MODERN ENGAGEMENTS WITH MEDIEVAL SCIENCE: SENSING AND PERCEIVING LIGHT AND SOUND

JOSHUA HARVEY



A thesis submitted to the Department of Engineering Science
in fulfillment of the requirements for the Degree of

Doctor of Philosophy

University of Oxford

Pembroke College

Trinity 2019

ABSTRACT

What information awaits us in the external world, and how is such information represented internally? Far from being solely modern reflections, these questions recur throughout intellectual history. With the present work, a step is taken towards a more accurate appraisal of medieval scientific thought and practice. The benefits of doing so lie not only in a more faithful picture of the past; by engaging with historical science we find insight and inspiration for novel research in diverse areas. This is reflected in the variety of topics and disciplines presented here, which would not surprise the well-rounded natural philosopher of the thirteenth century. The principal objects of study are light and acoustic waves, and their associated sensory phenomena of colour, rainbows, material appearance, sound, and speech.

In part i, the methodologies and concepts of medieval science, relevant to this work, are introduced and discussed. The writings of Robert Grosseteste on colour, the rainbow, and speech, are contextualised within the medieval world and explored.

Part ii presents research on the rainbow, as it was understood in the thirteenth century by Grosseteste. Evidence from physical experiments and optics simulations show it likely to be the first theory of its kind based on direct observation. Grosseteste's rainbow mechanism is then compared, and found analogous to, the geometric optics of another atmospheric halo—Parry's arc.

Part iii again engages with the writings of Grosseteste. First, the medieval conceptions of perception, and their commensurability with contemporary studies are discussed. This is done through the lens of *aspectus* and *affectus*, which we find in Grosseteste's treatise on the liberal arts. Then, his treatise on speech production, perception, and phonetics is explored with an interdisciplinary approach. His combinatorial, geometric scheme of figures—which he implicates in a multisensory

framework of mental imagery, vocal tract shaping, and the letter shapes of vowels—is interpreted through linear transformations of the two-dimensional plane. In response to this text, psychophysical experiments with artificial vowel stimuli find that such geometric figures can be incorporated into acoustic chambers giving rise to their associated vowels.

Part iv focusses on medieval art, specifically the use of materials in medieval polychrome sculpture, and on resultant questions of visual perception. The physical and psychophysical mechanisms of material optics and colour vision are discussed, and illustrated with a novel metamer visualisation and a public engagement installation. The material composition of polychrome imitation gold, its optical and perceptual qualities, are described.

Part v responds directly to part iv with a psychophysical investigation into material perception and the percept of metallicity. An experimental paradigm of conjoint measurement is used to unpick the contributions of metal roughness and coating properties to the metallic appearance of objects. Models employing low-level image statistics are found to be capable of replicating observer performance, suggesting a heuristics approach to metallicity judgement. An image-processing algorithm, for the automatic transformation of objects to a golden appearance, is also described.

In part vi a high-speed schlieren imaging system, with corresponding software, for the detection, visualisation, and audio recovery of acoustic waves in air is described, in response to Grosseteste's writings on the vibrational mechanics of sound.

Part vii concludes, evaluating the successes and challenges of interdisciplinary research and looking to future work.

PUBLICATIONS

Some of the content comprising this thesis has previously been published, or has been submitted for publication. Where this is the case, journals are indicated by grey marginalia. A list of works published or accepted for publication is found below.

Major publications

Harvey, J. S., Smithson, H. E., Siviour, C. R., Gasper, G. E. M., Sønnesyn, S. O., McLeish, T. C. B. and Howard, D. M. (2019). "A thirteenth-century theory of speech". *The Journal of the Acoustical Society of America* 146.2: 937–947. DOI: 10.1121/1.5119126.

Harvey, J. S., Smithson, H. E. and Siviour, C. R. (2018). "Visualization of acoustic waves in air and subsequent audio recovery with a high-speed schlieren imaging system: Experimental and computational development of a schlieren microphone". *Optics and Lasers in Engineering* 107: 182–193. DOI: 10.1016/j.optlaseng.2018.03.015.

Harvey, J. S., Smithson, H. E., Siviour, C. R., Gasper, G. E. M., Sønnesyn, S. O., Tanner, B. K. and McLeish, T. C. B. (2017). "Bow-shaped caustics from conical prisms: a 13th-century account of rainbow formation from Robert Grosseteste's De iride". *Applied Optics* 56.19: G197–G204. DOI: 10.1364/AO.56.00G197.

Minor publications

Farrant, T. and Harvey, J. S. (2017). "Understanding and feeling: theological and psychological reflections through the lens of Aspectus and Affectus". *Bulletin de Philosophie Médiévale* 59: 262–272.

Harvey, J. S. and Farrant, T. (2017). "Aspectus and Affectus: Robert Grosseteste, Understanding and Feeling. Georgetown University,Washington, D.C. (USA), 30 March-1 April 2017". *Revista Española de Filosofía Medieval* 24: 279–281. DOI: 10.21071/refime.v24i.10463.

Harvey, J. S., Kollandsrud, K. and Smithson, H. E. (2018). "Not All That Glitters: Probing Material Perception With a Physical Imitation Gold Stimulus". In: *Perception*. Vol. 47. SAGE Journals: 572. DOI: 10.1177/0301006618756416.

