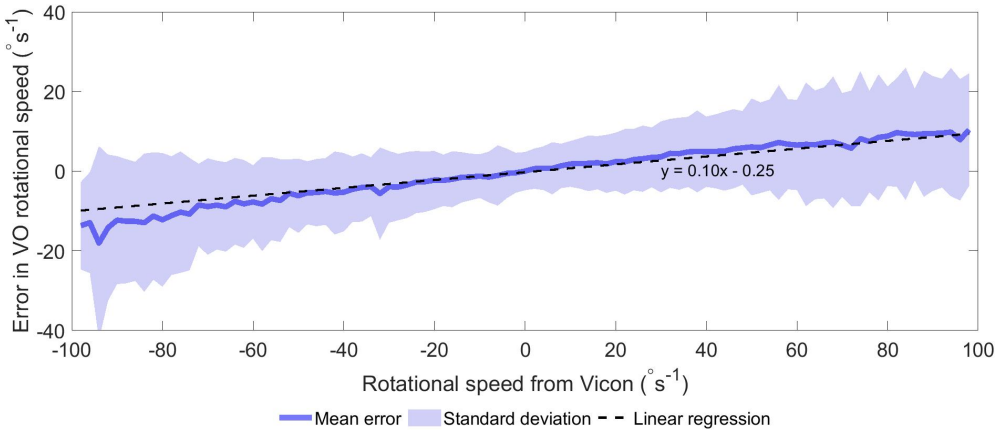
detect turns, and thus extract turn angle, mean speed, and max speed. Movement speed differs from walking speed. Walking speed can only occur while taking strides as per the definition in Section 5.2.5 but movement can occur while the feet are stationary.

6.3.1 Rotational speed & turns

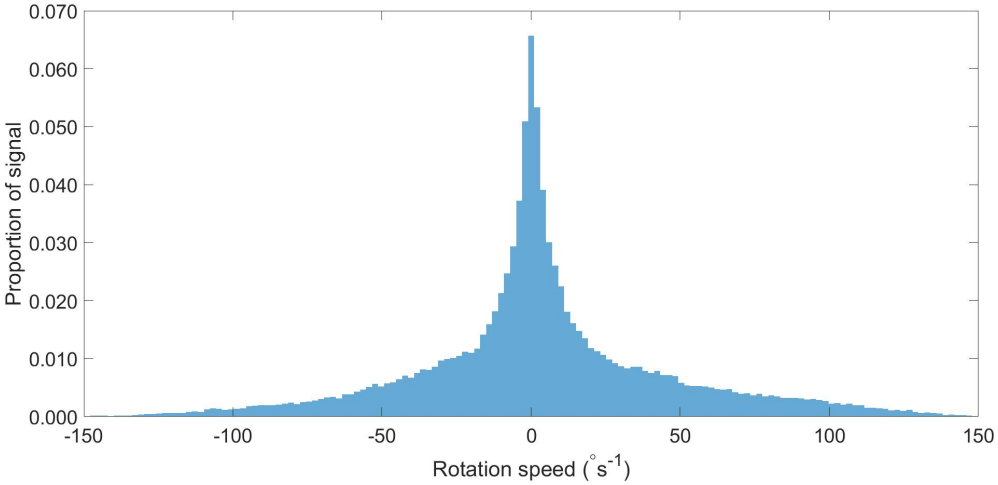
The accuracy of the rotational signal can be assessed before turns are extracted and measured. The accuracy of this signal is also important for future developments such as localisation and trunk sway. If every point on the signal is taken as a measurement of rotational speed, then Figure 6.5a demonstrates the error in measurement across the range of physiologically common speeds for all trials. There is some increase in the spread of error as the rotational speed increases, from a standard deviation of approximately $5°s^{-1}$ at $0°s^{-1}$, to $12°s^{-1}$ at $100°s^{-1}$. The difference is calculated as the Vicon value, minus the VO. Linear least squares regression of the data shows a consistent under estimation of rotational speed by about 10%. The certainty of the error distribution is substantially lower at more extreme rotational speeds due to the scarcity of their occurrence. Figure 6.5b demonstrates the distribution of rotational speeds, showing a sharp peak a $0ms^{-1}$ and high kurtosis.

For further insight into the sensitivity of the rotational speed accuracy to activity type, Fig. 6.5c shows the distributions of error by trial. It can be seen that there is a substantially greater range in error for both the Turns, ADL1 and ADL2 protocols. ADL3 however evidences less error than even straight walking.

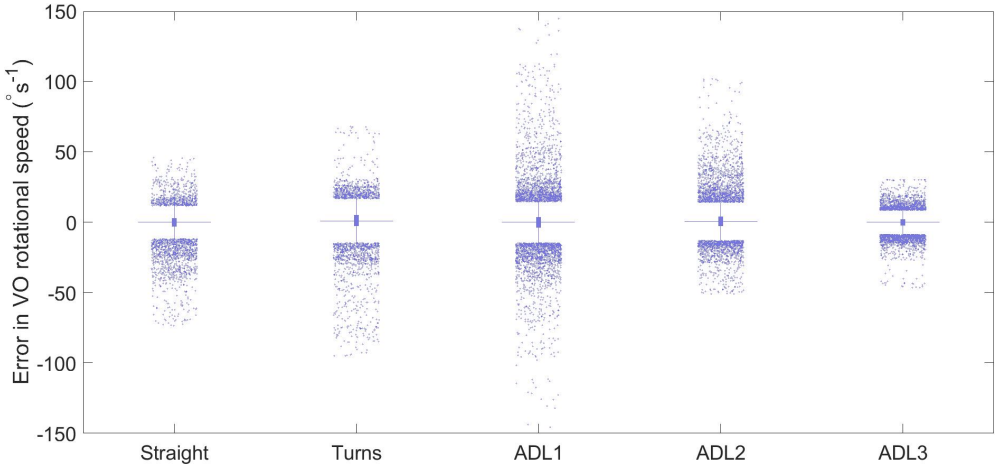
With the rotational speed signals processed in accordance with turn definition (Section 5.2.2), turns can be detected and measured. Table 6.1 provides the precision, recall and F-score for the turn detection. Other classification statistics cannot be provided as a turn event has no true negative. Turn events are detected at high precision across all subjects and trials. For further insight, each rotational speed signal can be converted to a binary 'turning' or 'not turning' signal based on the turning events that are detected. This enables the delineation of true negatives as any sample within the signal that is correctly identified as not during a turning event. The statistics provided by this further analysis are shown in Table 6.2. The precision with which the system can differentiate left turns and right turns is also provided in Table 6.3.



(a) Rotational speed accuracy



(b) Rotational speed distribution



(c) Distribution of rotational speed accuracy by trial

Figure 6.5: Rotational speed results. (a) demonstrates the dependence of rotational speed accuracy on the speed of the rotation by subtracting VO rotational speed from the Vicon value. The increasingly erratic results at the extreme ends of rotational speed are due to the relatively small number of samples at those speeds, as demonstrated by (b), bin size $2°s^{-1}$. (c) provides box plots to demonstrate the distribution of error in rotational speed across the different protocols.

Table 6.1: Turn event detection results. Event detection statistics for discrete turn events.

	Straight	Turns	ADL1	ADL2	ADL3	Overall
Precision	0.97	0.92	0.92	0.95	0.96	0.94
Recall	0.97	0.96	0.85	0.90	0.92	0.91
F-Score	0.98	0.94	0.89	0.92	0.94	0.93

Table 6.2: Continuous turn detection results. Classification statistics for the continuous binary turn classification of the rotational speed signal.

	Straight	Turns	ADL1	ADL2	ADL3	Overall
Sensitivity	0.89	0.90	0.78	0.82	0.92	0.86
Specificity	0.97	0.91	0.95	0.95	0.95	0.95
Precision	0.94	0.96	0.94	0.92	0.96	0.95
NPV	0.94	0.76	0.79	0.88	0.91	0.86

Table 6.3: Turn direction precision results. Turn direction precision results for each protocol by direction.

	Straight	Turns	ADL1	ADL2	ADL3	Overall
Left turn precision	1.00	0.99	0.99	1.00	1.00	1.00
Right turn precision	1.00	1.00	0.99	1.00	1.00	1.00

With turns detected, the most fundamental turn feature, turn angle, can be extracted (Figure 6.6). Correlation between the Vicon and VO systems of measurement is shown as well as Bland-Altman analysis that shows a mean error of 10% with standard deviation 18%. The proportional error in turn measurement appears to be independent of turn angle.

The two most basic additional turn features are mean turn speed and maximum turn speed. These have been analysed in a similar fashion to total turn angle, with correlations, Bland-Altman plots, and distributions of all and undetected turns in Figure 6.7. Interestingly, the mean error in both of these metrics is substantially less than 10%. This is acceptable for two reasons. Firstly, because studies have rendered similar error, even during substantially more constrained turning protocols (with some crucial turn studies neglecting to state error in measurement at all) [221, 340]. Secondly, because trunk sway before, during and after a turn confounds what is actually deemed to be the angle of the turn, and this sway often ranges between 5° and 7° [341].

Furthermore, the turn speed features are more normally distributed, implying consistent turn strategies regardless of turn angle. All turn features show a cut off at 30° as this was defined as the minimum threshold for a turn in Section 5.2.2.

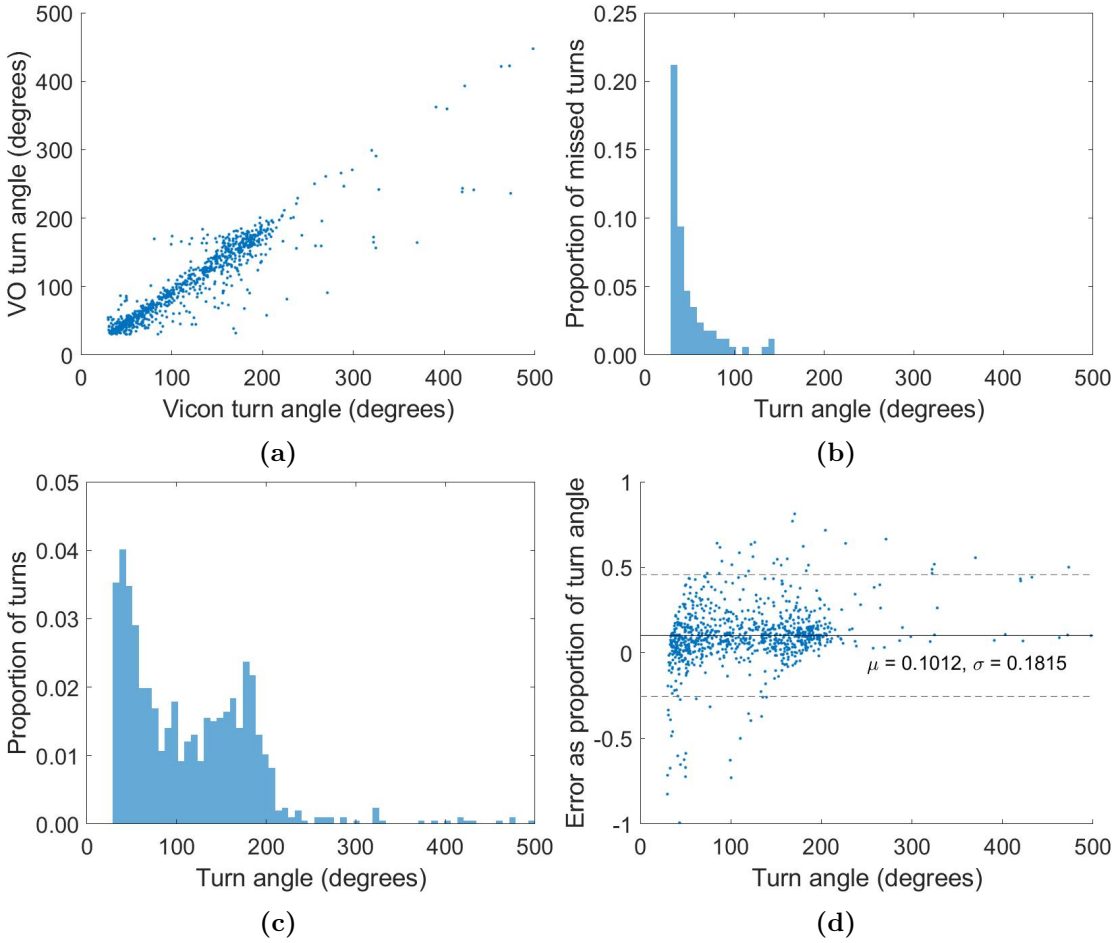


Figure 6.6: Turn angle. (a) plots the correlation between the VO and Vicon measured turn angles of all detected turns. The distribution of Vicon measured turn angle is shown in (c). (b) is the distribution of turn angles of only the turns that remained undetected. (d) shows a Bland-Altman plot of the proportional error in turn angle against the Vicon measured turn angle [342].

6.3.2 Translational speed

The distribution of the net error in measurement of translational speed appears to be consistent across the range of Vicon measured speeds (Figure 6.8c). Linear regression shows that there is a tendency for lower speeds to be over estimated by VO and higher speeds to be underestimated, with greatest accuracy occurring at approximately $0.64ms^{-1}$, however, this is also coincident with the highest range of error, $\sigma = 0.42$. The confidence in this accuracy is only lost at the extreme end of the translational speed spectrum, $> 1.4ms^{-1}$, where the number of data samples is low (Figure 6.8b).

Comparing the error in translational speed across different trial types, the distribution is relatively consistent around zero. The only exception is ADL3 which appears

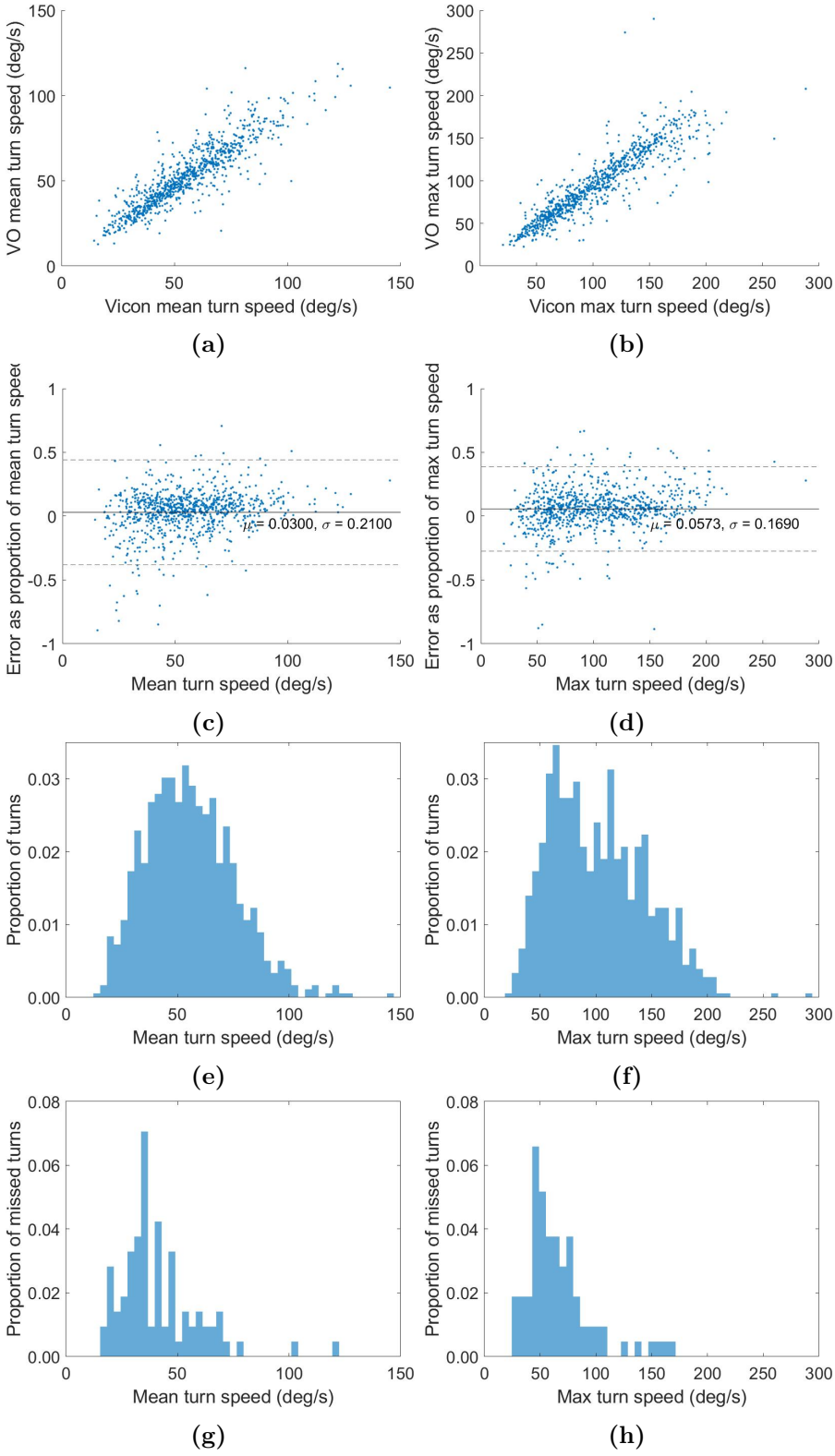


Figure 6.7: Mean and maximum turn speeds. (a) to (d) are correlations and Bland-Altman plots of the mean and maximum turn speed features. The distributions of the mean and maximum turn speeds for all turns and only undetected turns are shown in (e) to (f) and (g) to (h) respectively.

to be more accurate than all other protocols. In accordance with the aforementioned linear dependence of error on speed, the ADL protocols that involve more stationary time and relaxed movements, on average tend to slightly overestimate speed, whereas the Turns and Straight protocols slightly underestimate.

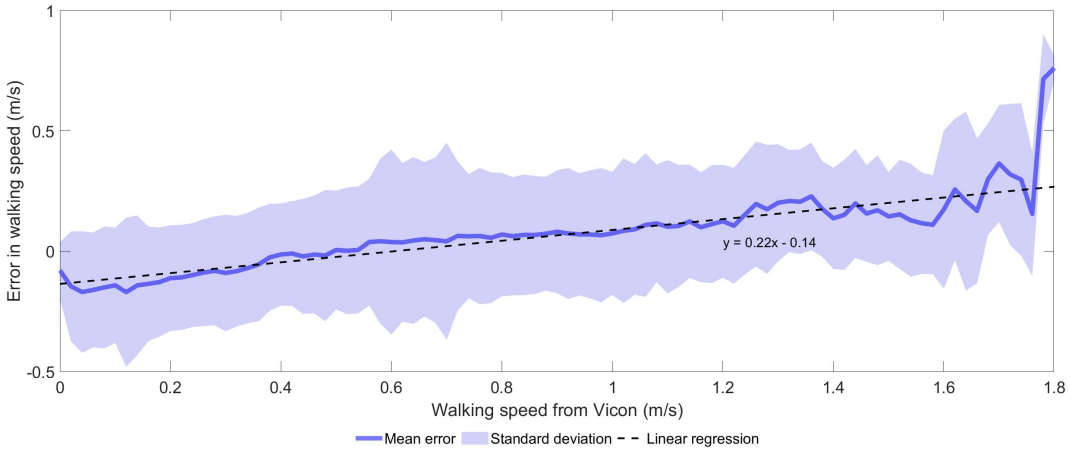
6.4 Discussion

Each of the key features, rotational speed, turns, and translational speed is discussed in turn. Further insight is then provided as to the dependency of the results on the individual subject, and the applicability of the VO methodology to real free-living data. With the methods discussed, the results are then compared to other devices and methodologies.

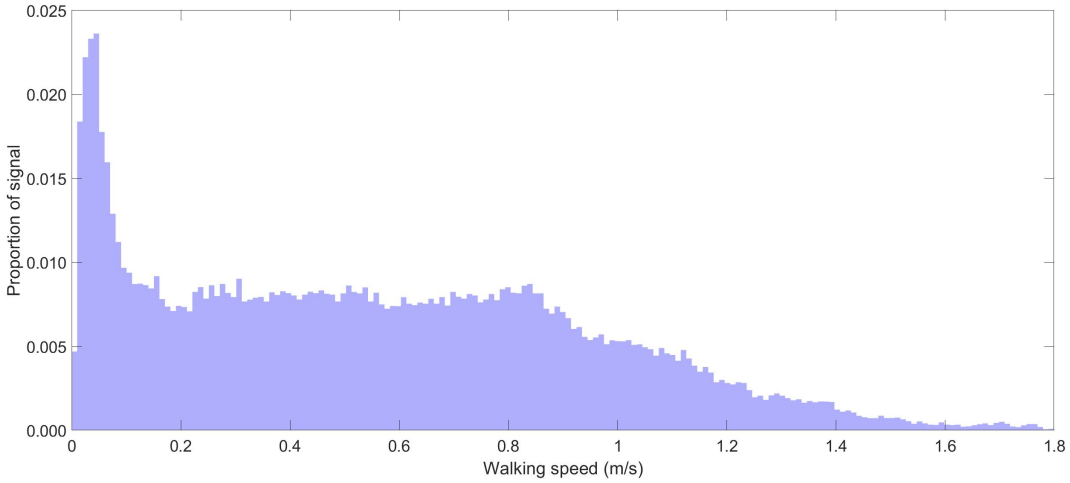
6.4.1 Measuring rotational speed

The results demonstrate that the VO system can measure rotational speed to reasonable accuracy and with limited sensitivity to the rotational speed itself. This will have a limit, as at sufficiently high rotational speed, blurring and the change in orientation of the target image relative to the search area will result in failure of the NCC for template matching. At what point this limit is reached is highly dependent on sampling frequency and can be increased with higher frequencies. However, it is clear from these results that it is not reached during these protocols: at the standard sampling frame rate of 30Hz and even during rapid, relatively unnatural, turns, such as during the Turns protocol. Figure 6.9d shows an example of the TM inliers after MSAC during a rapid about turn during the Turns protocol.

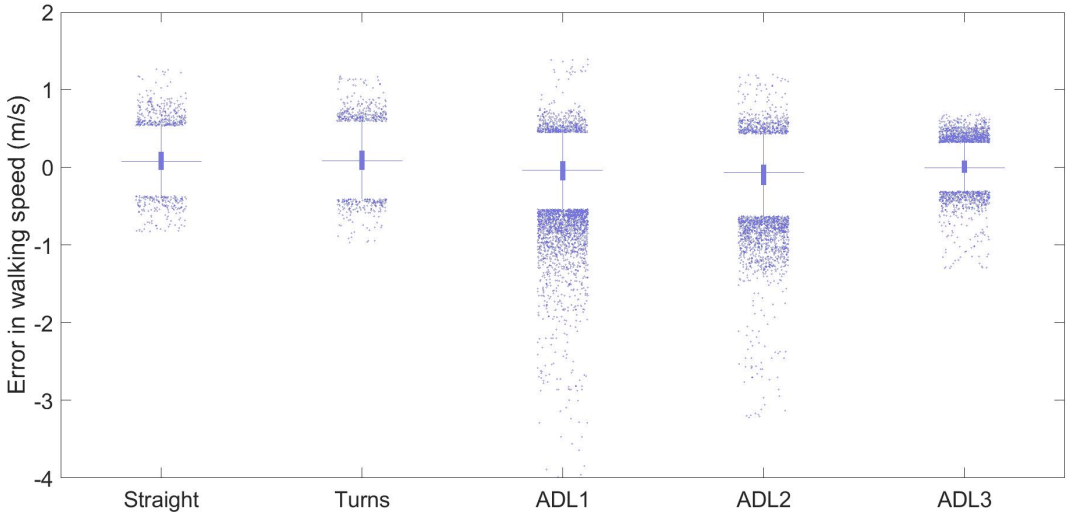
The consistent underestimation of the rotational speed that is demonstrated by gross statistics in Figure 6.10, can be clearly observed in the example shown in Figure 6.9b. Due to its consistency, it is most likely that this is the result of an error in camera calibration. Future improvements could remove this with more rigorous calibration methods or an increase in the number of parameters used. It is this underestimation that causes the larger range of errors in the Turns protocol (Fig. 6.5c); as during this protocol the subject is undergoing almost constant high rotational speeds.



(a) Walking speed accuracy



(b) Walking speed distribution



(c) Distribution of walking speed accuracy by trial

Figure 6.8: Translational speed results. (a) shows the sensitivity of the error in translational speed to its absolute value. Linear regression shows a trend of underestimation below $0.64ms^{-1}$ and overestimation above this threshold. (b) shows the distribution of the translational speeds in bins of $0.02ms^{-1}$, and (c) the distribution of error by different trial type.

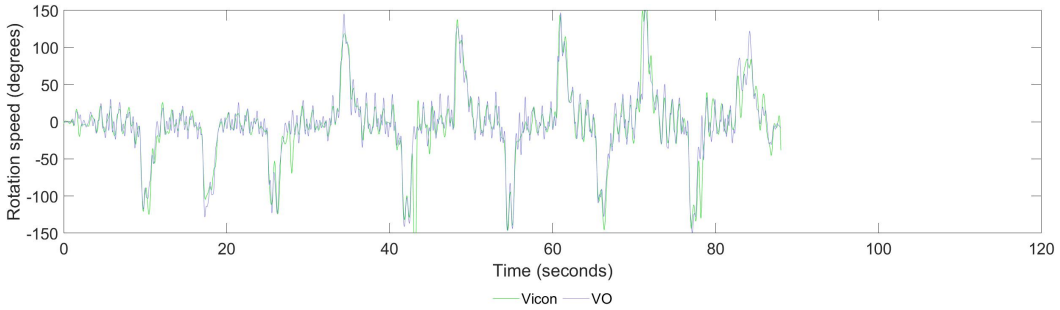
A concern for the further development of the VO system would be if low rotational speeds remained undetected. For example, a standard trunk sway of between 5° and 7° would remain within the bounds of mean error if the VO system simply did not detect this movement [343]. However, the example shown in Figure 6.9a shows that trunk sway is clearly detected during the periods of straight walking.

Overall, it appears that the chief confounding factor is that of view obstruction. This explains the greater distribution of speed accuracy in ADL1 and ADL2 when compared with ADL3 (Figure 6.5c). Furthermore, its effects can be observed in Figure 6.9c, where rotational speed is substantially underestimated as the majority of TMs are located on hands which fill the FOV and move with the body. Figure 6.9e shows an example of this occurring.

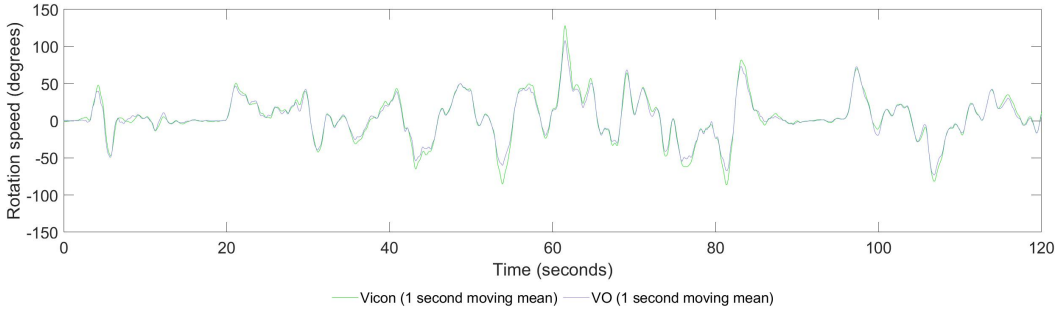
6.4.2 Turn detection & measurement

Overall, turn detection precision is high throughout and in fact is perfect for many trials - including those involving complex movements (Figures 6.10a and 6.10b). The main cause for it not reaching 100% precision appears to be that variability and a slight underestimation in the rotational speed causes the division of turns that are considered a single turn by definition in Section 5.2.2. This results in Vicon turns being split into two separate turns, thus resulting in one false positive. This can be clearly seen in Figure 6.10c. In fact these long turns that are split in the Turns protocol are actually the combination of what would be considered distinct protocol turns. This results in the upper outlier turn angles seen in Figure 6.6, and the substantial reduction in NPV value for protocols where long turns may be construed as the combination of a series of smaller turns (Table 6.2).

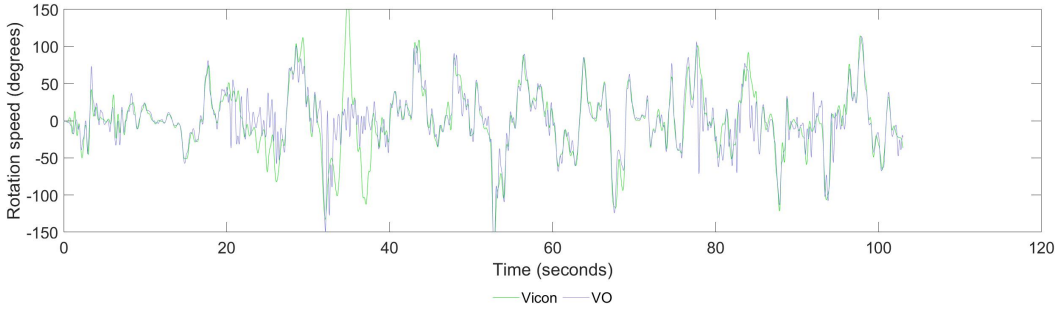
Although the turn detection has high precision across all types of trials. It is recall that appears to suffer and vary depending on the trial type. The fact that recall is lowest in ADL1, where obstruction is highest is unsurprising. However the fact that unlike rotation speed measurement accuracy, ADL3 also suffers, implies that reduced recall is partly due to the general features of everyday living turns, rather than simply view obstruction. Part of this reasoning would be that ADLs would have a greater number of turns around 30°, and thus at the limit, so if turn angle is systematically underestimated



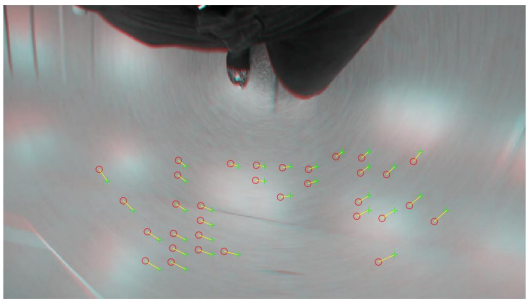
(a) Straight example



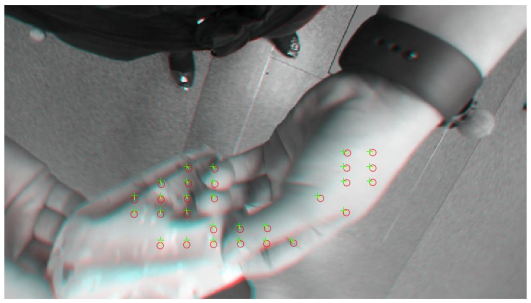
(b) ADL3 example



(c) ADL1 example



(d) TM inliers during rapid rotation



(e) TM inliers during FOV obstruction

Figure 6.9: Examples of rotational speed measurement. (a) shows the distinct signal of straight walking for approximately ten paces, followed by a 180° turn. Positive and negative rotational speeds correspond to left and right turning respectively. (b) demonstrates the lower frequency rotational movement during ADLs, as well as stationary periods. It is averaged over 1 second to clearly show the $\sim 10\%$ underestimation of rotational speed. (c) evidences rapid ADL movement and several examples of failure due to view obstruction. (d) and (e) show the MSAC TM inliers during a rapid turn that induces blurring, and an ADL with FOV obstruction that corresponds to the error at around 36 seconds in (c).

by 10%, turns of 30° to 35° are substantially more likely to remain undetected. This is heavily supported by the predominance of smaller turns that are shown to be those that remained undetected (Fig. 6.6b), the brevity of many of those missed in Figure 6.10d, and that Figures 6.7g and 6.7h show that missed turns are generally slower turns. Slower turns would be more affected by underestimation as a greater proportion of the overall turning process would fall beneath the threshold rotational speed.

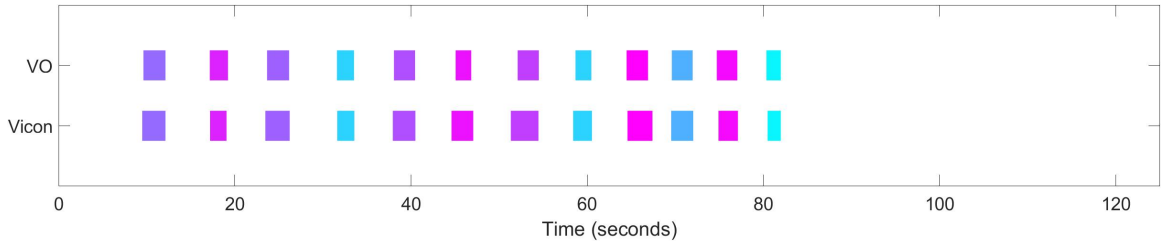
Overall, the VO turn detection method compares favourably the the most commonly utilised lumbar IMU method; with a sensitivity of 0.76 and specificity of 0.65 in healthy control subjects even during a prescribed turning exercises most comparable to the Turns protocol [221].

The accuracy of the VO system's capacity to measure turn angle appears to be largely independent of the length or speed of the turn. Therefore the main cause of error appears to be the consistent underestimation of the rotational speed by approximately 10% (Fig. 6.6d). However, there is a clear somewhat normal distribution of error on turn angle, with slight propensity for more extreme proportional error at shorter turn angles. One cause for this may be that unlike many previous studies, the turn angle ground truth measurement is made in reference to the subject's stance before and after the turn, and not a predefined track. Therefore, if the turn is determined to start and finish at different times for the different measurement systems (as can be seen in Fig. 6.10), then the turn will include different proportions of the trunk sway before and after a turn. With a mean trunk sway of approximately 6° in either direction, this timing difference alone could result in an absolute turn angle measurement error of 12°.

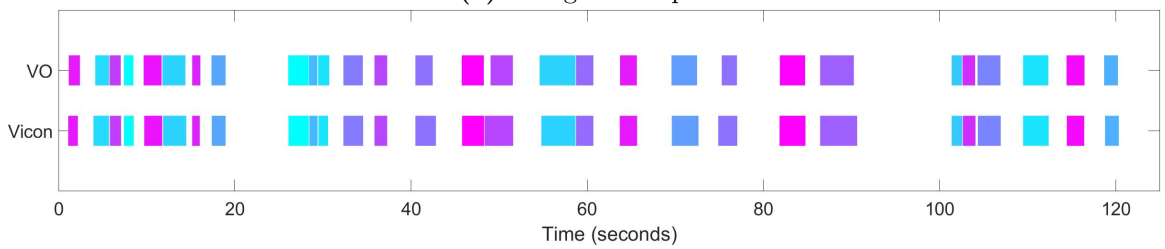
An interesting result is that of the accuracies of the mean and max turn speed features. Despite the consistent underestimation of rotational speed being evidenced in a range of results, this seems to have less of an effect on the turn features, where the mean proportional error is much closer to zero, although the standard deviation in proportional error is similar at around 18%.

When considering the accuracy of the turn features, it is also worth noting that there is reasonable uncertainty on the ground truth Vicon signal. This is mainly due to systematic uncertainty in the locations of the markers (Section 4.3). However, there are also instances of incorrect or ghost markers. For example, the sudden spikes in

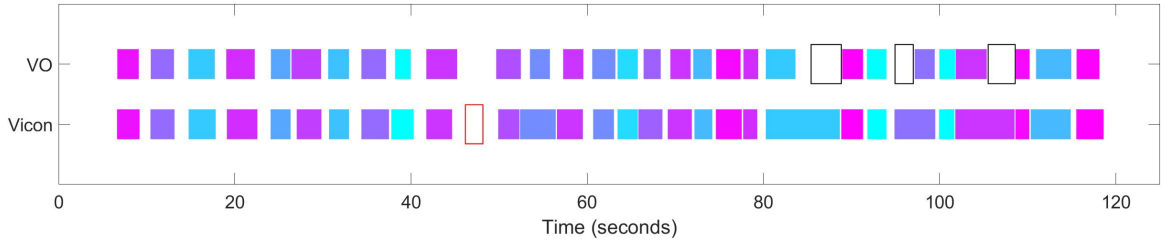
rotational speed to the right observed at approximately 43 and 77 seconds in Figure 6.9a are highly unlikely to be the actual true movements of the subject. It is therefore very possible that the smoother, more physiologically realistic, signals of the VO system at these moments are closer to the true subject movement.



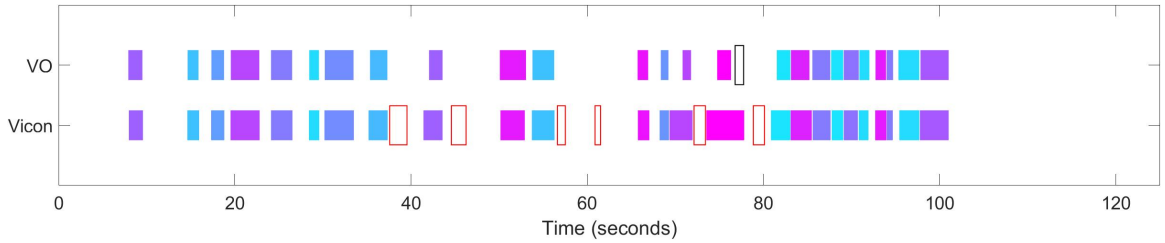
(a) Straight example



(b) ADL3 example

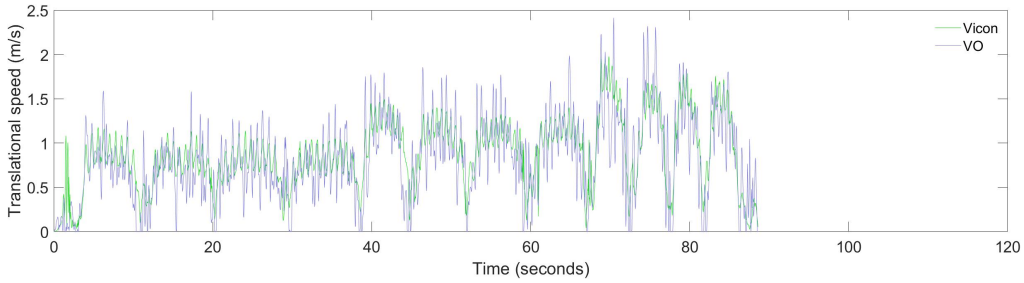


(c) Turns example

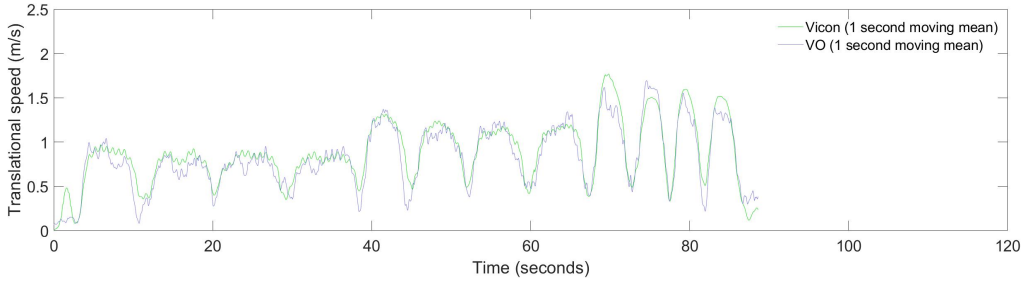


(d) ADL1 example

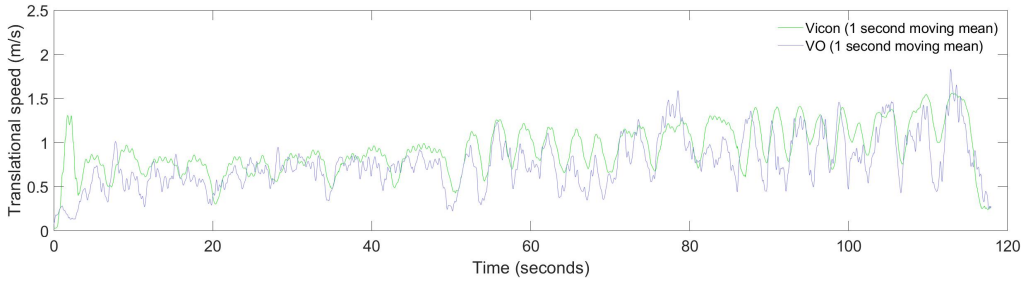
Figure 6.10: GET graphs for turn detection examples (Section 3.2.3.1). (a) and (b) show 100% binary turn detection agreement between Vicon and VO, with 85% and 92% accuracy in temporal alignment respectively, demonstrating that turn detection can be highly accurate with both clearly defined and complex turns. (c) demonstrates several examples of false positives as a result of long turning periods being split into two separate turns by the VO system. (d) shows examples of several temporally short false negative turns



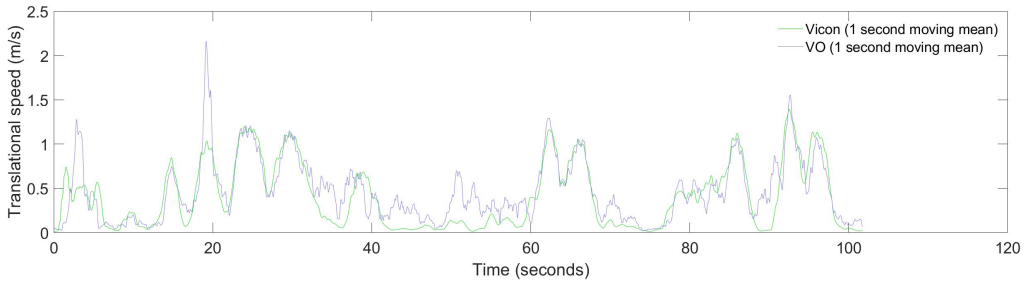
(a) Straight example (no averaging)



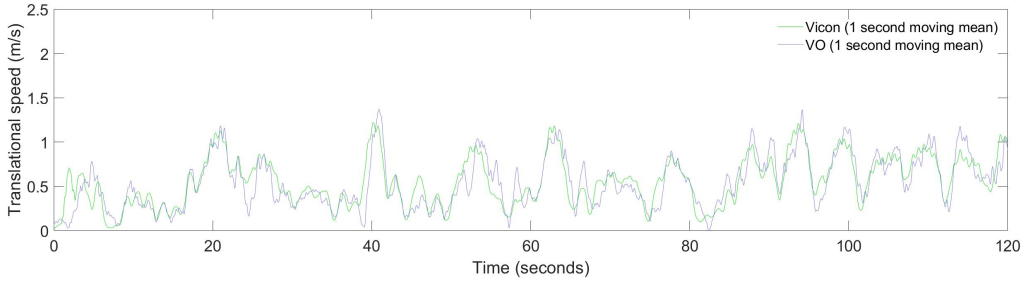
(b) Straight example



(c) Turns example



(d) ADL2 example



(e) ADL3 example

Figure 6.11: Examples of translational speed measurement. (a) shows an example before a moving average is applied. (b) to (e) show examples of various protocols and a range of discussed problems that affect translational speed measurement accuracy.

6.4.3 Translational speed

In general, the translational speed is shown to be measurable across a range of speeds and at reasonable accuracy during both simple and complex activities (Figures 6.11b and 6.11e). However, it has substantially greater inaccuracy than the rotational speed. This can be most clearly seen by direct comparison of the single trial examples in Figures 6.9 and 6.12. This is generally unsurprising as the rotation is extracted directly from the homography decomposition (Section 3.3.3.1), whereas the translation is a function of both the normalised translation and the scaling factor, the latter of which is itself a function of the various components of homography decomposition, each one of which harbours error.

The sensitivity of the translation to these variables is in part exhibited by the extremes in fluctuation of the signal as shown in Figure 6.11a. Here the oscillating variation in the subject's translational speed while walking in short straight bouts is observed, but magnified by measurement. The fact that this oscillation is observed is alone interesting, as it suggests that the system could be used to extract more complex gait speed features.

It is hypothesised that a substantial proportion of the error in translational speed measurement is caused by the relatively erratic nature of the movement in each of the trials. Figure 6.11b shows that there is greater consistency during the brief periods (2-6 steps) of consistent gait speed. This is supported by the consistent absolute error in speed measurement (Figure 6.8a). This means that the proportional error is substantially lower at higher speeds. These speeds correspond to the predominantly faster activities in the Straight and Turns protocols (Figure 6.12). This hypothesis is explored further in Section 6.4.5.

It is shown that low translational speeds are consistently overestimated and higher speeds underestimated (Figure 6.8a). The problem at low speeds can be readily seen during the stationary periods in the ADL2 example (Figure 6.11d). This is most likely caused by the compounding of the scale uncertainty, which is higher at very low speeds as the relative motion is very low and so errors are proportionally high. This could potentially be reduced by the implementation of some degree of keyframing and associated bundle adjustment. This adaptation would be especially effective

at reducing the error at low translational speeds - where the highest proportional error occurs and accrues.

The distribution of translational speeds in Figure 6.8b also shows that very little of the time is spent within $0.00 - 0.02ms^{-1}$, implying that subjects are constantly moving. This is likely due to postural sway as part of postural control. As such, it is possible that the VO system could be used to measure features of postural sway without cumbersome equipment such as Lord's sway meter or force plates [344, 345]. This peak in translational speed distribution in the range $0.02 - 0.10ms^{-1}$ is a little high to correspond directly with recorded sway speeds, which generally average at around $0.02ms^{-1}$, however in postural sway studies, the subjects are asked to deliberately stand upright and still [345–347]. Whereas in this study's protocols individuals stand naturally, converse, shift position and adjust posture constantly - much like in everyday life. This would increase the associated translational speeds [348].

Similar to rotational speed, it appears that a greater amount of extreme error is seen in ADL1 and ADL2 (Figure 6.8c). This is also likely a result of FOV obstruction. An example of this occurring can be seen in the erroneous spike at 19 seconds in the ADL2 examples (Figure 6.11d). It is also interesting that unlike with rotational speed, ADL3 exhibits a lower distribution in error than the other protocols. The most likely reason being that subjects tended to walk longer distances than in ADL1 and ADL2 (with less view obstruction), enabling stability of the speed to be attained, but at a slower average pace than the rapid speeds encouraged by the Straight and Turns protocols. FOV obstruction is not quantified throughout the trials as what constitutes an obstruction is wide and varied; such as the subject's hands, the subject's lower body (to varying amounts depending on stance and movement), objects and obstructions, furniture, etc. It would be too complicated to practically assess all of these obstructions, and even then the measurement would not be entirely meaningful as the number of template matching incidences obscured does not correspond with the accuracy of pose and scale estimation; as only five exact measurements are theoretically needed to estimate accurately (Section 3.3).

There is very little work on the instantaneous speed measured, and instead on average speed over long periods of linear walking. Even then however, speed and

distance measurements have a comparable, and often worse, error of between 5% and 20% [349]. Although there is room for improving the translational speed measurement, the results compare favourably to other reported methods - even though these are often measured during linear, stable, walking bouts [350].

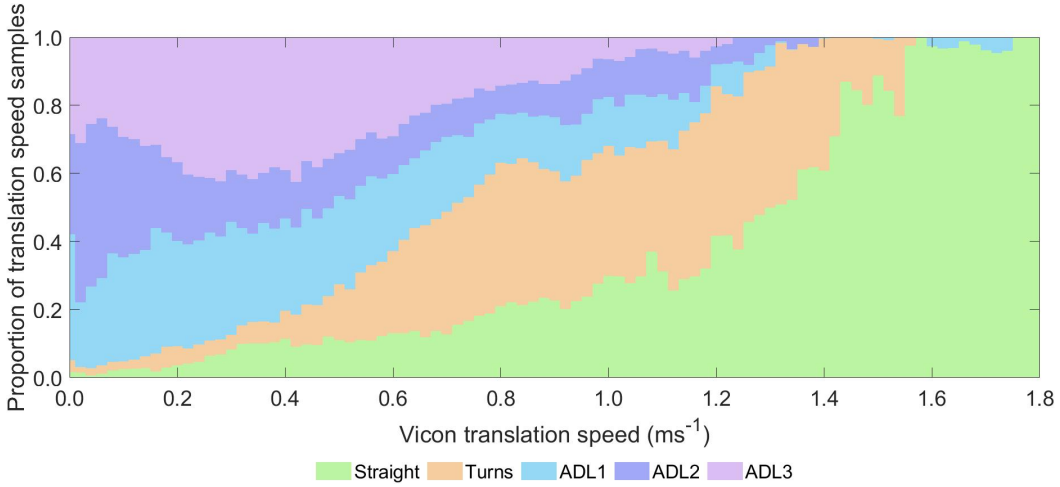
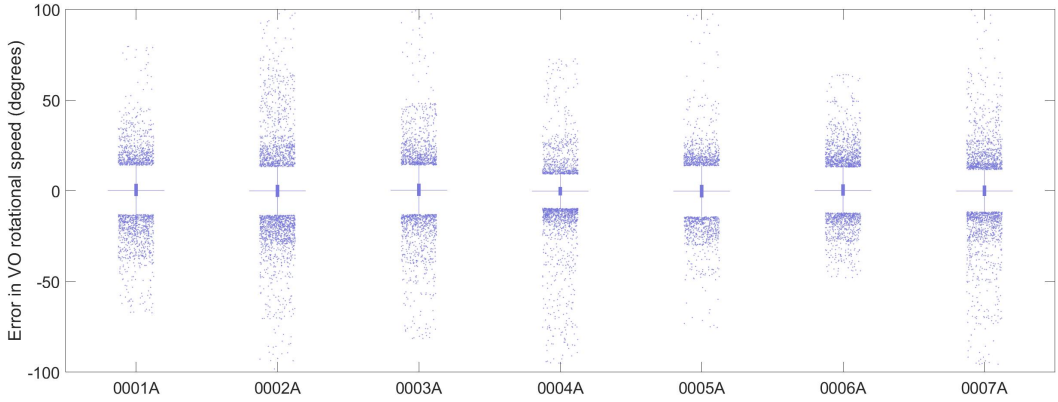


Figure 6.12: Relative distribution of translational speed by protocol; bins of $0.02ms^{-1}$. The Straight protocol shows hegemony over the higher translational speeds. Similarly Turns follows as although fast speeds were encouraged, constant turning limits the maximum speeds reached. ADL3 is dominant in consistent lower speeds, whereas ADL1 and ADL2 evidence periods of being stationary combined with mid-speed movements.

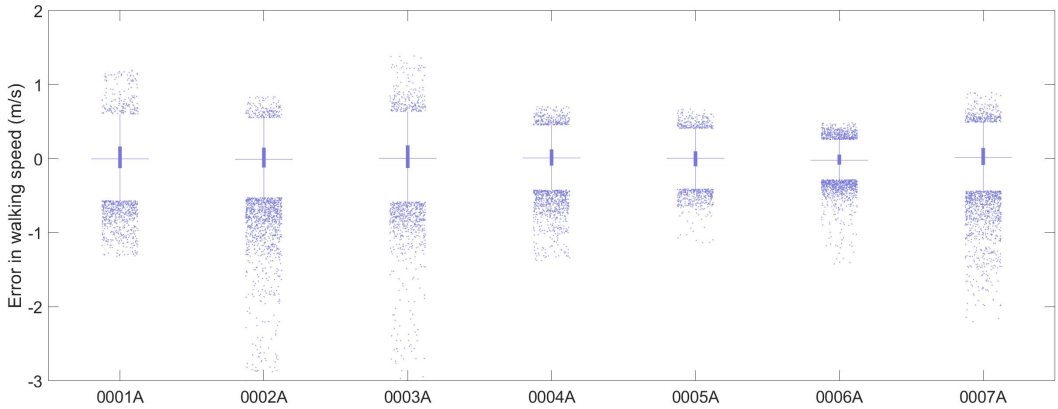
6.4.4 Subject dependency

Unlike rotational speed, due to the required input of measured camera height in order to calculate the scaling factor (Section 3.3.3.4), the accuracy of the translational speed varies between subject. For example, Figure 6.11c demonstrates consistent underestimation of the translational speed due to either error in the camera height measurement, or the camera height changing at some point during the study session. Furthermore, it is assumed that the camera height remains constant, although this is not true during subject movement. Therefore, subjects who have a greater range of vertical motion, or 'bobbing', during gait are likely to have less consistent translational speed measurements [351]. This in part explains the increased variability of translational speed accuracy between subjects as compared with rotational speed (Figure 6.13). An additional cause of inter-subject variablity would be the erraticness of subject movement,

whether sharp and energetic or smooth and calm. It remains interesting, but unclear, why a single subject (0004A) has significantly more accurate rotational speed results.



(a) Rotational speed accuracy



(b) Translational speed accuracy

Figure 6.13: Rotational and translational speed accuracy by subject.

6.4.5 Translational speed during linear walking

Despite the clear problems that ADL movements have on movement speed estimation, submitting the VO system to these erratic movements is important to the robustness of the technology in a free-living environment. However, the vast majority of comparable studies undertake periods of linear walking at constant speed and then compare the measured average speed with that of the walking bout. This type of data collection session was mimicked by walking the 400m in a single lap around an athletics track. As can be seen in Figure 6.14, the VO system works well in a relatively linear, steady-state walking environment. Although there is no motion capture ground truth for comparison, it seems likely from Figure 6.14a that it is actually picking up variations in walking speed, both throughout the recording and intra-step. A spatial ground truth can be

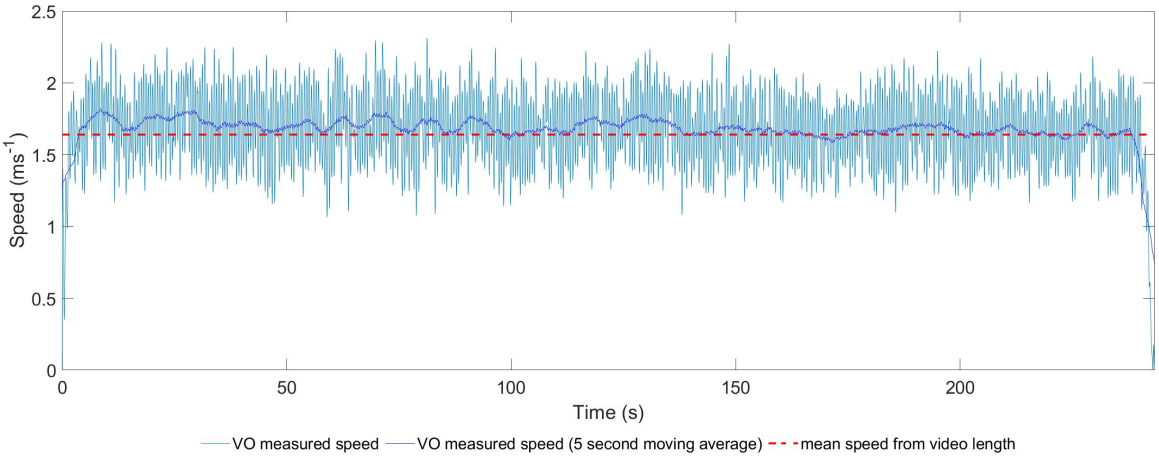
obtained however from satellite imagery. If the translational and rotational signals are combined, then a 3D path map can be created and compared with the known route. Figure 6.14b shows this for the 400m track. Drifting can be seen to occur up to 10 metres, however the majority of the time it keeps below 2 metres. Using satellite imagery to assess spatial accuracy will not work on local activities, such as most ADLs, due to: the spatial resolution being variable and between 2 and 10 metres, some locations being completely GPS denied, and the considerable GPS scatter in urban environments [352]. However, for movement over large outdoor areas it is sufficiently accurate and provides valuable insight into small amounts of drift that may accrue over time and distance.

6.5 Conclusion

In this chapter, a combined approach to visual odometry was implemented in order to extract the egomotion of a wearable camera and thus the person wearing it. Firstly, this allows the direct measurement of a range of gait features including movement speed, rotation speed, turn detection, turn angle, mean turn speed and maximum turn speed. Accuracy in these measurements is important, as these signals will then be used as part of the calculation of more complex step features; as covered by Chapter 7.

Promising results were observed for rotation, turns and linear walking. The translational speed during more complex movements, such as those involved in activities of daily living, was less reliable and could do with further improvement. These results compare favourably with comparable devices when accounting for the limitations of their validation studies. A fuller understanding is gleaned from the work in Chapter 8, where such alternative methods are tested under similar conditions.

The method was developed to account for a wide range of unique challenges and problems that are not necessarily common in the field of visual odometry. However, confronting these challenges was necessary to extract signals from data gathered in the way outlined in Chapter 4. One of the chief issues was occlusion due to movement and activity of the upper limbs. It is possible therefore that the method would be more successful in a slightly constrained free-living environment. As the biggest source of error is the translational scaling calculation, one such constraint might be the reduction of occlusion of the camera view by arm movements, perhaps focusing on more gait-heavy



(a)



(b)

Figure 6.14: Use of VO for steady-state gait around a running track. (a) suggests that the system provides a relatively accurate measure of walking speed during substantial periods of steady-state gait. In fact, the fluctuations detected are likely indicative of slight changes in speed and variation in pace across the gait cycle. (b) shows the route as mapped by combining the translation and rotation signals. This has been projected onto satellite imagery of the running track from Google Maps [353]. The track was walked anticlockwise from and to the same point, mainly on the inside lane. Green arrow indicates VO start, red, the finish. Overall, the mapping is within 2 metres, although some drifting occurs during the bends. It is positive that the path starts and ends within 2 metres of the same location. Similarly, half way down the first straight, an individual walking in the inside lane was overtaken and the second lane was used for the next 50m; this is also seemingly picked up.

activities than, for example, food preparation. Similarly, larger areas of activity with less erratic movement would also lead to more accurate results.

7

Foot tracking & gait features from a wearable camera

Contents

7.1 Introduction 141
7.2 Methods 151
7.3 Results 159
7.4 Discussion 163
7.5 Conclusion 169

7.1 Introduction

Cameras are a common device used in the assessment and measurement of gait; from the array of high resolution cameras used in the Vicon motion capture to the use of the consumer depth cameras such as those in the Microsoft Kinect [354]. Key benefits of using cameras are that they provide rich data, including context, 3D location and trajectories, and ground truth for binary features such as step count and walking bout detection.

However, the vast majority of cameras used in gait analysis are static third-person perspective cameras, either in a research facility or in the home. These suffer from only being able to capture gait bouts within a very limited space and environment. Additionally, there must be a clear line of sight between the static camera and the subject's legs and feet - which is unlikely in a realistic free-living environment [355, 356].

The increasing ubiquity of consumer cameras in the form of smartphone and action cameras means that these have been used increasingly to provide ground truth and contextual information for activity and gait recognition [14, 357]. However, actually

extracting spatio-temporal gait features from a wearable or portable camera is still a nascent field. There have been a small number of tentative attempts to use smartphone cameras or wearable action cameras to extract step features [358], detect steps [103], or improve the gait features extracted from accelerometers [102]. In one case a camera was attached to an automatic guided vehicle to follow the subject around and provide gait monitoring - which itself adds further complications, expense, and limitations to the method [359]. The unifying characteristic of these studies is that they are all tested using short bouts of linear gait; for example walking eight metres in a straight line at a constant speed [358]. Although they report good results for step feature extraction, validating solely in this way is contrary to the stated purpose of building a technology that utilises wearables to bring gait analysis to the free-living environment. Furthermore, the evidence presented throughout this thesis shows that free-living gait is considerably divergent from linear gait, and this difference corresponds to a substantial increase in difficulty of gait feature measurement.

In order to extract spatio-temporal gait features from videos of the feet, the location of the feet within the video frames must first be established. Following objects through videos is the goal of the object tracking field of computer vision. Object tracking is a burgeoning area of research due to its vast array of commercial and research applications. These include but are in no way limited to tracking: cars and pedestrians for traffic monitoring and autonomous vehicle guidance [360, 361]; individuals for activity recognition, security surveillance, and elderly fall detection [362, 363]; animals for ecological, zoological, behavioural and conservation research [364–368]; the user and their surroundings in virtual and augmented reality [369]; and an array of ordinance and targets in defense systems [370–372]. Like the other fields of computer vision, rapid recent progress is in part due to increasing computing power, camera quality, price, cultural adoption and general ubiquity.

The diversity in application is reflected in the plethora of methodologies for approaching the problem from broad scope to specific applications. It is therefore important to consider the application and associated data sets before developing a method. Therefore a thorough overview of the challenges is provided, both in general and with regards to this specific type of data. These are explored, their extent probed

using both the data sets, and the information used to inform development of the methods. The Newcastle data set is used to help inform the developed methods, and ensure that they are robust and applicable to free-living data.

After these challenges are explored, an explanation is given of the methods developed to track the feet and extract the gait features of stride length, step length, step width, and step count from wearable camera video. The results from the Oxford study are then provided along with critical analysis and a discussion on the accuracy of the method. Finally, recommendations for further improvements and conclusions are given.

7.1.1 Object tracking challenges

The success of an object tracking system is dependent on a range of parameters regarding the objects being tracked, the environment, and the camera being used. *Wu, Lim and Yang* identify 11 types of challenges that contribute to the complexity of a object tracking task [373], and the definitions for these are provided in Table 7.1.

The relative importance of each attribute is somewhat reflected by its prevalence in benchmark data sets and the relative use of these data sets. *Wu, Lim and Yang* also collated the most commonly used tracking sequences to establish their benchmark. The proportional representation of each attribute in that data set is shown in Fig. 7.1. The standard data sets are however developed to represent common use cases of object tracking - such as in autonomous vehicles. Therefore these attributes may not be representative of those relevant to wearable cameras. To account for this, the Oxford and Newcastle data sets were assessed for the importance of each attribute, and this is also shown in Figure 7.1. As can be seen, these are quite different from the standard data sets, and this is explained and justified in Sections 7.1.1.1 to 7.1.1.5, using the quantitative analysis provided in Figure 7.2 and the qualitative examples shown in Figures 7.3 and 7.4. It is worth noting that the variations in camera intrinsics and positioning may slightly affect the direct comparability of the two data sets (for example a more distorted lens making images of the foot smaller at some points in the FOV); however it was qualitatively observed that in general there was similarity and each of the following points holds in generality. Overall, the fundamental importance of these challenges is that they will affect and inform the methods developed in Section 7.2.

Attribute	Acronym	Definition
Illumination Variation	IV	The degree of change in the illumination of the object.
Scale Variation	SV	The ratio of the object size as represented by bounding box encompassing it in the frame of its first instance against the current frame, is outside of the range $[1/t_s, t_s], t_s > 1 (t_s = 2)]$.
Occlusion	OCC	The partial or full occlusion of the object.
Deformation	DEF	The non-rigid deformation of the object.
Motion Blur	MB	The object is blurred as a result of the movement of the object, the camera, or both.
Fast Motion	FM	The motion of the object relative to the camera is greater than t_m pixels ($t_m = 20$)
In-Plane Rotation	IPR	The rotation of the object within the image plane.
Out-of-Plane Rotation	OPR	The rotation of the object out of the image plane.
Out-of-View	OV	A proportion of the object leaves the field of view of the camera.
Background Clutter	BC	The background of the scene near the object has similarity to the object, or contains other objects that may be mistaken for the object.
Low Resolution	LR	The bounding box representing the object has fewer pixels than t_r, where $t_r = 400$)

Table 7.1: Object tracking attributes. The eleven attributes that contribute to the complexity of an object tracking task, as proposed by *Wu, Lim and Yang*[373]. This delineation and nomenclature has seen general uptake [374, 375]. Several of the attributes contain proposed values for thresholding.

7.1.1.1 Illumination, resolution & scale variation

The effect of illumination on object detection is highly dependent on the methodology that is used. Furthermore, both the absolute degree of illumination and its variation can have significant effects. The illumination of the feet in real-world applications can vary dramatically, both due to the occlusion of the light source and the ambient environment. Figure 7.2a gives the distribution of mean frame brightness in both the Oxford and Newcastle data sets. As the Oxford data set is collected exclusively in a well lit motion capture lab, the average brightness is higher and more consistent. The often stark variation in illumination in a free-living data set can be seen by comparing Figures 7.3a and 7.3b which are only 0.1 seconds apart. Figures 7.3a and 7.3c demonstrate how difficult it can be to view the feet in deep shadow and low ambient light.

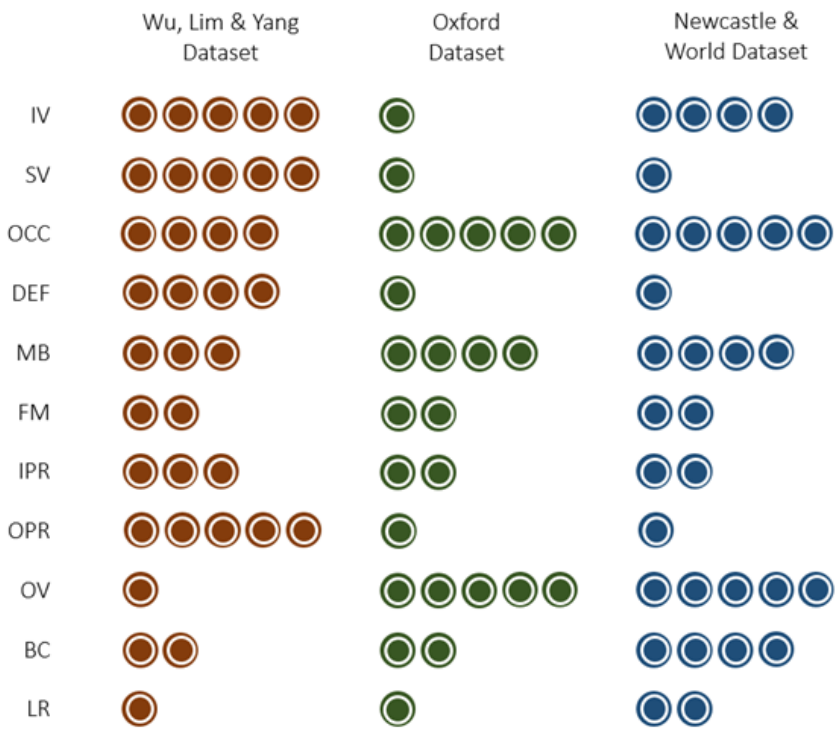


Figure 7.1: Object tracking attribute prevalence and importance. The prevalence of the object tracking attributes outlined in Table 7.1 in common data sets is shown as interpolated from *Wu, Lim and Yang 2013.* [373]. This is compared to the qualitatively determined importance of the attributes in both the Oxford data set, and the Newcastle data set (as representative of true free-living data).

Good illumination can also be problematic however, as can be seen in Figures 7.3d through 7.3f where the lighting causes shine and loss of information and texture. Illumination effects such as shine and shadows (Figures 7.3d through 7.3f) are highly problematic when trying to track movement as object tracking methods may end up tracking the movement of the shadow or shine, rather than the object itself. This is especially true when the object being tracked has weak discriminating features (as is the case with low textured or blurred objects), when compared to the high contrast edges caused by, for example, shadow (Figure 7.3e). Shadows can also cause inconsistent illumination even within a single frame (Figure 7.3d).

Resolution is generally a function of the device being used. The cameras used in the Oxford and Newcastle data sets both utilise 720p resolution. Figure 7.2b shows the distribution of the size of the bounding boxes for the feet in both data sets. As expected, the distributions are similar, as the foot remains relatively equidistant from

the chest. This varies between individuals although the variation will be limited, and furthermore, taller people generally have larger feet [376]. Additionally, the foot's size in the frame will remain relatively consistent when viewed from the chest-mounted camera as the feet predominantly move in a plane parallel to the camera's image plane. It is clear that the foot is always above the resolution threshold.

Following this logic, it may be surprising that there is so much variation in foot detection size. The true cause of this, as evidenced by Figure 7.3g, is partial occlusion of the foot rather than any actual variation in size. As such, scale variation is one of the biggest differences between those data sets used in developing object tracking models and those used in this work. A vehicle or pedestrian may be visible to an autonomous vehicle anywhere between the far distance and the immediate foreground, the feet remain at a reasonably consistent distance throughout.

7.1.1.2 Rotation & deformation

A key benefit to the data is that the rotation of the feet is highly constrained - the view is generally from immediately above with the foot pointing forwards. Figure 7.2c demonstrates the very limited variation in foot rotation in the Oxford data set. By far the greatest component of change in orientation is in-plane rotation caused by twisting of the torso; this can be seen towards its greatest extent in Figure 7.3h. The main component of out-of-plane rotation will be due to foot pitching caused by ankle plantarflexion and dorsiflexion during walking. Even then, this is greatest at push off and toe off which does not occur within the FOV anyway. Although the out-of-plane values in Figure 7.2c do not account for foot roll, this would be negligible in all but the most extreme cases of supination and pronation [377, 378].

As the shod foot is a semi-rigid entity, it undergoes very little deformation that affects its overall appearance. Similarly to out-of-plane rotation, what little deformation that occurs, does so during push-off which occurs outside of the FOV.

7.1.1.3 Occlusion & out-of-view

The feet not being within the FOV, or being occluded by other objects within the FOV is the greatest concern when developing a foot tracking system with this kind of data.

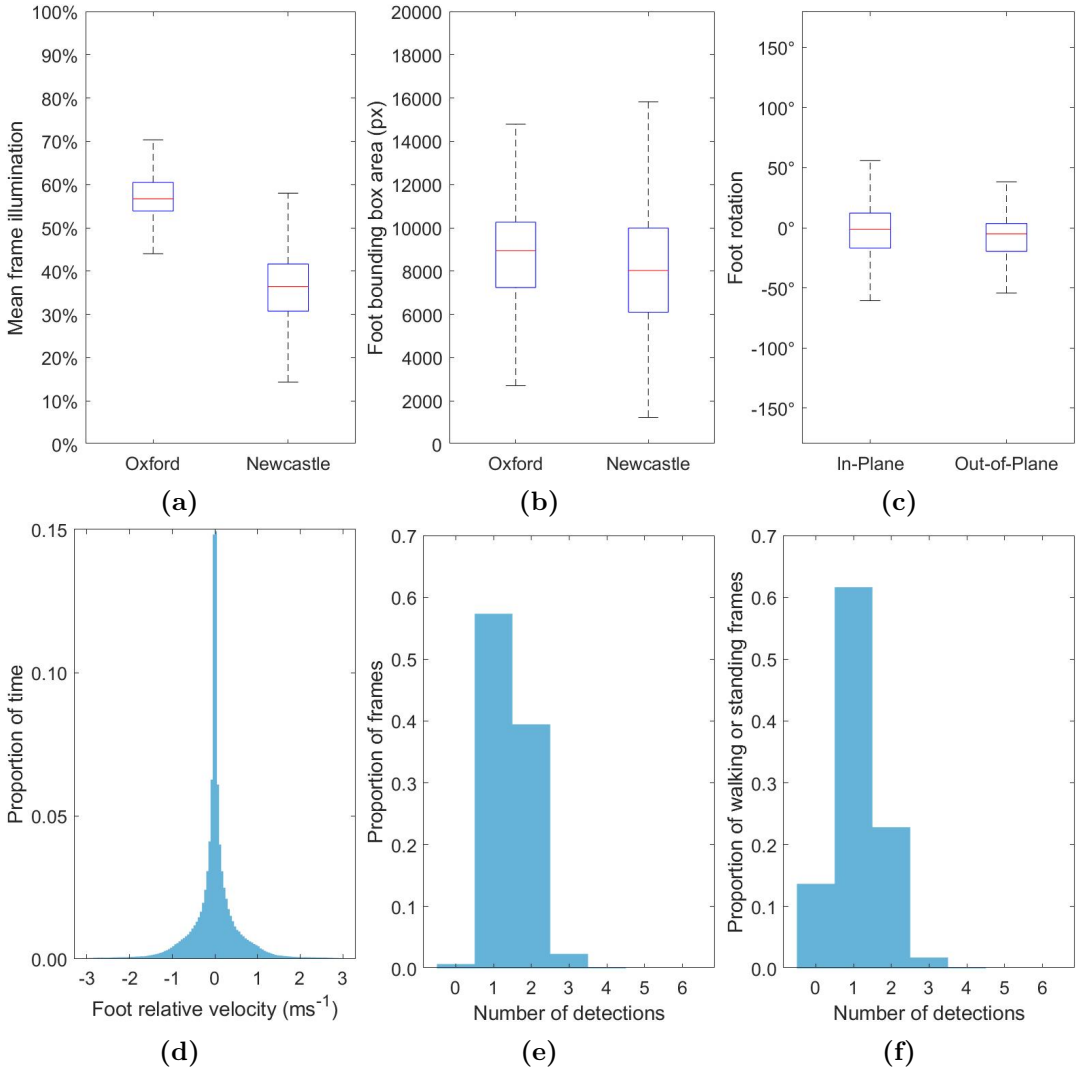


Figure 7.2: Analysis of the object tracking attributes in the data sets. (a) distributions of the brightness in both the data sets as a percentage of the maximum total brightness. (b) distributions of the bounding box area for all detections identified by the wearable camera system as either left or right foot detections (Section 7.2.2.2). (c) distributions of the in-plane and out-of-plane rotation of the feet relative to camera in the Oxford data set as extracted from Vicon. In-plane rotation is defined as the angle in the image plane between the foot vector (determined to be the vector between the foot's heel marker and it's toe marker), and the camera image plane's y-axis. Out-of-plane rotation is defined as the angle between the foot vector and the camera y-axis, in the camera's y-z plane. This does not account for foot roll, which is likely negligible. Furthermore, to accentuate the variation in out-of-plane rotation, only samples where the foot is moving are considered (as while the foot is on the ground, its out-of-plane rotation is negligible). This filters out 51% of the signal. If the stationary foot time is included, the mean out-of-plane rotation is $-5.2 \pm 19.9°$. (d) distribution of the foot velocity in the camera's reference frame; bins of $0.05ms^{-1}$. (e) and (f) distributions of unfiltered foot detections in the Oxford and Newcastle data sets respectively. Detections were considered above a confidence score of 0.9 (Section 7.2.1).

As mentioned in Section 7.1.1.1; it is the greatest cause for variation in bounding box size. An individual foot will actually be outside of the FOV for a substantial proportion,

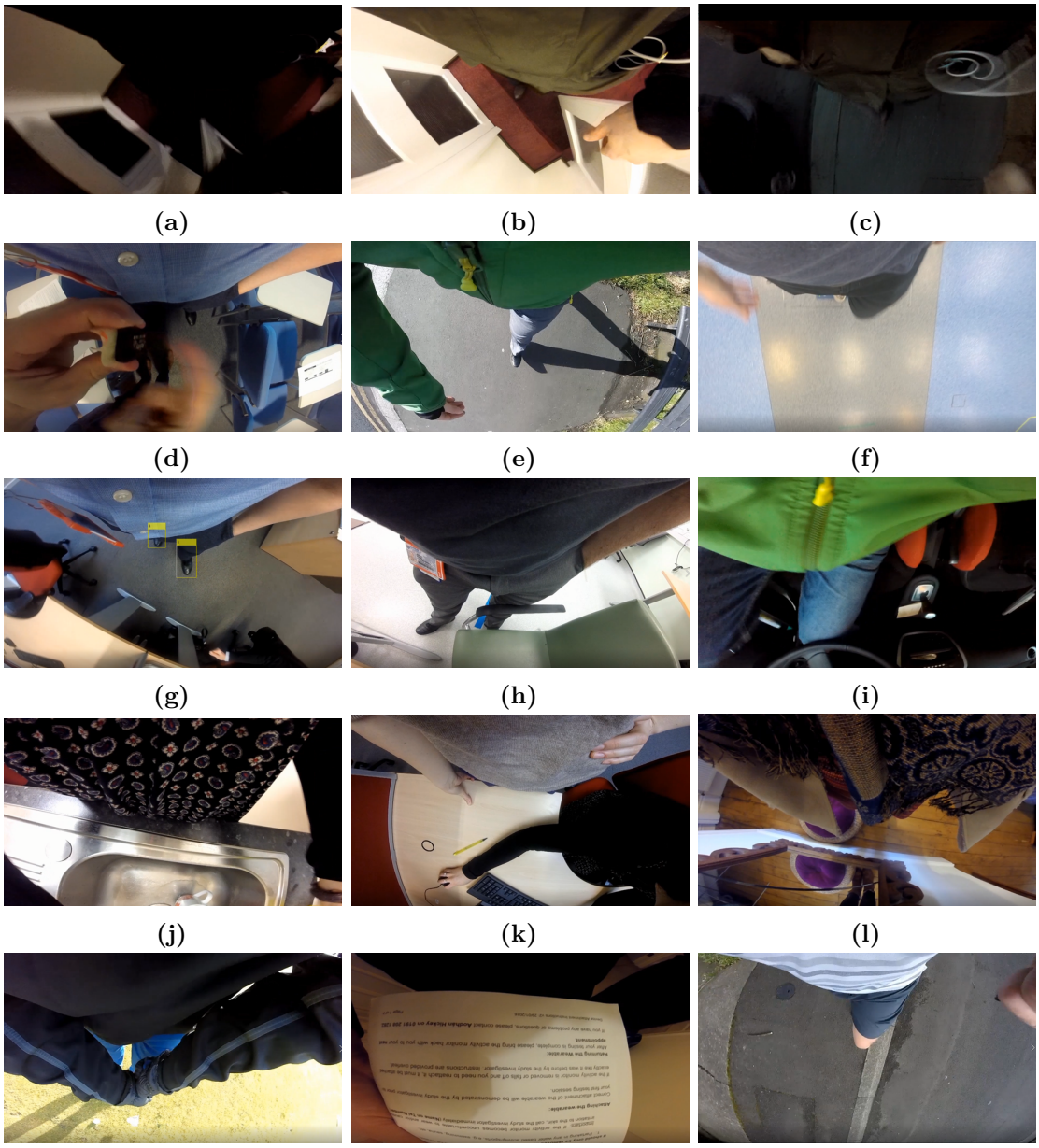


Figure 7.3: Object tracking attributes example frames. (a) and (b) demonstrate the sudden changes in illumination that can occur in a free-living environment. These frames are less than a tenth of a second apart. (a) exemplifies extreme shadow, whereas (c) very low ambient light. Both make detecting the feet incredibly difficult, even to the human observer. (d) exemplifies how illumination can vary in a single frame with one foot in bright light and the other complete shadow. (e) distinct foot-shaped shadow could potentially be mistaken for a foot (f) shine in the background could be mistaken as an object by a computer model, wheres shine on the desired object itself; such as the feet in (d) and (e) could cause a failure of foot detection. (g) demonstrates how bounding boxes, and thus perceived foot size, can vary due to occlusion. (h) In-plane rotation of the foot greater than 45°. (i) out-of-view due to sitting. (j-k) occlusion by work surface. (l-o) Occlusion by loose clothes, hands, held everyday objects, and the knee while running.

if not majority, of the time; as can be seen by the total detection counts per frame in Figures 7.2e and 7.2f. Although these figures include false positive detections, they are representative of the fact that during the stance phase of gait, the foot disappears from view as it passes under the body (as can be seen in many of the frames in Figure 7.3). This is also true in a range of standing stances. In free-living data this is even more of a concern as the majority of the time the subject is likely to be sitting or lying (Figure 7.3i). In the case of the Newcastle data set, this represented 70.6% of the data, and therefore Figure 7.2f only represents the other 29.4%. When compared to the Oxford data, a noticeable increase in frames with no detections is observed. This is due to the increased time spent with work surfaces or objects occluding the feet in the FOV. Standard examples can be seen in Figures 7.3j and 7.3k. A wide range of other objects can occlude the view of the feet, these include: loose clothing (Figure 7.3l); the subject's hands in an array of activities, swinging while walking, or gesticulating while talking (Figure 7.3m); held items that are interacted with (Figure 7.3n); and even the lower limbs themselves - the knee obscures the foot during walking, but this is exacerbated during running (Figure 7.3o).

Combined, the feet being occluded or out-of-view mean that rarely can they be tracked for more than ten to twenty frames. This is a considerable concern for robust tracking, discerning the left foot from the right foot, and distinguishing false positives from true positives. During linear gait, the feet pass in front of the camera in a moderately consistent pattern. However, building rules or using predictive models that are common in object tracking, such as Kalman filters, on this pattern would obfuscate the tracking during complex and erratic ADL movement [379]. For example, although the foot will generally enter the frame at the centre top, the torso may suddenly rotate or lean (Figure 7.4d), an occluding object may suddenly be removed, or the subject's stance may alter, or camera's position and orientation may change.

7.1.1.4 Fast motion & blur

Blurring is a key consideration with this type of data. Although the foot rarely travels particularly fast with respect to the camera (Figure 7.2d), the majority of its velocity component is parallel with the image plane, and at a reasonably close proximity.

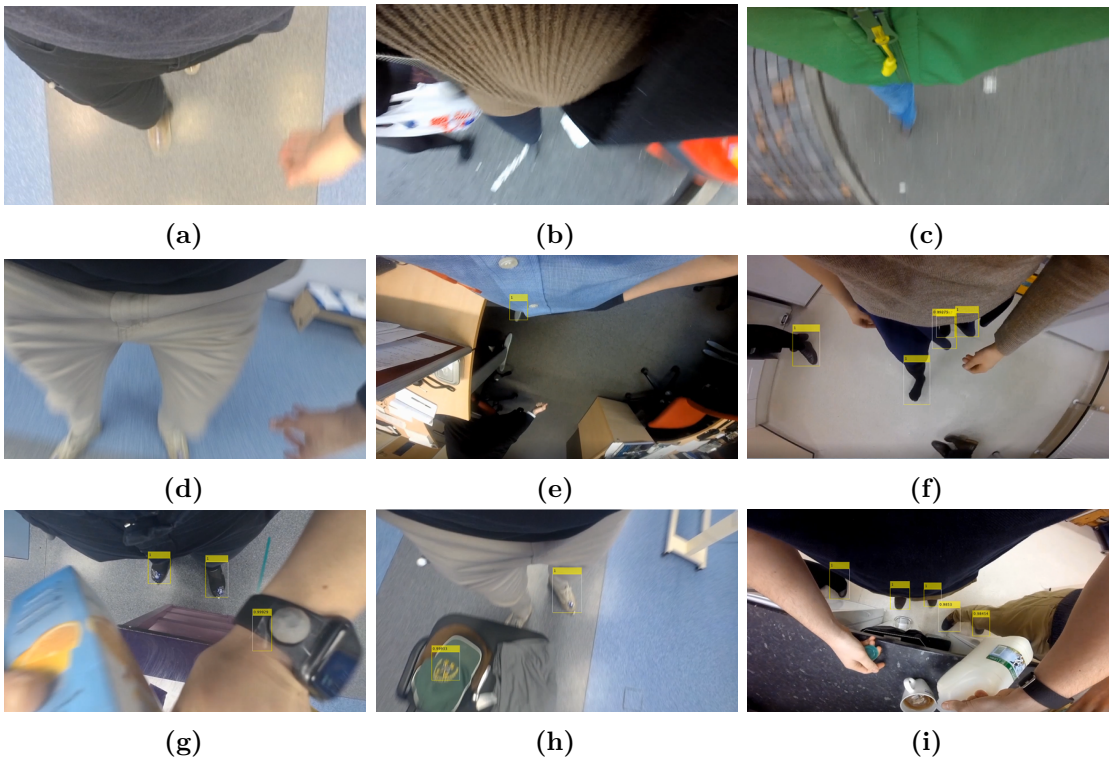


Figure 7.4: Blurring and false positives. (a-d) blurring occurs regularly during gait and movement, even at relatively slow speeds, due to the feet's vicinity to the camera. This blurring makes the feet objectively difficult to structurally discern from the false positives shown (e-i). (d) also demonstrates how the feet's location in the frame may be inconsistent and not always in the top centre. (f) and (i) show examples of multiple feet of other individuals that are not the subject also being detected in the FOV.

Therefore the foot travels across the FOV in a short period of time and will often exceed the recommended threshold. At 720p, twenty pixels is only 2.8% of the camera's image plane y-axis (the dimension of greatest foot movement). Figures 7.4a to 7.4c show several examples of foot blurring, and during walking bouts this is more rule than exception.

7.1.1.5 Clutter & false positives

General observation of the feet in the frames of Figures 7.3 and 7.4 will show that actually there is little structural or textural consensus between feet. Especially when considering the compounding issue of blurring or illumination reducing what limited shape or texture is consistent. Furthermore, the internal variation between shoes and shoe representation under blurring et cetera is seemingly greater than the difference between individual types of shoes and many everyday objects, shapes and visual patterns. For example, it is clear how the false positive detected in Figure 7.4e could

be construed as more similar to a shoe (such as in Figure 7.3e), than any of the shoes in Figures 7.4a to 7.4c. It is worth considering the ambiguity of these examples when questioning the high confidence false positives in Figures 7.4e to 7.4i. For example the curvilinear reflection in Figure 7.4g and the elongated elliptical shape formed by the pleats of the subject's dress in Figure 7.4f look very similar to other views of the feet. Similarly, the crest on the hat in Figure 7.4h can be misconstrued as the curved end of a sneaker with laces. These false positives are most concerning when detected on the subject themselves as they will provide perpetual ambiguity throughout the video. This is especially true when they are either detected in the near vicinity of the feet, as in the case of the dress, or when oscillating backwards and forwards and in and out of the FOV (as the feet also do), as in the case of the watch.

This ambiguity in foot structure is exacerbated by the cornucopia of clutter and background objects that appear in the free-living environment. Many examples of this can be seen in Figures 7.3 and 7.4. It is because of this, and a desire to make the study as applicable to free-living environments as possible, that the ADL protocols involve so many extra items (e.g. Figures 7.4g and 7.4h), whereas otherwise the gait lab would be devoid of potentially confounding clutter (e.g. Figures 7.4a and 7.3f).

An additional but uncommon cause of false positives is that of false images of the feet, for example in reflections (Figure 7.3l). A much greater concern however, and one which is fairly common and relatively unique to the data set, is the detection of other feet in the FOV that are not those of the subject (Figures 7.4f and 7.4i). Filtering these out will be a challenge, especially as sometimes they may perhaps be clearer and more exemplary examples of feet than the subject's own.

7.2 Methods

With the challenges of tracking feet explored and considered, the following section outlines: the choice in object detection technology, how detections in individual frames are filtered of errata and grouped into trajectories, how trajectories are filtered and assigned to the left or right foot, and how gait features can be extracted from these trajectories.

7.2.1 Model selection & training

Although there is a great variety of tools for object detection in an image [380, 381], and even for body-part tracking based on distinct shape and object geometry [382], it is clear from section 7.1.1 that the variation in feet is great, their unifying characteristics limited, and the free-living environment rife with confounding issues and clutter. Therefore, it was decided that a convolutional neural network (CNN) from the TensorFlow Object Detection API (application programming interface) would be used due to its proven robustness, and the hope that neural networks would be able to identify unifying features between feet that are not obvious to the human observer [383].

The model chosen for use from the TensorFlow 1 Detection Model Zoo was the SSD ResNet50 V1 FPN (Single-shot detector 50-layer residual network feature pyramid network version 1) [384–386]. This model was chosen due to its balance between high accuracy and speed; operating at a speed of 76 ms per 600×600 frame and mAP^{-1} of 35 on the COCO data set (common objects in context; Microsoft's collection of 2.5 million instances of 91 everyday object types) [383]. When balancing these factors, accuracy was considered the most important attribute as the foot detection is not being done in real time. However, a reasonable speed is necessary as detection must be conducted on every frame of a video. The Oxford study comprises approximately 20,000 frames per subject, and the Newcastle data set is composed of more than 3.6 million frames in total.

It was decided that transfer learning should be used, as many of the items in the COCO data set would be present in free-living data [387]. Although, 'shoe' is one of the 91 categories in the labelled COCO data set on which the model was trained, the category was not used. This was because the example images were from a range of viewpoints unlike that of the chest-worn camera, resulting in different rotations and scales - but without blurring or dramatic variations in illumination. Therefore, it was deemed better to leverage the consistencies in FOV and train the network on samples from our data sets.

For use in this chapter, the model was trained with a leave-one-out policy on the Oxford study subjects. To do this, frames were randomly extracted from the videos and kept if at least one of the subject's feet was visible. This was repeated until there were 90 frames per subject. These were proportionately downsized to approximate

600px in width and then labelled using the labelImg software [388]. This software converts the manually drawn labels to XML (extensible markup language) files of the PASCAL VOC format (pattern analysis, statistical modelling, and computational learning visual object classes; the format used by ImageNet), which can then be converted into TFRecord tensors and used for training.

On qualitative consideration of the data set, it was decided that much of what aids visual identification of feet, especially in the dark or when blurred, is context. The feet are generally observed immediately below the body, knee, or leg (depending on motion, torso orientation, occlusions, and so on). Also, the feet are generally upon a somewhat uniform surface. Therefore, when manually labelling the data, the bounding boxes where expanded beyond the limits of the feet, to include some context.

With a separate model trained for each subject, the object detector was run on all the Oxford data set videos; providing a list of the top 100 best object detections per frame, and the confidence score for each detection.

7.2.2 Foot tracking

At this stage, the inputs that can be combined to track the feet are the foot detections in each frame (Section 7.2.1), the camera translation and rotation as estimated by visual odometry (Chapter 6), and the measured height of the chest-worn camera as collected in the protocol (Section 4.2). To extract robust free-living step features, a step map needs to be constructed (Figure 5.9). The foot locations in a step map correspond to the foot being stationary with respect to the ground. Therefore, the foot can be assumed to be stationary when camera views at two separate times back project the image of the foot into the same location in the ground plane (Section 3.3.1.2).