Harvey, J. S., Morimoto, T. and Spitschan, M. (2019). "The Neon Fruit Illusion: A Fresh Recipe for Colour Science Demonstrations". *Perception* 48.3: 242–247. DOI: 10.1177/0301006618824484.

Harvey, J. S., Siviour, C. R. and Smithson, H. E. (2018). "From Spectra to Perceptual Color: Visualization Tools for the Dimensional Reduction Achieved by the Human Color Sense". In: *EG 2018 - Posters*. Ed. by E. Jain and J. Kosinka. The Eurographics Association. DOI: 10.2312/egp.20181025.

Harvey, J. S., Smithson, H. E. and Siviour, C. R. (2018). "Raw high-speed schlieren footage of acoustic waves in air for subsequent computational analysis and audio recovery". *Data in Brief* 19: 1647–1649. DOI: 10.1016/j.dib.2018.06.069.

Harvey, J. S., White, R. C., Smithson, H. E., McLeish, T. C. B., Howard, D. and Coleman, J. (2019). "Instrumental motions: Shaping and perceiving speech sounds". In: *Knowing and Speaking: Robert Grosseteste's De artibus liberalibus 'On the Liberal Arts' and De generatione sonorum 'On the Generation of Sounds'. Edition, Translation, and Interdisciplinary Analysis*. Vol. 1. The Scientific Works of Robert Grosseteste. Oxford: Oxford University Press: 336–366.

INDIVIDUAL CONTRIBUTION STATEMENT

Unless otherwise indicated, all work presented in this thesis is my own and authored by myself. Sections with significant input from other researchers are the historical overview of the rainbow in section 5.2, the co-authored letter in chapter 6, and the medieval zoological examples in chapter 8.

ACKNOWLEDGMENTS

This doctoral project has benefited from the generosity of the Andrew W. Mellon Foundation, and I sincerely hope that my contribution, however minimal, serves to advance their mission. To Vicky and the team at TORCH, thank you for all that you do.

I wish to thank Clive and Hannah for their contagious optimism, dedication, and good humour. This would not have been possible without their dynamic and enduring supervision.

I am fortunate to have a long list of collaborative partners to thank, many of whom I am glad to also count as friends. Giles, Tom, Cecilia, Brian, Sigbjørn, and the entire Ordered Universe team, it has been a joy and a privilege to share a place at the table. Tim, Kaja, Alice, and so many others at both the Oxford Perception Lab and the Solid Mechanics group, I have so enjoyed and valued our day-to-day interactions.

My deepest thanks go to my friends, in whose company Oxford has truly been a home. With you, I have discovered more than I hoped to find in libraries or laboratories. To my boyfriend Moritz, I owe more than words can say.

Last and foremost I must thank my parents, who have been and continue to be an unwavering source of support, encouragement, and love.

CONTENTS

I ENGAGING WITH HISTORICAL SCIENCE

1 INTRODUCTION 2

1.1 Medieval natural philosophy and the origins of science 3

1.2 Reading a medieval text in the twenty-first century 6

1.3 Motivations for the work presented here 8

2 INTELLECTUAL FOUNDATIONS OF MEDIEVAL NATURAL PHILOSOPHY 10

2.1 Colour in the medieval mind: new dimensions 11

2.2 Speech and the medieval sensorium 23

2.3 Summary 24

3 PUBLIC ENGAGEMENT 26

3.1 The accessible academy 26

3.2 Activities undertaken during the course of doctoral research 27

II OPTICS AND THE RAINBOW

4 THE RAINBOW AS A HISTORICAL MOTIF 31

4.1 What is special about rainbows? 32

4.2 A bridge between disciplines 33

5 BOW-SHAPED CAUSTICS FROM CONICAL PRISMS: A THIRTEENTH-CENTURY ACCOUNT OF RAINBOW FORMATION FROM ROBERT GROSSETESTE'S *DE IRIDE* 34

5.1 Introduction 35

5.2 Early history of the rainbow 37

5.3 Robert Grosseteste's theory of the rainbow 38

5.4 Historical re-enactment of cone-light interaction 41

5.5 Physics-based rendering of conical prisms 45

5.6 Results and discussion 47

5.7 Conclusion 51

6 PARRY'S ARC AND GROSSETESTE'S *DE IRIDE* 54

III PHONETICS AND SPEECH PERCEPTION

7 SPEECH PRODUCTION AND PERCEPTION 59

7.1 The liberal arts and the medieval curriculum 59

8 *ASPECTUS* AND *AFFECTUS* 61

9 A CATEGORICAL THEORY OF VOWELS FROM *DE GENERATIONE SONORUM* 66

9.1 Introduction 67

9.2 Robert Grosseteste's *De generatione sonorum* 69

9.3 A psychophysical interpretation of the text 80

9.4 Discussion 89

9.5 Conclusion 93

9.A Supplementary Material 95

IV COLOUR AND MATERIALS

10 VISUAL PERCEPTION OF MATERIAL QUALITIES 99

10.1 Human colour vision and its stimuli 100

10.2 The visual perception of material qualities 101

10.3 The passage of light through optical media 102

10.4 Surface reflectance properties and metallic materials 105

10.5 Simulating light-material interactions 108

10.6 Summary 109

11 FROM SPECTRA TO PERCEPTUAL COLOUR 110

11.1 Introduction 111

11.2 Two-wavelength radiant spectral surfaces 112

11.3 Polychromatic primaries 114

11.4 Visualising anomalous trichromacy 114

11.5 Three-wavelength radiant spectral set volumes 117
11.6 Conclusion 117
12 THE NEON FRUIT ILLUSION 118
12.1 Introduction 119
12.2 Ingredients 122
12.3 Recipe 122
12.4 Acknowledgements 124
13 MEDIEVAL VISUAL EFFECTS IN POLYCHROME SCULPTURE 125
13.1 Introduction 125
13.2 Medieval metallicity and gilded polychrome 126
13.3 Optical characteristics 128
13.4 Perceptual qualities 134
13.5 Embodied light or fool's gold? 139
V METALLICITY
14 PSYCHOPHYSICAL SCALING METHODS FOR SUPRATHRESHOLD MEASUREMENTS 141
14.1 Difference scaling 142
14.2 Conjoint measurement analysis 146
15 MATERIAL QUALITIES, IMAGE STATISTICS, AND VISUAL FEATURES DETERMINING THE PERCEPTION OF METALLICITY 152
15.1 Introduction 153
15.2 Experiment 1: Perceptual Scaling of Physical Parameters 161
15.3 Experiment 2: Conjoint Measurement of Roughness and Coating Bumpiness 166
15.4 Image Analysis and Modelling 169
15.5 Experiments with a High Dynamic Range display 181
15.6 Discussion 187
15.7 Conclusion 193

16 THREE-DIMENSIONAL COLOUR DISTRIBUTION WARPING FOR GOLD APPEARANCE TRANSFER 194

16.1 Introduction 194

16.2 Histograms as statistical correlates for material appearance 195

16.3 Three-dimensional colour histogram transformations 197

16.4 Results of the gold appearance transfer algorithm 199

16.5 Conclusion 200

VI HIGH-SPEED SCHLIEREN ACOUSTICS

17 INTRODUCTION AND LITERATURE REVIEW 202

18 EXPERIMENTAL AND COMPUTATIONAL DEVELOPMENT OF A SCHLIEREN MICROPHONE 204

18.1 Introduction 205

18.2 Background 206

18.3 Experimental setup 210

18.4 Image processing and data analysis 213

18.5 Data recombination through autocorrelation of signals 219

18.6 Simulated schlieren video of acoustic waves 221

18.7 Results 222

18.8 Discussion 229

18.9 Conclusion 231

18.J Appendix 232

VII CONCLUSION

19 MEETING IN THE MIDDLE 235

19.1 Summary of contributions 235

19.2 Future work 238

19.3 Beginnings, middles, and ends 239

BIBLIOGRAPHY 241

LIST OF FIGURES

Figure 1.1 A generalised routine of interdisciplinary medieval science studies. 7

Figure 2.1 Aristotle's one-dimensional colour theory. 13

Figure 2.2 Ibn Sina's expansion of Aristotle's colour theory. 17

Figure 2.3 Vocal tract diagram from Al-Sakkākī's *The Key to the Sciences*, reproduced from Bakalla, 1982. 25

Figure 5.1 An interpretation of the geometric optics in the *De iride*. 40

Figure 5.2 Conical glass beakers dated to the same period as Grosseteste. 42

Figure 5.3 Parameters for experimental and simulated caustics. 43

Figure 5.4 Physical experiment of cone-light interactions. 44

Figure 5.5 Light entering the top of a cone forms bow-shaped caustics. 44

Figure 5.6 Images of caustics resulting from cone-light interactions simulated using the Luxrender engine. 45

Figure 5.7 Parameter space of cone-produced caustics obtained from Luxrender simulations. 48

Figure 5.8 Simulations showing light entering the top of a cone collects into bow-shaped caustics. 49

Figure 5.9 Simulated dispersive effects of bow-shaped caustics. 50

Figure 5.10 Experimental results and simulation showing the formation of a double bow caustic. 51

Figure 6.1 The refraction mechanism responsible for Parry's arc compared with Grosseteste's rainbow mechanim. 56

Figure 8.1 Comparing the 'microscale' and 'macroscale' *aspectus-affectus* relationships at play in sensory perception and theological learning. 63

Figure 9.1 A diagrammatic depiction of one of the claims in Grosseteste's *De generatione sonorum*. 73

Figure 9.2 The simple, self-similar geometric movements that Grosseteste describes as the basis for vowel categorisation. 75

Figure 9.3 The combined movements that give rise to vowels in Grosseteste's model of phonetics. 76

Figure 9.4 Grosseteste describes a self-combination of circular movement, which he discounts as too complex for use in speech. 77

Figure 9.5 The configurations of the plate-type vocal tract model for used for vowel synethesis. 82

Figure 9.6 Spectrograms and spectral slices of synthetic vowels. 82

Figure 9.7 The mappings of the five synthetic vowels to acoustic and perceptual spacce. 83

Figure 9.8 Classifications obtained for each of the five samples from the second listening test. 87

Figure 9.9 Monte Carlo simulations of multidimensional scaling. 96

Figure 9.10 Mean ellipse areas of 100 Monte Carlo simulations. 97

Figure 10.1 A simple model of light passing between two media, using real and complex indices of refraction. 104