The first step is to filter the object detection results. Only detections with a model confidence score over 0.1 are selected. The reason for this is the relative uncertainty due to all the aforementioned confounding issues, and it is preferential to include more false positives than exclude potential steps. Later processing filters out these false positives.

The next step is to find which detections overlap on the ground plan when back projected. To reduce the search space within the sparse matrix of potential detection overlaps, a limit is applied to the number of frames forward the algorithm will search

for an overlapping foot detection instance. For example, if the camera, and thus the subject's torso has moved by at least 0.3 statures in any direction, then it is highly unlikely that their feet have remained in the same place and therefore frames beyond this spatial horizon need not be searched.

Therefore, for each frame a 'spatial search horizon' is established as the most number of frames to look forward, n_s, to find overlapping detections. n_s can be found for frame N as the maximum value that satisfies:

$$\sum_{N}^{N+n_s} \frac{v_t}{F_s} < 0.3h. \tag{7.1}$$

Where v_t is the translational speed signal of the camera in the transverse plane as established in Chapter 6, F_s is the sampling frequency, and h is the subject's height. Similarly, to reduce computational burden, a 'temporal search horizon' is also applied whereby only detections that occur within 3 seconds of frame N are considered. The number of frames to search forward, n_t, is therefore related to the video sample frequency, F_s, by:

$$n_s = 3F_s. \tag{7.2}$$

Due to the variability between feet, especially when considering blurring, it is possible that multiple detections may overlap within the same frame (Figure 7.5). If the OLC exceeds 0.5 (Section 3.4.2.1), only the detection with the greatest confidence is kept.

To remove some erroneous object detections, all the detections within a sliding 10 seconds window are considered (step size 0.2 seconds). Potential outliers are then flagged as those detections with centroids that fall beyond 1.5 interquartile ranges (IQR) of the upper or lower quartiles. These potential outlier detections are then removed if they fall within the central third of the time window. This time window is then passed along the temporal length of the sample with a step size of 1 second (Figure 7.6).

7.2.2.1 Tracklets

The next step is to group detection instances of the same foot across time into trajectory fragments, or tracklets [389]. Starting at frame N, all detections within the frame collect all the detections that fall within N and $min(n_s, n_t)$. They are searched through

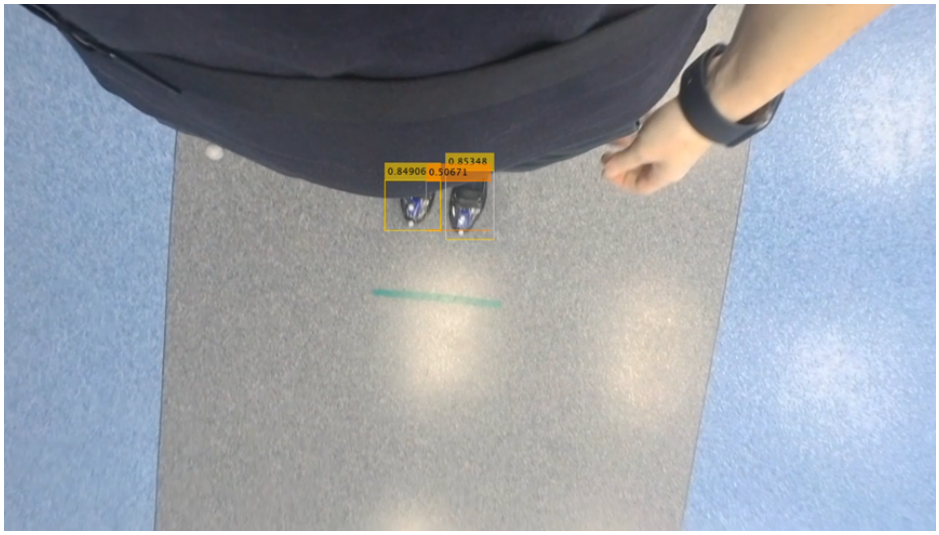


Figure 7.5: Example of detection overlap. Three foot detections are shown in the same frame, outlined by bounding boxes. Two detections are clearly of the same foot. The confidence score for each detection is shown.

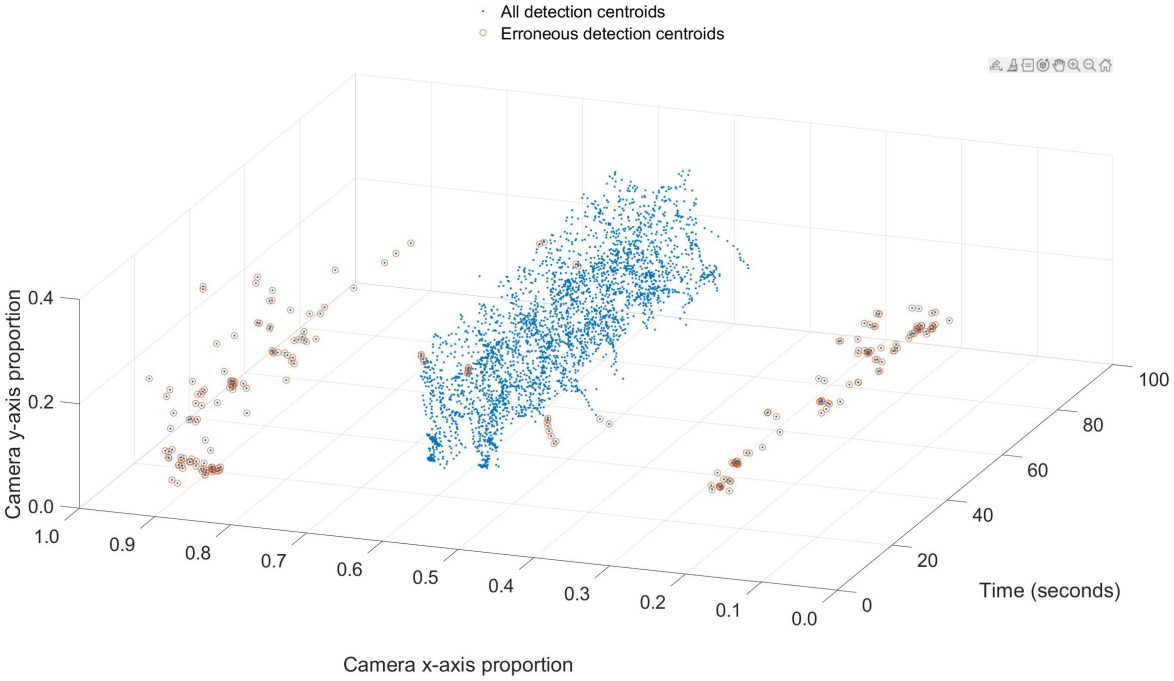


Figure 7.6: Erroneous detection removal. All the detections from an example Straight protocol trial are shown. The feet clearly remain near the centre of the camera frame, although this will not be the case for complex movement. A small number of errata are identified within the main collection of points. It is possible that these are indeed incorrect detections, however, if they are not, the redundancy in the rest of the foot tracking process ensures that it will not affect the step mapping.

chronologically, and if the detection back-projections overlap (OLC) greater than a threshold value, then the detection is considered to be part of the same foot trajectory and is thus added to the tracklet. The foot's location is then updated to be this instance

and the process is started again. This goes on until an instance has no projected detections overlapping within its spatial or temporal horizon. The algorithm then returns to frame n and does the same for any other detections in this frame and moves forwards chronologically. This method results in foot trajectory tracklets (Figure 7.7). The OLC threshold is considered to be 0.1 to account for foot movement, detection size inconsistencies, foot view obstruction, and uncertainty in the relative camera poses.

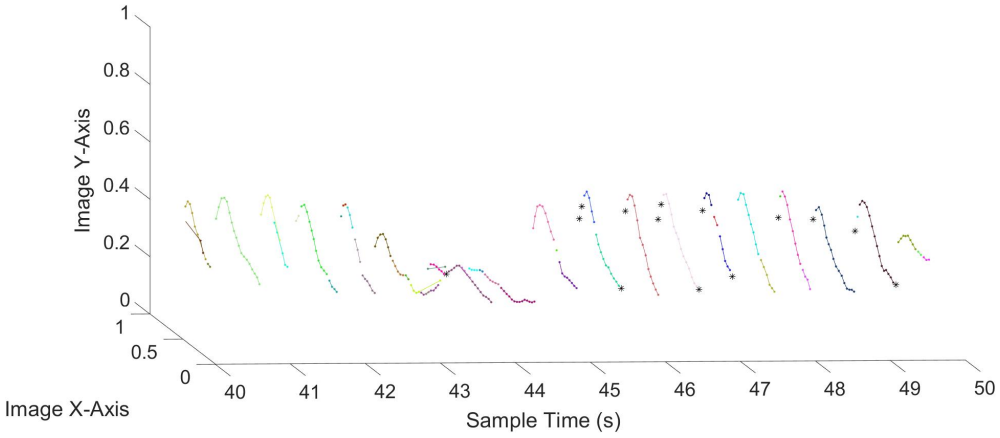
7.2.2.2 Left and right foot trajectory determination

With the detections filtered and grouped into tracklets, it must be determined which tracklets belong to which foot before a foot map can be constructed and features extracted. To achieve this, a five second moving average is applied to the x coordinate of the detection centroids. A tracklet is then scored based on the proportion of the detection centroids greater or less than the moving average. If more than half of the centroids are greater that the moving mean, the tracklet is allocated to the left foot; otherwise the right (Figure 7.8)

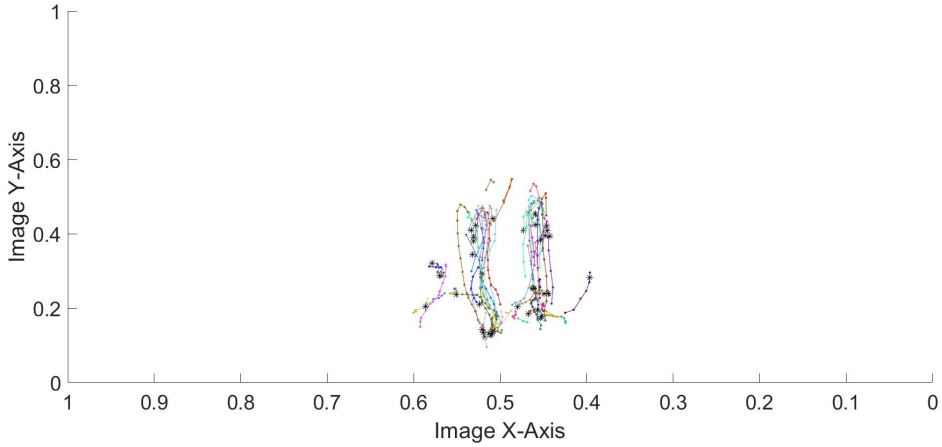
It is possible but very unlikely that the left and right feet will not be left and right of each other respectively from the camera's perspective. This could occur if the subject's torso rotated more than 90° relative to their legs (in which case the feet would almost certainly be outside of the FOV anyway), or if the subject crossed their feet while stationary, walking, or turning.

To remove further erroneous tracklets, any instance whereby the trajectories of two tracklets assigned to the same foot occur in the same frame are identified. The foot cannot be in more than one place at one time. Therefore, the tracklet composed of detections with the highest mean confidence is kept and the others discarded.

Additional erroneous tracklets are removed from the left and right paths by checking for trajectory changes that correspond to foot movement greater that a reasonable physiological speed of $5ms^{-1}$. This may occur, for example, if a foot is suddenly occluded by the subject's arms; whereby the left or right foot trajectory may jump to a false positive tracklet within the FOV. To determine whether it is the tracklet before or after the path discontinuity that is the erroneous tracklet, the ten previous and following tracklets of that foot side are considered, the detection image locations



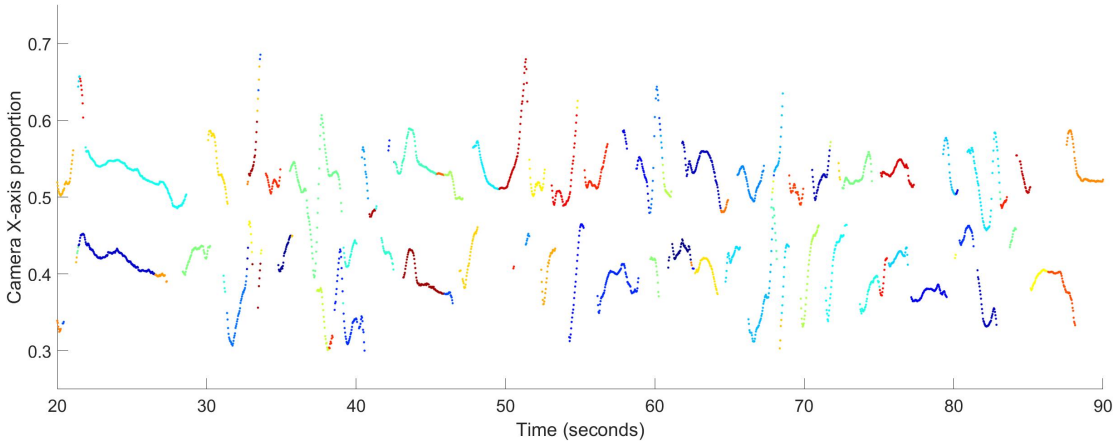
(a)



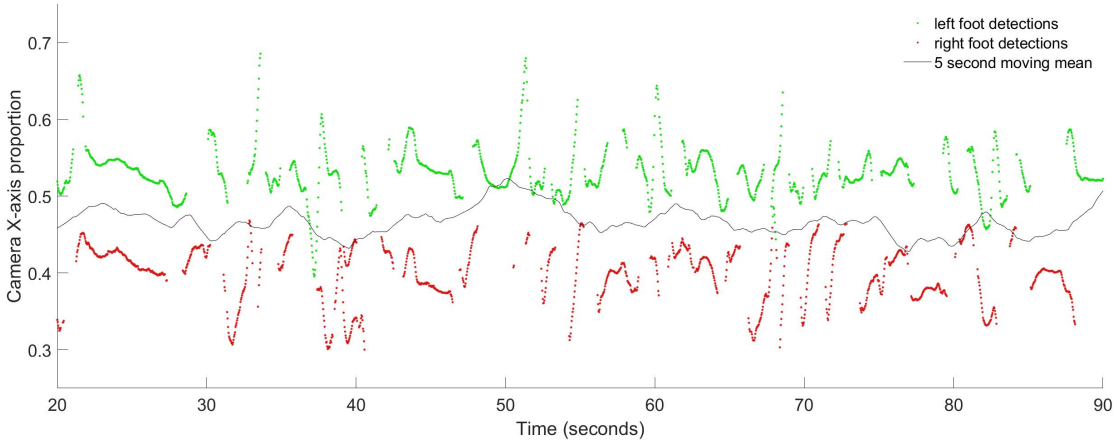
(b)

Figure 7.7: Tracklets example. An example of detection grouping into tracklets over two periods of linear walking connected by a 180° turn at 43 seconds; during the Straight protocol. Each point on a line indicates the centroid of a detection. Detections that are grouped together into a tracklet due to projection overlapping are linked by a line. Detections marked by a black asterisk indicate ungrouped detections. (a) accentuates the temporal dimension. Some of each foot in each step are not grouped together, this is expected as the foot is still moving as it enters the FOV from below the body, and swings forward into heel strike. The trajectories then head towards the top of the frame (image Y-axis value of 1) as the foot remains in the stance phase, stationary relative to the ground, and passes beneath the subject. It can be seen that some trajectories are broken into multiple tracklets. (b) is the same plot viewed along the time axis. It can be seen that during this protocol the feet generally pass through the middle of the frame, with the left and right foot being clearly distinguishable, although the foot relative position in the frame occasionally varies during turns.

k-means clustered, and the cluster centroids compared to the centroid of the tracklets either side of the discontinuity. The tracklets corresponding to the shortest distance between centroids is kept.



(a)



(b)

Figure 7.8: Left and right foot determination. Both subfigures show the camera x-axis centroid value for each detection in a subsample of an ADL3 trial. (a) shows the detections coloured by their tracklets. It can be seen that during complex everyday movements, such as leaning and twisting, the distinction between left and right foot in the x-axis is not as distinct as, for example, in Figure 7.7b. (b) shows the result of applying a 5 second moving average on the x-axis centroid component of all detection, and then allocating the tracklet as belonging to the left or right foot based on the proportion of detections in that tracklet on the respective side of the moving average. This enables robust left and right allocation during sudden movements and leans, such as at 47, 68 and 83 seconds.

7.2.3 Foot trajectories to step maps

To build a foot map, the location of the feet must be determined when they are stationary with respect to the ground. To do this, the same projection overlapping algorithm from Section 7.2.2.1 is used; but on each foot's tracklets independently, and with a required OLC threshold of 0.9. This greater threshold ensures that the foot must be in the same location in consecutive frames (and thus stationary), while accommodating for some uncertainty in the aforementioned parameters.

With the stationary detections identified, a foot movement signal is extracted (Figure 7.9). This is achieved again, with the same detection projection overlapping algorithm from Section 7.2.2.1. However, the distance travelled is calculated to be the Euclidean distance between the centroids of the two detection projections in the ground plane (for example, see Figure 7.10). This signal will show peaks when the stationary location of the foot changes.

A peak detection algorithm is then utilised to identify strides as a change in stationary location; with threshold peak height of 0.2m, as this is the minimum stride length as determined in Section 5.2.3.4 [144]. Additionally, a minimum peak separation is applied of 0.5s which accommodates physiologically likely cadences (Table 5.3), while filtering out occasions where continuous view of a foot is interrupted and drift or camera pose estimation error causes movement to be detected.

The peaks correspond approximately to heel strikes, as this is when the projection of the foot ceases to move with respect to the ground. The beginning of the stride is then considered to be the time of the previous heel strike of the opposing foot, or 0.5 seconds prior to the peak if no such opposing heel strike is detected within 1 second. The same conditions are applied to identify the end of the stride. The stride length, step length and step width are then established by creating a foot map from the projection of stationary detections of the same foot (Figure 7.10). For direct comparison between individual steps, Vicon detected and camera detected steps are matched using the algorithm outlined in Section 3.2.3.2. Errors are not propagated through the process as each measurement comprises a huge number of calculations - each of which has some error - and as such the error margins rapidly grow beyond the scale of the measurements themselves. This is especially true when camera movement is slow or near-stationary. It is for this reason that much redundancy is built into the process - such as including a large number of template matching instances, running multiple pose estimations, and re-projecting all foot instances.

7.3 Results

As the method of extracting step features is dependent on a hierarchy of processes, the results are split into the key steps involved in the process: the detection of feet within

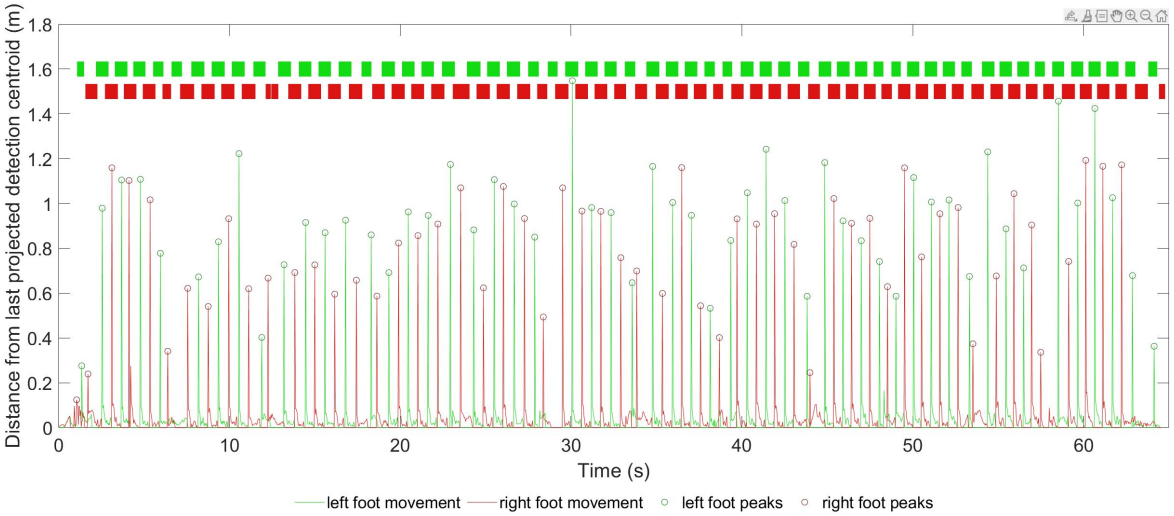


Figure 7.9: Foot movement tracking. The distance between consecutive detection projection centroids is shown for both the left and right foot during a Straight trial. The movement of a foot from one stationary location to another is detected as a peak in the movement signal. Small distances are registered while the foot is tracked as it passes beneath the body during the stance phase, whereupon it is occluded or leaves the FOV. For comparison, the stride times of both feet as extracted from Vicon are shown above the signals.

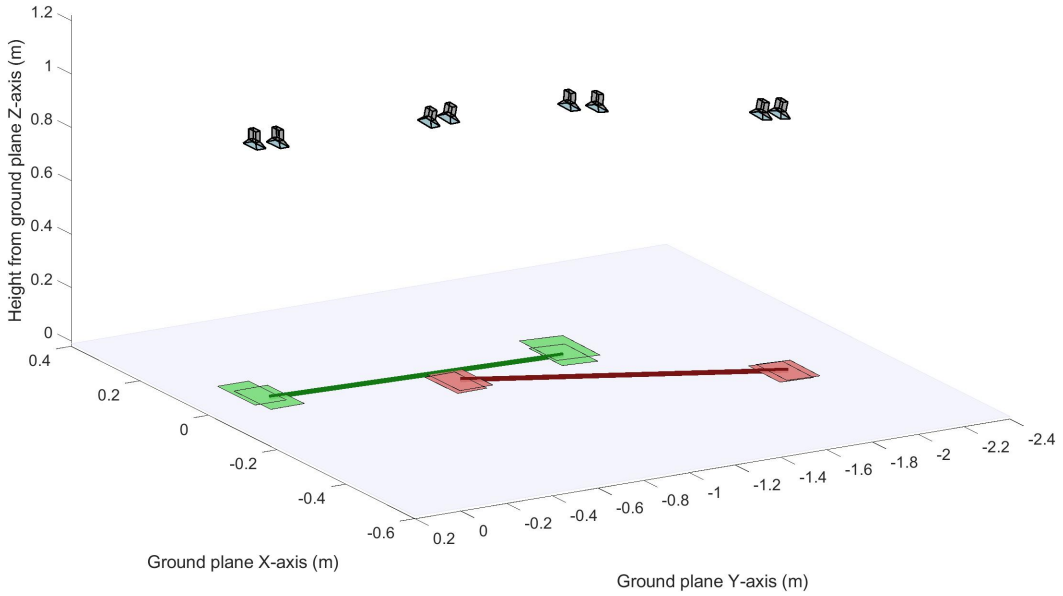


Figure 7.10: Foot map and step feature extraction. Each camera wireframe represents the 3D orientation and location of the camera at a series of non-linearly spaced time points. By projecting the stationary detections into the ground plane, a foot map can be constructed and step features extracted in accordance with Section 5.2.4. The stride vectors are created by connecting the centroids of the final stationary detection projection of each stationary tracklet. Only two overlapping instances are shown, however each stationary tracklet will likely have many such projections. The axes are scaled to emphasize movement and the ground plane is shaded.

video frames, the detection of steps from the combination of these foot detections and camera movement, and the extraction of step features from the detected steps.

7.3.1 Foot detection

As the foot detection model was trained on a leave-one-out policy, each subject had their own trained model. The performance of all these models was assessed and aggregated to create PR and mAP curves that express dependence on IoU threshold (Figure 7.11).

It can be seen that there is a heavy dependence on IoU, however even at very low IoU thresholds, there is an upper limit on recall at approximately 0.75. The sudden dropping off of precision after plateauing indicates relative insensitivity to the confidence threshold for discrimination between TPs and FPs. Figure 7.11b demonstrates a sharp falling off of precision beyond the IoU threshold of 0.2.

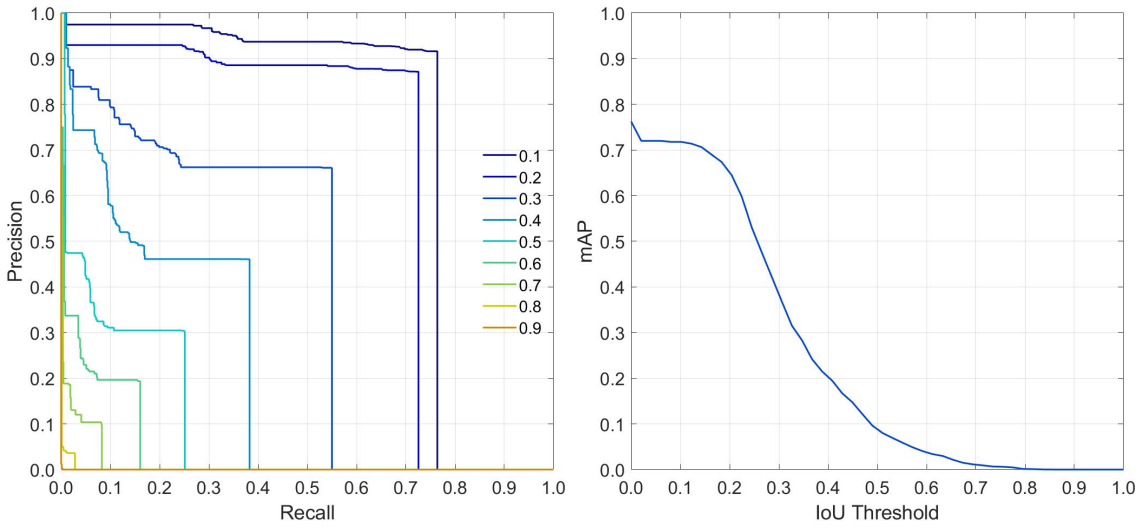


Figure 7.11: PR and mAP sensitivity to IoU for all subject foot detection models. (a) PR curves are shown at nine linearly spaced IoU thresholds. Both precision and recall are highly and similarly dependent on IoU threshold, with strong results shown at low IoUs. (b) clarifies this relationship with mAP as a function of IoU threshold.

7.3.2 Step detection

Table 7.2 shows that precision, recall and F-score are reasonable to excellent throughout the various protocols. There is however clear delineation between the capacity for detection during linear walking and ADLs. The capacity for the method to allocate steps correctly to the left and right foot is high throughout with some reduction during ADLs.

Table 7.2: Step detection results. Step detection statistics for all steps across all subjects and trials, including the precision of allocating a step to either the left or right foot.

	Straight	Turns	ADL1	ADL2	ADL3	Overall
Precision	0.98	0.97	0.80	0.78	0.85	0.90
Recall	0.96	0.91	0.86	0.86	0.80	0.89
F-Score	0.97	0.94	0.83	0.82	0.83	0.89
Left-right precision	0.99	0.97	0.90	0.92	0.93	0.95

7.3.3 Step features

The accuracy and correlation coefficients of the measurements of stride length, step length and step width are shown in Table 7.3. Although the mean error is relatively low, the standard deviation is substantial when compared with both the median features values observed in Figures 7.12i and the mean values. The mean values of $0.96 \pm 0.38m$ for stride length, $0.49 \pm 0.21m$ for absolute step length, and $0.17 \pm 0.10m$ for absolute step width show that the relative accuracy is lower for step length than stride length, and much worse for step width.

Figures 7.12a to 7.12c demonstrate the correlation coefficients of the results, and Figures 7.12d to 7.12f the distribution of the error in Bland-Altman plots. Pearson's was utilised over an intraclass correlation coefficient partly due to its regular use in similar studies, and partly as it accommodates for scaled correlations; as this is a first attempt at achieving such measurements through such methods - a strong but scaled correlation would be deemed a highly positive result that likely required a relatively simple change in calibration or some other consistent parameter. It is clear that shorter strides and steps have much higher uncertainty. Figures 7.12g to 7.12i further elucidate the distribution of the features across the five different protocols. From this it can be seen that steps during ADLs are much more evenly spread across the range of feature values.

The results can only be calculated for the steps that were detected as true positives. Therefore, to illuminate any relationship between feature values and failure for the method to detect a step, Figures 7.12j to 7.12l show the distribution of step features of false negatives. It is clear that there is a propensity for short strides, with short step lengths to remain undetected.

Table 7.3: Step features results. Step feature statistics errors are calculated as the the camera measurement minus the Vicon ground truth measurement, expressed in metres. Pearson correlation coefficients are also shown.

		Straight	Turns	ADL1	ADL2	ADL3	Overall
Stride Length	Error Mean	0.01	0.01	0.03	0.02	0.03	0.01
	Error StDv	0.20	0.23	0.30	0.33	0.21	0.25
	Corr. Coef.	0.80	0.67	0.56	0.55	0.72	0.73
Step Length	Error Mean	0.05	0.07	0.12	0.09	0.12	0.08
	Error StDv	0.25	0.26	0.34	0.32	0.31	0.29
	Corr. Coef.	0.63	0.57	0.34	0.39	0.39	0.54
Step Width	Error Mean	0.01	0.01	0.09	0.10	0.03	0.03
	Error StDv	0.11	0.18	0.47	0.57	0.18	0.30
	Corr. Coef.	0.17	0.06	0.07	-0.04	0.01	0.09

7.4 Discussion

The results show a broad range in the level of success at extracting gait features; from an excellent capacity to detect and measure strides in linear walking, to a complete inability to measure step width during ADLs. In general, the main factor that affects the results is the complexity of the calculation, and how far down the hierarchy of methods and calculations it is.

The next most confounding factor is the complexity of the movement. Similar to the results in Chapter 6; the Straight protocol gives slightly better results than Turns protocol due to the reduced complexity and unnaturalness of movement, and the ADL protocols are noticeably worse than Turns.

7.4.1 Foot detection

The foot detection results show high precision and recall at low IoU thresholds when compared to similar transfer learning attempts [390]. This is very positive considering all the complications outlined in Section 7.1.1. That the IoU threshold should be necessarily so low is not surprising considering the previously explained fast motion, blurring, regular partial occlusion, and the variation in labelling to include context as explained in Section 7.2.1. The required low IoU threshold and potential false positives are not a huge concern as robustness is built into the methods along with error filtering. These results help to justify the OLC threshold of 0.1 chosen in Section 7.2.2.1. Although such low values of IoU would generally be considered a poor result

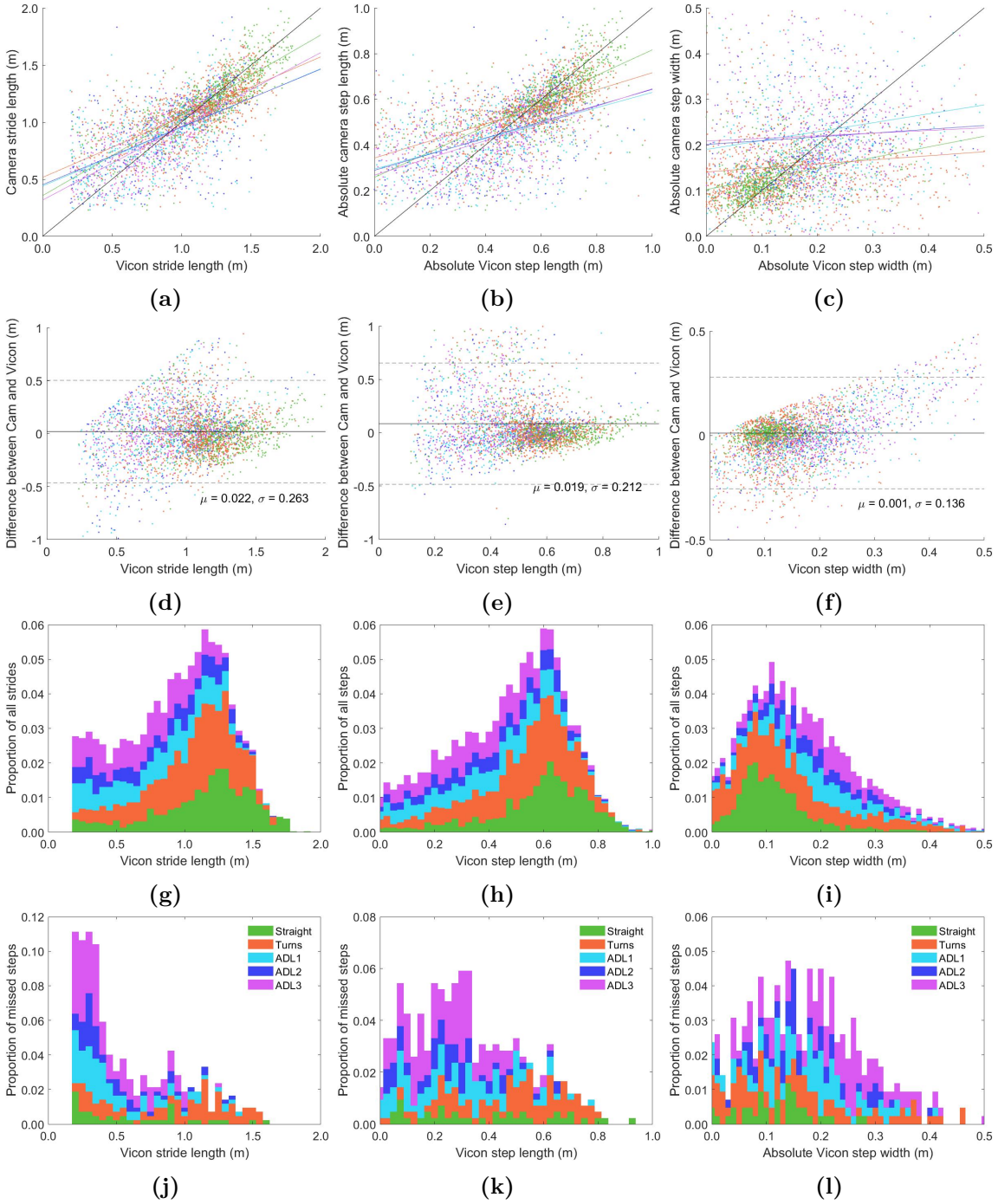


Figure 7.12: Step feature results. The columns of subfigures display the results for the respective step features: stride length, step length, and step width. The top two rows demonstrates the correlation and Bland-Atlman plots of errors between Vicon ground truth and measured step features. The bottom two rows show the distribution of the step feature values for all Vicon values, and all false negative detections respectively. The histograms utilise bins of 0.050, 0.025, and 0.010 metres for the stride length, step length and step width results respectively.

in the object detection field [391], the data set is fairly unique and the results sufficient for purpose. The step detection results suggest that this is the case.

7.4.2 Step detection

The substantial drop in precision during the ADLs is somewhat expected, partly due to occlusion of the feet by arms, objects and furniture. It appears that the main cause of this reduction however, is the increase in false positives caused by the various objects in the experimental area and thus within view of the camera. This is concluded from qualitative assessment, the fact that ADL3 is also considerably lower, and Figure 7.13 which shows that the majority of false positives are estimated to have short stride lengths. These are therefore likely during periods of slow movement, or of the left or right foot trajectories jumping to a false positive and then back to the feet again while stationary. The additional drop in precision between ADL3 and other ADLs is likely the result of obstruction of the view of the feet.

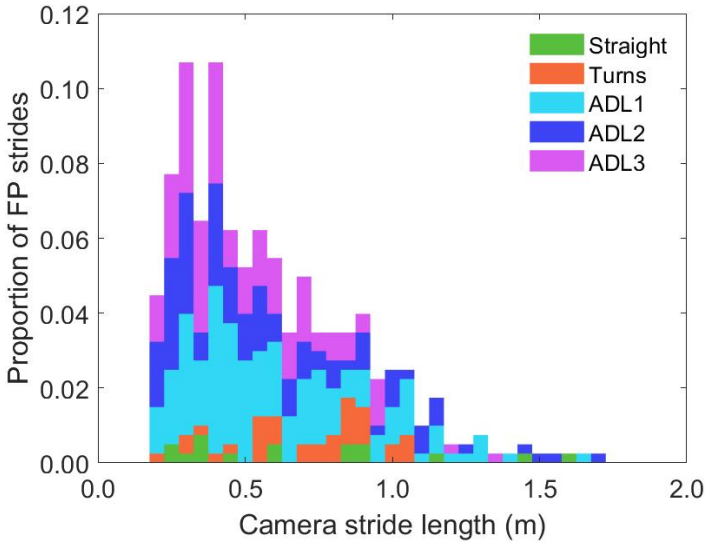


Figure 7.13: False positive strides as a function of stride length. Bins of $0.05m$ are used. The majority of false positives are associated with shorter strides of below $0.6m$.

Recall however, does not follow the same pattern. This is most noticeable as ADL3 exhibits the lowest value. This is most likely due to the large number of short strides within this protocol. Figure 7.12j shows that there is a substantial propensity for missed strides to be of short stride length.

The precision of the methods to determine left and right steps from one another is always high at over 0.90 throughout, and in fact over 0.97 for the simple protocols. Similar to precision however; complex movement, view obstruction and false positives from objects in the FOV cause errors in the allocation. There is little precedent in the

determination of the accuracy of which leg a step is made by against which to compare these results. Although several reported methods do delineate left steps from right steps (and this is certainly important for metrics of asymmetry), the veracity of the left and right assignment is not validated [392]. This is to some extent one of the many symptoms of not directly comparing gait step-to-step (Section 2.3.3); and only looking at linear gait whereby left and right steps are evenly distributed and predictable.

In general it can be said that the camera system to was able to consistently and accurately detect specific steps through the study. It is difficult however to compare these results to other studies as detected steps are very rarely compared to ground truth. The most commonly reported statistic is the step count correlation coefficient. This does not elucidate whether the detected steps are actually representative of real steps, and large numbers of false negatives can be offset by similarly large numbers of false positives. Therefore, for comparison with other studies the step count for each protocol was calculated and compared with ground truth (Figure 7.14). The resultant Pearson correlation coefficient of 0.96 is slightly higher than similar accelerometer methods in completely free-living data [14], and comparable to highly constrained linear walking using a range of wearables [263, 393–398]. Considering the protocols are semi-structured this suggests that the camera system is very competitive with other methods as far as step count is concerned.

It is important to note however that the ρ value in Figure 7.14 is for total steps detected and thus it is not representative of the real, true positive, steps. This offers the first evidence that averaged and gross gait features extracted from wearables may not be representative of ground truth step-by-step values.

7.4.3 Step features

The results for the step features are hugely variable. However, it seems that considering the complex movements; stride length and step length under some circumstances can be measured with reasonable reliability of between 1 and 7cm error. Although step length is worse than other reported values in the literature [399, 400], these studies involved highly constrained linear gait, where step length is generally only really half

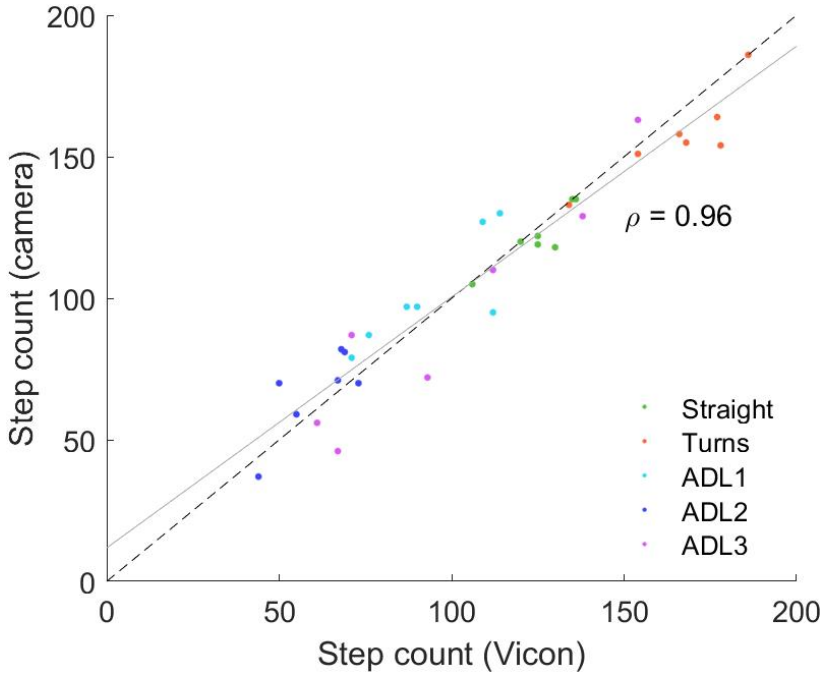


Figure 7.14: Step count results. The total number of steps detected shows strong agreement with ground truth across protocols with a range of activity amount and complexity. The grey line shows the result of linear regression, with the dashed line representing perfect correlation for comparison. The Pearson correlation coefficient is also shown.

of stride length. Therefore this study's stride length is more representative of these results, and in fact even this cannot be directly compared.

Several studies report ICC values for step length, stride length and step width between 0.80 and 0.98 [182, 401–403]. As such, the results from the Straight protocol are comparable, especially as the the protocol actually involves more complex gait than most studies, including: multiple turns and types of turns, rapid acceleration and deceleration, and varying walking speeds.

To investigate the effects of these factors, Figure 7.15 demonstrates the dependence of stride length estimation error on the mean walking speed across the stride time, the variance in the walking speed, and the acceleration across the stride time. It can be seen that lower walking speeds correspond to substantially increased uncertainty. This has also been observed with other wearables [399, 404]. This effect is likely exacerbated by the fact that slower walking speeds correspond to more erratic, ADL-style movement, which Figure 7.15b shows to also increase error. It is therefore likely that consistent low speed walking would actually render low errors. The effect of erratic movement can also be seen in Figure 7.15c; whereby the dumbbell shape of the percentage error

is likely caused by the greater accelerations and decelerations corresponding to the Straight walking protocol. This implies that change in direction is more important in causing error than change in velocity.

Overall, it is unsurprising that the further down the hierarchy of processes and calculations, the greater the compounding of uncertainty, and the lower the feature accuracy. Stride length is directly affected by camera rotation and translation estimation. Step length and step width are both calculated from the stride vectors during, before and after the specific step being measured (see Figure 5.9). Each of these harbours error and therefore propagation of errors inevitably leads to them being less accurate than stride length. A significant confounding factor is that of false negative and false positives. A single FP or FN result can, for example, shorten two strides, or lengthen one. This will then affect the foot map and thus measurements of four step length and step width measurements. This combines with the fact that the features, and the variation within them, results in an inability to provide the centimeter accuracy that is required to provide reliable step width estimations.

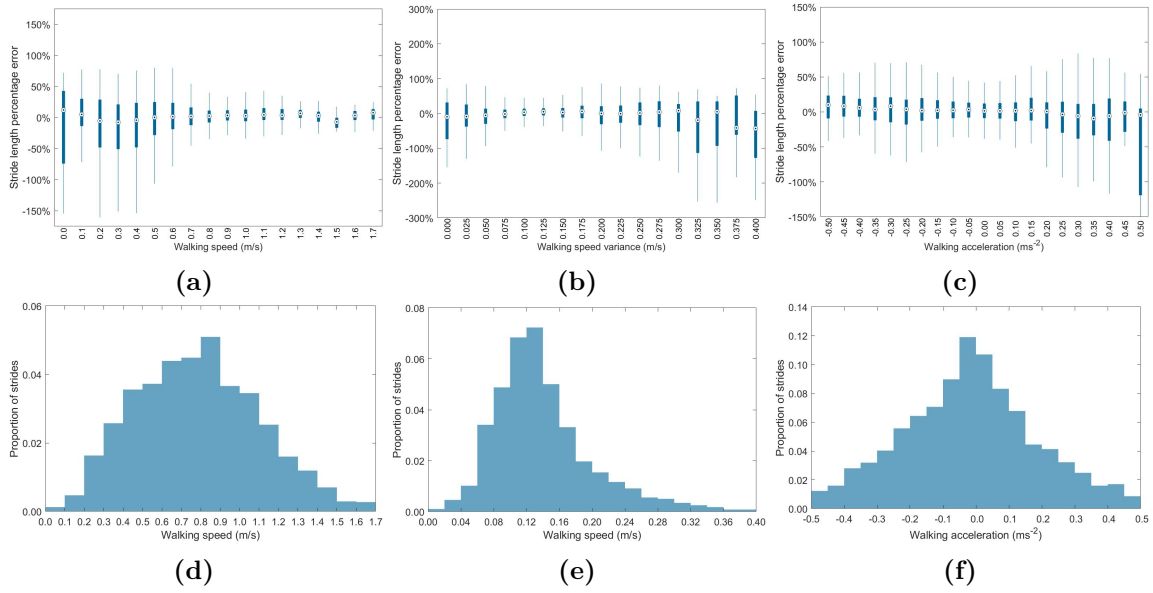


Figure 7.15: Step feature variation with walking speed and acceleration. (a-c) show the dependence of stride length estimation accuracy on the walking speed, speed variance and acceleration as measured by Vicon. (d-f) help to qualify the results by providing distributions of the steps by mean walking speed, speed variance and acceleration; bin sizes of $0.1ms^{-1}$, $0.02ms^{-1}$ and $0.05ms^{-2}$ respectively.