Figure 10.2 Conductor vs dielectric material appearance. 106

Figure 10.3 Renders of gold, silver, and copper objects. 107

Figure 10.4 Metallic appearance is not simply extreme gloss. 108

Figure 11.1 Projection of the $C\{610, 540\}$ and $C\{460, 540\}$ two-wavelength radiant spectral sets onto the $[w_1, w_2]$ plane. 112

Figure 11.2 Projection of two $C(w_1, w_2)$ surfaces onto a $[w_1, w_2, \text{hue}]$ mapping. 113

Figure 11.3 Example discontinuities due to hue and saturation in the $[w_1, w_2, \text{hue}]$ mapping. 113

Figure 11.4 A fully three-dimensional space such as HSV projects $C(w_1, w_2)$ spectral sets as continuous surfaces. 114

Figure 11.5 $C\{w_1, w_2\}$ colour sets projected to HSV and RGB spaces. 115

Figure 11.6 The surface generated by the spectral set $C\{\mathbf{D65}, 610\}$. 115

Figure 11.7 Two $C\{w_1, w_2\}$ spectral sets projected to RGB spaces defined by normal vision and anomalous trichromacy. 116

Figure 11.8 The three-primary radiant spectral set $C\{610, 540, 460\}$. 117

Figure 12.1 A sequence of frames taken from a video of the neon fruit illusion. 119

Figure 12.2 Typical stimuli of the neon fruit illusion. 120

Figure 12.3 A slice-through view of a neon fruit illusion video. 121

Figure 12.4 A simulation of the neon fruit illusion, using hyperspectral images of fruits and vegetables. 123

Figure 13.1 Optical interactions at play in the medieval polychrome 'imitation gold' material. 129

Figure 13.2 Spectroradiometric measurements of physical samples of the medieval polychrome 'imitation gold' material. 130

Figure 13.3 Chromaticity variations of the 'imitation gold' physical samples as a function of angle of incidence/reflectance. 131

Figure 13.4 The results of simulating the polychrome 'imitation gold' material with physically-based computer graphics rendering, for different angles of reflectance. 132

Figure 13.5 The mean chromaticities of images obtained by simulating the 'imitation gold' material. 133

Figure 13.6 Chromaticity and luminance as a function of reflectance angle in simulated images of the 'imitation gold' material 134

Figure 13.7 Shape-from-shading cues in medieval polychrome 'imitation gold'. 136

Figure 13.8 Renders of sculptures clad in smooth gold, rough gold, and polychrome 'imitation gold'. 137

Figure 13.9 Summary statistics of DeepGaze II saliency predictions, as a function of material. 138

Figure 14.1 A graphical overview of maximum likelihood difference scaling. 145

Figure 14.2 A graphical overview of maximum likelihood conjoint measurement. 149

Figure 14.3 The independent, additive, and saturated models of conjoint measurement as a nested hypothesis. 151

Figure 15.1 The parameter space used in experiments investigating the perception of metallicity. 157

Figure 15.2 An overview of the pipeline for the research into perception of metallicity. 160

Figure 15.3 Perceptual scale estimates for individual participants for roughness and bumpiness. 165

Figure 15.4 Mean perceptual scale estimates from aggregate data for roughness and bumpiness. 165

Figure 15.5 Conjoint measurement estimates for individual participants fit to the 'additive' model, for the SDR condition. 168

Figure 15.6 Three data sets from the conjoint measurement were best fit to the saturated model. 168

Figure 15.7 Swapping the amplitude spectra (A) and phase spectra (Φ) of the stimuli. 170

Figure 15.8 Roughness of metal has a direct effect on the spatial-frequency content of the stimuli. 172

Figure 15.9 Bumpiness of coating also has an effect on the spatial-frequency content of stimuli. 172

Figure 15.10 Amplitude spectra for the roughness and bumpiness stimuli used in the MLDS experiements. 173

Figure 15.11 Steerable pyramid synthesis, starting from an image at the center of the parameter space (P_0^0). 174

Figure 15.12 Heatmaps showing the magnitudes of filter responses from the steerable pyramid analysis. 176

Figure 15.13 The effects of parameters on the responses of oriented filters of the steerable pyramids. 176

Figure 15.14 Taking the ratio of the linear filter outputs at Level 1 and Level 4 of the steerable pyramid strongly correlates with coating bumpiness. 177

Figure 15.15 Simulation results of the maximum likelihood difference scaling experiment. 177

Figure 15.16 Conjoint measurement results from the models using different estimators of global contrast. 178

Figure 15.17 The models of roughness, bumpiness and metallicity using only the outputs of linear filters. 181

Figure 15.18 The computational model of metallicity. 182

Figure 15.19 A simple computational model of metallicity that compares the weighted outputs of Level 1 and Level 4 oriented filters accounts for the observer data. 182

Figure 15.20 Distribution of luminance values in the light probes used for both the SDR and HDR rendering conditions. 184

Figure 15.21 Perceptual scale estimates for individual participants for roughness and bumpiness in the HDR condition. 185

Figure 15.22 Mean perceptual scale estimates from aggregate data for roughness and bumpiness in the HDR condition. 186

Figure 15.23 Conjoint measurement estimates for individual participants fit to the 'additive' model, for the HDR condition. 187