7.5 Conclusion

In this chapter, computer vision object detection methods were utilised to detect feet in the videos collected from the studies outlined in Chapter 4. These detections were combined with the camera egomotion inferred by the work in Chapter 6, to filter the detections, group them into trajectories, and extract gait features from the videos. These gait features were then compared to those extracted from a Vicon motion capture system using the methods developed in Chapter 5.

The methods developed in this chapter, and the results obtained from them, are heavily affected by a range of complexities and challenges. These range from being unique to the data, to of general relevance in the field of object tracking. They are explored in length in the introduction, and their relevance referred to throughout.

The results range from excellent step detection precision of greater than 0.98, and stride length estimation error of of $0.01 \pm 0.20cm$, for the Straight protocol, to very poor estimation of step width during ADL protocols (for example mean error of $0.10 \pm 0.57cm$ in ADL2). This demonstrates that the capacity for the developed wearable camera system to extract accurate gait features is highly dependent on the complexity of the activity and the type of feature. In addition to the validation work on the wearable camera system, this chapter builds on the evidence that for accurate gait analysis to be conducted beyond highly constrained linear gait, much more in depth research needs to be conducted both regarding protocols and in step-by-step analysis. Chapter 8 goes on to investigate how other standard wearable devices fare under similar conditions.

8

Comparing gait analysis from multiple accelerometry-based wearable devices

Contents

8.1 Introduction 171
8.2 Methods 173
8.3 Results 184
8.4 Discussion 184
8.5 Conclusion 203

8.1 Introduction

The importance of measuring gait both within and beyond the clinic, in active tests and free-living environments is covered in depth in previous chapters. This is also true for the benefits of wearable over static devices, and consumer devices over bespoke research technologies. Whereas previous chapters have looked at the accuracy of a wearable camera-based system for gait analysis, this is actually one of the more unusual methods. By far the most common method is by using accelerometry provided by the inertial measurement units (IMUs) in both research-grade and consumer electronics.

The argument for the use of accelerometers in clinical, and especially free-living, gait analysis is clear: they are small and ergonomic, highly affordable and widely available, research-grade devices sample at the kilohertz level, the measurement of gravity enables orientation to be ascertained reducing the need for calibration and precise fixing to the subject, and they can sample and store data for long periods of time. Another crucial consideration is that many consumer electronics, such as

smartphones and smartwatches, already contain IMUs for various functions. This means that gait analysis could hypothetically be conducted constantly in everyday life. The benefits of this are manifest: indications could be identified before or as they clinically arise [405, 406], diseases and treatment regimens could be monitored as they progress [407, 408], the efficacy of therapies in clinical trials could be more rapidly assessed [114, 409], and problematic events such as falls could be detected [410–412].

The relatively recent but multitudinous benefits of this methodology have however resulted in a slew of problems in the literature corpus surrounding the practical use of such devices. These problems fundamentally arise from the desire for new methods to solve specific problems: from monitoring post operative recovery from cerebral palsy-related orthopaedic surgery to identifying differences in activity levels of community-dwelling older adults or even athletes [413–416]. To this end, one of a plethora of available devices is used, attached to the study's subjects in some manner, and submitted to a use-case specific protocol against which the accelerometry device and associated gait analysis algorithms are validated when compared to some ground truth. Problems arise at each stage of this process.

Not all IMU devices are created equal. Accelerometer range, sampling rate and sensitivity have all been shown to affect gait analysis results [417]. Therefore even identical experiments with different devices may not produce comparable results. This issue is confounded by the rapid evolution of the hardware and algorithms - especially in the consumer electronics space. An example of this might be the changes in the IMU and step counting algorithms between each generation of Apple Watch and iPhone. Similarly, how a device is attached to a subject, where, and in what specific orientation have all been shown to impact gait analysis accuracy [123, 418–422].

Potentially the biggest and most consistent issue when assuming the general validity of an accelerometer-based gait analysis system, is the protocols undertaken to prove validity. The vast majority of studies only undertake short episodes of linear walking; rarely involving anything more complex than an about-turn. As previously discussed, this is in no way representative of gait outside of the laboratory or clinic. The effect of natural gait on the accuracy of previously presumed validated gait analysis systems

cannot be understated, with even step counting accuracies ranging from 10% to 95% with the majority performing badly [423, 424].

This issue is confounded by the difficulty of obtaining a gait feature ground truth from free-living ADLs. Two main options at present are either to resort solely to manually counted steps from a wearable camera (hence the work conducted in Chapters 6 and 7), or to try and imitate everyday life in the lab (hence the protocols outlined in Chapter 4). Some studies have attempted to conduct imitation free-living activities of daily living (ADLs) within the lab - however few of these use motion capture (although this is common with linear protocols) [424–426]. Even those that do, do not utilise robust engineered methods for defining and extracting gait features (such as those outlined in Chapter 5).

The permutations and combinations of all the above variables have resulted in a vast literature corpus with limited homogeneity or validity for broad, generalised use. Therefore the two key aims of this Chapter are: firstly, to provide direct comparison between various consumer and research-grade IMU devices during both linear and ADL ambulation, and secondly, to illuminate issues that arise in the assumed general validity of gait analysis systems tested in a limited and specific way. This is all conducted in the context of the validation of the camera-based method devised in Chapters 6 and 7).

Unlike in Chapters 6 and 7, a range of gait features will not be considered. The validity of gait analysis will be solely tested by a device's efficacy at step detection. Partly this is due to the limited time and resources of the study, but mainly it is because the vast majority of accelerometry-based gait features require step detection as a crucial part of the process of extracting more in-depth gait features. This is perfectly understandable as it would be difficult to establish, for example, step length without first delineating where and when a step starts and finishes. Therefore, it is reasonable to assume that step feature accuracy will be highly dependent on step detection accuracy as an upstream process.

8.2 Methods

The devices used, experimental setup, and protocols undertaken are all outlined in Chapter 4. This resulted in five accelerometry traces for each trial and each subject:

one each for the smartwatch, smartphone, lumbar-IMU, and both foot-IMUs. All were processed independently, with the last two being combined at the end of the full process to establish an overall step detection signal for the feet as a single device.

In order to compare step detection results, step-by-step signals were synced manually. Subjects were asked to lightly jump up and down at the beginning and end of each protocol, and these accelerometry peaks were aligned. In order to then detect steps in the signals, the accelerometry traces were made orientation-independent, filtered, and steps detected using various algorithms.

8.2.1 Orientation

For both step detection and spatio-temporal features, it is necessary to know the orientation of the device. Furthermore, it is very possible that a sensor may alter in orientation relative to the wearer during recording (for example, a watch may move around the wrist or a phone within a pocket). Therefore it is beneficial to perform gait feature extraction on acceleration signals in a reference frame that is invariant to the orientation of the device.

To achieve this the acceleration signals $\boldsymbol{a} = (\boldsymbol{a_x}, \boldsymbol{a_y}, \boldsymbol{a_z})$ in the accelerometer's frame of reference undergo a transformation to make them orientation independent [427]. The transformation maps the accelerometer coordinate system $[\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}]$ to that of unit quaternions, or versors, $[\hat{\boldsymbol{\xi}}, \hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\psi}}]$. Here, $\hat{\boldsymbol{\xi}}$ is parallel to the direction of gravity, $\hat{\boldsymbol{\zeta}}$ is the projection of the direction of travel in the plane perpendicular to gravity, and $\hat{\boldsymbol{\psi}}$ is perpendicular to the other versors.

Given the assumption that gravity is the predominant component of the acceleration signal and that the signal has been lowpass filtered, over a period of n samples, the gravity vector, $\boldsymbol{G_n}$, can be estimated as $(\bar{\boldsymbol{a}}_{\boldsymbol{x}}, \bar{\boldsymbol{a}}_{\boldsymbol{y}}, \bar{\boldsymbol{a}}_{\boldsymbol{z}})^T$. Therefore,

$$\hat{\boldsymbol{\xi}} = \frac{\boldsymbol{G_n}}{\|\boldsymbol{G_n}\|}. \tag{8.1}$$

Over this same period of n samples, a $3 \times n$ acceleration matrix can be defined as $\boldsymbol{A} = [\boldsymbol{a_x}, \boldsymbol{a_y}, \boldsymbol{a_z}]$. Making the projected acceleration in the new coordinate system:

$$\boldsymbol{a_\zeta} = \boldsymbol{A} \cdot \hat{\boldsymbol{\zeta}}. \tag{8.2}$$

The acceleration in the plane perpendicular to gravity, $\boldsymbol{A^p}$, can therefore be described by:

$$\boldsymbol{A^p} = \boldsymbol{A} - \hat{\boldsymbol{\zeta}}\boldsymbol{a}_{\boldsymbol{\zeta}}^{\boldsymbol{T}}. \tag{8.3}$$

It can then be assumed, and has been demonstrated under experimental conditions [428], that the direction in the plane with the highest variance in acceleration represents the main direction of movement. Principal component analysis (PCA) is then used to identify this direction and the eigenvector, $\boldsymbol{v}$, of the covariance matrix of the normalised $\boldsymbol{A^p}$ matrix. The second versor, $\boldsymbol{\xi}$, is therefore this eigenvector normalised; $\hat{\boldsymbol{\xi}} = \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|}$. The final orthogonal versor can then be calculated as,

$$\hat{\boldsymbol{\psi}} = \hat{\boldsymbol{\zeta}} \times \hat{\boldsymbol{\xi}}. \tag{8.4}$$

Thus the final transformed acceleration components are, $\boldsymbol{a_\xi} = \boldsymbol{A} \cdot \hat{\boldsymbol{\xi}}$ and $\boldsymbol{a_\psi} = \boldsymbol{A} \cdot \hat{\boldsymbol{\psi}}$. In the work that proposes the rotation invariant transformation, n is defined as the number of samples in a gait cycle [427]. However, as the gait cycles are both heavily variable and ill defined in unconstrained free-living gait, a sliding window of the maximum integer samples in $0.5s$ is used.

An additional complication observed in free-living data that is not confronted in these papers is that of complex gait where the accelerometer may remain static or move erratically; rapidly changing direction. To account for this, a one second moving average is applied and any periods where the net velocity is less than $0.1ms^{-1}$ are removed (in accordance with minimum walking speeds explored in Section 5.2.5.3).

In aid of clarity, from this point forward, the reoriented accelerometer signals are referred to by their correspondence with the body axes (as this is how they approximately align during forward movement): a_{AP}, for the signal aligned with the anteroposterior axis, with positive being forwards relative to the subject; a_{SI}, for the signal along the superior-inferior axis, with positive being with the gravity vector; and a_{ML}, for the signal parallel to the mediolateral axis with positive being to the subjects left-hand side.

8.2.2 Signal filtering

A standard in gait signal analysis is to filter out the frequency components of the signal that are presumed not to be from ambulation. For example if a smartwatch is on the wrist of an individual who's arm is shaking while walking; one would like to remove the acceleration due to shaking. To achieve this a lowpass filter is traditionally chosen.

On looking at the range of cadences that are likely to be observed, a cut-off frequency of 2.5Hz would clearly include the vast majority of gait speeds (Table 5.3). This has been regularly stated in the literature although it is also recognised that children running fast may require a cut-off of 3Hz [429]. However, as this scenario is relatively unlikely in gait analysis compared to the predominant much slower movements, 2.5Hz is used. Further justification can be found in Section 2.4.1.

The type of filter by far most commonly used in the corpus of the research is that of a fourth-order zero lag Butterworth filter [430, 431]. Furthermore, it has been shown to be the superior choice specifically in gait kinematics [432], and the removal of lag caused by the filter is achieved by filtering in both directions consecutively [433]. This therefore brings the noteworthy limitation that zero-lag filtering can only be achieved on either offline, or delayed online signal segments.

8.2.3 Step detection

With bouts identified, individual steps must then be detected. This is done differently for each of the devices. Due to the use of different step detection algorithms, which themselves depend on the detection of different parts of the gait cycle, measured spatio-temporal gait feature results would vary dramatically between devices. However, this chapter solely deals with the binary existence of step detections, and so such variation, such as the temporal foreshortening of a step, will have little effect unless sufficiently drastic that it either make the step so short it is not detected at all, or results in no temporal overlap between the detected step and the Vicon ground truth.

8.2.3.1 Lumbar accelerometer & smartphone

This method was originally developed by *Dijkstra et al.* where they show that peaks in the anteroposterior (AP) acceleration that occur prior to its zero-crossing approximately

corresponded to foot contact as validated against ground reaction force (GRF) data [434]. This is justified by the observation that during gait, forward and downward acceleration occur after mid-stance as the body falls forward. After heel strike and foot contact, the body then begins to raise over the contralateral leg resulting in forward deceleration and upwards acceleration. This justification is grounded in the inverted pendulum model of gait [435].

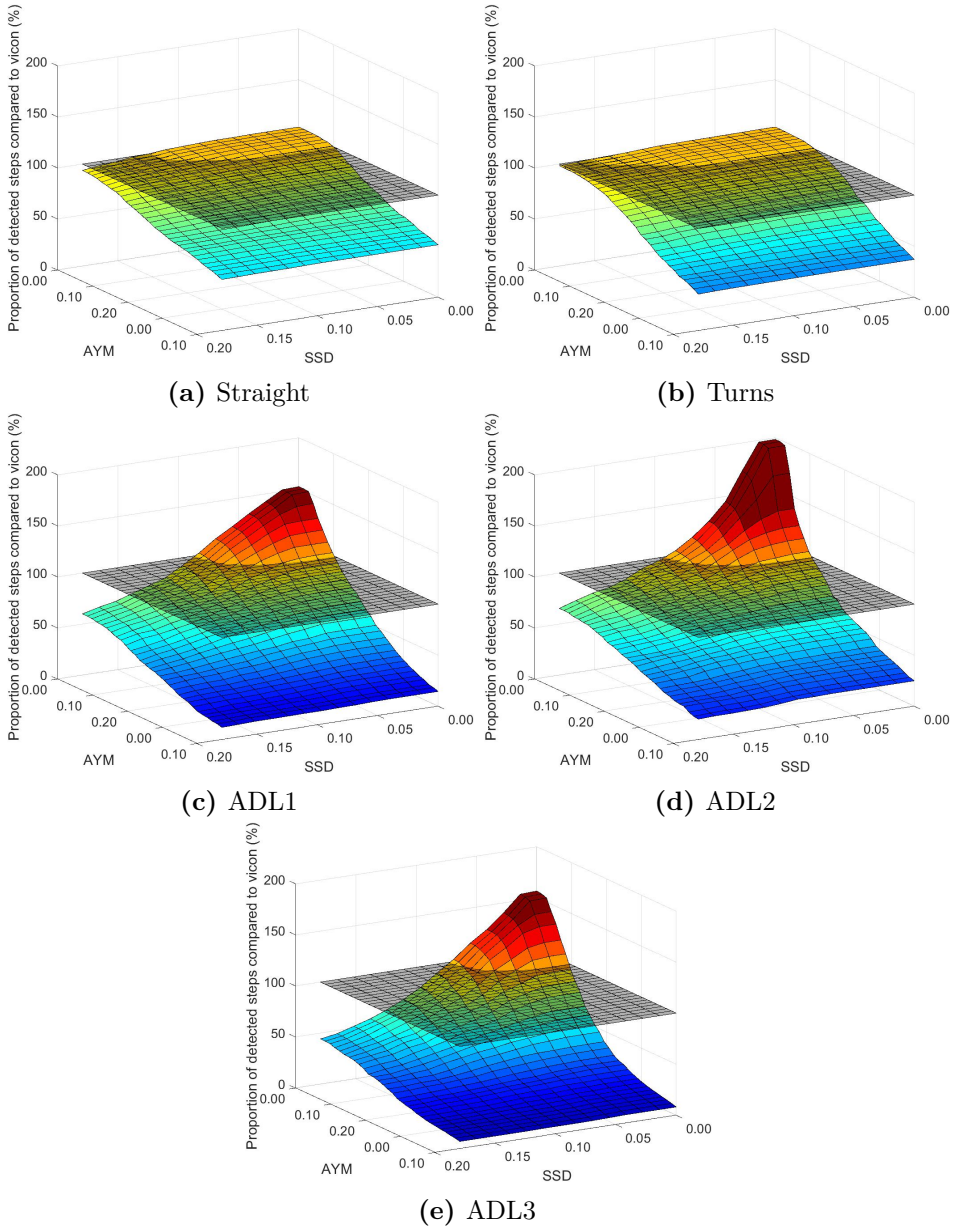
As such the filtered AP acceleration undergoes peak detection with the minimum time between peaks taken as $0.25s$. This is actually a divergence from the parameters suggested by *Zijlstra et al*, who use a minimum cut-off frequency of 2Hz [434]. However their validation study was on treadmills at restricted speeds. Whereas the preference here is to account for all gaits and maximum cadences (Table 5.3).

8.2.3.2 Bout detection

Once the signals are pre-processed and potential steps identified, periods of gait movement must also be identified and steps outside of these periods removed. This is especially important when considering free-living data as a substantial proportion of the signal is likely during stationary, or non-walking, periods.

To detect these periods of walking, *Hickey et al.* developed a method specifically for the case of free-living lumbar mounted accelerometry [14]. To identify when an individual is upright and moving, a window is taken of 0.1 seconds, and the sum of the standard deviations of the tri-axial acceleration, SSD, and the standard deviation of vertical acceleration, AYM, is calculated and compared to minimum thresholds of $0.77g$ and $0.05g$ respectively [436]. In addition, bouts of periods shorter than $0.5s$ were also excluded [437].

On inspection, although this method was relatively effective on periods of linear walking, it resulted in high false-positive and false-negative rates in free-living data. This was inspected by conducting sensitivity analysis with regard to the parameters AYM and SSD. Figure 8.1 demonstrates why this may be. Here we can observe that during continuous, linear walking, such as that seen during the Straight and Turns protocols, the total detected step count is rarely above the ground truth. This occurs even when the thresholds are reduced to zero. However, the sensitivity of the step



(a) Straight **(b)** Turns

(c) ADL1 **(d)** ADL2

(e) ADL3

Figure 8.1: Step count sensitivity to SSD and AYM for lumbar IMU. The surfaces demonstrate the sensitivity of total step count to variations in SSD and AYM across all subjects for each trial as compared to Vicon ground truth shown as the black surface (Chapter 5). A clear difference is observed between the linear walking and free-living activities. The black plane demonstrates the mean ground truth step count.

count accuracy to the parameters during free-living activity is much higher. This is due to the fact that a greater proportion of signal represents stepping during linear walking, therefore, if all of the signal is deemed a 'walking bout' then the step count upper limit is determined by the previously mentioned 'minimum time between peaks' parameter.

The intersection of the two surfaces in Figure 8.1 implies that at certain values of AYM and SSD, the observed step count does equate to the Vicon ground truth; however

these values clearly vary between trials. It is therefore useful to view the results as a function of the negative absolute percentage error (NAPE), as demonstrated in Figure 8.2. What this implies is that an optimum value of SSD and AYM is one that results in a total observed step count equal to the true step count, even if this is due to false positives balancing out false negatives. The validity of this decision is discussed in Section 8.4.4. However, given this, the threshold values decided upon are 0.05 and 0.05 for AYM and SSD respectively. As shown in Figure 8.3a this provides a high level of step count accuracy across all the subjects and trials. The reason however, for not choosing a more optimal value based on Figure 8.3a is that the actual optimum varies according to both subject and trial - and these values best provide robustness for future work (see Section 8.4.4 and Figure A.1).

The smartphone is kept in the subject's right trouser pocket, which is generally located high up on the right hip. This is anatomically close to the lumbar region and the pelvis. Therefore the phone accelerometry signal is very similar to that of the lumbar accelerometer (as demonstrated in Figure 8.4). Therefore the same methodology is used to extract step gait features from both devices. However, due to the increased movement compared to the lumbar-IMU, different values for AYM and SSD are required (as shown in Figure 8.3b). Therefore, similarly for consistency, and due to inter-subject and trial variability, values of 0.01 and 0.10 are chosen as the threshold for AYM and SSD respectively (Figures 8.3b and A.2).

8.2.3.3 Foot-mounted IMU

The method implemented to detect steps from foot dorsum-mounted accelerometers is a standard thresholding technique developed from investigating the samples. Similar methods use an acceleration threshold to determine when a foot is stationary [438, 439], however it was deemed more suitable to use the acceleration standard deviation to help reduce the dependence on gait speed and activity.

Here a moving window of size 0.1 seconds is applied to the tri-axial signals and the absolute magnitude of the acceleration vector, $\mathbf{a}$, where:

$$A3M = |\mathbf{a}| = \sqrt{\mathbf{a}_{AP}^2 + \mathbf{a}_{SI}^2 + \mathbf{a}_{ML}^2}, \tag{8.5}$$

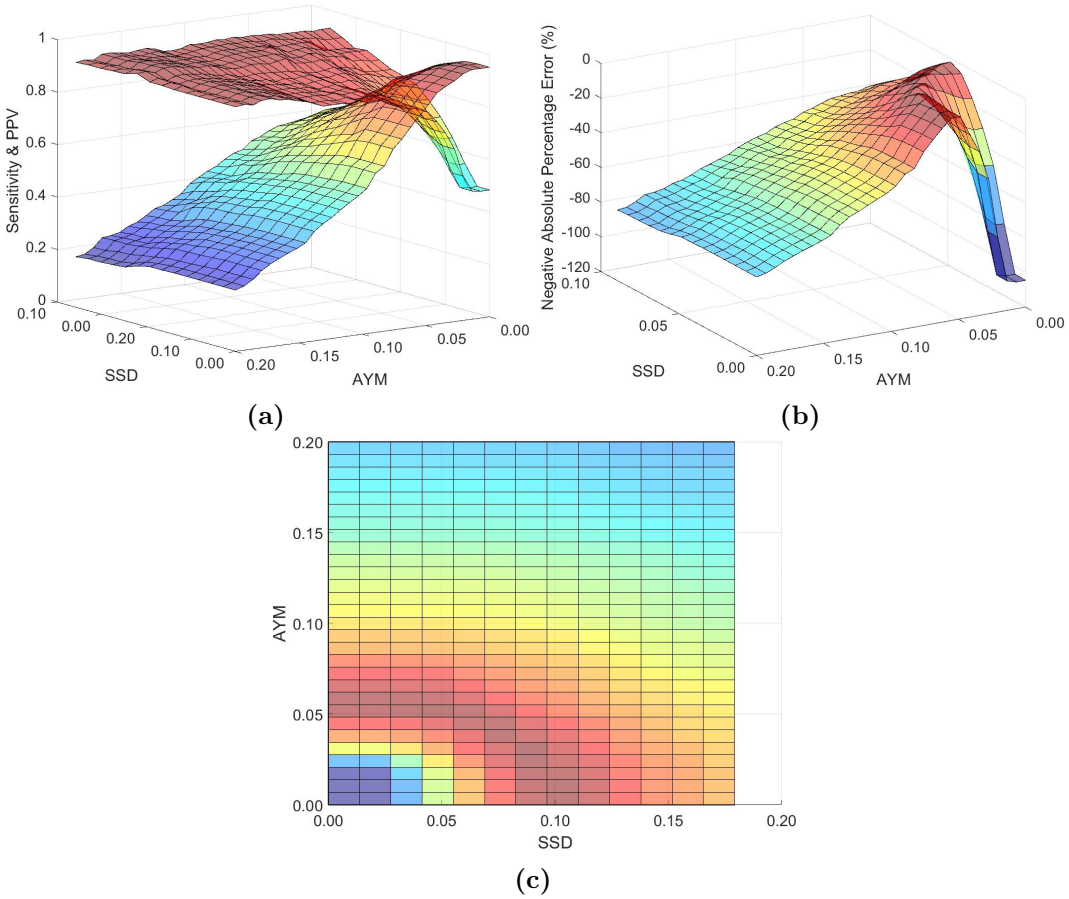
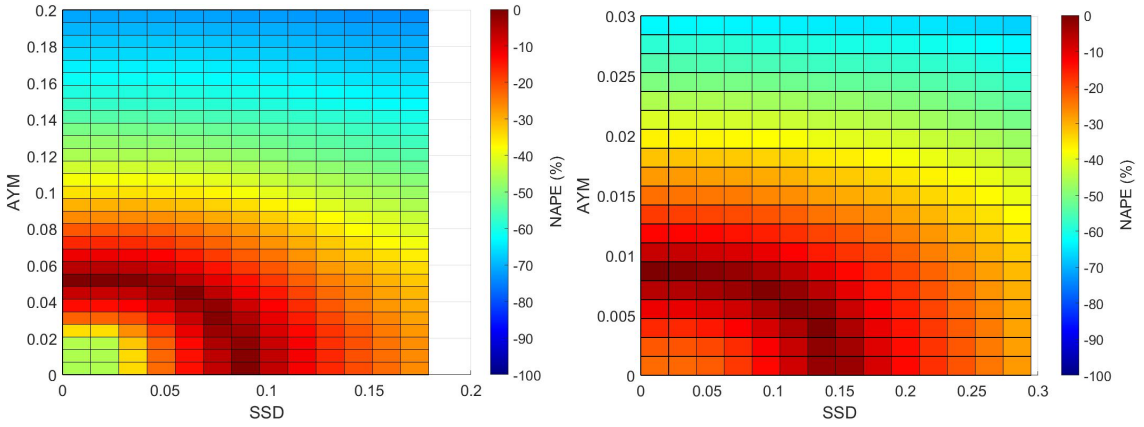


Figure 8.2: Step count percentage error. The figures show the sensitivity of step count to the variables SSD and AYM for the ADL2 protocol. (a) the predominantly red and top surface is PPV, the other is sensitivity. It shows that as they reduce to zero, the false positives begin to outweigh the true positives. This is more concisely demonstrated by the negative absolute percentage error (NAPE) in figure (b) which is most usefully observed from along the Z-axis as in figure (c). Therefore the dark red maxima represent a point where false positives are balanced by false negatives to achieve a correct overall step count.

and the combined sum of the standard deviation in the three dimensions calculated, SSD. This ensures that any dragging of the foot will still be detected as stride movement. This may not be true if only the vertical acceleration of the foot is considered - which has traditionally been a common method for identifying gait phases [440, 441]. Both signals are then thresholded to determine periods of movement.

Fluctuations across the threshold of less than $0.2s$ in either direction are considered to be errata or signal noise as a foot movement of such speed would be incongruous with observed cadences during gait (Table 5.3). Therefore, these are removed. Erroneous fluctuations that drop below the threshold are removed before those that exceed it. All periods above the threshold are then considered strides. This is different to the



(a) Lumbar-IMU **(b)** Smartphone

Figure 8.3: Lumbar-IMU and Smartphone step count sensitivity to AYM and SSD. Sensitivity to AYM and SSD of step count for both lumbar-IMU (a), and a smartphone in the right pocket (b), across all protocols and subjects.

required bout detection for the other devices as it is much less likely that the feet will be moving if the subject is not walking - whereas plenty of torso, leg and arm movement can occur when standing stationary.

Using this method, there is no requirement for bout detection as with lumbar or smartphone accelerometry, as the movement of feet outside of gait is negligible. Therefore, observation of the step count NAPE (Figure 8.5) suggests that a suitable value for the thresholds is 0.5 and 0.5 for A3M and SSD respectively (further justified in Section 8.4.4 and Figure 8.9). For comparison with ground truth and step-to-step matching, the step results for both feet are combined while maintaining discriminatory labelling for each.

8.2.3.4 Smartwatch

Evidence has shown that accelerometers attached to the wrist provide more inaccurate gait features than those on the feet, legs or trunk [442, 443]. This is expected, as the connection of the wrist to the legs actually performing the gait, is through a series of articulations and so the signal is confounded; plus the arms may be performing other activities while walking.

Most methods are dependent on detecting the rhythmic swing of the arm - whereby each arm swings alongside the contralateral leg during locomotion. This movement improves stability and energy efficiency while walking, by reducing the angular

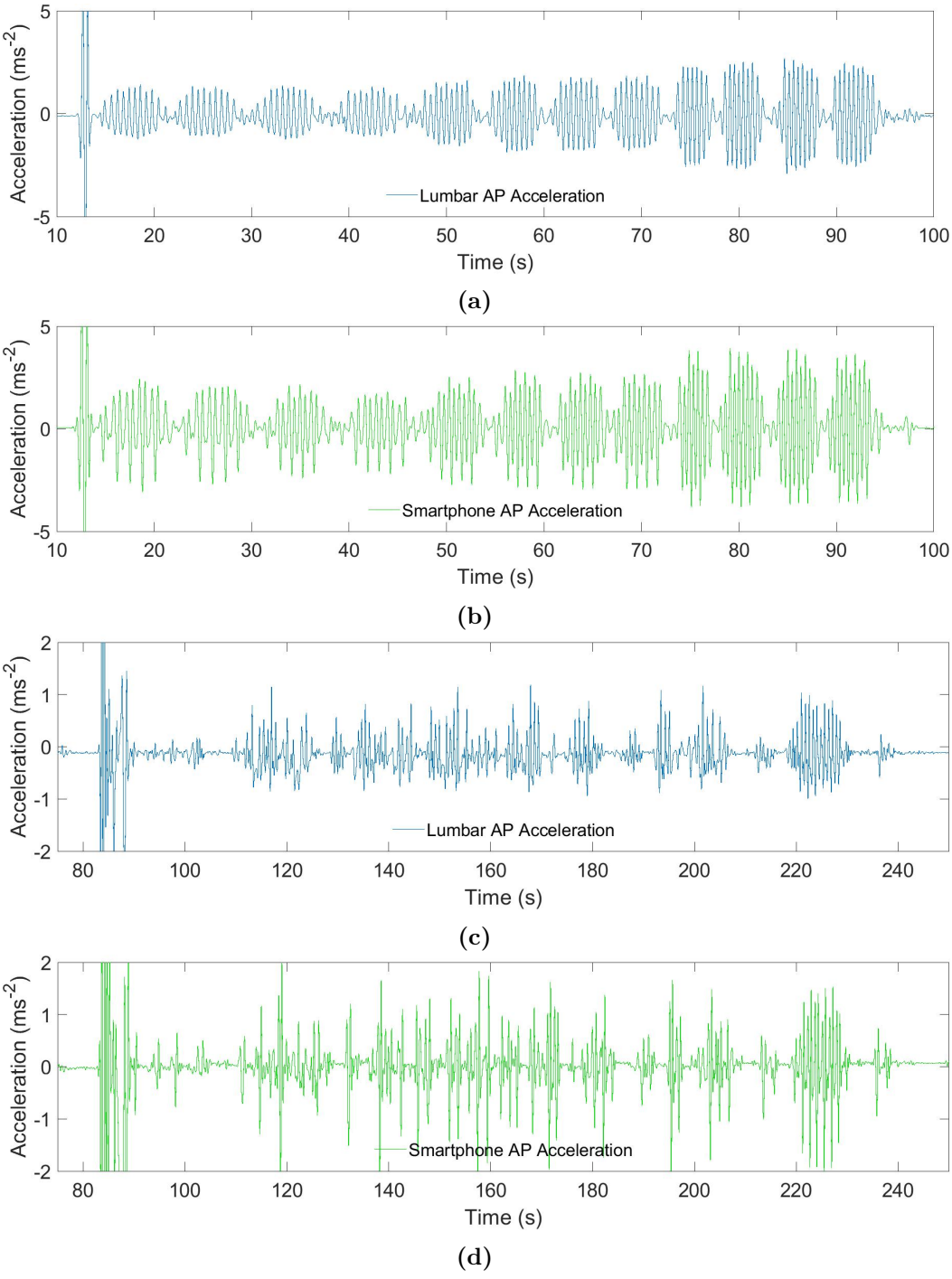


Figure 8.4: Comparing lumbar-mounted IMU and hip-pocket Smartphone AP accelerometry signals. Similarity between the two signals is clearly shown. This is more so in the example of linear walking from the Straight protocol (a-b) than in the free-living walking example corresponding to the same subject's ADL3 protocol (c-d). Noteworthy, but expected, differences are the asymmetry and increased magnitude of the phone signal compared to the lumbar signal. This is due to the lateral positioning of the phone compared to the medial location of the lumbar IMU, and the increased swinging of the legs compared to the spine.

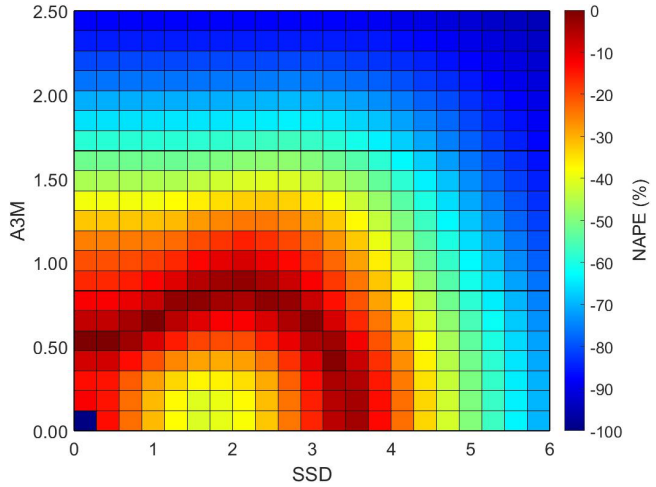


Figure 8.5: Feet-IMU step count sensitivity to A3M and SSD. Foot mounted IMUs show a distinct function of sensitivity when compared to the other devices.

momentum of the body and compensating for deviations in uprightness off the vertical axis, which itself enables a narrower and more efficient gait stance [444, 445]. Due to the confounding movement of the arm, independent of the legs, many of these methods (especially the more recent) use an adaptive threshold and windowing for peak detection in the accelerometry signal; as opposed to static as with the previous devices [424, 446, 447].

The implemented method was one adapted from these other techniques. An adaptive threshold value, λ, for step detection at sample, t, was established as:

$$\lambda = \begin{cases} \sigma\sqrt{\dfrac{\sum_{t-n}^{t+n}(|a_t| - |\bar{a_t}|)^2}{2n+1}} & \text{when } |a_t| > \lambda_{min} \\ \lambda_{min} & \text{otherwise,} \end{cases} \tag{8.6}$$

where,

$$\tau = \frac{2n+1}{F_s}, \tag{8.7}$$

and τ is the window length of 0.25 seconds based on the maximum cadence (see Section 5.3), F_s is the IMU sampling frequency, σ is a scaling parameter, and λ_{min} is a defined minimum possible value of the threshold. The minimum value is required to ensure that small movements, for example while standing still, do not result in step detection.

Similar to the bout detection parameters for the other IMUs, step detection sensitivity to λ and σ needed investigating before assigning values. Figures 8.6 and A.3 demonstrate why the values of σ and λ_{min} were assigned as 1.50 and 0.22 respectively.

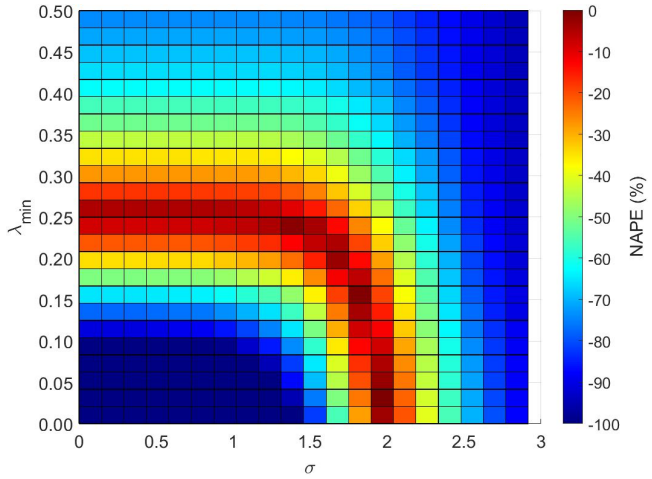


Figure 8.6: Smartwatch step count sensitivity to λ and σ. The smartwatch shows the greatest sensitivity to parameter values of all devices.

8.3 Results

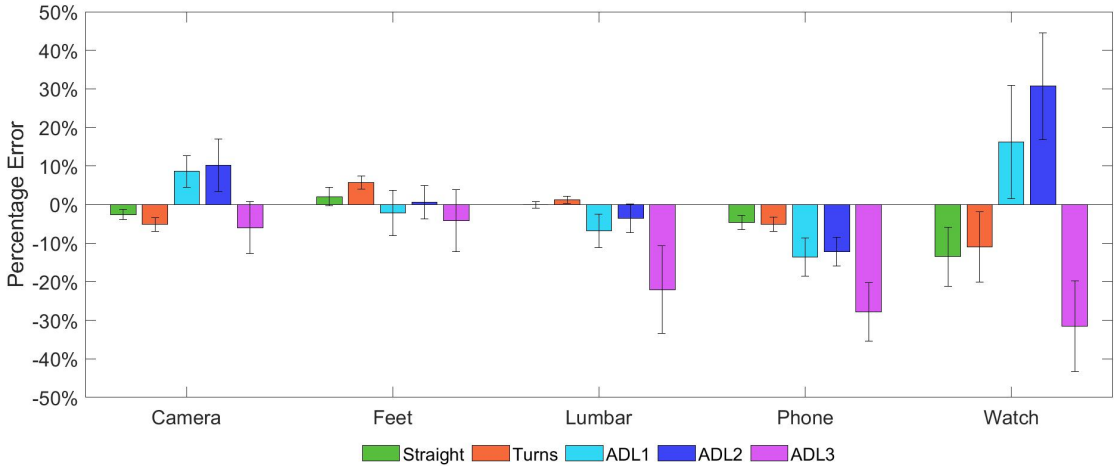
The step count results were calculated for each accelerometry device, as well as using the wearable camera as outlined in Chapters 6 and 7. The results are shown in Figure 8.7. The percentage error against the ground truth is shown, however to help elucidate the causes in variation, sensitivity and positive predictive value (PPV) are also shown.

8.4 Discussion

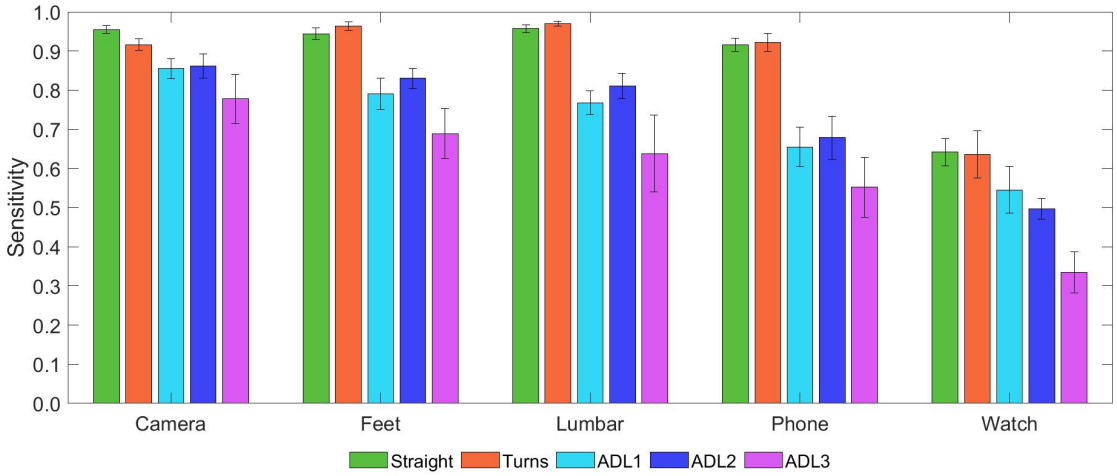
The results show that there is significant variation in step count accuracy between both devices and protocols. The following section explores the causes of these variations, and what these variations mean with regards to orientation, location, device selection, protocol development, subject dependence, parameter selection, and various other parameters.

8.4.1 Inter-device comparison

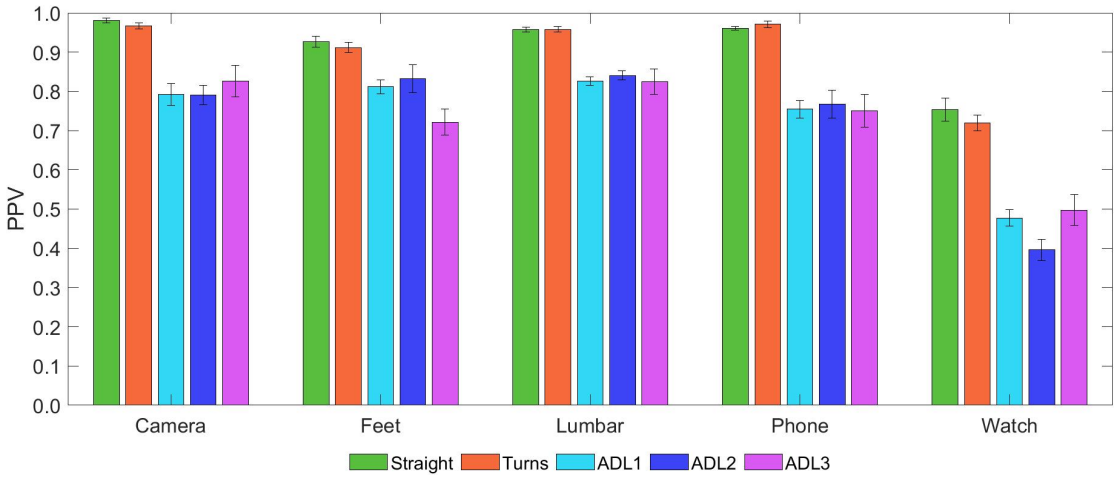
It is immediately clear that step count is most accurate for the foot-mounted IMUs; consistently being below 5% error. This is unsurprising as the movement of the



(a) Percentage Error



(b) Sensitivity



(c) PPV

Figure 8.7: A comparison of step count accuracy of multiple accelerometry-based devices, and a wearable camera-based system. (a) demonstrates the mean percentage error in step count against the ground truth as ascertained by Vicon motion capture. (b) and (c) show the sensitivity and PPV of the step count for each device and trial type, to help elucidate the cause of the variation in percentage error.

feet is that most directly associated with walking; and its superior reliability has been previously acknowledged in the literature [448]. Sensors attached to any other part of the body (even ankle or flank) will be submitted to confounding movements and non-zero accelerations when the feet are stationary. An interesting outcome is the high-accuracy of the foot-mounted IMUs during the ADL protocols - especially when compared to the other devices.

For the foot-mounted IMUs, the true step count is generally within standard error; with the exception of during the Turns protocol. The over-estimation of step count in the Turns and Straight protocols is most likely due to the rotation of the feet in the transverse plane during rapid changes of direction (e.g. during a 180° pivot turn) being detected as a step. This is supported by the fact that the Turns protocol includes a much higher proportion of rapid turns than the Straight protocol. Further, this likely contributes to the relatively high accuracy of the lumbar-IMU during these protocols, as only the actual movement of the trunk is detected. Despite extremely high accuracy in the continual walking protocols ($< 2\%$ error), the lumbar-IMU accuracy substantially drops off during the ADL protocols (ranging from 5% to 22% error).

It is important to note that both the lumbar and smartphone IMUs substantially undercount during the ADL protocols. This is by far worst in the ADL3 protocol which best imitated frenetic movement about a small space. In this scenario, undercounting is likely caused by turning, pivoting and irregular walking; none of which would cause the pelvis to undulate in the distinctive periodic fashion for which most step counting algorithms are designed. This would also account for the undercounting during the Straight and Turns protocols for the smartphone when compared to the lumbar IMU. Here, the inside leg during a sharp turn may move only slightly but the lumbar will still move in accordance with the outside leg. However, the ADL1 and ADL2 protocols also result in undercounting. This is most likely as they tend to imitate short bouts between stationary periods. In this scenario, bouts of sufficiently few steps may be missed, or longer bouts may be divided into excluded shorter bouts by brief pauses or small step movements. This is a documented concern about many of the accelerometry algorithms and devices. In one paper it was demonstrated in ten commercial devices (lumbar, hip and ankle-located); where often a minimum bout length of 6 to 10 steps

was required for the bout to be detected [449]. Undercounting may be ameliorated in future studies by the lowering of threshold values. However, this comes with it's own dangers with the increased detection of non-step movements.

Much like the lumbar and smartphone algorithms, the smartwatch algorithms are generally dependent on detecting the cyclical contralateral swinging of the arms. However, this movement was disrupted by sharp turns (in Straight and Turns protocols), and moving short irregular distances or turning rapidly where the arms are used predominantly for balance (in ADL3). This is most reasonably why substantial underestimation of step count occurs in these trials. The significant overcounting during ADL1 and ADL2 trials might be expected as both of these protocols involve the use of the arms in preparation of food, tidying, putting on clothes, etc. This is often done immediately before, during, or after ambulation - thus confounding and lengthening the true periods of gait.

From Figures 8.7b and 8.7c it can be deduced that for at least the camera, foot-IMUs, lumbar-IMU and smartphone, the general variation in accuracy in the ADL protocols tends to be more from an increase in false negatives than false positives, and in fact the PPV is remarkably consistent across devices. This further supports the assumption that the algorithms struggle to pick up shorter bouts of less periodic movement.

It is also notable that when considering the accuracy of the device to pick up specific step events, rather than just an estimation of the total number of steps (as is vital for the calculation of most further gait features, such as step length), the camera system actually provides the most reliable results. For the accelerometry methods, a higher proportion of false positives are masked by an increased number of false negatives. This is especially true for the smartwatch. This is to some extent a result of the designed algorithms, but demonstrates the importance of investigating individual step events beyond gross statistics.

As a caveat to all of the above discussion, it should be reiterated that the foot-mounted and lumbar-IMUs are of a much higher research grade than those of the consumer smartphone or smartwatch, with an almost 40-fold higher sampling rate (Section 4.3). That being said, consumer devices have been show to be generally in agreement with research-grade devices, even in free-living environments [450].

Finally, it is worth considering how the camera-based system compared to its accelerometry counterparts. Overall it provides accuracy roughly comparable to both the lumbar and feet-mounted IMUs - and much greater accuracy than either of the other two consumer devices. Furthermore, it actually provides more accurate detection of specific step events, rather than just gross step counts. This will be vital in the future for precise clinical free-living gait assessment. The overcounting in ADL1 and ADL2 protocols is likely due to feet being obscured by work surfaces and arms during activities that involve the upper body - causing sudden jumps in perceived location which could be construed as a stride. This is an issue that could be resolved with further work and the integration of methods such as loop closure.

8.4.2 Comparison to published results

Due to the bespoke implementation of the accelerometry algorithms, it is worth considering the results in the context of prior published work. To support the validity of these results more generally, the protocols most similar to those conducted in other studies should render similar results. As previously mentioned, the vast majority of other studies use protocols of regular, linear walking at moderate speeds; whether in a straight line, on a treadmill, or back and forth over a distance of generally more than 5-10 metres. Therefore direct comparison can only really be achieved with the results from the Straight protocol - but it must be remembered throughout that even this involves a much higher ratio of abnormal turning, pivoting, accelerating and decelerating steps to linear walking steps than most protocols. Where previous work using ADL-style protocols has been conduced, these are also considered. Additionally, only validation studies in healthy control adults are considered as this is our subject population - unless results on patient populations provide additional insight.

8.4.2.1 Lumbar-IMU

Using an accelerometer attached to the waist, lumbar or sacrum is the most common method for accelerometry-based gait analysis in research. As such this is where the most work has been done and the vast majority of device-algorithm setups report high accuracy and mean average percentage error (MAPE) of just a few percent

[424]. In fact many purport excellent validity, with errors of below 1%, but again, these are almost always gross step count over short periods of linear, moderate speed walking [210, 451]. In ADL-like or free-living protocols however, the step counting accuracy appears to be similar to those observed, with error between 10% and 30% [14]; if not substantially better [424].

8.4.2.2 Smartphone

Although, as discussed, the upper thigh is similar the the pelvis and lower back, the importance of incorporating the methods into ergonomic consumer electronics means that considerable research has been done on smartphones and in-pocket devices specifically. A range of consumer step counting devices, in-pocket, provide around 90% accuracy but ranging between 75% and 95% with some bespoke algorithms providing greater than 95% [263, 452–456]. Specifically looking at smartphones, we see similar results with consumer devices around $90-95\%$ accuracy and bespoke advanced algorithms substantially improving on that [399, 424, 457, 458].

A confounding issue in using a smartphone is that often it may be moved from the pocket, and carried in the hand (whereas a smartwatch is almost certainly always attached to the wrist). In one of the studies above they demonstrate how this causes the results to more closely resemble the smartwatch results; ranging between 65% and 90% accuracy [399, 452, 457].

Most attempts to imitate daily activities is limited to travelling up and down stairs, standing and sitting, and running. Although these are arguably not complex or representative of true ADL, even then *Dirican et al* got accuracies ranging from 41% to 99% [459].

8.4.2.3 Foot-IMU

Although usually acknowledged as the most accurate method, foot-mounted IMUs are generally the least studied due to the complexity of setup and lack of ergonomics (i.e. they require non-consumer technology and often need syncing to one another if both feet are used). Therefore it should be unsurprising that often studies report greater than 99% accuracy during linear walking protocols. Indeed, both *Jimenez et*

al and *Liu et al* report accuracies of around 99.9% [460, 461]. However the fact that this then drops to sub 90% when tested against even rudimentary ADL such as stair climbing or running, suggests that it has been over-engineered for the linear walking protocol [460]. It is therefore reasonable to conclude that the foot-IMU results obtained from this study are similar to prior work considering the complexity of the protocols. This is supported by *Anwary et al* who demonstrated more robust testing of multiple foot-mounted IMU setups; which resulted in accuracies of around 95% for most setups, and around 90% when the IMU was placed over the toes [462].

8.4.2.4 Smartwatch

As expected, the published results from smartwatches range massively. Some of the most accurate results are from simple protocols and even then consumer devices have been shown to vary from 79.8% to 99.1% [463, 464], and substantially more so from subject-to-subject and trial-to-trial [465]. In fact, it is this wide ranging variability that somewhat validates the quite wild range of MAPE in this study. For example, seven consumer devices under relatively simplistic protocols provided a range of MAPEs from 1% to 40% with most around the 10% mark - in accordance with the Straight and Turns protocol results [424, 454, 466]. On the other hand, under similar circumstances, *Fermiano et al* achieved sub 10% MAPE, even though their subjects were cardiovascular patients [467].

Attempts at ADL or free-living protocols have rendered hugely varying results depending mainly on protocol; with accuracy medians as low as 33%, and many individual subject results in an ADL-like protocol at slow gait velocities demonstrating 0% agreement [18, 263] .

Further validation of this study's results can be found in a similar paper that compared multiple devices and algorithms. The accuracy of the results are highly comparable - although all devices and algorithms slightly under-perform their counterparts in this study [468]. Similarly, although published results are often more accurate than those from this study (even if we exclude the confounding influence of the protocols) the results here are on a step-by-step basis, and this has been shown to reduce accuracy [263].

In summary, although it is hard to directly compare like-for-like given the wide range of parameters, from device and positioning to protocols and subjects, it does however seem reasonable to presume that the results of this study and the bespoke algorithms implemented are comparable to those of previously published works. As such, the results of this chapter can be considered relevant to the corpus of accelerometry-based gait analysis generally.

8.4.3 Algorithm selection

It is acknowledged in the previous section that there are numerous bespoke algorithms that can be used to achieve better accuracy results, although often within a constrained context. Indeed, methods can be found representing all the classic signal processing and machine learning techniques, including but not limited to: peak detection [14, 265, 469], wavelet-transforms [470–475], dynamic time warping [476], template matching [477], Hidden Markov Models (HMM) [478–482], random forests [483], Bayesian methods [484, 485], traditional CNNs [486], and deep learning [487, 488]. Therefore the choice of implementing novel algorithms for this study must be justified. The decision primarily came from simplicity and reproduciblity. Due to the range of protocols, devices, and the focus on the camera-based computer vision method it was important to implement methods that were robust and relatively simple. This necessity is compounded by the fact that many more complex methods lack comprehensive explanation in their respective publications. More advanced methods were in fact attempted, such as the use of continuous wavelet transforms (CWT) [475], however they tended to struggle with complex movement or lacked full explanation. Furthermore, peak detection is by far the most commonly used method and is sufficiently simple that it can run live on most wearable devices.

Even when implementing and adapting established, validated methods, it became rapidly apparent that the figures proposed for parameters such as thresholding values would not work for this study's devices and protocols. It is because of this that new parameter values had to be decided upon.

8.4.4 Parameter selection

As explained in Section 8.2.3, each of the algorithms has been developed to be dependent on two variables, the values of which have been selected to provide near-optimal step count results. An additional consideration was for the threshold values not be over-optimised or too contrived. On initial inspection of Figures 8.3, 8.5 and 8.6, it is clear that the chosen parameter values are not optimised. This is because these figures show all the steps across all the subjects and all the trials. This gives a natural bias towards the protocols with a greater number of steps, or those subjects that move around more than others. Therefore, in selection of the parameter values, similar graphs were created for each device across each trial and each subject independently. An example of this is shown for the foot-IMU algorithm in Figure 8.8, and the same is done for each of the other devices in Appendix A).

Clearly variation between both the trials and subjects is considerable and an optimal combination of parameter values for one may result in very poor results for another. Therefore values were chosen that provided reasonable results throughout without biasing one subject or protocol. The relative importance of this decision for each device is best shown by Figure 8.9.

These figures may also explain why the importance of the threshold values in the literature is rarely discussed. Since the majority of work is done using pelvis, lower-back or hip-located IMUs, and the vast majority of validation studies are conducted using linear walking protocols, it is clear from Figures 8.9a and 8.9c that the perceived sensitivity of the step counting to the threshold would be negligible. However this is clearly not the case for the other devices, or the more complex protocols. This explains the increased prevalence of adaptive thresholds in the literature - although this is likely a bandage for a more profound problem [263, 424, 446, 447]. The severity of the problem is most highly pronounced in the smartwatch during ADL protocols (Figure 8.9h) where relatively small under-, or over-estimations of the threshold have dramatic effects on accuracy.

It may seem somewhat contrived to have optimised for NAPE, especially as in much of these circumstances false positives are being outweighed by false negatives; however, values had to be selected, and generally studies focus on the gross step count

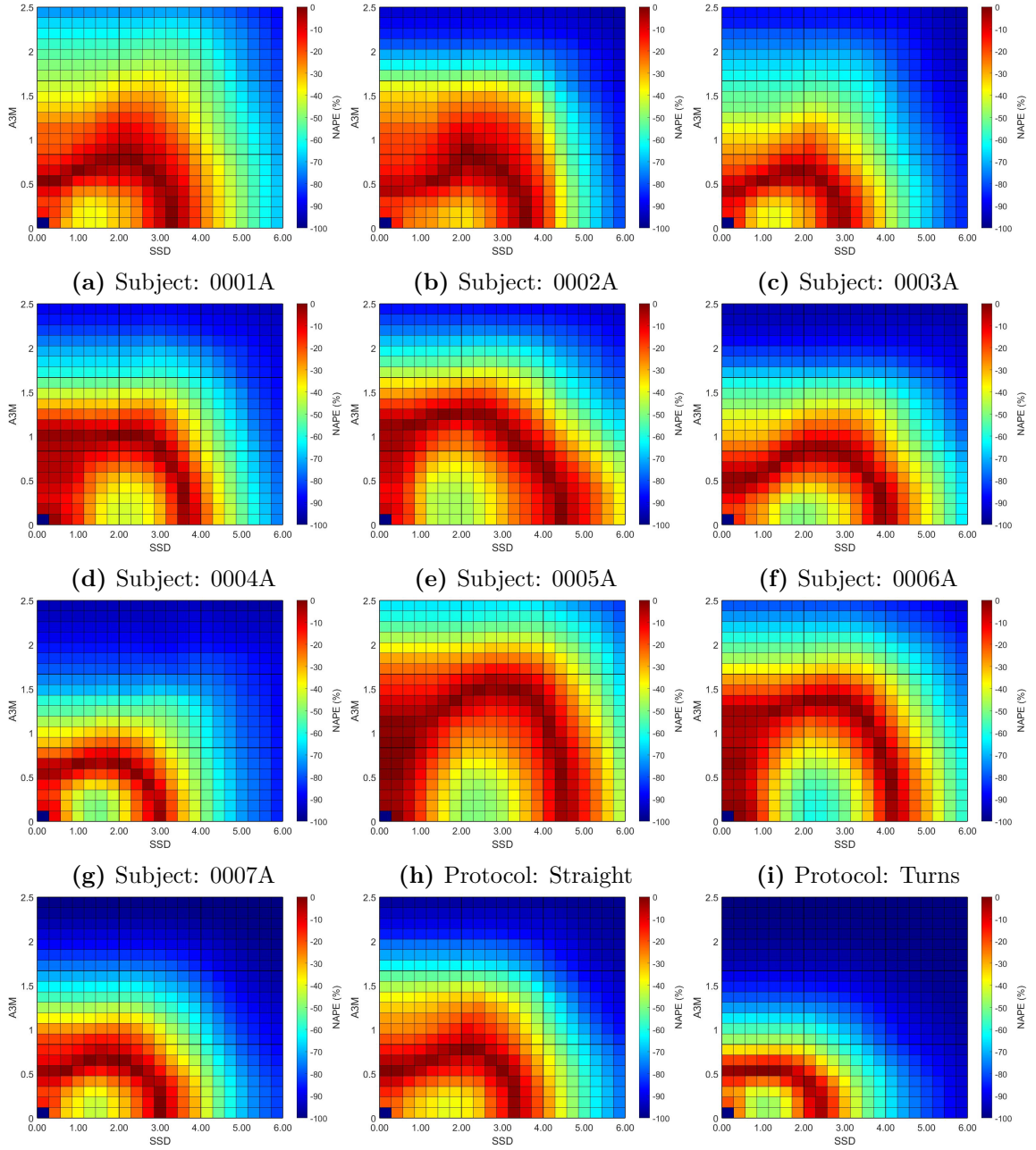
regardless of whether individual detections represent steps or not. This is certainly something that needs investigating in more depth.

Finally, it is also worth mentioning that the bandpass cutoff frequency limits of the signal filtering step were also changed from the literature consensus - although there is variation in the literature as well. They were generally expanded in order to account for a wider range of gait speeds and cadences as might be expected in free-living gait, and in accordance with the definitions in Chapter 5.

8.4.5 The importance of the individual

The literature that validates gait analysis methods generally only reports gross statistics and does not investigate the effects of various subject-specific characteristics or sub-groups within the experimental cohorts. This is remarkable considering the acknowledged idiosyncratic and potentially individually unique biometric nature of an person's gait [125, 489].

One of the key reasons for this oversight is the small cohort sizes that are used in most validation studies. This is also the case with this study, however out of interest, provisional insight is provided by Figures 8.10 and 8.11. Figure 8.10a demonstrates the accuracy of step counting regardless of device. This suggests that perhaps there are characteristics about the individual that heavily affect the results. Subjects 0004A, 0005A and 0007A for example, are remarkably consistent, with two generally undercounting and the other overcounting; regardless of protocol. The other subjects have great variability; both between protocols and each other. Explanations could include, for example, that Subject 0002A and 0003A performed the tasks all with greater exuberance than the others. This would potentially result in a clear and consistent signal for the Straight and Turns protocols, increasing accuracy, but more sporadic, frenetic and thus noisy signal for the ADL protocols. Similarly, Figure 8.10b demonstrates accuracy regardless of protocol. Here we can see that the watch especially, varies between individuals - possibly the result of some individuals being more or less expressive with their arms while talking, or more frenetic in their ADL activities that use the arms. An additional interesting observation is that although, as expected, the lumbar and phone results are generally consistent with one another



(a) Subject: 0001A **(b)** Subject: 0002A **(c)** Subject: 0003A

(d) Subject: 0004A **(e)** Subject: 0005A **(f)** Subject: 0006A

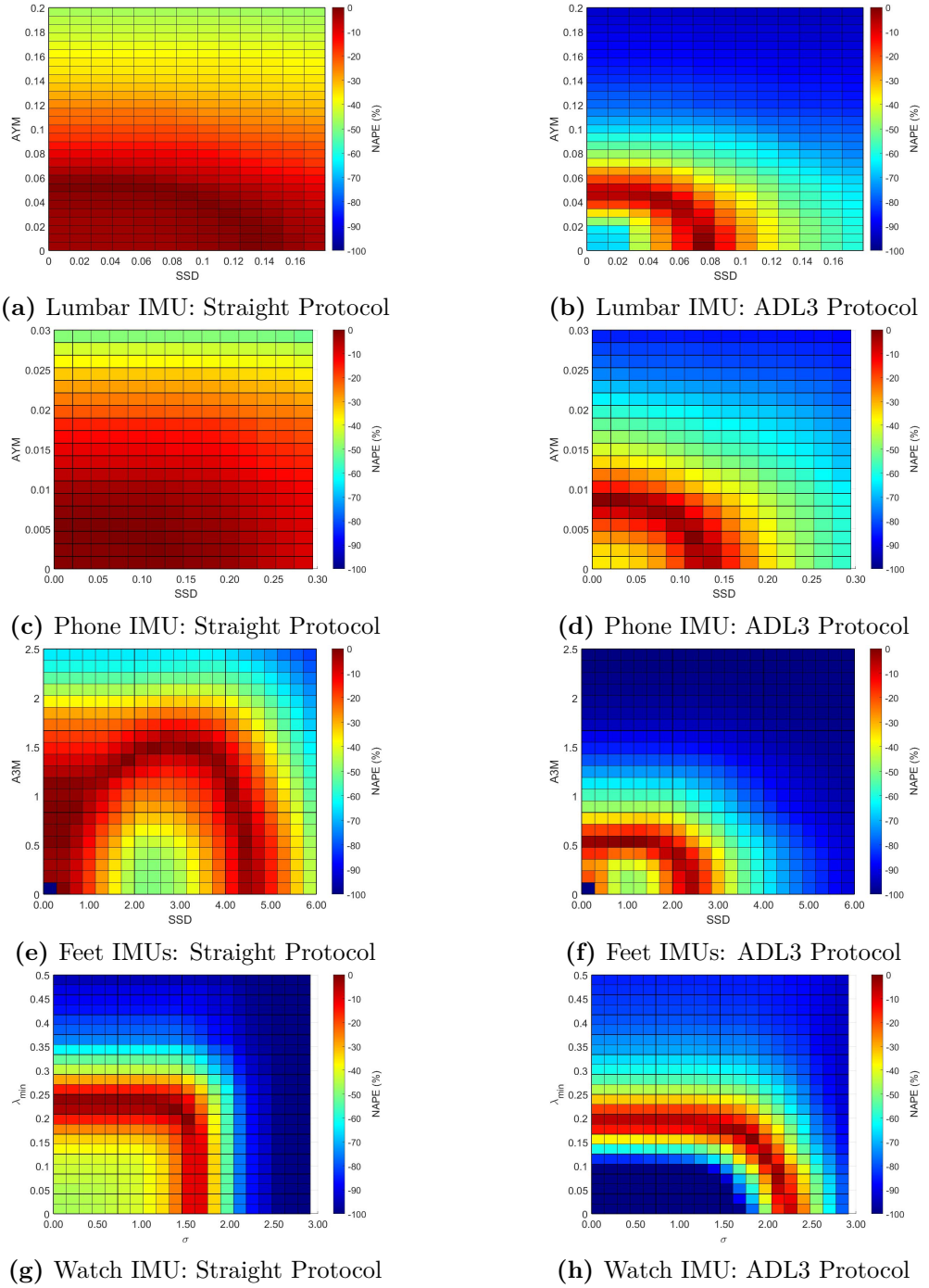
(g) Subject: 0007A **(h)** Protocol: Straight **(i)** Protocol: Turns

(j) Protocol: ADL1 **(k)** Protocol: ADL2 **(l)** Protocol: ADL3

Figure 8.8: Threshold sensitivity for foot-IMU algorithm. The sensitivity of the step counting algorithm for the foot-IMU is demonstrated by the negative absolute percentage error across a range of values of both A3M and SSD

as assumed in Section 8.2.3.2 while developing the methods; this is not the case for subject 0006A. This is possibly due to the phone not being rigidly attached to the hip; for example in a loose-fitting trouser pocket.

All these explanations are entirely speculative, and are simply proposed as conjecture partly based on qualitative observation of the variations in subject's movement during



(a) Lumbar IMU: Straight Protocol

(b) Lumbar IMU: ADL3 Protocol

(c) Phone IMU: Straight Protocol

(d) Phone IMU: ADL3 Protocol

(e) Feet IMUs: Straight Protocol

(f) Feet IMUs: ADL3 Protocol

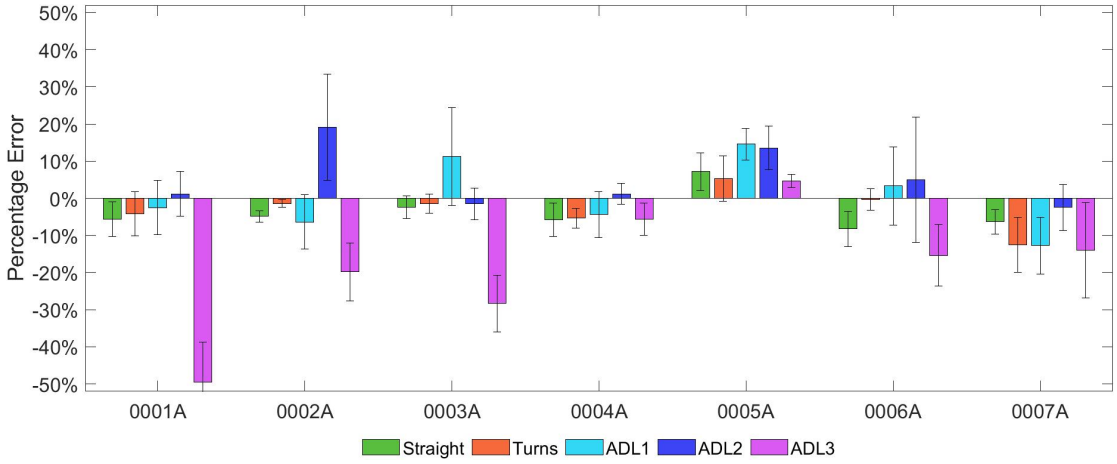
(g) Watch IMU: Straight Protocol

(h) Watch IMU: ADL3 Protocol

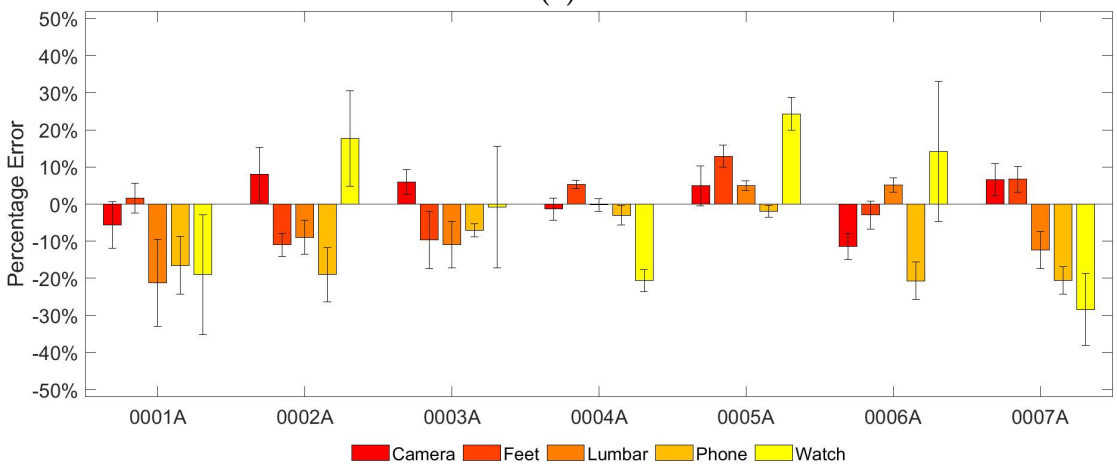
Figure 8.9: Threshold sensitivity for all devices as compared between the Straight and ADL3 protocols across all subjects. Differences between Straight and ADL protocols are clear.

the experiments. It is hard to say without comparing the results to both physiological characteristics and metrics of personality and movement-style (the latter of these could perhaps utilise energetic measurements as a proxy [490, 491]). Although the number of subjects is small, some characteristic data was gathered and the results

compared against this - and although conclusions cannot be drawn, comparisons are still worth investigating. One hypothesis might be that an individuals height or limb length affects the results; as shorter limbs would require faster changes in direction of movement to achieve similar speeds. Figures 8.11a and 8.11b show that there does not seem to be any clear relationship between subject height and either the accuracy or uncertainty in accuracy.



(a)



(b)

Figure 8.10: Dependency of step counting accuracy on subject. a) demonstrates the consistency in step counting accuracy between subjects regardless of device, and b) regardless of protocol.

It has been long acknowledged that both age and gender have a significant impact on gait variability and features generally [126, 492, 493]. Indeed, both *Riaz et al* and *Van Hamme et al* have shown the predictive power of gait accelerometry in estimating these characteristics [494, 495]. Therefore these are also explored in Figures 8.11c -

8.11d and Figures 8.11e - 8.11f respectively. There is a slight suggestion that age does indeed contribute to an increase in the accuracy of step detection, however it is by no means the whole story as is reinforced by the 44 year old subject. No conclusions can be drawn from a cohort this small, but it would make sense as a contributory factor in step counting accuracy. This is because gait variability tends to increase with age [496], and the accelerometry-based step counting methods tend to work best with consistent, repetitious gait cycles. Figure 8.11e suggests that gender does not seem to play a roll in the accuracy of the step counting. Difference between genders in this graph are arguably the result of the threshold of the definition of outlier, which is significant with so few data points per group. This is supported by the similarity in error variability as demonstrated by Figure 8.11f

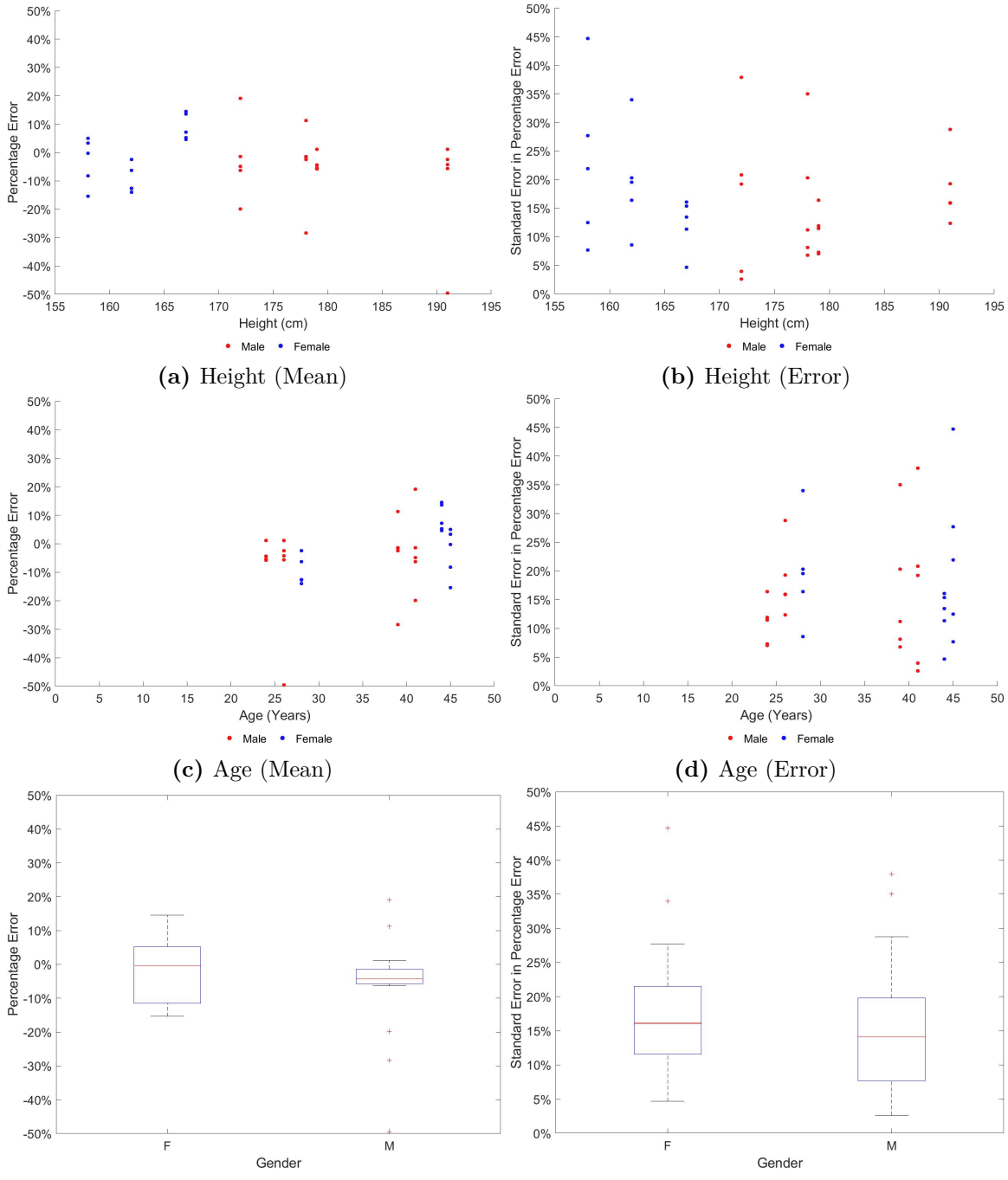
This focus on the individual highlights another issue: how to express and thus account for the idiosyncratic nature of an individual's gait. Some progress has been made in this area, especially by using DNNs (deep neural networks) [137]. Use of such methods may enable an easier representation of the variation and nuance given the complexity and lack of robust definitions; especially in a free-living environment (as outlined by Chapter 5).

Further work beyond this thesis may wish to look at relationships between the optimum parameter values as discussed in Section 8.4.4 and various subject characteristics. It is clear from Figures 8.8a - 8.8g that there is substantial variation; even amongst this small cohort of healthy individuals of similar ages.

Overall, it seems that personalised algorithms, that account to some degree for an individual's physiological and movement-style characteristics may yield more accurate gait analysis from accelerometry signals.

8.4.6 The importance of ADL protocols and free-living gait

It is clear from the results that non-linear walking, short bouts and ADLs have a significant impact on the accuracy of step counting. It is therefore reasonable to suggest that methods only validated with simple protocols should not be assumed validated beyond that, and especially not for use in free-living environments. The importance of this has been increasingly recognised in more recent literature [5, 75, 478, 497]. In



(a) Height (Mean) **(b)** Height (Error)

(c) Age (Mean) **(d)** Age (Error)

(e) Gender (Mean) **(f)** Gender (Error)

Figure 8.11: Impact of subject characteristics on step counting accuracy. In Figures (a) to (d) each point represents a single trial, and the figures show the percentage error, and standard error in the percentage error against both subject height and age. Points are coloured by subject gender. The impact of gender is independently demonstrated in Figures (e) and (f).

acknowledgement of the issue, bodies such as the Mobilise-d consortia have undergone efforts to homogenise and unify ADL protocols to ensure more robust applicability [208].

Unfortunately, few studies incorporate both for comparison, however those that

do clearly highlight the issue of only validating with basic protocols. For example, *Genovese et al* acknowledged the importance of the limitations of moderate speed linear walking. They demonstrated that although a consumer smartwatch, lumbar-IMU and their bespoke watch algorithm all performed well during moderate speed linear walking, with a MAPE of below 3%, during slower or intermittent walking the accuracy massively dropped. This was exacerbated even more so with their single ADL-like protocol where MAPE dropped to $17 - 61\%$. [424].

The stark contrast between linear walking and ADLs is best demonstrated in this study by the lumbar-IMU results. If an investigator was to only subject the device to linear walking then it would seem fair to ascribe the device and algorithm as having extremely high accuracy. This then results in further studies using the device-algorithm combination as a validated ground truth in free-living studies.

Another important feature of the results is the variation in both accuracies and optimum threshold values between ADL protocols. It is clear that even between these relatively similar activities, recorded in the same limited and somewhat contrived manner and space, ADL type has a huge impact. Much of what likely causes this variation between ADLs has been discussed before, and includes: rapidity of movements, use of arms, turning, bending and changing stance. One of the key elements however, and one that is relevant to most ADL types, is the bout length and definition. Many traditional accelerometry methods fail to pick up short bouts of under 6 to 10 steps, or may split up longer bouts into undetected shorter bouts if the ambulation is sporadic. This has little impact on, for example treadmill walking, but can render the methods almost useless in activities around the home where very short bouts have been shown to be the majority [260]. This problem is understandably exacerbated even more so in elderly or clinical populations [5, 478, 498].

Overall, this suggests that future work should endeavour to truly validate methods by conducting trials in a wide range of free-living or ADL-style environments. This is not something that is seen in the literature. When looking across the corpus of the literature that claims ADL protocols, we see a collection of activities that are both inhomogeneous and limited in scope. More often than not, activities of daily living will be limited to simple isolated actions taken out of context such as: standing,

sitting, stair climbing and descending, different speeds of walking, and sometimes obstacle avoidance [423, 425]. Even when outside of these limited motions and when the importance of ADLs has been acknowledged, protocols can seem questionably representative of true free-living ADLs; such as walking along a specified complex path, up and down a corridor, or arbitrarily through a field [465, 466].

8.4.7 Sensor location & orientation

The importance of sensor location and orientation on the body has always been recognised [419–421]. This can even be observed within a relatively small and rigid body-area. For example, *Anwary et al* compared the stride detection accuracy of accelerometers located in five different locations on the foot, over the: metatarsal, proximal phalange, side metatarsal, talus and Achilles tendon. Results ranged from 86% at the proximal phalange to 100% for the Achilles tendon, with the rest at around 96%. All this variability was observed over only 25 steps of linear walking [462].

This is especially important when considering that the lumbar and flank pocket were deemed sufficiently similar to utilise the same algorithm - however this of course is of limited validity. For example, very loose pockets, or highly asymmetric gait exacerbate differences in the algorithm efficacy for both placements.

In this study, the effect of variation in sensor location and orientation was ameliorated by transforming the signals to be orientation independent. Although the algorithm to achieve this was based on one used in previous work, the key assumption - that the direction in the plane with the highest variance in acceleration represents the main direction of movement - may not hold true; especially during ADLs [428]. It has however been used in heading estimation for in-pocket smartphones demonstrating an accuracy of $5.0 \pm 2.5°$ under simple linear walking conditions [428, 499]. Future work should include robust testing of such algorithms in a motion capture lab. Regardless, it seems that more advanced de-orientation methods, and perhaps those using machine learning, are likely the best solution to this problem - rather than testing devices and algorithms in evermore orientations and to some extent locations [419]. If this is the case, then the step detection and gait feature extraction algorithms need to be designed specifically for these de-oriented signals - which is not generally the case at present.

8.4.8 Other methods, limitations and uncontrolled parameters

It is important to acknowledge the greater context of this work, and that the experimental methods here are only a small segment of the wide variety that have been explored for providing accurate gait analysis - even from these specific devices. Even when using the same experimental set up; window size, sampling rate and bandpass frequencies may be varied. More importantly, there are a range of other possible methods for extracting the gait features from the accelerometry signals. For example increasingly promising results are being found with the use of various machine learning methods [500, 501], including: CNNs [502, 503], DNNs [488], and LSTMs (long short term memory networks) [504]. Similarly, more of the MEMS sensors within the devices could be used. Promising results have been attained using magnetometers [505], gyroscopes [423, 461], and combinations of the two [506]. Two slight problems with these methods are gimbal lock [507], and the impact of magnetic interference, especially in urban or indoor environments [508, 509]. It is therefore likely that the best solutions will eventually utilise all the available sensors, and combine them using sensor fusion [510, 511], as well as by using multiple sensors in tandem [512]. This of course however brings about difficulties in on-board processing capacity and device syncing.

Beyond methodological limitations, here we have only considered the most common research (lumbar and foot) or consumer (smartphone and smartwatch) locations and devices. There is however, a vast range of others that have been looked into. For example, head [513], ear (utilising the IMU in some earphones) [514], knee [515], ankle [516], and chest [517] - to name just a few.

Crucial considerations in the adoption of these methods in truly free-living environments are how the sensor is attached to the individual. Common carrying methods have been shown to affect results, including: in a shirt pocket, front or rear trouser pocket, arm band, hand bag, belt carrying case, or in the hand [422, 518, 519]. Even the tightness of the pocket must be considered [123]. The effect of a smartphone being in an individuals hand while walking must be seriously considered given the relative inaccuracy of the smartwatch results. This is increasingly relevant as people use their smartphones while walking more and more [520]. Similarly, and predictably from the impact of ADL1 and ADL2 protocols on watch step counting, the dominant handedness of the individual

and on which arm they choose to wear their watch will also heavily affect results [394, 521, 522]. The impact of this could not be investigated with the cohort acquired.

The importance of walking speed has been well documented, with the biggest impact being at very slow velocities [187, 523]. For a sense of severity, *Benzuidenhout* found that step detection accuracy ranged from $12 - 78\%$ at below $1.0ms^{-1}$ to greater than 87% above $1.0ms^{-1}$ [524]. This is by no means a unique example either [525]. It is worth noting however, that even the slowest speeds tested in any study for step counting are still substantially faster than the minimum decided upon in this thesis (Section 5.2.4), and about which the camera and accelerometer algorithms were developed (to ensure robustness in all scenarios). Robust and truly correct validation of a method must therefore include slow gait speeds before use in free-living environments. Indeed, it is likely that the reduced gait speed observed in the ADL protocols is a major source of their increased inaccuracy. The impact of gait speed could be directly investigated using the data collected in this study; given that participants were asked to walk at three different speeds for the Straight and Turns protocol. Although we would still not see the full range of walking speeds, especially the more problematic very low or high speeds, this could be interesting future work. It is also at the highly problematic low gait speeds that the wearable camera-based system could be the most useful by comparison - as movement is directly inferred from the camera's frame-to-frame relationship to the ground. Figures 6.5a and 6.8a show that the camera-based system performs consistently well at measuring translational and rotational speed at slow gait paces (crucial elements of the step detection system).

Finally, it was outside of the scope and resources of this study to consider pathological or clinically relevant gait. However, much of the wearable-based gait analysis research is conducted from the context of wanting to help monitor, diagnose and help-rehabilitate those with indications that result in some form of gait pathology. Therefore an important parameter not considered in this research, and rarely in validation studies generally, is the impact of such gait pathologies. This is most likely due to the added ethical complications and resource demand of such studies, but those that have been conducted show that many of the devices provide substantially worse results. With MAPEs as consistently high as $60 - 80\%$ for wrist and shank

work devices in subject who had undergone lumbar fusion surgery [526]. Similarly, accuracy has been shown to be worse in in-hospital and post-discharge cardiac disease patients than in healthy controls (although it is noteworthy that the subjects were not ideally age matched in this study) [527].

8.5 Conclusion

Overall, this study raises a number of questions and issues with regards to the validity of accelerometers to measure gait in both clinical and free-living environments. On one hand, foot-mounted IMUs provide robust step detection even in imitation free-living environments (never greater than 6% error). On the other, even heavily validated methods show high inaccuracy when tested against ADLs; let alone smartwatches and smartphones which regularly ranged from 10% to $>$ 30% error. Fundamentally the key take-home message is that step detection, and to a broader extent gait analysis technologies, must be tested more robustly in order to provide generalised validation. This must include a great breadth of free-living-like protocols (well beyond what is tested here), as well as sensitivity to exact sensor location and orientation, and subject-specific characteristics. Additional insights from this study are the need to further investigate the impact of parameter selection in algorithms and to tease apart the difference between gross gait features and step-to-step gait features.

The issues discussed in this chapter highlight the biggest overall concern in the field, which is the complexity of the problem at hand - and the difficulty of comparing studies due to the great number of impactful but uncontrolled, or unaccounted for, parameters. There is precedent for comparing algorithms, devices and locations in a similar manner to this study; and the accuracy results are highly comparable. However, all devices and algorithms slightly under-perform their counterparts [468]. It seems from this paper that device placement is more salient than window size, which itself has a greater effect than sampling frequency. This is reassuring and the reduced accuracy likely due to the parameter optimisation conducted as part of this work (Section 8.4.4). Finding a way to compare between protocols considering the importance of each of the outlined parameters will be tricky, this is demonstrated by the range of values provided exclusively in studies highly constrained to treadmill walking [454].

Despite all this, it seems that foot-mounted IMUs are likely the best device for free-living step detection and thus probably also gait analysis generally. However, wearable camera-based methods are certainly competitive, especially when: improved beyond this implementation; considering slow or intermittent gait; or desiring devices that are commercially available and easily set up.

Finally, a consideration of the results in the context of the literature corpus leads to a inevitable conclusion that an additional confounding factor can be identified in the rapid adoption of such technologies [528]. Consumer devices will continue to evolve and improve in ways including: on-board processing power, storage, sensitivity, connectivity, size and ergonomics. Although this is commonplace in science, it is especially problematic here due to the pace of adoption and evolution, combined with the validation methodologies being so sensitive to device and setup variables. This will result in already questionable validation studies becoming rapidly outdated. Therefore, there should be a push to a generalisation of the methods rather than the ever-aspiring improvement in setup-specific accuracy.

Looking forward, this plethora of opened questions provides fertile ground for future work. Although this is discussed in greater depth in Chapter 9, it includes: improving algorithms, generalising algorithms, utilising machine learning, investigating the impact of subject characteristics and parameter selection, broadening and diversifying the battery of lab-based ADL activities and subject population, identifying ways to compare studies, independent validation of the de-orientation algorithms, identifying what causes under- and over-counting, improving methods for step-by-step analysis rather than just gross step statistics, and investigating more gait features beyond step detection and count.

9 Conclusions & Future Work

Contents

9.1 Summary of results & contributions 205
9.2 Future work 208

The aims of this work were to investigate the reliability of gait analysis tools beyond standard restricted validation protocols, define new protocols and methods for more robust and broadly applicable free-living gait analysis, develop a computer vision-based method for extracting gait features from a wearable camera, and compare this method to a range of wearable accelerometer devices. Each of these aims was confronted and this chapter summarises the key results and contributions obtained during this process. Opportunities for further work are then explored and discussed.

9.1 Summary of results & contributions

This work can be broken down into three key parts that follow the chronological order of the thesis.

9.1.1 Defining and assessing complex gait

Before new technological techniques of gait assessment could even be developed, it was necessary to construct more robust methods of defining and assessing complex gait. The majority of validation studies utilise motion capture as their ground truth. Therefore, in **Chapter 4**, a novel set of protocols is proposed to test wearables in a way comparable, but more thorough, than previous studies, while also emulating free-living activities. The Turns protocol adds a broader range of turn angles and

speeds than previous protocols, highlighting the need for a more universal definition of turning. The activity of daily living (ADL) protocols simulate some of the key components of everyday movement. Together, these protocols constitute the Oxford study, whereby seven individuals undertook each protocol while also wearing an action camera, smartwatch, smartphone, foot-mounted IMUs, and a lumbar IMU. Therefore this work has generated a novel data set that can be used to compare these various devices as well as assess the validity of ADL protocols in a motion capture lab.

In order to assess gait during ADL protocols - and by extension free-living activities - it is necessary to develop a new, robust and universal set of definitions for gait features that can be used for any individual in any circumstance. This is done in **Chapter 5** for walking bouts, strides, turns, stride length, step length and step width; with consideration given to: edge cases, compound movements, and pathological gait. As part of this process new concepts were explored, such as: the minimum possible walking speed, a foot velocity model, and issues with defining strides by gait cycle events such as toe-off or foot-strike.

By applying these definitions to the motion capture data from the study protocols, it was elucidated that gait features exist on a spectrum - and as such threshold values within the definitions have significant effect on results. This is a substantially under-acknowledged issue in the field and may considerably affect inter-study comparability and validity.

On a more pragmatic front, it was also shown that gait features can be extracted from ADL protocols with high fidelity in a motion capture lab. This being said, confounding issues and considerations for future such studies are also highlighted. Interestingly, the study also illustrated concern with the reliability of motion capture to accurately assess all movement. The accuracy of instantaneous measurements of movement and speed are questionable at high sampling frequencies and this issue is heavily compounded by additional calculations such as those required for orientation or instantaneous angular velocity. This is interesting considering its somewhat indisputable reliability. Similarly, the Vicon software provides limited suitable options for gap-filling during complex movement.

With the protocols and definitions outlined and explored, two new and adapted graphical concepts were introduced in **Chapter 3** to aid with the forthcoming analysis and discussion of complex gait; namely step maps and GET graphs. These are especially pertinent for step-by-step gait analysis - which is assisted by the use of linear assignment.

It is also important to note both the similarities and differences between the definitions proposed in this work, and those proposed from expert consensus [208]. Some key differences are apparent, such as the minimum number of strides in a walking bout or degrees in a turn. The impact of these choices needs to be assessed, and a conclusion come to. However, the fact that similar protocols and definitions are being suggested and pushed for alone provides validation in the value of this work.

9.1.2 Gait analysis with a wearable camera

With the protocols and free-living gait feature definitions established, a novel computer-vision based method was established for extracting gait features from videos collected from a wearable camera. The first step of this was in the extraction of the egomotion of the wearer using visual odometry. This is achieved in **Chapter 6**, where it is shown that walking and turns can be detected and measured with relatively high accuracy across a range of movements. Turns are detected with precision and recall of 0.94 and 0.91 respectively when assessed across all protocols, and turn speeds are consistently underestimated by approximately 10%. Walking speed is estimated with lower accuracy and greater variability due to the scale estimation element of visual odometry.