Figure 15.24 Conjoint measurement estimates for individual participants fit to the 'saturated' model, for the HDR condition. 187

Figure 15.25 Silver and glass objects, rendered with different physical roughness, shown with luminance profiles. 189
Figure 15.26 HDR conjoint measurement simulation results, using different models to estimate metallicity. 192
Figure 16.1 One-dimensional luminance histograms of an object rendered with a matte and shiny material appearance. 196
Figure 16.2 Single-channel RGB histograms fail to describe perceptual colour distributions. 198
Figure 16.3 The algorithm for transferring a gold-like appearance to objects in images. 199
Figure 16.4 Results of the gold material transfer algorithm. 200
Figure 18.1 The basic principle of schlieren imaging 207
Figure 18.2 The single-mirror double-coincident setup for schlieren imaging. 211
Figure 18.3 Different post-processing for a 20 kHz sine wave, 115 dB. 214
Figure 18.4 Filtering a frame of schlieren video with 2D Gaussian filters. 215
Figure 18.5 Raw schlieren footage and after processing with a high-pass temporal filter. 216
Figure 18.6 Visibility of acoustic waves is dependent on a high enough SPL for a given schlieren sensitivity. 217
Figure 18.7 A visualisation of footage taken of a 20 kHz sine wave. 218
Figure 18.8 An illustration of the effective sampling rates for intra-frame spatial resolution and inter-frame temporal resolution. 219
Figure 18.9 Simulated schlieren images, produced with different waveforms and signal-to-noise ratios. 222
Figure 18.10 The general work flow for extracting audio signals from schlieren footage. 223
Figure 18.11 Examples of different frequencies and waveforms recovered by the schlieren microphone. 225

Figure 18.12 Summary of the visibility and recoverability of acoustic waves by schlieren imaging. 226
Figure 18.13 Variations of pixel RMS values according to acoustic wave amplitude. 227
Figure 18.14 Plots of SNR to evaluate the quality of recovered audio signal. 227
Figure 18.15 The human visual system is limited by a contrast sensitivity function. 228
Figure 18.16 The $(\cos\theta)^2$ scaling function for gradients of refractive index along the optical path of the schlieren system. 233

LIST OF TABLES

Table 9.1 Our interpretation of phonemes from the vowel letters Grosseteste uses in the *DGS*. 95
Table 9.2 Diameters of the plate-type models for vowel synthesis. 95
Table 9.3 Results from the vowel classification experiment. 95
Table 15.1 Global contrast estimators used in modelling metallicity. 179
Table 15.2 The optimal weights for a computational model of metallicity fit to observer data. 183
Table 18.1 Parameters varied for schlieren imaging of acoustic waves. 212

ABBREVIATIONS

AHRC Arts and Humanities Research Council

BRDF bi-directional reflectance distribution function

cdf cumulative distribution function

DMLBS Dictionary of Medieval Latin from British Sources

DAL *De artibus liberalibus* [*On the Liberal Arts*]

DC *De colore* [*On Colour*]

DGS *De generatione sonorum* [*On the Generation of Sounds*]

DI *De iride* [*On the Rainbow*]

DL *De luce* [*On Light*]

DoF degrees of freedom

EMD earth mover's distance

GLM generalised linear model

HDR high dynamic range

IoR index of refraction

JND just noticeable difference

MDS multidimensional scaling analysis

MLCM maximum likelihood conjoint measurement

MLDS maximum likelihood difference scaling

RMS root mean square

SDR standard dynamic range

SPL sound pressure level

SNR signal-to-noise ratio

TORCH the Oxford Research Centre in the Humanities

Part I

ENGAGING WITH HISTORICAL SCIENCE

1

INTRODUCTION

Sensation is the foundation of every branch of philosophy.

—*The Metalogicon*, John of Salisbury (1152)

PART ABSTRACT

This part serves as an introduction for the work presented in this thesis, but also for the motivations, literatures, and methodologies that underpin it. The overarching question inspiring this work has been to what extent the mental and material medieval worlds have relevance to, or application for, modern science. This question is asked both from the perspectives of different disciplines, and from a consideration of modern science in general.

This thesis is divided into six self-contained parts, each responding to a particular text, theory, or material facet of the medieval world. In general, there is a progression from highly interdisciplinary work in tandem with history and philosophy, towards exclusively experimental or computational work. This first part will introduce the reader to the broader methodology of research that has resulted in this thesis, namely sustained collaboration across disciplinary and divisional divides, and public engagement. Both of these activities are removed from the typical laboratory setting of experimental psychology and engineering science, but, as explored here, can provide a stimulating starting points for scientific investigations.

Chapter 2 gives a review of the literature foundational for this thesis, namely the historical developments of scientific theory surrounding light, colour, rainbows, speech, and sound. The introductory part closes with chapter 3, which discusses the public engagement activities undertaken during the period of doctoral research.

1.1 MEDIEVAL NATURAL PHILOSOPHY AND THE ORIGINS OF SCIENCE

A common question posed when examining intellectual history is on the origins of science. Who invented science, the scientific method, or, 'Who was the first scientist'? Such questions, often tackled in popular literature, invariably fail to produce convincing answers, for several reasons.

First, despite the critical writings of Thomas Kuhn (1922 - 1996) and their influence, principally his seminal text *The Structure of Scientific Revolutions* (Kuhn, 1970) wherein he advocates for a 'historiographic revolution' of science, we have a tendency to project Popperian standards onto the past. Positivist criteria such as Karl Popper's (1902 - 1994)—that to be demarcated as science requires falsifiable hypotheses, testable predictions, and direct observation—aim to give clear and unambiguous evaluations of historical methods and figures. But the historical reality is far more complex. Writings on the natural world throughout history exhibit claims readily falsifiable by modern science—should these, then, be evaluated as scientific? Moreover, as Kuhn argues, the falsification criterion is itself infeasible due to the very structure of scientific theories, their incompleteness, and the nature of anomalous observations. Rather than the clinical criteria of Popper, Kuhn asserts that science is a complex, amorphous, and socially constructed system, more about contributing to a societal consensus about the natural world than a distinct methodology. It is not surprising that this view of science, which Kuhn introduces as "a sketch of the quite different concept of science that can emerge from the historical record of the research activity itself", should affect how we read and understand historical science. In doing so, we cannot speak of an 'invention' of science (except perhaps in the limited case of

pre-paradigmatic inquiries), but rather the establishment of paradigms—each one a variety of 'normal science'—which are naturally incommensurable to our own.

Second, even if we do approach the past with defined criteria for scientific demarcation, in attempting to taxonomise theories and influential figures of the past we run into problems. The evolution of intellectual history is a gradual and non-linear progression, so questions such as 'Who is the first scientist' can be as problematic as asking 'Who was the first *Homo sapiens*'. Isaac Newton (1643 - 1727), a person touted as leading the 'scientific revolution' to victory with the publication of his *Opticks* in 1704 (Wootton, 2015), was an avid alchemist and strove for scientific readings of the Bible (Snobelen, 2016). The economist John Maynard Keynes (1883 - 1946) addressed Newton's seemingly contradictory intellectual activities in his research, categorising Newton as "...not the first of the age of reason. He was the last of the magicians", even saying that "[h]is experiments were always, I suspect, a means, not of discovery, but always of verifying what he knew already" (Keynes, 1947; Kuehn, 2013). Controversies such as these, and variations on criteria, can result in wildly differing ideas as to who the first scientist was; a list of prospective candidates includes Thales of Miletus (624 - c. 545 BCE) (O'Grady, 2002), Aristotle (384 - 322 BCE) (Leroi, 2015), Ibn Al-Haytham (c. 965 - 1040) (Steffens, 2006), Robert Grosseteste (c. 1175 - 1253) (Crombie, 2002), Roger Bacon (c. 1219 - c. 1292) (Clegg, 2004), William Gilbert (1544 - 1603) (Zilsel, 1941), Francis Bacon (1561 - 1626) (Urbach, 1987), Galileo Galilei (1564 - 1642) (Weidhorn, 2005), René Descartes (1596 - 1650) (Grosholz, 1991), and, of course, Newton. To ask 'Who is the first scientist' is perhaps better understood as asking who, as documented by history, was the first to share our (i.e. my) conception of science. As this list of ten men, spanning more than two-thousand two-hundred years illustrates, the potential answers are varied and ultimately uninformative.

Third, such approaches to science focus nearly exclusively on the conscious (and literally documented) thoughts and actions of a particular person or group of individuals, overlooking the unconscious norms, beliefs and realities diffused throughout the societies within which they were embedded. These are integral to

understanding scientific discourses of the past; without them we cannot even wager which kinds of questions would be meaningful in a given epoch, or what forms of evidence might be permitted in support of conjectures about the natural world. To engage with historical science, we must therefore study not just individuals, or ideas, or experiments, but also the cultural and intellectual reality within which they exist, the epistemological unconscious, or *épistémè*, which Michel Foucault (1926 - 1984) speaks of in his excavations of life, labour, and language (biology, political economy, and philology) in *The Order of Things: An Archaeology of the Human Sciences* (Foucault, 1971).

Needless to say, an approach to history of such sophistication is beyond the remit of the modern scientist, regardless of how earnestly she approaches historical science. For this reason, interdisciplinarity is of paramount importance for scientists who wish to engage with the classics (which may indeed be celebrated, or overlooked, intellectual products of a historical epoch). There are also practical considerations. While the thirteenth-century scholar would be proficient in Latin, optics, geometry, music, astronomy, and many other fields, the modern academy is normally situated within a 'specialisationist' knowledge economy (Cooke, 2014). This results in a steep methodological hurdle: a psychologist would be unable to read and make sense of medieval texts on cognition without first learning medieval Latin, and familiarising herself with the key figures of that writer's intellectual inheritance, such as Aristotle, Euclid, and others without whom a great many Latin terms would remain as meaningless jargon, before finally undertaking what could be termed a 'functional analysis' (Smithson et al., 2014b) for the diverse disciplines contained within possibly a single short text.