Chapter 7 then focuses on the next step: detecting the wearer's feet in the video samples. With these two information streams established, step maps can be built, and gait features extracted. This resulted in step detection precision ranging from ~ 0.80 in ADL protocols to ~ 0.98 in linear gait protocols. However, the accuracy of gait features was then dependent on the compounding inaccuracy of the various up-stream processing steps. This resulted in feature measurement errors varying massively; from the stride lengths being comparable to other methods during the Straight protocol, to the step widths being completely unreliable during ADL protocols.

Explored in both of these chapters are in-depth considerations of the particular problems and nuances when attempting to perform gait-related visual odometry and

object detection on videos produced by a wearable camera in both the gait lab and a free-living environment. This is important as an increasing number of studies are using these methods.

9.1.3 Accelerometry-based wearables during complex gait

To provide context, and investigate how accelerometry-based wearables stand up in the study protocols, **Chapter 8** describes the implementation of algorithms to extract step count from the study smartwatch, smartphone, lumbar-IMU and feet-IMUs. The most pertinent result is that success during linear protocols does not confer ubiquitous validity to ADL protocols (or thus free-living gait). At generally greater than 95% accuracy, results for the Straight and Turns protocols were similar to published results conferring applicability of the methods to the research corpus. However the lumbar-IMU, and especially consumer devices, were greatly inaccurate during the ADL protocols. This is particularly pertinent given the ever-increasing adoption of such devices. Furthermore, the accuracy varied substantially by individual and protocol. A range of insights into why this may be are provided as well as analysis of the importance of algorithm thresholds on the results.

Although comparison studies between multiple devices exist, rarely are they incorporated into the validation studies, or are assessed step-by-step. What is observed here is that the wearable-camera computer vision system is comparable to the research-grade IMU results, and in fact supersedes them when considering step-by-step analysis. It is superior to the other consumer devices even when only considering gross step counts.

9.2 Future work

Given the new methods developed, and the novel data set collected, there are plenty of potential avenues for future work that could be investigated given more time and resources.

9.2.1 Expansion and additional use of the study data

There is a wide range of additional investigations that could be relatively easily conducted given the novel data set. An important and relatively simple piece of

work would be on how gait features differ between ADL and linear protocols - this is already provisionally shown throughout.

As is clear throughout, the ADL protocols are not only different to the constrained protocols, but to each other. Therefore a crucial next step would be the expansion of the battery of ADL protocols to include as many true free-living activities of daily living as possible to ensure validation is genuinely representative - this is particularly necessary when considering patient populations. This could also potentially involved the expansion to turns and movements outside of walking bouts. To aid comparability to other studies, it would be worth conducting these protocols with both this work's reduced marker set, and more traditional marker sets, and comparing. Similarly, it would be important and interesting to expand the subject cohort to be larger, more diverse, and include gait pathologies. This could be combined with investigating the importance of the individual variation both for each of the wearables and in ADL protocols generally, including concepts previously unexplored such as vigour of movement and freneticism.

It would also be useful to compare the results of the study using both the gait feature definitions here, and those proposed from a expert consensus; as well as using the Vicon data to investigate the robustness and sensitivity of the consensus definitions [208].

An additional area of particular interest is in the use of the Vicon ground truth to label gait features or spatio-temporal movements on the accelerometry signals. This could then be used to train deep learning models to better assess free-living gait.

9.2.2 Further development of the wearable-camera system

The wearable camera system has shown great promise, even in this initial attempt at automatically extracting gait features from video. First and foremost, due to its novel nature, further similar studies should be conducted in order to test the retest reliability of of both the experimental and computational methods. This would be especially true of investigating the more generalised applicability of the computer vision methods to other similar video recording devices. However, given its superior step detection when compared to other consumer devices, improving and expanding the system is certainly recommended. This is reinforced by the ever increasing use of wearable cameras in real-world applications such as policing and mining, and as a

free-living gait ground truth [14, 529]. These trends will only continue as cameras are miniaturised, memory expanded, and battery life lengthened.

It would be important in future studies, including those retesting the reliability, to account for the degree of gap filling that was required in this study. This could ameliorate it by reducing the protocol space or increasing the capture volume so as to create a greater buffer for variability in subject movement during turns and the beginning and end of protocols.

Expansion of methods could be: in the use of loop closure to improve the movement estimation, by expanding on the gait feature set extracted, and of object detection and scene recognition to provide more contextual gait information such as: surface, location, lighting, weather conditions, shoe-type, and many more. After all, the biggest benefit of video is its richness in information. Similarly, if the system was to utilise transfer learning and initial calibration for foot detection, it would likely substantially improve foot detection positive predictive value - and thus all gait feature results.

An important improvement for the more widespread use of the camera-based method would be in speeding up the analysis pipeline. Currently, as accuracy was deemed more imperative than computational burden, the processing time is many fold the sample length. Ideally, the whole system could be run on a mobile device, potentially even live to provide instant analysis.

9.2.3 Expanding on the multiple accelerometry device comparison

A clear expansion of the accelerometry work would be in the addition of many more gait features beyond step detection. However the options are manifold, including: more advanced algorithms could be implemented, current thresholds could be changed to optimise for specificity instead of gross step count, signal integration to improve gait detection, and investigating the impact of a subject's characteristics or protocol's specific actions of results.

However, it would be best if future work tends towards the generalisation rather than increased specialisation of methods to reduce the impact of set up, gait complexity, pathology, and so on. As part of this it would be beneficial to conduct

independent and rigorous verification of the orientation-transformation algorithm in unconstrained movement.

Similarly, a crucial line of inquiry given the range of IMU devices, is in the impact of algorithm thresholds and parameters - especially when combined with the aforementioned importance of gait feature definition parameters.

Overall, it is likely that the most accurate future systems will incorporate multiple sensors simultaneously, therefore the combining of devices or the inclusion of the additional sensors in the devices may be a fruitful endeavour. A final flourish may be the investigation of gait identification and biometrics methods, and what insight they can provide as to the variability in gait analysis accuracy.

9.2.4 General observations & trends

It is noted that despite the great corpus of literature on gait analysis devices, there is limited homogeneity in definition, terminology, focus, and generalisation. This is exacerbated and epitomised by the relative lack of cross-referencing between the clinical, pedestrian dead reckoning, industrial, activity monitoring and sport science fields. This was observed while reviewing literature, and so future work should endeavour to unite these areas and generalise the discipline. Some steps towards this are already being undertaken. For example, the Mobilise-D consortium of 34 institutions that is aiming to unify and and validate digital mobility outcomes (DMOs) by agreeing on definitions, protocols, and outcome metrics [208, 530]. Similarly, collaborative interdisciplinary groups such as DiMe (Digital Medicine Society) are developing best practices and advocating for unification within the field of digital health to ensure the practical, efficacious, and ethical use of the technologies in the real world.

In accordance with such efforts, it is crucial going forward that ergonomics, affordability and accessibility be considered when developing wearable gait analysis methods. If this is adhered to, then it will enable the true power of remote, passive gait analysis to be realised in applications from disease diagnosis and management to injury recovery, fitness and sporting excellence. As such, it is certainly worth investigating methods that utilise consumer electronics to achieve the same goals - even if they currently achieve less accurate or reliable results. Similarly, it will be

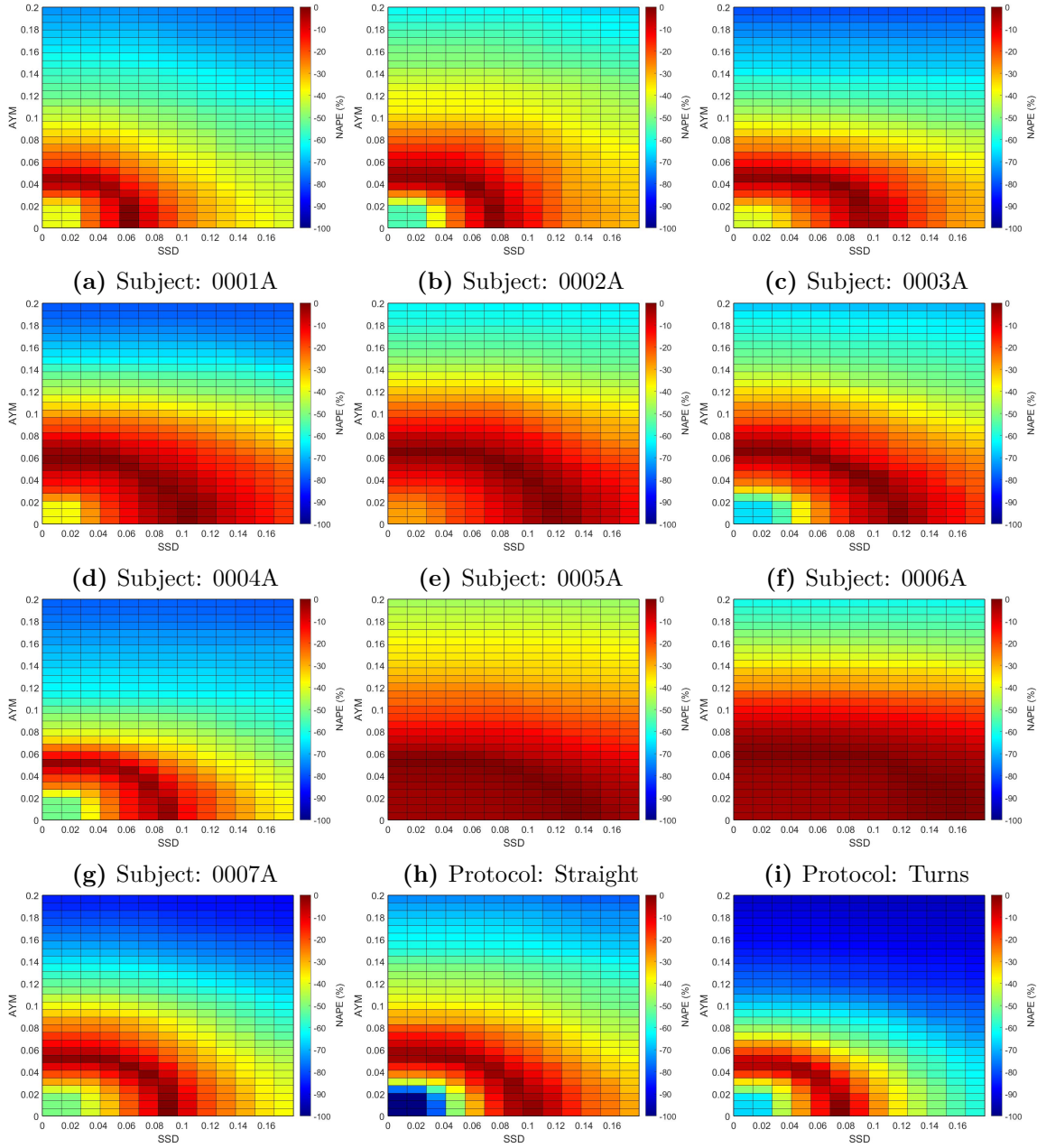
important to increase the focus on step-by-step analysis as only here can the issues with methods truly be understood. This is exemplified by the fact that it is often at specific unique moments that major gait events occur; such as falling or freezing of gait [6, 410]. Finally, it is vital that all work push towards free-living gait as we must not lose the fundamental focus that these technologies will hopefully help as many people as possible, in the real-world.

A

Appendix: Accelerometry Thresholds

The thresholds decided upon in the development of the step detection algorithms for the various accelerometer devices in Chapter 8, were chosen upon inspection of step count sensitivity to the thresholds across both subjects and protocols. This was necessary as although Figures 8.3a, 8.3b, 8.5 and 8.6 show the sensitivity across all subjects and trials - this obfuscates the person-to-person and between-protocol variation. Therefore inspection of each subject and trial was deemed necessary to choose thresholds that would not bias one trial or individual more than another. Only looking at the overall results would bias the thresholds strongly towards the Straight and Turns protocol as they contain substantially more steps. However, only inspecting the results aggregated by trial may bias the thresholds towards those individuals who, for example, perhaps moved more around more, for longer, and thus undertook substantially more steps in their ADL3 protocol.

The more detailed overview of threshold sensitivity for the foot-IMU is provided in Figure 8.9. Here, Figure A.1 provides the same breakdown for the lumbar-IMU, Figure A.2 for the smartphone, and Figure A.3 for the smartwatch.



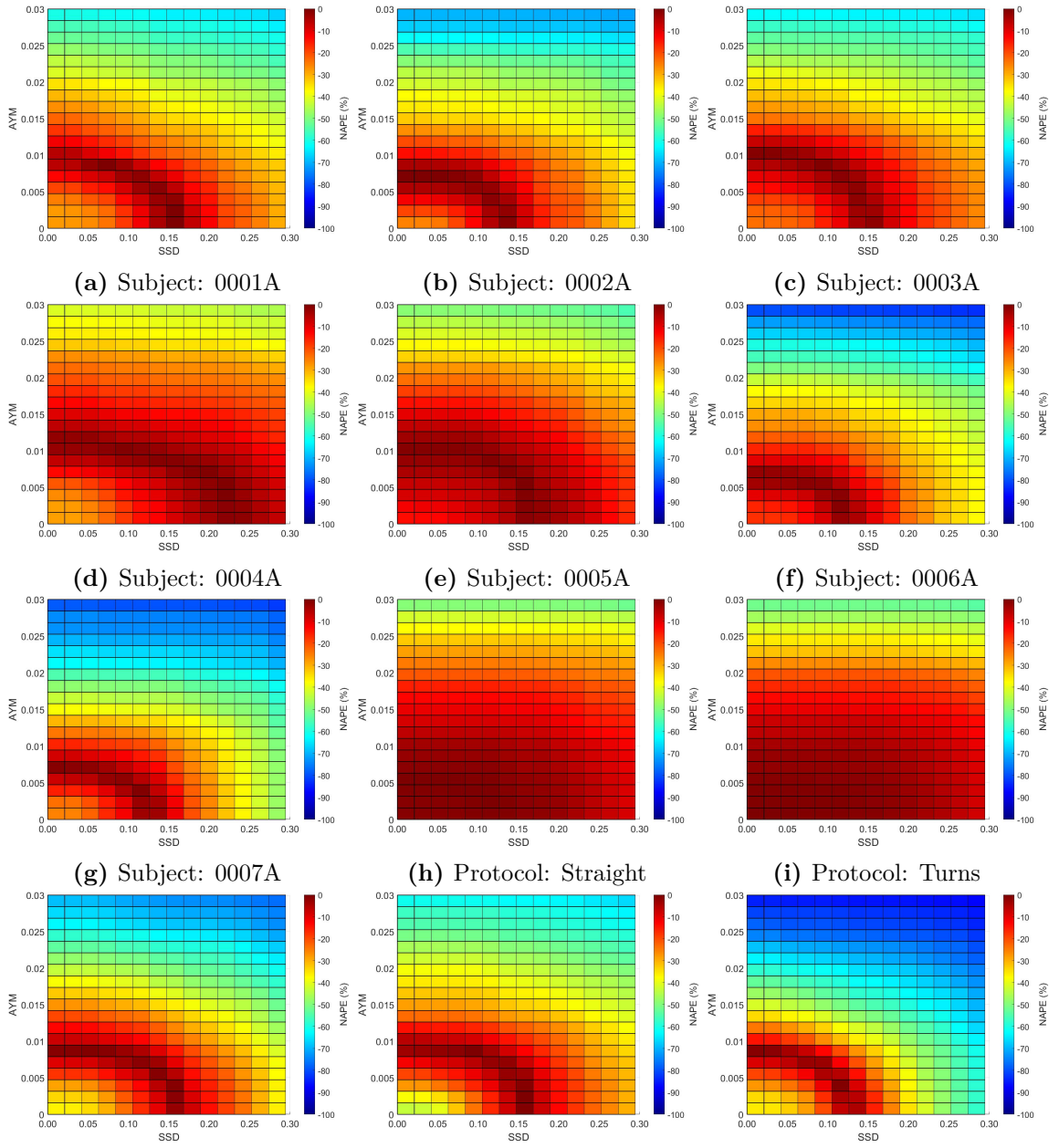
(a) Subject: 0001A **(b)** Subject: 0002A **(c)** Subject: 0003A

(d) Subject: 0004A **(e)** Subject: 0005A **(f)** Subject: 0006A

(g) Subject: 0007A **(h)** Protocol: Straight **(i)** Protocol: Turns

(j) Protocol: ADL1 **(k)** Protocol: ADL2 **(l)** Protocol: ADL3

Figure A.1: Threshold sensitivity for lumbar-IMU algorithm. Lumbar and smartphone (Figure A.2) are shown to be more robust to threshold variation than foot-IMUs or the smartwatch. There appears to be a much greater sensitivity to AYM than SDD, however the optimal SSD varies massively between subjects and protocols. The Straight and Turns protocols show that the time spent on these trials was almost entirely spent stepping - as overcounting is barely achieved even when the thresholds are set to zero. This makes the minimum time between the steps the limiting factor.



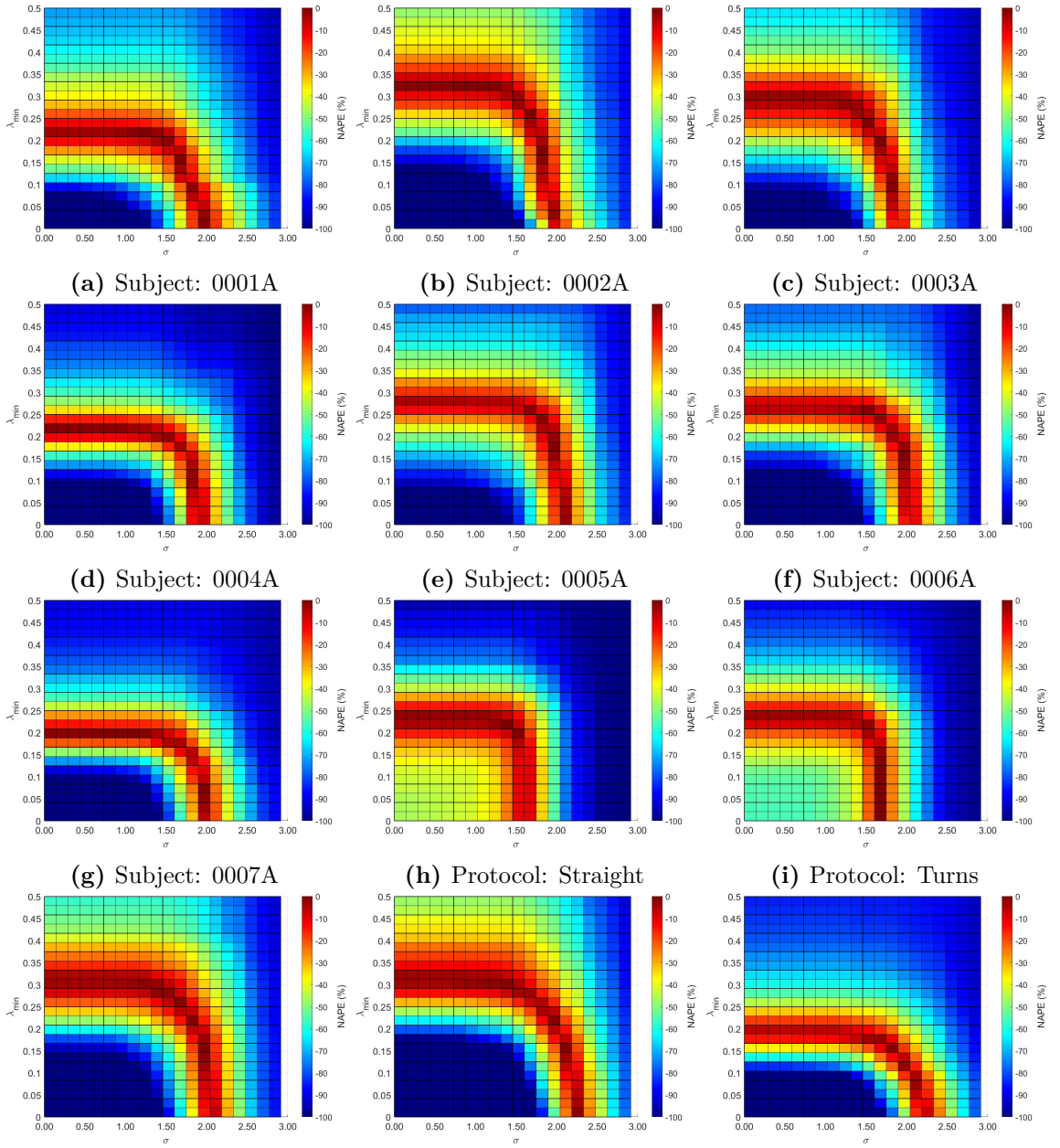
(a) Subject: 0001A **(b)** Subject: 0002A **(c)** Subject: 0003A

(d) Subject: 0004A **(e)** Subject: 0005A **(f)** Subject: 0006A

(g) Subject: 0007A **(h)** Protocol: Straight **(i)** Protocol: Turns

(j) Protocol: ADL1 **(k)** Protocol: ADL2 **(l)** Protocol: ADL3

Figure A.2: Threshold sensitivity for smartphone-IMU algorithm. Threshold sensitivity is shown to be similar to that for the lumbar-IMU - further supporting the case that the same algorithm can be used for both given their vicinity to one another on the body. Overall however, the smartphone seems to be less sensitive to the thresholds. This is likely due to the fact that, being on the hip, the smartphone undergoes greater movement during ambulation, giving a clearer and more discerning signal from which to detect steps.



(a) Subject: 0001A **(b)** Subject: 0002A **(c)** Subject: 0003A

(d) Subject: 0004A **(e)** Subject: 0005A **(f)** Subject: 0006A

(g) Subject: 0007A **(h)** Protocol: Straight **(i)** Protocol: Turns

(j) Protocol: ADL1 **(k)** Protocol: ADL2 **(l)** Protocol: ADL3

Figure A.3: Threshold sensitivity for smartwatch-IMU algorithm. As expected, the smartwatch is by far the most sensitive to the threshold values. Clearly no values would be suitable for all subjects and protocols. This is an important learning given the popularity of consumer step counting watches. The perpendicular nature of the maximum contour demonstrates that the clear importance of the threshold value on the individual. Overall this is evidence that more intelligent algorithms will be needed to make smartwatches a reliable gait analysis tool (both within and outside of the lab).

References

[1] Veronica Cimolin and Manuela Galli. "Summary measures for clinical gait analysis: A literature review". In: *Gait and Posture* 39.4 (2014), pp. 1005–1010.

[2] Richard Baker. "The history of gait analysis before the advent of modern computers". In: *Gait and Posture* 26.3 (Sept. 2007), pp. 331–342.

[3] Bjørn Lofterød et al. "Preoperative gait analysis has a substantial effect on orthopedic decision making in children with cerebral palsy: Comparison between clinical evaluation and gait analysis in 60 patients". In: *Acta Orthopaedica* 78.1 (Feb. 2007), pp. 74–80.

[4] Deepak Sharan. "Orthopedic surgery in cerebral palsy: Instructional course lecture". In: *Indian Journal of Orthopaedics*. Vol. 51. 3. Medknow Publications, May 2017, pp. 240–255.

[5] Silvia Del Din et al. "Free-living gait characteristics in ageing and Parkinson's disease: Impact of environment and ambulatory bout length". In: *Journal of NeuroEngineering and Rehabilitation* 13.1 (May 2016).

[6] Martina Mancini et al. "Measuring freezing of gait during daily-life: an open-source, wearable sensors approach". In: *Journal of NeuroEngineering and Rehabilitation* 18.1 (Dec. 2021).

[7] Mina Nouredanesh et al. "Chasing Feet in the Wild: A Proposed Egocentric Motion-Aware Gait Assessment Tool". In: *Computer Vision – ECCV 2018 Workshops*. Springer, Cham, Sept. 2019, pp. 176–192.

[8] Rebecca I. Spain et al. "Body-worn sensors capture variability, but not decline, of gait and balance measures in multiple sclerosis over 18 months". In: *Gait and Posture* 39.3 (Mar. 2014), pp. 958–964.

[9] Shanshan Chen et al. "Toward Pervasive Gait Analysis With Wearable Sensors: A Systematic Review". In: *IEEE Journal of Biomedical and Health Informatics* 20.6 (Nov. 2016), pp. 1521–1537.

[10] William Johnston et al. "Recommendations for determining the validity of consumer wearable and smartphone step count: expert statement and checklist of the INTERLIVE network". In: *Br J Sports Med* 0 (2020), pp. 1–14.

[11] David R. Bassett et al. "Step Counting: A Review of Measurement Considerations and Health-Related Applications". In: *Sports Medicine* 47.7 (July 2017), pp. 1303–1315.

[12] Salvatore Tedesco et al. "Validity evaluation of the fitbit charge2 and the garmin vivosmart HR+ in free-living environments in an older adult cohort". In: *JMIR mHealth and uHealth* 7.6 (June 2019), e13084.

[13] Yangjian Huang et al. "Validity of FitBit, Jawbone UP, Nike+ and other wearable devices for level and stair walking". In: *Gait & Posture* 48 (July 2016), pp. 36–41.

[14] Aodhán Hickey et al. "Detecting free-living steps and walking bouts: validating an algorithm for macro gait analysis". In: *Physiological Measurement* 38.1 (2017), N1–N15.

[15] Lindsay P. Toth et al. "Video-Recorded Validation of Wearable Step Counters under Free-living Conditions". In: *Medicine and Science in Sports and Exercise* 50.6 (June 2018), pp. 1315–1322.

[16] Janelle Dedic. “Step Counter Accuracy in Free-Living Environment Over a 24-Hour Period”. PhD thesis. Knoxville: University of Tennessee, 2017.

[17] Rebekah Lynn et al. “Step-Counting Validity of Wrist-Worn Activity Monitors During Activities With Fixed Upper Extremities”. In: *Journal for the Measurement of Physical Behaviour* 3.3 (Apr. 2020), pp. 197–203.

[18] Morten Bilde Simonsen, Mikkel Jacobi Thomsen, and Rogerio Pessoto Hirata. “Validation of different stepping counters during treadmill and over ground walking”. In: *Gait and Posture* 80 (July 2020), pp. 80–83.

[19] Frederik Rose Svarre et al. “The validity of activity trackers is affected by walking speed: The criterion validity of Garmin Vivosmart HR and StepWatch 3 for measuring steps at various walking speeds under controlled conditions”. In: *PeerJ* 2020.7 (July 2020), e9381.

[20] Werner L. Popp et al. “Wearable Sensors in Ambulatory Individuals With a Spinal Cord Injury: From Energy Expenditure Estimation to Activity Recommendations”. In: *Frontiers in Neurology* 10 (Nov. 2019), p. 1092.

[21] Alan K. Bourke, Espen A. F. Ihlen, and Jorunn L. Helbostad. “Validation of the activPAL3 in Free-Living and Laboratory Scenarios for the Measurement of Physical Activity, Stepping, and Transitions in Older Adults”. In: *Journal for the Measurement of Physical Behaviour* 2.2 (), pp. 58–65.

[22] Abdul Saboor et al. “Latest Research Trends in Gait Analysis Using Wearable Sensors and Machine Learning: A Systematic Review”. In: *IEEE Access* 8 (Sept. 2020), pp. 167830–167864.

[23] Hari Prasanth et al. “Wearable sensor-based real-time gait detection: A systematic review”. In: *Sensors* 21.8 (Apr. 2021), p. 2727.

[24] Giovanni Marco Scalera, Maurizio Ferrarin, and Marco Rabuffetti. “Gait regularity assessed by wearable sensors: Comparison between accelerometer and gyroscope data for different sensor locations and walking speeds in healthy subjects”. In: *Journal of Biomechanics* 113 (Dec. 2020), p. 110115.

[25] David. Levine, Jim. Richards, and Michael. Whittle. *Whittle’s gait analysis*. Churchill Livingstone/Elsevier, 2012, p. 177.

[26] James Parkinson. “An Essay on the Shaking Palsy”. In: *J Neuropsychiatry Clin Neurosci* 14.2 (2002).

[27] Bastiaan R. Bloem et al. “Falls and freezing of gait in Parkinson’s disease: A review of two interconnected, episodic phenomena”. In: *Movement Disorders* 19.8 (Aug. 2004), pp. 871–884.

[28] Sheng Li, Gerard E. Francisco, and Ping Zhou. “Post-stroke Hemiplegic Gait: New Perspective and Insights”. In: *Frontiers in Physiology* 9.AUG (Aug. 2018), p. 1021.

[29] JE Esnouf et al. “Impact on activities of daily living using a functional electrical stimulation device to improve dropped foot in people with multiple sclerosis, measured by the Canadian Occupational Performance Measure”. In: *Multiple Sclerosis Journal* 16.9 (Sept. 2010), pp. 1141–1147.

[30] C.G. Goetz. “Unified Parkinson’s Disease Rating Scale (UPDRS) and The Movement-Disorder Society Sponsored-unified Parkinson’s Disease Rating Scale (MDS-UPDRS)”. In: *Encyclopedia of Movement Disorders*. Elsevier, 2010, pp. 307–309.

[31] John F. Kurtzke. “Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS)”. In: *Neurology* 33.11 (Nov. 1983), pp. 1444–1452.

[32] Marcus Richards et al. “Interrater reliability of the unified Parkinson’s disease rating scale motor examination”. In: *Movement Disorders* 9.1 (Jan. 1994), pp. 89–91.

[33] Soichiro Koyama et al. “Intra- and inter-rater reliability and validity of the tandem gait test for the assessment of dynamic gait balance”. In: *European Journal of Physiotherapy* 20.3 (July 2018), pp. 135–140.

[34] Sophia Giannitsi et al. “6-minute walking test: a useful tool in the management of heart failure patients”. In: *Therapeutic advances in cardiovascular disease* 13 (Jan. 2019).

[35] Robert W. Motl et al. “Validity of the timed 25-foot walk as an ambulatory performance outcome measure for multiple sclerosis”. In: *Multiple Sclerosis* 23.5 (Apr. 2017), pp. 704–710.

[36] Talia Herman, Nir Giladi, and Jeffrey M. Hausdorff. “Properties of the ‘Timed Up and Go’ test: More than meets the eye”. In: *Gerontology* 57.3 (Apr. 2011), pp. 203–210.

[37] Kate E. Webster, Joanne E. Wittwer, and Julian A. Feller. “Validity of the GAITRite® walkway system for the measurement of averaged and individual step parameters of gait”. In: *Gait and Posture* 22.4 (Dec. 2005), pp. 317–321.

[38] Joe Verghese et al. “Abnormality of Gait as a Predictor of Non-Alzheimer’s Dementia”. In: *New England Journal of Medicine* 347.22 (Nov. 2002), pp. 1761–1768.

[39] Anat Mirelman et al. “Gait alterations in healthy carriers of the LRRK2 G2019S mutation”. In: *Annals of Neurology* 69.1 (Jan. 2011), pp. 193–197.

[40] Rosie Morris et al. “Gait Rather Than Cognition Predicts Decline in Specific Cognitive Domains in Early Parkinson’s Disease”. In: *The Journals of Gerontology: Series A* 72.12 (Nov. 2017), pp. 1656–1662.

[41] Conrad Tucker et al. “A data mining methodology for predicting early stage Parkinson’s disease using non-invasive, high-dimensional gait sensor data”. In: *IIE Transactions on Healthcare Systems Engineering* 5.4 (Oct. 2015), pp. 238–254.

[42] Raffaele Di Gregorio and Lucas Vocenas. “Identification of Gait-Cycle Phases for Prosthesis Control”. In: *Biomimetics* 6.2 (Mar. 2021), p. 22.

[43] Erin E. Butler et al. “Clinical motion analyses over eight consecutive years in a child with crouch gait: A case report”. In: *Journal of Medical Case Reports* 10.1 (June 2016).

[44] Sławomir Winiarski and Alicja Rutkowska-Kucharska. “Estimated ground reaction force in normal and pathological gait.” In: *Acta of bioengineering and biomechanics* 11.1 (2009), pp. 53–60.

[45] Ervin Sejdić et al. “A comprehensive assessment of gait accelerometry signals in time, frequency and time-frequency domains.” In: *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society* 22.3 (May 2014), pp. 603–12.

[46] Nancye M. Peel, Suzanne S. Kuys, and Kerenaftali Klein. *Gait speed as a measure in geriatric assessment in clinical settings: A systematic review.* Jan. 2013.

[47] Stephanie Studenski et al. “Physical performance measures in the clinical setting”. In: *Journal of the American Geriatrics Society* 51.3 (Mar. 2003), pp. 314–322.

[48] Walter Pirker and Regina Katzenschlager. “Gait disorders in adults and the elderly: A clinical guide”. In: *Wiener Klinische Wochenschrift* 129.3-4 (Feb. 2017), pp. 81–95.

[49] Stacy Fritz and Michelle Lusardi. “Walking speed: The sixth vital sign”. In: *Journal of Geriatric Physical Therapy* 32.2 (2009), pp. 2–5.

[50] Addie Middleton, Stacy L. Fritz, and Michelle Lusardi. “Walking speed: The functional vital sign”. In: *Journal of Aging and Physical Activity* 23.2 (Apr. 2015), pp. 314–322.

[51] Jeroen H.M. Bergmann, Caroline Alexiou, and Ian C.H. Smith. “Procedural Differences Directly Affect Timed Up and Go Times”. In: *Journal of the American Geriatrics Society* 57.11 (Nov. 2009), pp. 2168–2169.

[52] Ríona Mc Ardle et al. “The impact of environment on gait assessment: Considerations from real-world gait analysis in dementia subtypes”. In: *Sensors (Switzerland)* 21.3 (Feb. 2021), pp. 1–15.

[53] Brian C. Glaister et al. “Video task analysis of turning during activities of daily living”. In: *Gait and Posture* 25.2 (Feb. 2007), pp. 289–294.

[54] John G. Nutt et al. “Freezing of gait: Moving forward on a mysterious clinical phenomenon”. In: *The Lancet Neurology* 10.8 (Aug. 2011), pp. 734–744.

[55] Jooeun Song et al. “Altered dynamic postural control during step turning in persons with early-stage Parkinson’s disease”. In: *Parkinson’s Disease* (2012).

[56] Martina Mancini et al. “Continuous Monitoring of Turning Mobility and Its Association to Falls and Cognitive Function: A Pilot Study”. In: *Journals of Gerontology - Series A Biological Sciences and Medical Sciences* 71.8 (Aug. 2016), pp. 1102–1108.

[57] Julia M. Leach et al. “Natural turn measures predict recurrent falls in community-dwelling older adults: a longitudinal cohort study”. In: *Scientific Reports* 8.1 (Dec. 2018), p. 4316.

[58] Wei Yi Cheng et al. “U-turn speed is a valid and reliable smartphone-based measure of multiple sclerosis-related gait and balance impairment”. In: *Gait and Posture* 84 (Feb. 2021), pp. 120–126.

[59] Barry R. Greene et al. “Assessment and Classification of Early-Stage Multiple Sclerosis with Inertial Sensors: Comparison Against Clinical Measures of Disease State”. In: *IEEE Journal of Biomedical and Health Informatics* 19.4 (July 2015), pp. 1356–1361.

[60] Umar A.R. Chaudhry et al. “The effects of step-count monitoring interventions on physical activity: Systematic review and meta-analysis of community-based randomised controlled trials in adults”. In: *International Journal of Behavioral Nutrition and Physical Activity* 17.1 (Oct. 2020), pp. 1–16.

[61] Pedro F. Saint-Maurice et al. “Association of Daily Step Count and Step Intensity with Mortality among US Adults”. In: *JAMA - Journal of the American Medical Association* 323.12 (Mar. 2020), pp. 1151–1160.

[62] Nitin Ohri et al. “Daily Step Counts: A New Prognostic Factor in Locally Advanced Non-Small Cell Lung Cancer?” In: *International Journal of Radiation Oncology Biology Physics* 105.4 (Nov. 2019), pp. 745–751.

[63] Jennifer M. Yentes et al. "Patients with chronic obstructive pulmonary disease walk with altered step time and step width variability as compared with healthy control subjects". In: *Annals of the American Thoracic Society* 14.6 (June 2017), pp. 858–866.

[64] Galit Yogev et al. "Gait asymmetry in patients with Parkinson's disease and elderly fallers: When does the bilateral coordination of gait require attention?" In: *Experimental Brain Research* 177.3 (Mar. 2007), pp. 336–346.

[65] Neelesh K. Nadkarni et al. "Spatial and temporal gait parameters in Alzheimer's disease and aging". In: *Gait and Posture* 30.4 (Nov. 2009), pp. 452–454.

[66] Ginette Azizah Mbourou, Yves Lajoie, and Normand Teasdale. "Step length variability at gait initiation in elderly fallers and non-fallers, and young adults". In: *Gerontology* 49.1 (2003), pp. 21–26.

[67] Rana Zia Ur Rehman et al. "Selecting Clinically Relevant Gait Characteristics for Classification of Early Parkinson's Disease: A Comprehensive Machine Learning Approach". In: *Scientific Reports* 9.1 (Dec. 2019), pp. 1–12.

[68] European Medicines Agency. *Qualification opinion on stride velocity 95th centile as a secondary endpoint in Duchenne Muscular Dystrophy measured by a valid and suitable wearable device*. Tech. rep. European Medicines Agency, 2019.

[69] Felix Kluge et al. "Consensus based framework for digital mobility monitoring". In: *PLoS ONE* 16.8 August (2021).

[70] Vrutangkumar V. Shah et al. "Laboratory versus daily life gait characteristics in patients with multiple sclerosis, Parkinson's disease, and matched controls". In: *Journal of NeuroEngineering and Rehabilitation* 17.1 (Dec. 2020), pp. 1–12.

[71] Diego Orcioli-Silva et al. "Double obstacles increase gait asymmetry during obstacle crossing in people with Parkinson's disease and healthy older adults: A pilot study". In: *Scientific Reports* 10.1 (Dec. 2020), pp. 1–8.

[72] Janeesata Kuntapun et al. "Smartphone Monitoring of Gait and Balance During Irregular Surface Walking and Obstacle Crossing". In: *Frontiers in Sports and Active Living* 2 (Nov. 2020), p. 560577.

[73] W. A. Sparrow et al. "Characteristics of gait in stepping over obstacles". In: *Human Movement Science* 15.4 (Aug. 1996), pp. 605–622.

[74] Anke H. Snijders et al. "Obstacle avoidance to elicit freezing of gait during treadmill walking". In: *Movement Disorders* 25.1 (Jan. 2010), pp. 57–63.

[75] Fabio A. Storm et al. "Free-living and laboratory gait characteristics in patients with multiple sclerosis". In: *PLoS ONE* 13.5 (May 2018), e0196463.

[76] Jim McCambridge, John Witton, and Diana R. Elbourne. "Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects". In: *Journal of Clinical Epidemiology* 67.3 (2014), pp. 267–277.

[77] Marzieh M. Ardestani and T. George Hornby. "Effect of investigator observation on gait parameters in individuals with stroke". In: *Journal of Biomechanics* 100 (Feb. 2020), p. 109602.

[78] Verónica Robles-García et al. "Spatiotemporal gait patterns during overt and covert evaluation in patients with Parkinson's disease and healthy subjects: Is there a Hawthorne effect?" In: *Journal of Applied Biomechanics* 31.3 (June 2015), pp. 189–194.

[79] Erik K. Antonsson and Robert W. Mann. “The frequency content of gait”. In: *Journal of Biomechanics* 18.1 (Jan. 1985), pp. 39–47.

[80] Ming Sun and J. O. Hill. “A method for measuring mechanical work and work efficiency during human activities”. In: *Journal of Biomechanics* 26.3 (Mar. 1993), pp. 229–241.

[81] J. Muybridge. *The horse in motion [8]*. 1882.

[82] Pierre Merriaux et al. “A study of vicon system positioning performance”. In: *Sensors (Switzerland)* 17.7 (July 2017), p. 1591.

[83] Markus Windolf, Nils Götzen, and Michael Morlock. “Systematic accuracy and precision analysis of video motion capturing systems-exemplified on the Vicon-460 system”. In: *Journal of Biomechanics* 41.12 (Aug. 2008), pp. 2776–2780.

[84] Patric Eichelberger et al. “Analysis of accuracy in optical motion capture – A protocol for laboratory setup evaluation”. In: *Journal of Biomechanics* 49.10 (July 2016), pp. 2085–2088.

[85] Nobuyasu Nakano et al. “Evaluation of 3D Markerless Motion Capture Accuracy Using OpenPose With Multiple Video Cameras”. In: *Frontiers in Sports and Active Living* 2 (May 2020), p. 50.

[86] Javier Marýn et al. “Integrating a gait analysis test in hospital rehabilitation: A service design approach”. In: *PLoS ONE* 14.10 (Oct. 2019), e0224409.

[87] Bruce Carse et al. “Affordable clinical gait analysis: An assessment of the marker tracking accuracy of a new low-cost optical 3D motion analysis system”. In: *Physiotherapy (United Kingdom)* 99.4 (Dec. 2013), pp. 347–351.

[88] Gergely Nagymáté and Rita M. Kiss. “Affordable gait analysis using augmented reality markers”. In: *PLoS ONE* 14.2 (Feb. 2019), e0212319.

[89] Steffi L. Colyer et al. “A Review of the Evolution of Vision-Based Motion Analysis and the Integration of Advanced Computer Vision Methods Towards Developing a Markerless System”. In: *Sports Medicine - Open* 4.1 (Dec. 2018).

[90] V. Bonnet et al. “Toward an affordable and user-friendly visual motion capture system”. In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*. Institute of Electrical and Electronics Engineers Inc., Nov. 2014, pp. 3634–3637.

[91] Adso Fernández-Baena, Antonio Susín, and Xavier Lligadas. “Biomechanical validation of upper-body and lower-body joint movements of kinect motion capture data for rehabilitation treatments”. In: *Proceedings of the 2012 4th International Conference on Intelligent Networking and Collaborative Systems, INCoS 2012*. 2012, pp. 656–661.

[92] Stepan Obdrzalek et al. “Accuracy and robustness of Kinect pose estimation in the context of coaching of elderly population”. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*. Vol. 2012. Annu Int Conf IEEE Eng Med Biol Soc, 2012, pp. 1188–1193.

[93] Moataz Eltoukhy et al. “Microsoft Kinect can distinguish differences in over-ground gait between older persons with and without Parkinson’s disease”. In: *Medical Engineering and Physics* 44 (June 2017), pp. 1–7.

[94] Gaddi Blumrosen et al. “A Real-Time Kinect Signature-Based Patient Home Monitoring System”. In: *Sensors (Basel, Switzerland)* 16.11 (Nov. 2016), p. 1965.

[95] Jorge Latorre et al. “Gait analysis with the Kinect v2: Normative study with healthy individuals and comprehensive study of its sensitivity, validity, and reliability in individuals with stroke”. In: *Journal of NeuroEngineering and Rehabilitation* 16.1 (July 2019), pp. 1–11.

[96] Alexandra Pfister et al. “Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis”. In: *Journal of Medical Engineering & Technology* 38.5 (July 2014), pp. 274–280.

[97] Panagiotis Giannakeris et al. “Activity recognition from wearable cameras”. In: *Proceedings - International Workshop on Content-Based Multimedia Indexing*. Vol. 2018-September. IEEE Computer Society, Oct. 2018.

[98] Kai Zhan, Fabio Ramos, and Steven Faux. “Activity recognition from a wearable camera”. In: *2012 12th International Conference on Control Automation Robotics & Vision (ICARCV)*. IEEE, Dec. 2012, pp. 365–370.

[99] Marc Bolaños, Mariella Dimiccoli, and Petia Radeva. “Towards Storytelling from Visual Lifelogging: An Overview”. In: *IEEE Transactions on Human-Machine Systems* 47.1 (July 2015), pp. 77–90.

[100] Cassim Ladha et al. “Dog’s life: Wearable activity recognition for dogs”. In: *UbiComp 2013 - Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 2013, pp. 415–418.

[101] Christopher P. Hurt et al. “Assessing a novel way to measure step count while walking using a custom mobile phone application”. In: *PLoS ONE* 13.11 (July 2019), e0206828.

[102] Ferenc Aubeck, Carsten Isert, and Dominik Gusenbauer. “Camera based step detection on mobile phones”. In: *2011 International Conference on Indoor Positioning and Indoor Navigation, IPIN 2011*. 2011.

[103] Koray Ozcan and Senem Velipasalar. “Robust and reliable step counting by mobile phone cameras”. In: *ACM International Conference Proceeding Series*. Vol. 08-11-Sep-2015. Association for Computing Machinery, Sept. 2015, pp. 164–169.