Interdisciplinary collaboration with modern science is increasingly being recognised as a valuable, at times essential, tool for researching historical science. Any functional collaboration must span the disciplinary divide imposed on contemporary academia, comprising not only the expected gaps in knowledge and skills, but also broader differences in research culture and purpose. For this reason, collaborative projects can result in a separateness (working with little interaction in a multidiscip-

linary manner) or unevenness (researchers departing from one area to visit another in a cross-disciplinary manner). The Ordered Universe Project—an international, interdisciplinary research group investigating thirteenth-century natural philosophy and science—has, from its inception, positioned itself to avoid these two pitfalls. The shared goal of publishing volumes on Grosseteste's treatises addressing natural phenomena, which include critical editions and translations, as well as functional analyses or experimental interpretations, ensures a common purpose across the disciplinary spectrum. The regular symposia and collaborative reading sessions provide a space for knowledge exchange and skill development outside individuals' fields of expertise. The embedding of this doctorate within the research activities of the Ordered Universe Project has facilitated the uniquely interdisciplinary approach required.

1.2 READING A MEDIEVAL TEXT IN THE TWENTY-FIRST CENTURY

The first object of study for the work presented here lies in the thirteenth century—the scientific opuscula of Robert Grosseteste. There are significant hurdles to translating the content of these treatises into a format amenable to contemporary scientific analysis and inquiry, which could not be overcome without interdisciplinary collaboration with other researchers within the Ordered Universe Project. A generalised routine of how to go from extant manuscripts written in Medieval Latin to modern research is presented in Figure 1.1.

It may seem that some of the stages of this routine are achievable by individuals or individual disciplines. But surprisingly, it has been increasingly found within the Ordered Universe Project that though this may be the case, all stages of the process benefit from interdisciplinary interaction as well as interaction with the other stages. The routine is therefore not exclusively feed-forward from the manuscripts to the science. Rather, interrogation of the scientific content of the text may in some cases assist in translation, or even in the generation of a critical edition. Where manuscripts

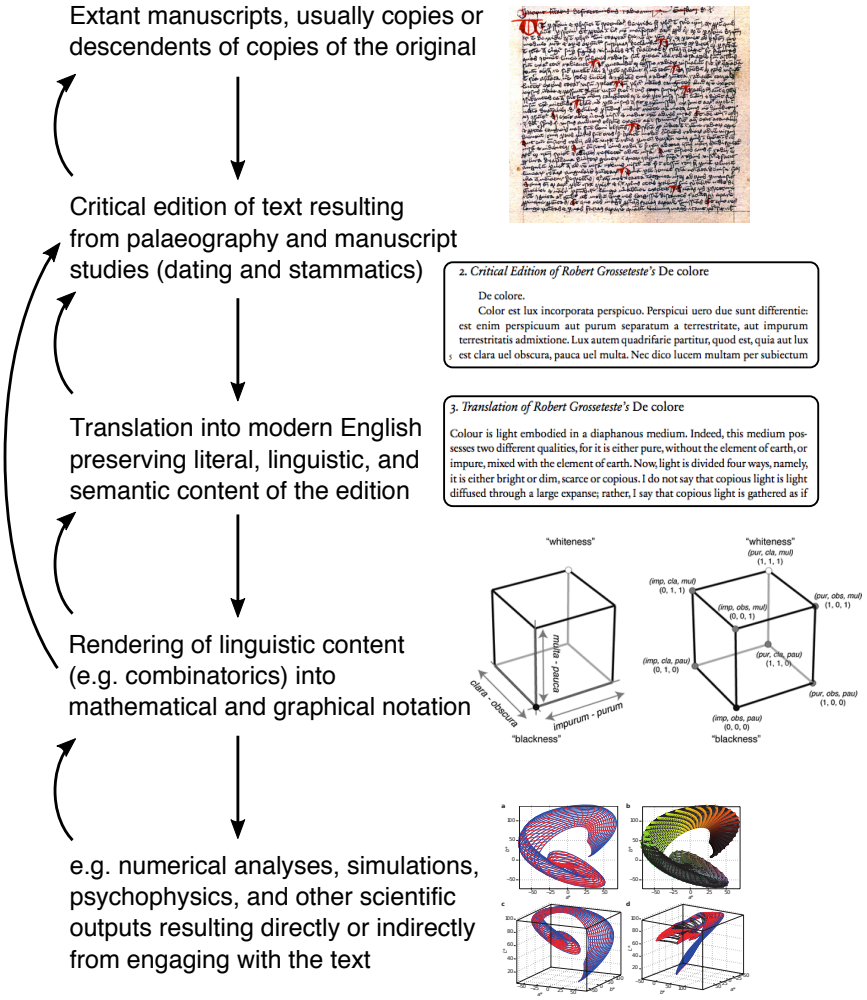


Figure 1.1: A generalised routine of going from surviving manuscripts of medieval scientific treatises to modern scientific outputs. The routine is not exclusively feed-forward from manuscript through to modern science, but rather all stages of the routine feedback to and benefit one another. Not shown are all of the 'skip connections' between non-consecutive different layers, in both directions.

disagree, analysis of the content may inform which prospective rendering of the text is more plausible. The carrying out of modern science that engages with a treatise's content can uncover new interpretations and presentations of the theories described within, as well as further elucidating the nature of the phenomena which form its focus.

1.3 MOTIVATIONS FOR THE WORK PRESENTED HERE

There are three principal motivations for the work presented here. The first motivation is to participate in a movement towards a richer understanding of medieval science in the context of the history of science. In practice this entails fresh theoretical analyses of treatises, often initiated concurrently with their translation from Latin to English.