[104] Milica D. Djurić-Jovičić, Nenad S. Jovičić, and Dejan B. Popović. “Kinematics of gait: New method for angle estimation based on accelerometers”. In: *Sensors* 11.11 (Nov. 2011), pp. 10571–10585.

[105] Jung Ah Lee et al. “Wearable accelerometer system for measuring the temporal parameters of gait”. In: *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*. 2007, pp. 483–486.

[106] Chaitanya Nutakki et al. “Classification and Kinetic Analysis of Healthy Gait using Multiple Accelerometer Sensors”. In: *Procedia Computer Science*. Vol. 171. Elsevier B.V., Jan. 2020, pp. 395–402.

[107] M. C. Fennema et al. “Repeatability of measuring knee flexion angles with wearable inertial sensors”. In: *Knee* 26.1 (Jan. 2019), pp. 97–105.

[108] Kieron Jie Han Ngoh et al. “Estimation of vertical ground reaction force during running using neural network model and uniaxial accelerometer”. In: *Journal of Biomechanics* 76 (July 2018), pp. 269–273.

[109] Qingsong Ai et al. “Research on lower limb motion recognition based on fusion of sEMG and accelerometer signals”. In: *Symmetry* 9.8 (Aug. 2017), p. 147.

[110] E. Maranesi et al. “A goniometer-based method for the assessment of gait parameters”. In: *2014 IEEE/ASME 10th International Conference on Mechatronic and Embedded Systems and Applications (MESA)*. IEEE, Sept. 2014, pp. 1–4.

[111] Alessandro Tognetti et al. “Wearable goniometer and accelerometer sensory fusion for knee joint angle measurement in daily life”. In: *Sensors (Switzerland)* 15.11 (Nov. 2015), pp. 28435–28455.

[112] Christopher Moufawad el Achkar. *Instrumented shoes for daily activity monitoring in healthy and at risk populations*. Tech. rep. 2016.

[113] Marco Avvenuti et al. “Smart shoe-assisted evaluation of using a single trunk/pocket-worn accelerometer to detect gait phases”. In: *Sensors (Switzerland)* 18.11 (Nov. 2018), pp. 14–15.

[114] Delaram Jarchi et al. “A Review on Accelerometry-Based Gait Analysis and Emerging Clinical Applications”. In: *IEEE Reviews in Biomedical Engineering* 11 (Feb. 2018), pp. 177–194.

[115] Qi Zhao et al. “Improved method of step length estimation based on inverted pendulum model”. In: *International Journal of Distributed Sensor Networks* 13.4 (Apr. 2017).

[116] Lucia Pepa et al. “Real time step length estimation on smartphone”. In: *2016 IEEE International Conference on Consumer Electronics, ICCE 2016*. Institute of Electrical and Electronics Engineers Inc., Mar. 2016, pp. 315–316.

[117] Marcin Straczkiewicz, Nancy W. Glynn, and Jaroslaw Harezlak. “On Placement, Location and Orientation of Wrist-Worn Tri-Axial Accelerometers during Free-Living Measurements”. In: *Sensors (Basel, Switzerland)* 19.9 (May 2019), p. 2095.

[118] Sietse M. Rispens et al. “Consistency of gait characteristics as determined from acceleration data collected at different trunk locations”. In: *Gait and Posture* 40.1 (May 2014), pp. 187–192.

[119] Diana Trojaniello, Andrea Cereatti, and Ugo Della Croce. “Accuracy, sensitivity and robustness of five different methods for the estimation of gait temporal parameters using a single inertial sensor mounted on the lower trunk”. In: *Gait and Posture* 40.4 (2014), pp. 487–492.

[120] Yu Zhong and Yunbin Deng. “Sensor orientation invariant mobile gait biometrics”. In: *IJCB 2014 - 2014 IEEE/IAPR International Joint Conference on Biometrics*. Institute of Electrical and Electronics Engineers Inc., Dec. 2014.

[121] Ravichandran Subramanian. “Orientation Invariance Methods for Inertial Gait”. PhD thesis. University of South Florida, June 2018.

[122] Zachary Arnold, Danielle Larose, and Emmanuel Agu. “A factorial experiment to investigate naturalistic factors affecting smartphone gait analysis”. In: *2015 17th International Conference on E-Health Networking, Application and Services, HealthCom 2015*. Institute of Electrical and Electronics Engineers Inc., 2015, pp. 451–454.

[123] Brad Manor et al. “Smartphone app–based assessment of gait during normal and dual-task walking: Demonstration of validity and reliability”. In: *JMIR mHealth and uHealth* 6.1 (Jan. 2018).

[124] Johannes C. M. Schlachetzki et al. “Wearable sensors objectively measure gait parameters in Parkinson’s disease”. In: *PLOS ONE* 12.10 (Oct. 2017). Ed. by Mathias Toft, e0183989.

[125] Mohammad H. Ghaeminia, Shahriar B. Shokouhi, and Abdollah Amirkhani. "Biometric Gait Identification Systems: From Spatio-Temporal Filtering to Local Patch-Based Techniques". In: *Biometric Identification Technologies Based on Modern Data Mining Methods* (2021), pp. 19–37.

[126] A. Gabell and U. S.L. Nayak. "The effect of age on variability in gait". In: *Journal of gerontology* 39.6 (1984), pp. 662–666.

[127] Lesley A. Brown et al. "Central set influences on gait". In: *Experimental Brain Research 2002 145:3* 145.3 (June 2002), pp. 286–296.

[128] Erin Smith, Tara Cusack, and Catherine Blake. "The effect of a dual task on gait speed in community dwelling older adults: A systematic review and meta-analysis". In: *Gait & Posture* 44 (Feb. 2016), pp. 250–258.

[129] Daekyoo Kim, Cara L. Lewis, and Simone V. Gill. "Effects of obesity and foot arch height on gait mechanics: A cross-sectional study". In: *PLOS ONE* 16.11 (Nov. 2021), e0260398.

[130] Birol Balaban and Fatih Tok. "Gait Disturbances in Patients With Stroke". In: *PM&R* 6.7 (July 2014), pp. 635–642.

[131] Johannes Michalak et al. "Embodiment of sadness and depression-gait patterns associated with dysphoric mood". In: *Psychosomatic Medicine* 71.5 (June 2009), pp. 580–587.

[132] Fabio Augusto Barbieri et al. "Systematic review of the effects of fatigue on spatiotemporal gait parameters". In: *Journal of Back and Musculoskeletal Rehabilitation* 26.2 (Jan. 2013), pp. 125–131.

[133] Kiwon Park et al. "Assessing gait changes in firefighters due to fatigue and protective clothing". In: *Safety Science* 49.5 (June 2011), pp. 719–726.

[134] Aditya Viswakumar et al. "Human Gait Analysis Using OpenPose". In: *Proceedings of the IEEE International Conference Image Information Processing* 2019-November (Nov. 2019), pp. 310–314.

[135] Shaohua Wan et al. "Deep Multi-Layer perceptron classifier for behavior analysis to estimate Parkinson's disease severity using smartphones". In: *IEEE Access* 6 (July 2018), pp. 36825–36833.

[136] Alejandro Rodríguez-Molinero et al. "Analysis of correlation between an accelerometer-Based algorithm for Detecting Parkinsonian gait and UPDRS subscales". In: *Frontiers in Neurology* 8.SEP (Sept. 2017), p. 431.

[137] Fabian Horst et al. "Explaining the unique nature of individual gait patterns with deep learning". In: *Scientific Reports* 9.1 (Dec. 2019), pp. 1–13.

[138] Ying Zhang et al. "Review of gait recognition methods". In: *ICEIEC 2019 - Proceedings of 2019 IEEE 9th International Conference on Electronics Information and Emergency Communication*. Institute of Electrical and Electronics Engineers Inc., July 2019, pp. 166–169.

[139] William Shakespeare. *Julius Caesar*. Wordsworth Classics, 2004.

[140] Felix Laufer et al. "On Expressive Features for Gait Analysis using Lower Limb Inertial Sensor Data". In: *IFAC-Online* (July 2020).

[141] F. Huxham et al. "Defining spatial parameters for non-linear walking". In: *Gait & Posture* 23.2 (Feb. 2006), pp. 159–163.

[142] Björn Müller et al. “Validation of enhanced kinect sensor based motion capturing for gait assessment”. In: *PLOS ONE* 12.4 (Apr. 2017), e0175813.

[143] Caterina Rosano et al. “Gait variability is associated with subclinical brain vascular abnormalities in high-functioning older adults”. In: *Neuroepidemiology* 29.3-4 (Feb. 2007), pp. 193–200.

[144] *MATLAB version 9.8 (R2020a)*. Natick, Massachusetts: The Mathworks Inc., 2020.

[145] I. S. Duff and J. Koster. “On algorithms for permuting large entries to the diagonal of a sparse matrix”. In: *SIAM Journal on Matrix Analysis and Applications* 22.4 (Jan. 2001), pp. 973–996.

[146] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. 3rd. New York: Cambridge University Press, 2003.

[147] Richard Szeliski. *Computer Vision: Algorithms and Applications*. 2nd. Springer, 2021.

[148] Lai Kang et al. “A highly accurate dense approach for homography estimation using modified differential evolution”. In: *Engineering Applications of Artificial Intelligence* 31 (May 2014), pp. 68–77.

[149] Anubhav Agarwal et al. *A Survey of Planar Homography Estimation Techniques*. Tech. rep. Hyderabad: Centre for Visual Information Technology International Institute of Information Technology, 2005.

[150] I.E. Sutherland. “Three-dimensional data input by tablet”. In: *Proceedings of the IEEE* 62.4 (1974), pp. 453–461.

[151] David Kriegman. *Computer Vision I CSE 252A, Winter 2007 Homography Estimation*. Tech. rep. University of California, San Diego, 2007.

[152] Keijo Inkilä. “Homogeneous Least Squares Problem”. In: *The Photogrammetric Journal of Finland* 19.2 (2005).

[153] Finny G. Kuruvilla, Peter J. Park, and Stuart L. Schreiber. “Vector algebra in the analysis of genome-wide expression data.” In: *Genome biology* 3.3 (Feb. 2002), research0011.1.

[154] E. Anderson et al. *LAPACK Users’ Guide*. 3rd. Philadelphia: Society for Industrial and Applied Mathematics, Jan. 1999.

[155] Martin Fischler and Robert Bolles. “Random Sample Consensus: A paradigm for model fitting with applications to image analysis and automated cartography”. In: *Communications of the JCM* (1980).

[156] P.H.S. Torr and A. Zisserman. “MLESAC: A New Robust Estimator with Application to Estimating Image Geometry”. In: *Computer Vision and Image Understanding* 78.1 (Apr. 2000), pp. 138–156.

[157] Peter J. Huber. “Projection Pursuit”. In: *The Annals of Statistics* 13.2 (June 1985), pp. 435–475.

[158] O.D. Faugeras and F. Lustman. “Motion and structure from motion in a piecewise planar environment”. In: *International Journal of Pattern Recognition and Artificial Intelligence* 02.03 (Sept. 1988), pp. 485–508.

[159] Zhongfei Zhang and Allen R. Hanson. “3D Reconstruction Based on Homography Mapping”. In: *In ARPA Image Understanding Workshop*. 1996, pp. 0249–6399.

[160] Ezio Malis and Manuel Vargas. *Deeper understanding of the homography decomposition for vision-based control*. Tech. rep. INRIA, Sept. 2007.

[161] G. Bradski. *The OpenCV Library*. 2000.

[162] Yaqing Ding et al. "An efficient solution to the homography-based relative pose problem with a common reference direction". In: *Proceedings of the IEEE International Conference on Computer Vision*. Vol. 2019-Octob. Institute of Electrical and Electronics Engineers Inc., Oct. 2019, pp. 1655–1664.

[163] Richard I. Hartley and Peter Sturm. "Triangulation". In: *Computer Vision and Image Understanding* 68.2 (Nov. 1997), pp. 146–157.

[164] He Liu. "Deeper Understanding on Solution Ambiguity in Estimating 3D Motion Parameters by Homography Decomposition and its Improvement". PhD thesis. University of Fukui, Mar. 2012, pp. 31–33.

[165] Dingfu Zhou, Yuchao Dai, and Hongdong Li. "Ground Plane based Absolute Scale Estimation for Monocular Visual Odometry". In: *IEEE Transactions on Intelligent Transportation Systems* 21.2 (Mar. 2019), pp. 791–802.

[166] Radu Horaud et al. "An analytic solution for the perspective 4-point problem". In: *Computer Vision, Graphics and Image Processing* 47.1 (July 1989), pp. 33–44.

[167] Bernd Kitt et al. "Monocular Visual Odometry using a Planar Road Model to Solve Scale Ambiguity". In: *Proceedings of the European Conference on Mobile Robots*. Sept. 2011.

[168] Jiawei Mo and Junaed Sattar. "Extending Monocular Visual Odometry to Stereo Camera Systems by Scale Optimization". In: *IEEE International Conference on Intelligent Robots and Systems* (May 2019), pp. 6921–6927.

[169] Paul Gakne and Kyle O'Keefe. *Tackling The Scale Factor Issue In A Monocular Visual Odometry Using A 3D City Model*. Tech. rep. Toulouse, France: International Technical Symposium on Navigation and Timing, Aug. 2018.

[170] Daniel P. Shepard and Todd E. Humphreys. "High-precision globally-referenced position and attitude via a fusion of visual SLAM, carrier-phase-based GPS, and inertial measurements". In: *Record - IEEE PLANS, Position Location and Navigation Symposium*. Institute of Electrical and Electronics Engineers Inc., 2014, pp. 1309–1328.

[171] Zhencheng Hu and Keiichi Uchimura. "Real-time data fusion on tracking camera pose for direct visual guidance". In: *IEEE Intelligent Vehicles Symposium, Proceedings*. 2004, pp. 842–847.

[172] Ji Zhang, Sanjiv Singh, and George Kantor. "Robust monocular visual odometry for a ground vehicle in undulating Terrain". In: *Springer Tracts in Advanced Robotics*. Vol. 92. Springer, Berlin, Heidelberg, 2014, pp. 311–326.

[173] Zhou Dingfu, Yuchao Dai, and Hongdong Li. "Reliable scale estimation and correction for monocular Visual Odometry". In: *IEEE Intelligent Vehicles Symposium, Proceedings*. Vol. 2016-August. Institute of Electrical and Electronics Engineers Inc., Aug. 2016, pp. 490–495.

[174] Rafael Padilla, Sergio L. Netto, and Eduardo A.B. Da Silva. "A Survey on Performance Metrics for Object-Detection Algorithms". In: *International Conference on Systems, Signals, and Image Processing*. Vol. 2020-July. IEEE Computer Society, July 2020, pp. 237–242.

[175] Patima Silsupadol et al. “Smartphone-Based Assessment of Gait during Straight Walking, Turning, and Walking Speed Modulation in Laboratory and Free-Living Environments”. In: *IEEE Journal of Biomedical and Health Informatics* 24.4 (Apr. 2020), pp. 1188–1195.

[176] Abena Primo et al. “Context-aware active authentication using smartphone accelerometer measurements”. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. IEEE Computer Society, Sept. 2014, pp. 98–105.

[177] Bernd Thomas. *SensorLog v3.8*. Stuttgart, 2020.

[178] Martin Gjoreski et al. “How accurately can your wrist device recognize daily activities and detect falls?” In: *Sensors (Switzerland)* 16.6 (June 2016).

[179] Susan Park et al. “Dominant vs. Non-Dominant Wrist Placement of Activity Monitors: Impact on Steps per Day”. In: *Journal for the Measurement of Physical Behaviour* 2.2 (June 2019), pp. 118–123.

[180] Xiangyu Liu et al. “When good becomes evil: Keystroke inference with smartwatch”. In: *Proceedings of the ACM Conference on Computer and Communications Security*. Vol. 2015-October. New York, NY, USA: Association for Computing Machinery, Oct. 2015, pp. 1273–1285.

[181] Eline van der Kruk and Marco M. Reijne. “Accuracy of human motion capture systems for sport applications; state-of-the-art review”. In: *European Journal of Sport Science* 18.6 (July 2018), pp. 806–819.

[182] Niveditha Muthukrishnan, James J. Abbas, and Narayanan Krishnamurthi. “A Wearable Sensor System to Measure Step-Based Gait Parameters for Parkinson’s Disease Rehabilitation”. In: *Sensors* 20.22 (Nov. 2020), p. 6417.

[183] Marcello Fusca, Paolo Perego, and Giuseppe Andreoni. “Method for wearable kinematic gait analysis using a harmonic oscillator applied to the center of mass”. In: *Journal of Sensors* 2018 (2018).

[184] Wei Zhang et al. “Gait symmetry assessment with a low back 3d accelerometer in post-stroke patients”. In: *Sensors (Switzerland)* 18.10 (Oct. 2018).

[185] Dylan Drover et al. “Faller classification in older adults using wearable sensors based on turn and straight-walking accelerometer-based features”. In: *Sensors (Switzerland)* 17.6 (June 2017).

[186] Johnny Farah, Edward Lemaire, and Natalie Baddour. “Comparison of inertial sensor data from the wrist and mid-lower back during a 2-minute walk test”. In: *2016 IEEE EMBS International Student Conference: Expanding the Boundaries of Biomedical Engineering and Healthcare, ISC 2016 - Proceedings*. Institute of Electrical and Electronics Engineers Inc., July 2016.

[187] Kristi L. Storti et al. “Gait speed and step-count monitor accuracy in community-dwelling older adults”. In: *Medicine and Science in Sports and Exercise* 40.1 (Jan. 2008), pp. 59–64.

[188] Wonil Park et al. “Effect of walking speed and placement position interactions in determining the accuracy of various newer pedometers”. In: *Journal of Exercise Science and Fitness* 12.1 (June 2014), pp. 31–37.

[189] Diego Alvarez et al. “Comparison of Step Length Estimators from Wearable Accelerometer Devices”. In: *Encyclopedia of Healthcare Information Systems*. IGI Global, Jan. 2008, pp. 244–250.

[190] Rana Zia Ur Rehman et al. "Turning Detection During Gait: Algorithm Validation and Influence of Sensor Location and Turning Characteristics in the Classification of Parkinson's Disease". In: *Sensors* 20.18 (Sept. 2020), p. 5377.

[191] P. Crenna et al. "The association between impaired turning and normal straight walking in Parkinson's disease". In: *Gait and Posture* 26.2 (July 2007), pp. 172–178.

[192] Ava D. Segal et al. "Local dynamic stability in turning and straight-line gait". In: *Journal of Biomechanics* 41.7 (2008), pp. 1486–1493.

[193] Baptiste Ulrich et al. "Gait events during turning can be detected using kinematic features originally proposed for the analysis of straight-line walking". In: *Journal of Biomechanics* 91 (June 2019), pp. 69–78.

[194] Silvia Del Din et al. "Free-living monitoring of Parkinson's disease: Lessons from the field". In: *Movement Disorders* 31.9 (Sept. 2016), pp. 1293–1313.

[195] J. Merrilees. "Activities of Daily Living". In: *Encyclopedia of the Neurological Sciences*. Elsevier Inc., Jan. 2014, pp. 47–48.

[196] Clarissa M. Giebel, Caroline Sutcliffe, and David Challis. "Activities of daily living and quality of life across different stages of dementia: A UK study". In: *Aging and Mental Health* 19.1 (Jan. 2015), pp. 63–71.

[197] Stephanie A. Bridenbaugh and Reto W. Kressig. "Motor cognitive dual tasking: Early detection of gait impairment, fall risk and cognitive decline". In: *Zeitschrift fur Gerontologie und Geriatrie* 48.1 (Jan. 2015), pp. 15–21.

[198] Nathan Theill et al. "Simultaneously measuring gait and cognitive performance in cognitively healthy and cognitively impaired older adults: The basel motor-cognition dual-task paradigm". In: *Journal of the American Geriatrics Society* 59.6 (June 2011), pp. 1012–1018.

[199] Stephanie A. Bridenbaugh and Reto W. Kressig. "Laboratory review: The role of gait analysis in seniors' mobility and fall prevention". In: *Gerontology* 57.3 (Apr. 2011), pp. 256–264.

[200] Stephanie Bridenbaugh and Reto Kressig. "Quantitative Gait Disturbances in Older Adults with Cognitive Impairments". In: *Current Pharmaceutical Design* 20.19 (May 2014), pp. 3165–3172.

[201] Stephanie A. Bridenbaugh et al. "Association between dual task-related decrease in walking speed and real versus imagined Timed Up and Go test performance". In: *Aging Clinical and Experimental Research* 25.3 (2013), pp. 283–289.

[202] Jyotirmay Mathur et al. "Experimental investigations on solar chimney for room ventilation". In: *Solar Energy* 80.8 (Aug. 2006), pp. 927–935.

[203] Martin Gérin-Lajoie, Carol L. Richards, and Bradford J. McFadyen. "The negotiation of stationary and moving obstructions during walking: Anticipatory locomotor adaptations and preservation of personal space". In: *Motor Control* 9.3 (2005), pp. 242–269.

[204] Wagner Souza Silva et al. "Healthy young adults implement distinctive avoidance strategies while walking and circumventing virtual human vs. non-human obstacles in a virtual environment". In: *Gait and Posture* 61 (Mar. 2018), pp. 294–300.

[205] Christopher R. Kovacs. "Age-Related Changes in Gait and Obstacle Avoidance Capabilities in Older Adults: A Review". In: *Journal of Applied Gerontology* 24.1 (Feb. 2005), pp. 21–34.

[206] Jamie Shotton, Andrew Blake, and Roberto Cipolla. “Contour-based learning for object detection”. In: *Proceedings of the IEEE International Conference on Computer Vision*. Vol. I. 2005, pp. 503–510.

[207] Blaine Reeder and Alexandria David. “Health at hand: A systematic review of smart watch uses for health and wellness”. In: *Journal of Biomedical Informatics* 63 (Oct. 2016), pp. 269–276.

[208] Claudia Mazzà et al. “Technical validation of real-world monitoring of gait: a multicentric observational study”. In: *BMJ Open* 11.12 (Dec. 2021), e050785.

[209] Kelly R. Evenson, Michelle M. Goto, and Robert D. Furberg. “Systematic review of the validity and reliability of consumer-wearable activity trackers”. In: *International Journal of Behavioral Nutrition and Physical Activity* 12.1 (Dec. 2015), p. 159.

[210] Marcello Fusca et al. “Validation of a Wearable IMU System for Gait Analysis: Protocol and Application to a New System”. In: *Applied Sciences* 8.7 (July 2018), p. 1167.

[211] A. E.J.M. Cavelaars et al. “Persistent variations in average height between countries and between socio-economic groups: An overview of 16 European countries”. In: *Annals of Human Biology* 27.4 (2000), pp. 407–421.

[212] Shiqi Yu, Daoliang Tan, and Tieniu Tan. “A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition”. In: *Proceedings - International Conference on Pattern Recognition*. Vol. 4. 2006, pp. 441–444.

[213] Elmar C. Kal, William R. Young, and Toby J. Ellmers. “Face masks, vision, and risk of falls”. In: *BMJ* 371 (Oct. 2020).

[214] Keith J. Ruskin et al. “COVID-19, Personal Protective Equipment, and Human Performance”. In: *Anesthesiology* 134.4 (Apr. 2021), pp. 518–525.

[215] F. Taia Alaoui, David Betaille, and Valerie Renaudin. “Pedestrian dead reckoning navigation with the help of A* -based routing graphs in large unconstrained spaces”. In: *Wireless Communications and Mobile Computing* 2017 (Jan. 2017).

[216] Sehoon Ha et al. “Learning to Walk in the Real World with Minimal Human Effort”. In: *arXiv* (Feb. 2020).

[217] Tuomas Haarnoja et al. “Learning to Walk via Deep Reinforcement Learning”. In: *arXiv* (Dec. 2018).

[218] Andrew H. Hansen, Dudley S. Childress, and Margrit R. Meier. “A simple method for determination of gait events”. In: *Journal of Biomechanics* 35.1 (Jan. 2002), pp. 135–138.

[219] Diana Trojaniello et al. “Estimation of step-by-step spatio-temporal parameters of normal and impaired gait using shank-mounted magneto-inertial sensors: application to elderly, hemiparetic, parkinsonian and choreic gait”. In: *Journal of NeuroEngineering and Rehabilitation* 11.1 (Nov. 2014), p. 152.

[220] Diana Trojaniello, Ugo Croce, and Andrea Cereatti. “Gait direction of progression changes using shank worn MIMUs. Validation on healthy and TBI subjects”. In: *Conference: IV National Congress of National Group of Bioengineering (GNB), 2014*. June 2014.

[221] Mahmoud El-Gohary et al. “Continuous Monitoring of Turning in Patients with Movement Disability”. In: *Sensors* 14.1 (Dec. 2013), pp. 356–369.

[222] Vrutangkumar V. Shah et al. “Quantity and quality of gait and turning in people with multiple sclerosis, Parkinson’s disease and matched controls during daily living”. In: *Journal of Neurology* 267.4 (Apr. 2020), pp. 1188–1196.

[223] Hung Nguyen et al. “Using Inertial Sensors to Automatically Detect and Segment Activities of Daily Living in People with Parkinson’s Disease”. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26.1 (Jan. 2018), pp. 197–204.

[224] Nicholas Smith and Nikos Mourselas. “It’s your turn: The measurement of turning in free-living stepping”. In: *International Conference on Ambulatory Monitoring of Physical Activity and Movement (ICAMPAM)*. The International Society for the Measurement of Physical Behaviour (ISMBP), July 2019.

[225] Alice Nieuwboer et al. “The Short-Term Effects of Different Cueing Modalities on Turn Speed in People With Parkinson’s Disease”. In: *Neurorehabilitation and Neural Repair* 23 (2009), pp. 831–836.

[226] Anke H. Snijders et al. “Neurological gait disorders in elderly people: clinical approach and classification”. In: *Lancet Neurology* 6.1 (Jan. 2007), pp. 63–74.

[227] Harald Böhm et al. “Stiff-knee gait in cerebral palsy: How do patients adapt to uneven ground?” In: *Gait and Posture* 39.4 (Apr. 2014), pp. 1028–1033.

[228] M.J.D. Taylor, P. Dabnichki, and S.C. Strike. “A three-dimensional biomechanical comparison between turning strategies during the stance phase of walking”. In: *Human Movement Science* 24.4 (Aug. 2005), pp. 558–573.

[229] Richard W. Bohannon. “Comfortable and maximum walking speed of adults aged 20—79 years: reference values and determinants”. In: *Age and Ageing* 26.1 (Jan. 1997), pp. 15–19.

[230] Taina Rantanen and Janne Avela. “Leg extension power and walking speed in very old people living independently”. In: *Journals of Gerontology - Series A Biological Sciences and Medical Sciences* 52.4 (1997).

[231] Jama L. Purser et al. “Walking speed predicts health status and hospital costs for frail elderly male veterans”. In: *Journal of Rehabilitation Research and Development* 42.4 (July 2005), pp. 535–545.

[232] Susan E. Hardy et al. “Improvement in usual gait speed predicts better survival in older adults”. In: *Journal of the American Geriatrics Society* 55.11 (Nov. 2007), pp. 1727–1734.

[233] Jean A. Langlois et al. “Characteristics of older pedestrians who have difficulty crossing the street”. In: *American Journal of Public Health* 87.3 (1997), pp. 393–397.

[234] Michele L. Callisaya et al. “Cognitive status, fast walking speed and walking speed reserve—the Gait and Alzheimer Interactions Tracking (GAIT) study”. In: *GeroScience* 39.2 (Apr. 2017), pp. 231–239.

[235] L. Miller et al. “The impact of walking speed on the effects of functional electrical stimulation for foot drop in people with multiple sclerosis”. In: *Disability and Rehabilitation: Assistive Technology* 11.6 (Mar. 2015), pp. 1–6.

[236] Yea Ru Yang et al. “Dual-Task Exercise Improves Walking Ability in Chronic Stroke: A Randomized Controlled Trial”. In: *Archives of Physical Medicine and Rehabilitation* 88.10 (Oct. 2007), pp. 1236–1240.

[237] Inácio Teixeira Da Cunha-Filho et al. "Differential Responses to Measures of Gait Performance among Healthy and Neurologically Impaired Individuals". In: *Archives of Physical Medicine and Rehabilitation* 84.12 (Dec. 2003), pp. 1774–1779.

[238] Cristina Soaz and Klaus Diepold. "Step Detection and Parameterization for Gait Assessment Using a Single Waist-Worn Accelerometer". In: *IEEE Transactions on Biomedical Engineering* 63.5 (May 2016), pp. 933–942.

[239] Inès A. Kramers-De Quervain et al. "Gait analysis in patients with idiopathic scoliosis". In: *European Spine Journal* 13.5 (Apr. 2004), pp. 449–456.

[240] Meir Plotnik et al. "Gait asymmetry, and bilateral coordination of gait during a six-minute walk test in persons with multiple sclerosis". In: *Scientific Reports* 10.1 (Dec. 2020).

[241] Kara K. Patterson et al. "Evaluation of gait symmetry after stroke: A comparison of current methods and recommendations for standardization". In: *Gait and Posture* 31.2 (Feb. 2010), pp. 241–246.

[242] Chitralakshmi K. Balasubramanian et al. "Relationship Between Step Length Asymmetry and Walking Performance in Subjects With Chronic Hemiparesis". In: *Archives of Physical Medicine and Rehabilitation* 88.1 (Jan. 2007), pp. 43–49.

[243] Catrine Tudor-Locke et al. "Preliminary Comparison of Clinical and Free-Living Measures of Stepping Cadence in Older Adults". In: *Journal of Physical Activity and Health* 10.8 (Nov. 2013), pp. 1175–1180.

[244] Marco Monticone et al. "Reliability of spatial-temporal gait parameters during dual-task interference in people with multiple sclerosis. A cross-sectional study". In: *Gait and Posture* 40.4 (Sept. 2014), pp. 715–718.

[245] Rodger J. Elble, Larry Hughes, and Constance Higgins. "The syndrome of senile gait". In: *Journal of Neurology* 239.2 (Feb. 1992), pp. 71–75.

[246] Andrew W. Gardner et al. "Gender differences in daily ambulatory activity patterns in patients with intermittent claudication". In: *Journal of Vascular Surgery* 52.5 (Nov. 2010), pp. 1204–1210.

[247] Suzie Mudge et al. "Test-retest reliability of the StepWatch activity monitor outputs in healthy adults". In: *Journal of Physical Activity and Health* 7.5 (Sept. 2010), pp. 671–676.

[248] Suzie Mudge and Susan N. Stott. "Test-retest reliability of the StepWatch Activity Monitor outputs in individuals with chronic stroke". In: *Clinical Rehabilitation* 22.10-11 (Oct. 2008), pp. 871–877.

[249] James Oat Judge, Roy B Davis, and Sylvia Ounpuu. "Step Length Reductions in Advanced Age The Role of Ankle and Hip Kinetics". In: *Journal of Gerontology: Medical Sciences* 51.6 (1996), pp. 3–3.

[250] David. Levine, Jim. Ricahrds, and Michael. Whittle. *Whittle's Gait Analysis*. Fifth Edit. Churchill Livingstone/Elsevier, 2012, p. 29.

[251] Shehroz S. Khan and Jesse Hoey. *Review of fall detection techniques: A data availability perspective*. Jan. 2017.

[252] Niranjan Bidargaddi et al. "Detecting walking activity in cardiac rehabilitation by using accelerometer". In: *Proceedings of the 2007 International Conference on Intelligent Sensors, Sensor Networks and Information Processing, ISSNIP*. 2007, pp. 555–560.

[253] Paulo Lopez-Meyer, George D. Fulk, and Edward S. Sazonov. "Automatic detection of temporal gait parameters in poststroke individuals". In: *IEEE Transactions on Information Technology in Biomedicine* 15.4 (July 2011), pp. 594–601.

[254] M. I. Dallas et al. "Using pedometers to monitor walking activity in outcome assessment for pulmonary rehabilitation". In: *Chronic Respiratory Disease* 6.4 (Nov. 2009), pp. 217–224.

[255] Shyamal Patel et al. "A review of wearable sensors and systems with application in rehabilitation". In: *Journal of NeuroEngineering and Rehabilitation* 9.1 (Apr. 2012), p. 21.

[256] Marcos Coronado et al. "Walking activity measured by accelerometry during respiratory rehabilitation". In: *Journal of Cardiopulmonary Rehabilitation* 23.5 (Sept. 2003), pp. 357–364.

[257] M. Baghianimoghadam et al. "Current education versus peer-education on walking in type 2 diabetic patients based on Health Belief Model: a randomized control trial study." In: *Rom J Intern Med* 50.2 (2012), pp. 165–172.

[258] Alexis Le Faucheur et al. "Study of Human Outdoor Walking with a Low-Cost GPS and Simple Spreadsheet Analysis". In: *Med. Sci. Sports Exerc* 39.9 (2007), pp. 1570–1578.

[259] Bénédicte Noury-Desvaux et al. "The Accuracy of a Simple, Low-Cost GPS Data Logger/Receiver to Study Outdoor Human Walking in View of Health and Clinical Studies". In: *PLoS ONE* 6.9 (Sept. 2011). Ed. by Conrad P. Earnest, e23027.

[260] Michael S Orendurff et al. "How humans walk: Bout duration, steps per bout, and rest duration". In: *The Journal of Rehabilitation Research and Development* 45.7 (2008), pp. 1077–1090.

[261] Bijan Najafi et al. "Ambulatory system for human motion analysis using a kinematic sensor: Monitoring of daily physical activity in the elderly". In: *IEEE Transactions on Biomedical Engineering* 50.6 (June 2003), pp. 711–723.

[262] A. Godfrey et al. "Instrumenting gait with an accelerometer: A system and algorithm examination". In: *Medical Engineering & Physics* 37.4 (Apr. 2015), pp. 400–407.

[263] Emma Fortune et al. "Validity of using tri-axial accelerometers to measure human movement - Part II: Step counts at a wide range of gait velocities". In: *Medical Engineering and Physics* 36.6 (June 2014), pp. 659–669.

[264] Matthew A.D. Brodie et al. "Wearable pendant device monitoring using new wavelet-based methods shows daily life and laboratory gaits are different". In: *Medical and Biological Engineering and Computing* 54.4 (Apr. 2016), pp. 663–674.

[265] Baukje Dijkstra, Ype P. Kamsma, and Wiebren Zijlstra. "Detection of gait and postures using a miniaturized triaxial accelerometer-based system: Accuracy in patients with mild to moderate Parkinson's disease". In: *Archives of Physical Medicine and Rehabilitation* 91.8 (Aug. 2010), pp. 1272–1277.

[266] Karen M. Culhane et al. "Long-term mobility monitoring of older adults using accelerometers in a clinical environment". In: *Clinical Rehabilitation* 18.3 (May 2004), pp. 335–343.

[267] James T. Cavanaugh et al. "Using step activity monitoring to characterize ambulatory activity in community-dwelling older adults". In: *Journal of the American Geriatrics Society* 55.1 (Jan. 2007), pp. 120–124.

[268] Johannes B.J. Bussmann et al. “Quantification of physical activities by means of ambulatory accelerometry: A validation study”. In: *Psychophysiology* 35.5 (Sept. 1998), pp. 488–496.

[269] Margaret A. Roos, Katherine S. Rudolph, and Darcy S. Reisman. “The structure of walking activity in people after stroke compared with older adults without disability: A cross-sectional study”. In: *Physical Therapy* 92.9 (Sept. 2012), pp. 1141–1147.

[270] Jennifer G. Lyons et al. “Assessing the agreement between 3-meter and 6-meter walk tests in 136 community-dwelling older adults”. In: *Journal of Aging and Health* 27.4 (June 2015), pp. 594–605.

[271] Denise M. Peters et al. “Concurrent validity of walking speed values calculated via the GAITRite electronic walkway and 3 meter walk test in the chronic stroke population”. In: *Physiotherapy Theory and Practice* 30.3 (Apr. 2014), pp. 183–188.

[272] James W. Youdas et al. “Agreement Between the GAITRite Walkway System and a Stopwatch-Footfall Count Method for Measurement of Temporal and Spatial Gait Parameters”. In: *Archives of Physical Medicine and Rehabilitation* 87.12 (Dec. 2006), pp. 1648–1652.

[273] Brice T. Cleland, Haris Arshad, and Sangeetha Madhavan. “Concurrent validity of the GAITRite electronic walkway and the 10-m walk test for measurement of walking speed after stroke”. In: *Gait and Posture* 68 (Feb. 2019), pp. 458–460.

[274] Alexander M. Aurand, Jonathan S. Dufour, and William S. Marras. “Accuracy map of an optical motion capture system with 42 or 21 cameras in a large measurement volume”. In: *Journal of Biomechanics* 58 (June 2017), pp. 237–240.

[275] Bill McArthur. *Human Factors/Ergonomics Handbook for the Design for Ease of Maintenance (DOE-HDBK-1140-2001)*. Tech. rep. United States of America Department of Energy, 2001.

[276] Edward Hall. *The Hidden Dimension: Man’s Use of Space in Public and Private*. Anchor Books, 1969.

[277] Michael S.A. Graziano and Dylan F. Cooke. “Parieto-frontal interactions, personal space, and defensive behavior”. In: *Neuropsychologia* 44.6 (2006), pp. 845–859.

[278] Claudio Brozzoli et al. “Peripersonal space: A multisensory interface for body-object interactions”. In: *The Neural Bases of Multisensory Processes*. CRC Press, Aug. 2011, pp. 449–465.

[279] James E. Graham et al. “Walking speed threshold for classifying walking independence in hospitalized older adults”. In: *Physical Therapy* 90.11 (Nov. 2010), pp. 1591–1597.

[280] K. Aminian et al. “Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes”. In: *Journal of Biomechanics* 35.5 (May 2002), pp. 689–699.

[281] Vicon Motion Systems Limited. *Vicon Nexus 1.8.5 Product Guide*. Oxford, UK, 2015.

[282] Annapuma Sharma et al. “Frequency based classification ofactivities using accelerometer data”. In: *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*. 2008, pp. 150–153.

[283] Aikaterini Pacha and Tianjian Ji. “Frequency and velocity of people walking”. In: *The Structural Engineer* (Feb. 2005), pp. 36–40.

[284] Yasuhiro Osaki et al. "Relative contribution of walking velocity and stepping frequency to the neural control of locomotion". In: *Experimental Brain Research* 185.1 (Feb. 2008), pp. 121–135.

[285] Khalid Al-Naime, Adnan Al-Anbuky, and Grant Mawston. "Human Movement Monitoring and Analysis for Prehabilitation Process Management". In: *Journal of Sensor and Actuator Networks* 9.1 (Jan. 2020), p. 9.

[286] Christian Meinhardt, Volkmar Zabel, and Hernando Gonzalez. "An in-siotu case study of human induced vibrations on slender staircases". In: *Proceedings of the 9th International Conference on Structural Dynamics*. Ed. by A Cunha et al. Porto, Portugal, June 2014.

[287] Federico Bianchi et al. "Barometric pressure and triaxial accelerometry-based falls event detection". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 18.6 (Dec. 2010), pp. 619–627.

[288] Dean M. Karantonis et al. "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring". In: *IEEE Transactions on Information Technology in Biomedicine* 10.1 (Jan. 2006), pp. 156–167.

[289] Rinat Khusainov et al. "Real-time human ambulation, activity, and physiological monitoring: Taxonomy of issues, techniques, applications, challenges and limitations". In: *Sensors (Switzerland)* 13.10 (Sept. 2013), pp. 12852–12902.

[290] Carlijn V.C. Bouten et al. "A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity". In: *IEEE Transactions on Biomedical Engineering* 44.3 (Mar. 1997), pp. 136–147.

[291] Kenneth A. Mann, Frederick W. Wernere, and Andrew K. Palmer. "Frequency spectrum analysis of wrist motion for activities of daily living". In: *Journal of Orthopaedic Research* 7.2 (1989), pp. 304–306.

[292] Kevin Supakkul. "Using Positional Heel-marker Data to More Accurately Calculate Stride Length for Treadmill Walking: A Step Length Approach". In: *arXiv* (Oct. 2017).

[293] Ciara M. O'Connor et al. "Automatic detection of gait events using kinematic data". In: *Gait and Posture* 25.3 (Mar. 2007), pp. 469–474.

[294] Eric Desailly et al. "Foot contact event detection using kinematic data in cerebral palsy children and normal adults gait". In: *Gait and Posture* 29.1 (Jan. 2009), pp. 76–80.

[295] Valerie J. Block et al. "Association of Continuous Assessment of Step Count by Remote Monitoring With Disability Progression Among Adults With Multiple Sclerosis". In: *JAMA network open* 2.3 (Mar. 2019), e190570.

[296] E. Flux et al. "The Human Body Model versus conventional gait models for kinematic gait analysis in children with cerebral palsy". In: *Human Movement Science* 70 (Apr. 2020), p. 102585.

[297] Felix Stief. "Variations of marker sets and models for standard gait analysis". In: *Handbook of Human Motion*. Vol. 1-3. Springer International Publishing, Apr. 2018, pp. 509–526.

[298] A. Sokolova and A. Konushin. "Methods of Gait Recognition in Video". In: *Programming and Computer Software* 45.4 (July 2019), pp. 213–220.

[299] Jeremy Jones and Duncan Hunter. “Qualitative Research: Consensus methods for medical and health services research”. In: *BMJ* 311.7001 (1995).

[300] Weijun Tao et al. “Gait Analysis Using Wearable Sensors”. In: *Sensors* 12.12 (Feb. 2012), pp. 2255–2283.

[301] Kenji Matsuo et al. “An attention-based activity recognition for egocentric video”. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. IEEE Computer Society, Sept. 2014, pp. 565–570.

[302] Yumi Iwashita et al. “First-person animal activity recognition from egocentric videos”. In: *Proceedings - International Conference on Pattern Recognition*. Institute of Electrical and Electronics Engineers Inc., Dec. 2014, pp. 4310–4315.

[303] Girmaw Abebe, Andrea Cavallaro, and Xavier Parra. “Robust multi-dimensional motion features for first-person vision activity recognition”. In: *Computer Vision and Image Understanding* 149 (Aug. 2016), pp. 229–248.

[304] Xuanhan Wang et al. “Deep appearance and motion learning for egocentric activity recognition”. In: *Neurocomputing* 275 (Jan. 2018), pp. 438–447.

[305] K. Konsolakis. “Physical Activity Recognition using Wearable Accelerometers in Controlled and Free-Living Environments”. PhD thesis. Delft University of Technology, 2018.

[306] Irvin Hussein Lopez-Nava and Angélica Muñoz-Meléndez. “Human action recognition based on low- and high-level data from wearable inertial sensors”. In: *International Journal of Distributed Sensor Networks* 15.12 (Dec. 2019).

[307] Georgios Meditskos et al. “Multi-modal activity recognition from egocentric vision, semantic enrichment and lifelogging applications for the care of dementia”. In: *Journal of Visual Communication and Image Representation* 51 (Feb. 2018), pp. 169–190.

[308] Davide Scaramuzza and Friedrich Fraundorfer. “Visual Odometry [Tutorial]”. In: *IEEE Robotics & Automation Magazine* 18.4 (Dec. 2011), pp. 80–92.

[309] D. Nister, O. Naroditsky, and J. Bergen. “Visual odometry”. In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.* Vol. 1. IEEE, pp. 652–659.

[310] Khalid Yousif, Alireza Bab-Hadiashar, and Reza Hoseinnezhad. “An Overview to Visual Odometry and Visual SLAM: Applications to Mobile Robotics”. In: *Intelligent Industrial Systems* 1.4 (Dec. 2015), pp. 289–311.

[311] Muhammed Tahsin Rahman et al. “Improving Multisensor Positioning of Land Vehicles with Integrated Visual Odometry for Next-Generation Self-Driving Cars”. In: *Journal of Advanced Transportation* 2018 (2018).

[312] Vitor Guizilini and Fabio Ramos. “Visual odometry learning for unmanned aerial vehicles”. In: *Proceedings - IEEE International Conference on Robotics and Automation*. 2011, pp. 6213–6220.

[313] Kyung M Kim. “Monocular Visual Odometry for Fixed Wing Unmanned Aircraft Systems”. PhD thesis. Air Force Institute of Technology, 2019.

[314] Mohammad O.A. Aqel et al. *Review of visual odometry: types, approaches, challenges, and applications*. Dec. 2016.

[315] Yang Cheng, Mark Maimone, and Larry Matthies. “Visual odometry on the Mars Exploration Rovers”. In: *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*. Vol. 1. 2005, pp. 903–910.

[316] Tung Sing Leung and Gerard Medioni. "Visual navigation aid for the blind in dynamic environments". In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. IEEE Computer Society, Sept. 2014, pp. 579–586.

[317] Young Hoon Lee and Gérard Medioni. "Wearable RGBD indoor navigation system for the blind". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 8927. Springer Verlag, 2015, pp. 493–508.

[318] Taragay Oskiper et al. "Visual odometry system using multiple stereo cameras and inertial measurement unit". In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2007.

[319] Francisco J. Perez-Grau et al. "An architecture for robust UAV navigation in GPS-denied areas". In: *Journal of Field Robotics* 35.1 (Jan. 2018), pp. 121–145.

[320] Andres Gomez et al. "Wearable, energy-opportunistic vision sensing for walking speed estimation". In: *SAS 2017 - 2017 IEEE Sensors Applications Symposium, Proceedings*. Institute of Electrical and Electronics Engineers Inc., Apr. 2017.

[321] Bradley Schneider et al. "Comparison of gait speeds from wearable camera and accelerometer in structured and semi-structured environments". In: *Healthcare Technology Letters* 7.1 (Feb. 2020), pp. 25–28.

[322] Marius Muja and David G. Lowe. "Fast matching of binary features". In: *Proceedings of the 2012 9th Conference on Computer and Robot Vision, CRV 2012*. 2012, pp. 404–410.

[323] Marius Muja and David G. Lowe. "Fast approximate nearest neighbors with automatic algorithm configuration". In: *VISAPP International Conference on Computer Vision Theory and Applications* (2009), pp. 331–340.

[324] Stefan Leutenegger et al. "Keyframe-based visual–inertial odometry using nonlinear optimization". In: *The International Journal of Robotics Research* 34.3 (Mar. 2015), pp. 314–334.

[325] Bill Triggs et al. "Bundle adjustment – a modern synthesis". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 1883. Springer Verlag, 2000, pp. 298–372.

[326] Stéphane Bazeille et al. "Combining Odometry and Visual Loop-Closure Detection for Consistent Topo-Metrical Mapping Stéphane". In: *Conference on Cognitive systems with Interactive Sensors (COGIS2009)*. France, 2009, pp. 1–7.

[327] Sen Wang et al. "DeepVO: Towards End-to-End Visual Odometry with Deep Recurrent Convolutional Neural Networks". In: *Proceedings - IEEE International Conference on Robotics and Automation* (Sept. 2017), pp. 2043–2050.

[328] Basam Musleh et al. "Estimation and prediction of the vehicle's motion based on visual odometry and Kalman filter". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 7517 LNCS. Springer, Berlin, Heidelberg, 2012, pp. 491–502.

[329] Manolis Lourakis and Xenophon Zabulis. "Accurate scale factor estimation in 3D reconstruction". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 8047 LNCS. PART 1. 2013, pp. 498–506.

[330] John Goodman. “Apollo 13 Guidance, Navigation, and Control Challenges”. In: *AIAA SPACE 2009 Conference & Exposition*. Reston, Virigina: American Institute of Aeronautics and Astronautics, Sept. 2009.

[331] A. Alaimo et al. “Comparison between Euler and quaternion parametrization in UAV dynamics”. In: *AIP Conference Proceedings*. Vol. 1558. 1. American Institute of PhysicsAIP, Oct. 2013, pp. 1228–1231.

[332] Pranjay Shyam, Antyanta Bangunharcana, and Kyung-Soo Kim. “Retaining Image Feature Matching Performance Under Low Light Conditions”. In: *International Conference on Control, Automation and Systems* 2020-October (Sept. 2020), pp. 1079–1085.

[333] Saman Sarraf, Cristina Saverino, and Ali Mohammad Golestani. “A robust and adaptive decision-making algorithm for detecting brain networks using functional MRI within the spatial and frequency domain”. In: *2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*. IEEE, Feb. 2016, pp. 53–56.

[334] Roberto Brunelli. *Template Matching Techniques in Computer Vision: Theory and Practice*. Wiley, Mar. 2009.

[335] Mohammad O.A. Aqel et al. “Adaptive-search template matching technique based on vehicle acceleration for monocular visual odometry system”. In: *IEEJ Transactions on Electrical and Electronic Engineering* 11.6 (Nov. 2016), pp. 739–752.

[336] Ramon Gonzalez et al. “Combined visual odometry and visual compass for off-road mobile robots localization”. In: *Robotica* 30.6 (Oct. 2012), pp. 865–878.

[337] Saman Sarraf et al. “Brain network extraction from probabilistic ICA using functional Magnetic Resonance Images and advanced template matching techniques”. In: *2014 IEEE 27th Canadian Conference on Electrical and Computer Engineering (CCECE)*. IEEE, May 2014, pp. 1–6.

[338] S. J. Maybank. “Properties of essential matrices”. In: *International Journal of Imaging Systems and Technology* 2.4 (24 1990), pp. 380–384.

[339] Ronald K. Pearson et al. “Generalized Hampel Filters”. In: *Eurasip Journal on Advances in Signal Processing* 2016.1 (Dec. 2016), p. 87.

[340] Diliang Chen et al. “Bring Gait Lab to Everyday Life: Gait Analysis in Terms of Activities of Daily Living”. In: *IEEE Internet of Things Journal* 7.2 (Feb. 2020), pp. 1298–1312.

[341] Se Won Lee et al. “Trunk sway during walking among older adults: norms and correlation with gait velocity.” In: *Gait & posture* 40.4 (Sept. 2014), pp. 676–81.

[342] Davide Giavarina. “Understanding Bland Altman analysis”. In: *Biochemia Medica* 25.2 (2015), pp. 141–151.

[343] Se Won Lee et al. “Trunk sway during walking among older adults: Norms and correlation with gait velocity”. In: *Gait and Posture* 40.4 (Sept. 2014), pp. 676–681.

[344] S. R. Lord, R. D. Clark, and I. W. Webster. “Postural stability and associated physiological factors in a population of aged persons”. In: *Journals of Gerontology* 46.3 (1991).

[345] Thierry Paillard and Frédéric Noé. “Techniques and Methods for Testing the Postural Function in Healthy and Pathological Subjects”. In: *BioMed Research International* 2015 (2015).

[346] G. R. Fernie et al. “The relationship of postural sway in standing to the incidence of falls in geriatric subjects”. In: *Age and Ageing* 11.1 (1982), pp. 11–16.

[347] Junhong Zhou et al. “The complexity of standing postural sway associates with future falls in community-dwelling older adults: The MOBILIZE Boston study”. In: *Scientific Reports* 7.1 (Dec. 2017).

[348] DA A. Winter. “Human balance and posture control during standing and walking”. In: *Gait and Posture* 3.4 (Dec. 1995), pp. 193–214.

[349] Seung Hyuck Shin and Chan Gook Park. “Adaptive step length estimation algorithm using optimal parameters and movement status awareness”. In: *Medical Engineering and Physics* 33.9 (Nov. 2011), pp. 1064–1071.

[350] Michaela Schimpl, Christian Lederer, and Martin Daumer. “Development and Validation of a New Method to Measure Walking Speed in Free-Living Environments Using the Actibelt® Platform”. In: *PLoS ONE* 6.8 (Aug. 2011). Ed. by Pablo Villoslada, e23080.

[351] Firas Massaad, Thierry M. Lejeune, and Christine Detrembleur. “The up and down bobbing of human walking: A compromise between muscle work and efficiency”. In: *Journal of Physiology* 582.2 (July 2007), pp. 789–799.

[352] Krista Merry and Pete Bettinger. “Smartphone GPS accuracy study in an urban environment”. In: *PLOS ONE* 14.7 (July 2019), e0219890.

[353] *Iffley Road Sports Centre*. 2020.

[354] Erik Stone and Marjorie Skubic. “Evaluation of an inexpensive depth camera for in-home gait assessment”. In: *Journal of Ambient Intelligence and Smart Environments* 3.4 (Jan. 2011), pp. 349–361.

[355] Farnood Gholami et al. “A Microsoft Kinect-Based Point-of-Care Gait Assessment Framework for Multiple Sclerosis Patients”. In: *IEEE Journal of Biomedical and Health Informatics* 21.5 (Sept. 2017), pp. 1376–1385.

[356] Shmuel Springer and Galit Yogev Seligmann. “Validity of the Kinect for Gait Assessment: A Focused Review”. In: *Sensors* 16.2 (Feb. 2016), p. 194.

[357] Shian-Ru Ke et al. “A Review on Video-Based Human Activity Recognition”. In: *Computers* 2.2 (June 2013), pp. 88–131.

[358] Albert Kim et al. “A wearable smartphone-enabled camera-based system for gait assessment”. In: *Gait & Posture* 42.2 (July 2015), pp. 138–144.

[359] Jeffery Young, Milena Simic, and Milan Simic. “Gait Monitoring by Automatic Guided Vehicle”. In: *Procedia Computer Science*. Vol. 112. Elsevier B.V., 2017, pp. 2137–2146.

[360] Akshay Rangesh and Mohan Manubhai Trivedi. “No Blind Spots: Full-Surround Multi-Object Tracking for Autonomous Vehicles Using Cameras and LiDARs”. In: *IEEE Transactions on Intelligent Vehicles* 4.4 (Dec. 2019), pp. 588–599.

[361] Kantip Kiratiratanapruk and Supakorn Siddhichai. “Vehicle detection and tracking for traffic monitoring system”. In: *IEEE Region 10 Annual International Conference, Proceedings/TENCON*. Institute of Electrical and Electronics Engineers Inc., 2006.

[362] Shugang Zhang et al. “A Review on Human Activity Recognition Using Vision-Based Method”. In: *Journal of Healthcare Engineering* 2017 (2017).

[363] Kun-Lin Lu and Edward Chu. "An Image-Based Fall Detection System for the Elderly". In: *Applied Sciences* 8.10 (Oct. 2018), p. 1995.

[364] Jake Wall et al. "Novel opportunities for wildlife conservation and research with real-time monitoring". In: *Ecological Applications* 24.4 (June 2014), pp. 593–601.

[365] Prashanth C. Ravoor and Sudarshan T.S.B. "Deep Learning Methods for Multi-Species Animal Re-identification and Tracking a Survey". In: *Computer Science Review* 38 (Nov. 2020), p. 100289.

[366] Akanksha Rathore et al. "Multi-object tracking in heterogeneous environments (MOTHe) for animal space-use studies". In: *bioRxiv* (Jan. 2020), p. 2020.01.10.899989.

[367] Anthony I. Dell et al. "Automated image-based tracking and its application in ecology". In: *Trends in Ecology and Evolution* 29.7 (July 2014), pp. 417–428.

[368] Alice A. Robie et al. "Machine vision methods for analyzing social interactions". In: *Journal of Experimental Biology* 220.1 (Jan. 2017), pp. 25–34.

[369] Youngmin Park, Vincent Lepetit, and Ntack Woo. "Multiple 3D object tracking for augmented reality". In: *Proceedings - 7th IEEE International Symposium on Mixed and Augmented Reality 2008, ISMAR 2008*. 2008, pp. 117–120.

[370] Jong Sue Bae et al. "An imaging target tracking software for a precision guided missile application". In: *13th Conference on Information Fusion, Fusion 2010*. 2010.

[371] Christoforos Kanellakis and George Nikolakopoulos. "Survey on Computer Vision for UAVs: Current Developments and Trends". In: *Journal of Intelligent and Robotic Systems: Theory and Applications* 87.1 (July 2017), pp. 141–168.

[372] Andreas Nussberger, Helmut Grabner, and Luc Van Gool. "Aerial object tracking from an airborne platform". In: *2014 International Conference on Unmanned Aircraft Systems, ICUAS 2014 - Conference Proceedings*. IEEE Computer Society, 2014, pp. 1284–1293.

[373] Yi Wu, Jongwoo Lim, and Ming Hsuan Yang. "Online object tracking: A benchmark". In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2013, pp. 2411–2418.

[374] Peng Gao et al. "Siamese attentional keypoint network for high performance visual tracking". In: *Knowledge-Based Systems* 193 (Apr. 2020).

[375] Matthias Mueller, Neil Smith, and Bernard Ghanem. "A benchmark and simulator for UAV tracking". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 9905 LNCS. Springer Verlag, 2016, pp. 445–461.

[376] Daniel M.T. Fessler, Kevin J. Haley, and Roshni D. Lal. "Sexual dimorphism in foot length proportionate to stature". In: *Annals of Human Biology* 32.1 (Jan. 2005), pp. 44–59.

[377] Yvonne M. Golightly et al. "Foot disorders associated with overpronated and oversupinated foot Function: The johnston county osteoarthritis project". In: *Foot and Ankle International* 35.11 (Nov. 2014), pp. 1159–1165.

[378] D. R. Wenger and J. Leach. "Foot deformities in infants and children". In: *Pediatric Clinics of North America* 33.6 (Dec. 1986), pp. 1411–1427.

[379] Wenhan Luo et al. "Multiple Object Tracking: A Literature Review". In: *Artificial Intelligence* 293 (Sept. 2014).

[380] Kemal Oksuz et al. *Imbalance problems in object detection: A review*. Aug. 2019.

[381] Wang Zhiqiang and Liu Jun. "A review of object detection based on convolutional neural network". In: *Chinese Control Conference, CCC*. IEEE Computer Society, Sept. 2017, pp. 11104–11109.

[382] Tjokorda Agung Budi Wirayuda et al. "Real-time hand-tracking on video image based on palm geometry". In: *2013 International Conference of Information and Communication Technology, ICoICT 2013*. 2013, pp. 241–246.

[383] Jonathan Huang et al. "Speed/accuracy trade-offs for modern convolutional object detectors". In: *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*. Vol. 2017-January. Institute of Electrical and Electronics Engineers Inc., Nov. 2017, pp. 3296–3305.

[384] Kaiming He et al. "Deep residual learning for image recognition". In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Vol. 2016-December. IEEE Computer Society, Dec. 2016, pp. 770–778.

[385] Tsung-Yi Lin et al. "Feature Pyramid Networks for Object Detection". In: *arXiv* (Dec. 2016).

[386] Martín Abadi et al. "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems". In: *arXiv* (Mar. 2016).

[387] Tsung Yi Lin et al. "Microsoft COCO: Common objects in context". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 8693 LNCS. PART 5. Springer Verlag, May 2014, pp. 740–755.

[388] Tzutalin. *LabelImg*. 2015.

[389] Tao Hu, Lichao Huang, and Han Shen. "Multi-object Tracking via End-to-end Tracklet Searching and Ranking". In: *arXiv* (Mar. 2020).

[390] Hiya Roy et al. "Toward Better Planetary Surface Exploration by Orbital Imagery Inpainting". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14 (2021), pp. 175–189.

[391] Jake Cowton, Ilias Kyriazakis, and Jaume Bacardit. "Automated Individual Pig Localisation, Tracking and Behaviour Metric Extraction Using Deep Learning". In: *IEEE Access* 7 (2019), pp. 108049–108060.

[392] Nicole A. Capela, Edward D. Lemaire, and Natalie C. Baddour. "A smartphone approach for the 2 and 6-minute walk test". In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Vol. 2014. IEEE, Aug. 2014, pp. 958–961.

[393] Randal C. Foster et al. "Precision and accuracy of an ankle-worn accelerometer-based pedometer in step counting and energy expenditure". In: *Preventive Medicine* 41.3-4 (Sept. 2005), pp. 778–783.

[394] François Modave et al. "Mobile device accuracy for step counting across age groups". In: *JMIR mHealth and uHealth* 5.6 (June 2017), e7870.

[395] Meredith A. Case et al. "Accuracy of Smartphone Applications and Wearable Devices for Tracking Physical Activity Data". In: *JAMA* 313.6 (Feb. 2015), p. 625.

[396] Meng Shiuan Pan and Hsueh Wei Lin. "A step counting algorithm for smartphone users: Design and implementation". In: *IEEE Sensors Journal* 15.4 (Apr. 2015), pp. 2296–2305.

[397] Agata Brajdic and Robert Harle. "Walk detection and step counting on unconstrained smartphones". In: *UbiComp 2013 - Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 2013, pp. 225–234.

[398] Fabio A. Storm, Ben W. Heller, and Claudia Mazzà. "Step Detection and Activity Recognition Accuracy of Seven Physical Activity Monitors". In: *PLOS ONE* 10.3 (Mar. 2015). Ed. by Lei Ren, e0118723.

[399] Ngoc Huynh Ho, Phuc Huu Truong, and Gu Min Jeong. "Step-detection and adaptive step-length estimation for pedestrian dead-reckoning at variouswalking speeds using a smartphone". In: *Sensors (Switzerland)* 16.9 (Sept. 2016).

[400] Estefania Munoz Diaz and Ana Luz Mendiguchia Gonzalez. "Step detector and step length estimator for an inertial pocket navigation system". In: *IPIN 2014 - 2014 International Conference on Indoor Positioning and Indoor Navigation*. Institute of Electrical and Electronics Engineers Inc., 2014, pp. 105–110.

[401] M. A. van Loo et al. "Test-re-test reliability of walking speed, step length and step width measurement after traumatic brain injury: A pilot study". In: *Brain Injury* 18.10 (2004), pp. 1041–1048.

[402] Taufique Sayeed et al. "Adapted step length estimators for patients with Parkinson's disease using a lateral belt worn accelerometer". In: *Technology and Health Care* 23.2 (Jan. 2015), pp. 179–194.

[403] Steven Diaz, Shawn Disdier, and Miguel A. Labrador. "Step Length and Step Width Estimation using Wearable Sensors". In: *2018 9th IEEE Annual Ubiquitous Computing, Electronics and Mobile Communication Conference, UEMCON 2018*. Institute of Electrical and Electronics Engineers Inc., Nov. 2018, pp. 997–1001.

[404] Haifeng Xing et al. "Pedestrian Stride Length Estimation from IMU Measurements and ANN Based Algorithm". In: *Journal of Sensors* 2017 (2017).

[405] James R. Williamson et al. "Detecting Parkinson's Disease from Wrist-Worn Accelerometry in the U.K. Biobank". In: *Sensors (Basel, Switzerland)* 21.6 (Mar. 2021), pp. 1–18.

[406] Vijay R. Varma et al. "Continuous gait monitoring discriminates community-dwelling mild Alzheimer's disease from cognitively normal controls". In: *Alzheimer's & Dementia : Translational Research & Clinical Interventions* 7.1 (Jan. 2021).

[407] Massimiliano Pau et al. "Smoothness of gait detects early alterations of walking in persons with multiple sclerosis without disability". In: *Gait & posture* 58 (Oct. 2017), pp. 307–309.

[408] Jens Barth et al. "Biometric and mobile gait analysis for early diagnosis and therapy monitoring in Parkinson's disease". In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, Aug. 2011, pp. 868–871.

[409] Michael J. Bradshaw et al. "Wearable biosensors to monitor disability in multiple sclerosis". In: *Neurology: Clinical Practice* 7.4 (Aug. 2017), pp. 354–362.

[410] Ana Lígia Silva de Lima et al. "Freezing of gait and fall detection in Parkinson's disease using wearable sensors: a systematic review". In: *Journal of Neurology* 264.8 (Aug. 2017), pp. 1642–1654.

[411] Zhuo Wang et al. "Possible Life Saver: A Review on Human Fall Detection Technology". In: *Robotics 2020, Vol. 9, Page 55* 9.3 (July 2020), p. 55.

[412] Rohit Tanwar et al. “Pathway of Trends and Technologies in Fall Detection: A Systematic Review”. In: *Healthcare 2022, Vol. 10, Page 172* 10.1 (Jan. 2022), p. 172.

[413] Zmnako Amen, Omer Rafiq, and Sherwan Moer. “Gait analysis in ambulatory children with cerebral palsy using a tri-axial accelerometer”. In: *Acta Neuropsychologica* 18 (Dec. 2020), pp. 465–476.

[414] Paulo Roberto Fonseca Junior. “Gait Analysis Using Wearable Sensor Inertial in a Child with CP after Orthopedic Surgery: Case Report”. In: *Clin Case Rep Int.* 2.1 (2018).

[415] Serene S. Paul et al. “Validity of the Fitbit activity tracker for measuring steps in community-dwelling older adults”. In: *BMJ Open Sport & Exercise Medicine* 1.1 (July 2015).

[416] Carlos D. Gómez-Carmona et al. “Accelerometry as a method for external workload monitoring in invasion team sports. A systematic review”. In: *PLoS ONE* 15.8 (Aug. 2020).

[417] Michael V. Potter et al. “Effect of IMU Design on IMU-Derived Stride Metrics for Running”. In: *Sensors (Basel, Switzerland)* 19.11 (June 2019).

[418] Vânia Guimarães, Inês Sousa, and Miguel Velhote Correia. “Orientation-Invariant Spatio-Temporal Gait Analysis Using Foot-Worn Inertial Sensors”. In: *Sensors 2021, Vol. 21, Page 3940* 21.11 (June 2021), p. 3940.

[419] Aliénor Vienne et al. “Inertial sensors to assess gait quality in patients with neurological disorders: A systematic review of technical and analytical challenges”. In: *Frontiers in Psychology* 8.MAY (May 2017), p. 817.

[420] Sheila M. Korpan et al. “Effect of ActiGraph GT3X+ Position and Algorithm Choice on Step Count Accuracy in Older Adults”. In: *Journal of Aging and Physical Activity* 23.3 (July 2015), pp. 377–382.

[421] Sylvia Horvath et al. “The effect of pedometer position and normal gait asymmetry on step count accuracy”. In: *Applied physiology, nutrition, and metabolism = Physiologie appliquee, nutrition et metabolisme* 32.3 (June 2007), pp. 409–415.

[422] Aleksandra Dybus et al. “Validation of smartphone step count algorithm used in STARFISH smartphone application”. In: *Technology and Health Care* 25.6 (Jan. 2017), pp. 1157–1162.

[423] Xiaomin Kang, Baoqi Huang, and Guodong Qi. “A Novel Walking Detection and Step Counting Algorithm Using Unconstrained Smartphones”. In: *Sensors* 18.1 (Jan. 2018), p. 297.

[424] Vincenzo Genovese, Andrea Mannini, and Angelo M. Sabatini. “A Smartwatch Step Counter for Slow and Intermittent Ambulation”. In: *IEEE Access* 5 (2017), pp. 13028–13037.

[425] Parastoo Alinia et al. “How Accurate Is Your Activity Tracker? A Comparative Study of Step Counts in Low-Intensity Physical Activities”. In: *JMIR Mhealth Uhealth* 5.8 (Aug. 2017).

[426] Vahid Abdollah et al. “Investigating the validity of a single tri-axial accelerometer mounted on the head for monitoring the activities of daily living and the timed-up and go test”. In: *Gait & Posture* 90 (Oct. 2021), pp. 137–140.

[427] Matteo Gadaleta and Michele Rossi. “IDNet: Smartphone-based Gait Recognition with Convolutional Neural Networks”. In: *Pattern Recognition* 74 (June 2016), pp. 25–37.

[428] Kai Kunze et al. “Which way am I facing: Inferring horizontal device orientation from an accelerometer signal”. In: *Proceedings - International Symposium on Wearable Computers, ISWC* (2009), pp. 149–150.

[429] Jonatan Fridolfsson, Mats Börjesson, and Daniel Arvidsson. “A Biomechanical Re-Examination of Physical Activity Measurement with Accelerometers”. In: *Sensors 2018, Vol. 18, Page 3399* 18.10 (Oct. 2018), p. 3399.

[430] David A Winter. *Biomechanics and Motor Control of Human Movement*. 4th. Vol. 7. 2009, pp. 6–22.

[431] Jonathan Sinclair, Paul John Taylor, and Sarah Jane Hobbs. “Digital Filtering of Three-Dimensional Lower Extremity Kinematics: an Assessment”. In: *Journal of Human Kinetics* 39.1 (2013), p. 25.

[432] Bing Yu et al. “Estimate of the optimum cutoff frequency for the Butterworth low-pass digital filter”. In: *Journal of Applied Biomechanics* 15.3 (1999), pp. 318–329.

[433] Robertson DG and Dowling JJ. “Design and responses of Butterworth and critically damped digital filters”. In: *Journal of electromyography and kinesiology : official journal of the International Society of Electrophysiological Kinesiology* 13.6 (2003), pp. 569–573.

[434] Wiebren Zijlstra and At L Hof. “Assessment of spatio-temporal gait parameters from trunk accelerations during human walking”. In: *Gait & Posture* 18.2 (Oct. 2003), pp. 1–10.

[435] W. Zijlstra and A. L. Hof. “Displacement of the pelvis during human walking: experimental data and model predictions”. In: *Gait & Posture* 6.3 (Dec. 1997), pp. 249–262.

[436] G. M. Lyons et al. “A description of an accelerometer-based mobility monitoring technique”. In: *Medical Engineering & Physics* 27.6 (July 2005), pp. 497–504.

[437] John H. Hollman, Eric M. McDade, and Ronald C. Petersen. “Normative spatiotemporal gait parameters in older adults”. In: *Gait & Posture* 34.1 (May 2011), pp. 111–118.

[438] Zebo Zhou et al. “Behaviors classification based distance measuring system for pedestrians via a foot-mounted inertial sensor”. In: *Asian Journal of Control* 21.4 (July 2019), pp. 1483–1495.

[439] Liqiang Zhang et al. “Learning-based stance-phase detection for a pedestrian dead-reckoning system with dynamic gait speeds”. In: *Measurement Science and Technology* 32.10 (June 2021), p. 105108.

[440] Alvarez JC et al. “Multisensor approach to walking distance estimation with foot inertial sensing”. In: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference* 2007 (2007), pp. 5719–5722.

[441] Wang JS et al. “Walking pattern classification and walking distance estimation algorithms using gait phase information”. In: *IEEE transactions on bio-medical engineering* 59.10 (2012), pp. 2884–2892.

[442] Yang Bai et al. “Comparison of Consumer and Research Monitors under Semistructured Settings”. In: *Medicine and science in sports and exercise* 48.1 (2016), pp. 151–158.

[443] Dimitrios Psaltos et al. “Multimodal Wearable Sensors to Measure Gait and Voice”. In: *Digital Biomarkers* 3.3 (Sept. 2019), pp. 133–144.

[444] F. Prince et al. “Anticipatory control of upper body balance during human locomotion”. In: *Gait & Posture* 2.1 (Mar. 1994), pp. 19–25.

[445] Christopher J. Arellano and Rodger Kram. “The effects of step width and arm swing on energetic cost and lateral balance during running”. In: *Journal of Biomechanics* 44.7 (Apr. 2011), pp. 1291–1295.

[446] Duong Trong Bui, Nhan Duc Nguyen, and Gu Min Jeong. “A robust step detection algorithm and walking distance estimation based on daily wrist activity recognition using a smart band”. In: *Sensors (Switzerland)* 18.7 (July 2018).

[447] Hwan Hee Lee, Suji Choi, and Myeong Jin Lee. “Step detection robust against the dynamics of smartphones”. In: *Sensors (Switzerland)* 15.10 (Oct. 2015), pp. 27230–27250.

[448] Patrick Robertson, Michael Angermann, and Bernhard Krach. “Simultaneous localization and mapping for pedestrians using only foot-mounted inertial sensors”. In: *UbiComp ’09: Proceedings of the 11th international conference on Ubiquitous computing*. 2009, pp. 93–96.

[449] Lindsay P. Toth et al. “Effects of Brief Intermittent Walking Bouts on Step Count Accuracy of Wearable Devices”. In: *Journal for the Measurement of Physical Behaviour* 2.1 (Mar. 2019), pp. 13–21.

[450] Ty Ferguson et al. “The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study”. In: *International Journal of Behavioral Nutrition and Physical Activity* 12.1 (Dec. 2015), p. 42.

[451] Yuanyuan Feng et al. “Comparison of tri-axial accelerometers step-count accuracy in slow walking conditions”. In: *Gait & Posture* 53 (Mar. 2017), pp. 11–16.

[452] Kinh Tran, Tu Le, and Tien Dinh. “A high-accuracy step counting algorithm for iPhones using Accelerometer”. In: *2012 IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2012* (2012), pp. 213–217.

[453] Zehra Naqvi et al. “Step Counting Using Smartphone-Based Accelerometer”. In: *International Journal on Computer Science and Engineering* 4.5 (May 2012).

[454] Christopher C. Moore et al. “Toward Harmonized Treadmill-Based Validation of Step-Counting Wearable Technologies: A Scoping Review”. In: *Journal of Physical Activity and Health* 17.8 (July 2020), pp. 840–852.

[455] Van Thanh Pham et al. “Highly Accurate Step Counting at Various Walking States Using Low-Cost Inertial Measurement Unit Support Indoor Positioning System”. In: *Sensors (Basel, Switzerland)* 18.10 (Oct. 2018).

[456] Pham Van Thanh et al. “A novel step counter supporting for indoor positioning based on inertial measurement unit”. In: *Proceedings of 2017 7th International Conference on Integrated Circuits, Design, and Verification, ICDV 2017* 2017-December (Dec. 2017), pp. 69–74.

[457] Fuqiang Gu et al. “Robust and accurate smartphone-based step counting for indoor localization”. In: *IEEE Sensors Journal* 17.11 (June 2017), pp. 3453–3460.

[458] Hani Muhsen, Odeh Al-Amaydeh, and Rakan Al-Hamlan. “Algorithm design for accurate steps counting based on smartphone sensors for indoor applications”. In: *Advances in Science, Technology and Engineering Systems* 5.6 (2020), pp. 811–816.

[459] Ahmet Dirican and Selim Aksoy. “ Step Counting Using Smartphone Accelerometer and Fast Fourier Transform”. In: *Innovations on Intelligent Systems and Applications Symposium* 8.2 (2017).

[460] Xin Liu et al. “A Novel Robust Step Detection Algorithm for Foot-Mounted IMU”. In: *IEEE Sensors Journal* 21.4 (Feb. 2021), pp. 5331–5339.

[461] A.R. Jimenez et al. “A comparison of Pedestrian Dead-Reckoning algorithms using a low-cost MEMS IMU”. In: *2009 IEEE International Symposium on Intelligent Signal Processing*. IEEE, Aug. 2009, pp. 37–42.

[462] Arif Reza Anwary, Hongnian Yu, and Michael Vassallo. “Optimal Foot Location for Placing Wearable IMU Sensors and Automatic Feature Extraction for Gait Analysis”. In: *IEEE Sensors Journal* 18.6 (Mar. 2018), pp. 2555–2567.

[463] Fatema El-Amrawy and Mohamed Ismail Nounou. “Are Currently Available Wearable Devices for Activity Tracking and Heart Rate Monitoring Accurate, Precise, and Medically Beneficial?” In: *Healthcare Informatics Research* 21.4 (Oct. 2015), pp. 315–320.

[464] Daniel Fuller et al. “Reliability and Validity of Commercially Available Wearable Devices for Measuring Steps, Energy Expenditure, and Heart Rate: Systematic Review”. In: *JMIR Mhealth Uhealth 2020;8(9):e18694 https://mhealth.jmir.org/2020/9/e18694* 8.9 (Sept. 2020), e18694.

[465] Jeffrey Montes et al. “Step count reliability and validity of five wearable technology devices while walking and jogging in both a free motion setting and on a treadmill”. In: *International Journal of Exercise Science* 13.7 (2020), pp. 410–426.

[466] Lei Wang et al. “Evaluation on Step Counting Performance of Wristband Activity Monitors in Daily Living Environment”. In: *IEEE Access* 5 (2017), pp. 13020–13027.

[467] Riccardo Femiano et al. “Validation of open-source step-counting algorithms for wrist-worn tri-axial accelerometers in cardiovascular patients”. In: *Gait & Posture* 92 (Feb. 2022), pp. 206–211.

[468] Buke Ao et al. “Context Impacts in Accelerometer-Based Walk Detection and Step Counting”. In: *Sensors 2018* 18.11 (Oct. 2018), p. 3604.

[469] Markus Lueken et al. “Peak detection algorithm for gait segmentation in long-Term monitoring for stride time estimation using inertial measurement sensors”. In: *2019 IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2019 - Proceedings*. Institute of Electrical and Electronics Engineers Inc., May 2019.

[470] John McCamley et al. “An enhanced estimate of initial contact and final contact instants of time using lower trunk inertial sensor data”. In: *Gait & posture* 36.2 (June 2012), pp. 316–318.

[471] Mohamed Boutaayamou et al. “Segmentation of gait cycles using foot-mounted 3D accelerometers”. In: *2015 International Conference on 3D Imaging, IC3D 2015 - Proceedings* (Jan. 2016).

[472] Mohamed Boutaayamou et al. “Algorithm for temporal gait analysis using wireless foot-mounted accelerometers”. In: *Communications in Computer and Information Science* 690 (2017), pp. 236–254.

[473] G V Prateek et al. “Gait Cycle Validation and Segmentation Using Inertial Sensors”. In: *IEEE Transactions on Biomedical Engineering* 67.8 (2020).

[474] Ning Ji et al. “Appropriate Mother Wavelets for Continuous Gait Event Detection Based on Time-Frequency Analysis for Hemiplegic and Healthy Individuals”. In: *Sensors (Basel, Switzerland)* 19.16 (Aug. 2019).

[475] Minh H. Pham et al. “Validation of a Step Detection Algorithm during Straight Walking and Turning in Patients with Parkinson’s Disease and Older Adults Using an Inertial Measurement Unit at the Lower Back”. In: *Frontiers in Neurology* 8.SEP (Sept. 2017), p. 457.

[476] Jens Barth et al. “Subsequence dynamic time warping as a method for robust step segmentation using gyroscope signals of daily life activities”. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (2013), pp. 6744–6747.

[477] Aliénor Vienne-Jumeau et al. “Personalized Template-Based Step Detection From Inertial Measurement Units Signals in Multiple Sclerosis”. In: *Frontiers in Neurology* 11 (Apr. 2020), p. 261.

[478] Nils Roth et al. “Hidden Markov Model based stride segmentation on unsupervised free-living gait data in Parkinson’s disease patients”. In: *Journal of NeuroEngineering and Rehabilitation* 18.1 (Dec. 2021), pp. 1–15.

[479] Thilo Pfau et al. “A hidden Markov model-based stride segmentation technique applied to equine inertial sensor trunk movement data”. In: *Journal of biomechanics* 41.1 (2008), pp. 216–220.

[480] A. Mannini and A. M. Sabatini. “A hidden Markov model-based technique for gait segmentation using a foot-mounted gyroscope”. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society.* Vol. 2011. IEEE, Aug. 2011, pp. 4369–4373.

[481] Andrea Mannini and Angelo Maria Sabatini. “Gait phase detection and discrimination between walking-jogging activities using hidden Markov models applied to foot motion data from a gyroscope”. In: *Gait & posture* 36.4 (Sept. 2012), pp. 657–661.

[482] Andrea Mannini et al. “Hidden Markov model-based strategy for gait segmentation using inertial sensors: Application to elderly, hemiparetic patients and Huntington’s disease patients”. In: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference* 2015 (Nov. 2015), pp. 5179–5182.

[483] Yingbiao Yao et al. “A Robust Step Detection and Stride Length Estimation for Pedestrian Dead Reckoning Using a Smartphone”. In: *IEEE Sensors Journal* 20.17 (Sept. 2020), pp. 9685–9697.

[484] I. H. López-Nava et al. “Estimation of temporal gait parameters using Bayesian models on acceleration signals”. In: *Comput Methods Biomech Biomed Engin.* 19.4 (Mar. 2016), pp. 396–403.

[485] Johan Wahlstrom et al. “Zero-Velocity Detection-A Bayesian Approach to Adaptive Thresholding”. In: *IEEE Sensors Letters* 3.6 (June 2019).

[486] Tobias Steinmetzer et al. “Comparison of algorithms and classifiers for stride detection using wearables”. In: *Neural Computing and Applications* 32.24 (Dec. 2020), pp. 17857–17868.

[487] Matteo Gadaleta et al. “Deep Learning Techniques for Improving Digital Gait Segmentation”. In: *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (July 2019), pp. 1834–1837.

[488] Wenhua Shao et al. “DePedo: Anti periodic negative-step movement pedometer with deep convolutional neural networks”. In: *IEEE International Conference on Communications* 2018-May (July 2018).

[489] Andrea Ferlini et al. “EarGate: Gait-based User Identification with In-ear Microphones”. In: *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking* (2022).

[490] Ju Han and Bir Bhanu. “Individual recognition using gait energy image”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28.2 (2006), pp. 316–322.

[491] Andrew P. Creagh et al. “Interpretable deep learning for the remote characterisation of ambulation in multiple sclerosis using smartphones”. In: *Scientific Reports 2021 11:1* 11.1 (July 2021), pp. 1–14.

[492] Tommy Oberg, Alek Karsznia, and Kurt Oberg. “Basic gait parameters : Reference data for normal subjects, 10-79 years of age”. In: *Department of Veterans Affairs Journal of Rehabilitation Research and Development* 30.2 (1993), pp. 210–223.

[493] T. M. Steffen and L. A. Mollinger. “Age- and gender-related test performance in community-dwelling adults”. In: *Journal of neurologic physical therapy : JNPT* 29.4 (2005), pp. 181–188.

[494] Qaiser Riaz et al. “One Small Step for a Man: Estimation of Gender, Age and Height from Recordings of One Step by a Single Inertial Sensor”. In: *Sensors 2015, Vol. 15, Pages 31999-32019* 15.12 (Dec. 2015), pp. 31999–32019.

[495] Tim Van Hamme et al. “A systematic comparison of age and gender prediction on IMU sensor-based gait traces”. In: *Sensors (Switzerland)* 19.13 (July 2019).

[496] Michele L. Callisaya et al. “Ageing and gait variability–a population-based study of older people”. In: *Age and ageing* 39.2 (Jan. 2010), pp. 191–197.

[497] Paola Tamburini et al. “Moving from laboratory to real life conditions: Influence on the assessment of variability and stability of gait”. In: *Gait & Posture* 59 (Jan. 2018), pp. 248–252.

[498] Vrutangkumar V. Shah et al. “Effect of Bout Length on Gait Measures in People with and without Parkinson’s Disease during Daily Life”. In: *Sensors 2020, Vol. 20, Page 5769* 20.20 (Oct. 2020), p. 5769.

[499] Zhi-An Deng et al. “Heading Estimation for Indoor Pedestrian Navigation Using a Smartphone in the Pocket”. In: *Sensors 2015, Vol. 15, Pages 21518-21536* 15.9 (Aug. 2015), pp. 21518–21536.

[500] Théo Jourdan, Noëlie Debs, and Carole Frindel. “The Contribution of Machine Learning in the Validation of Commercial Wearable Sensors for Gait Monitoring in Patients: A Systematic Review”. In: *Sensors 2021, Vol. 21, Page 4808* 21.14 (July 2021), p. 4808.

[501] Rafael Caldas et al. “A systematic review of gait analysis methods based on inertial sensors and adaptive algorithms”. In: *Gait & Posture* 57 (Sept. 2017), pp. 204–210.

[502] Basil Lin. “Machine Learning and Pedometers: An Integration-Based Convolutional Neural Network for Step Counting and Detection”. PhD thesis. Clemson University, Dec. 2020.

[503] Jan Kupke et al. “Development of a step counter based on artificial neural networks”. In: *Journal of Location Based Services* 10.3 (July 2016), pp. 161–177.

[504] Ziyi Chen. “An LSTM Recurrent Network for Step Counting”. In: *arXiv* (Feb. 2018).

[505] Estelle Raffin, Stéphane Bonnet, and Pascal Giraux. “Concurrent validation of a magnetometer-based step counter in various walking surfaces”. In: *Gait & Posture* 35.1 (Jan. 2012), pp. 18–22.

[506] Maan Khedr and Nasser El-Sheimy. “A Smartphone Step Counter Using IMU and Magnetometer for Navigation and Health Monitoring Applications”. In: *Sensors* 17.11 (Nov. 2017), p. 2573.

[507] László Kundra and Péter Ekler. “The summary of indoor navigation possibilities considering mobile environment”. In: *Proceedings - 2013 IEEE 3rd Eastern European Regional Conference on the Engineering of Computer Based Systems, ECBS-EERC 2013* (2013), pp. 165–166.

[508] Dayu Yan, Chuang Shi, and Tuan Li. “An improved PDR system with accurate heading and step length estimation using handheld smartphone”. In: *The Journal of Navigation* (2021), pp. 1–19.

[509] Zijun Zhou et al. “Pedestrian Navigation Method Based on Machine Learning and Gait Feature Assistance”. In: *Sensors 2020, Vol. 20, Page 1530* 20.5 (Mar. 2020), p. 1530.

[510] Sen Qiu et al. “Inertial/magnetic sensors based pedestrian dead reckoning by means of multi-sensor fusion”. In: *Information Fusion* 39 (Jan. 2018), pp. 108–119.

[511] Shanzhi Gu et al. “The Improvement of Traditional Indoor Localization Model Using Magnetic Field Based on Smartphone”. In: *Proceedings of IEEE 14th International Conference on Intelligent Systems and Knowledge Engineering, ISKE 2019* (Nov. 2019), pp. 694–700.

[512] F. Salis et al. “A wearable multi-sensor system for real world gait analysis”. In: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society.* Vol. 2021. NLM (Medline), Nov. 2021, pp. 7020–7023.

[513] Tong-Hun Hwang et al. “Validation of Real Time Gait Analysis Using a Single Head-Worn IMU”. In: *EKC 2019 Conference Proceedings* (July 2019), pp. 87–97.

[514] Jay Prakash et al. “STEAR: Robust Step Counting from Earables”. In: *Proceedings of the 1st International Workshop on Earable Computing, EarComp 2019* (Sept. 2019), pp. 36–41.

[515] Teng Chia Wang et al. “IMU-based Smart Knee Pad for Walking Distance and Stride Count Measurement”. In: *Proceedings - International Symposium on Quality Electronic Design, ISQED* 2020-March (Mar. 2020), pp. 173–178.

[516] Matthew B. Rhudy and Joseph M. Mahoney. “A comprehensive comparison of simple step counting techniques using wrist- and ankle-mounted accelerometer and gyroscope signals”. In: *J Med Eng Technol* 42.3 (Apr. 2018), pp. 236–243.

[517] A. Godfrey et al. “Activity classification using a single chest mounted tri-axial accelerometer”. In: *Medical Engineering & Physics* 33.9 (Nov. 2011), pp. 1127–1135.

[518] Arun Kumar Siddanahalli Ninge Gowda, Swarna Ravindra Babu, and Dhineshkumar Chandra Sekaran. “UMOISP: Usage Mode and Orientation Invariant Smartphone Pedometer”. In: *IEEE Sensors Journal* 17.3 (Feb. 2017), pp. 869–881.

[519] Andrew P. Creagh et al. “Smartphone-and Smartwatch-Based Remote Characterisation of Ambulation in Multiple Sclerosis during the Two-Minute Walk Test”. In: *IEEE Journal of Biomedical and Health Informatics* 25.3 (Mar. 2021), pp. 838–849.

[520] Shigeru Haga et al. “Effects of using a Smart Phone on Pedestrians’ Attention and Walking”. In: *Procedia Manufacturing* 3 (Jan. 2015), pp. 2574–2580.

[521] Rohini R. Rao, Akriti Singh, and Vikhyath Kamath. “Assessment of step accuracy and usability of activity trackers”. In: *Proceedings of the 3rd International Conference on Contemporary Computing and Informatics, IC3I 2018* (Oct. 2018), pp. 38–42.

[522] Tomas Vetrovsky et al. “Validity of six consumer-level activity monitors for measuring steps in patients with chronic heart failure”. In: *PLOS ONE* 14.9 (Sept. 2019), e0222569.

[523] Jessica J. Chow et al. “Accuracy of step count measured by physical activity monitors: The effect of gait speed and anatomical placement site”. In: *Gait & Posture* 57 (Sept. 2017), pp. 199–203.

[524] Lucian Bezuidenhout et al. “Validity of hip and ankle worn actigraph accelerometers for measuring steps as a function of gait speed during steady state walking and continuous turning”. In: *Sensors* 21.9 (May 2021).

[525] Lorraine J. Phillips, Gregory F. Petroski, and Natalie E. Markis. “A comparison of accelerometer accuracy in older adults”. In: *Research in Gerontological Nursing* 8.5 (Nov. 2015), pp. 213–219.

[526] Sarah J. Gilmore et al. “The validity of using activity monitors to detect step count after lumbar fusion surgery”. In: *Disability and rehabilitation* 42.6 (Mar. 2020), pp. 863–868.

[527] Charlotte Brun Thorup et al. “Accuracy of a step counter during treadmill and daily life walking by healthy adults and patients with cardiac disease”. In: *BMJ Open* 7.3 (Mar. 2017).

[528] Marina Koytcheva. *Market Forecast: Wearables Worldwide, 2021-2025 (August 2021 Update)*. Tech. rep. CSS Insight, Aug. 2021.

[529] Mary D. Fan. “Body Cameras, Big Data, and Police Accountability”. In: *Law & Social Inquiry* 43.4 (Sept. 2018), pp. 1236–1256.

[530] Lynn Rochester et al. “A Roadmap to Inform Development, Validation and Approval of Digital Mobility Outcomes: The Mobilise-D Approach”. In: *Digital Biomarkers* 4.1 (Nov. 2020), pp. 13–27.