The second motivation is to perform a rigorous examination of both the mental and material worlds in which the medieval thinker and observer was working. This requires applying the theories, models, and experimental methodologies of modern science to the content of medieval texts and contexts. In so doing, it may be that modern and medieval theory find some limited degree of agreement, while our current theoretical frameworks provide more satisfying explanations. This is not a presentist approach; Grosseteste knew nothing of the three classes of cone photoreceptors when hypothesising a three-dimensional theory of colour perception, and equally did not learn anything about them by constructing his theory, although his vision would have been a product of these biological processes that we now understand. It is not so important whether medieval scientific theory was correct by today's standards, but whether it best accommodated the observations of the physical world available and permissible at the time. It is therefore necessary to evaluate the limitations and potentials of medieval sensory data, as well as how these data, and other sorts of evidence, could plausibly inform theory.

The third motivation is to produce novel research prompted by questions, theories and observations described in medieval texts. The role of creativity in science is often

neglected, though it is the engine driving progress. As this work illustrates, it can be refreshing and inspiring to engage in academic and historical contexts different from ones own. A large part of this has been the benefit of engaging with medieval thought, with its different axioms and assumptions about the physical world. But there was also considerable benefit from being based at two different scientific departments. Engineering and perceptual psychology have conceptual overlaps, as well as common and complementing methodologies. There are also practical overlaps; the use of physically-based computer graphics to generate stimuli for psychophysical experiments, and evaluating the limits of flow visualisation techniques imposed by the human visual system. One, perhaps less expected, result of engaging with medieval science has been to foster modern interdisciplinary science.

2

INTELLECTUAL FOUNDATIONS OF MEDIEVAL NATURAL PHILOSOPHY

Death itself does not interrupt the scientists's work.
Theories once unfolded are eternally living and acting.

—*History of Science*, Sarton, 1916

This chapter presents an introductory overview to scientific theories that developed throughout antiquity and the medieval period, which lays the foundations for work presented in later chapters. The natural phenomena in focus are colour, the rainbow, sound, and speech. The writings of Robert Grosseteste take centre stage, as these have formed the basis for sustained engagement via interdisciplinary collaborative readings during symposia organised by the Ordered Universe Project. To understand his theories, expressed in the treatises *De colore* [*On Colour*] (*DC*), *De iride* [*On the Rainbow*] (*DI*), *De artibus liberalibus* [*On the Liberal Arts*] (*DAL*), and *De generatione sonorum* [*On the Generation of Sounds*] (*DGS*), it is necessary to provide the, often Aristotelian, intellectual foundations and sources. These sources serve in both interpreting the texts, and in understanding the contribution that Grosseteste makes.

2.1 COLOUR IN THE MEDIEVAL MIND: NEW DIMENSIONS

2.1.1 *The Aristotelian line of thinking*

Colour is perhaps our most immediate sensation. Across all human cultures, colour signals are used not only for survival, but for aesthetic, religious, and emotional practice, showing prominence in language, craft, and economics. In science too, and natural philosophy, colour enjoys a particular importance. In writings by or attributed to Aristotle, and their initial attempt to taxonomise the natural world, we find what is widely considered to be the first theory of colour. In his texts *De sensu et sensato* (*De sensu*) [*Sense and Sensibilia*], and *De anima* [*On the Soul*] (Aristotle, 1984b, 1987), Aristotle makes bold claims about sensory perception, the physical world, and their interrelations. As with many other natural phenomena, he reasons that physical colours are instantiated through the states and combinations of the four classical elements: fire, air, water, and earth. We might then expect to find a four-dimensional theory of colour, whereby perceptually separable colours form a one-to-one mapping with unique combinations of the four elements. However, we find instead a one-dimensional, linear mapping of the perceptual space of colours, between two extrema. In this way, colour is a particular case of the general principle of perception that Aristotle identifies, that all sensory perception is an act of comparison between extremes. Writing in the *De sensu* (translation by Beare), he says:

> But contraries are extremes, and every object of sense-perception involves contrariety; e.g. in colour, white and black; in savour, sweet and bitter, and in all the other sensibles also the contraries are extremes.

These extremes are white and black, and are generated in both determinate bodies (i.e. possessing a fixed colour of their own) and transparent bodies, such as air and water, as a function of how much light is present, with black correspondingly a privation of light, darkness. These colours are associated with elemental fire and earth, respectively. Between these extrema, he identifies seven 'species' of colours, a

technical term that denotes classifications of difference due to the simplest and most immediate *differentiae*:

> Savours and colours contain respectively about the same number of species. For there are seven species of each, if, as is reasonable, we regard grey as a variety of black (for the alternative is that yellow should be classed with white, as rich with sweet); while crimson, violet, leek-green, and deep blue, come between white and black, and from these all others are derived by mixture.

Although it is not explicitly stressed how these species of colour may be arranged between white and black, it is commonly assumed that Aristotle lists them in the intended order: white, yellow, crimson, violet, leek-green, deep blue, and black, as shown in Figure 2.1.a. Plainly, Aristotle has decided on there being seven species of colour independent of any empirical observations, and in fact against the direct observation of grey being different to black, and is making no effort to conceal the logic of his thinking. His motivation for there being seven species of colour comes from the empirical and mathematical studies of music, where seven notes may be found in the musical scale. As with the study of musical harmony, the species of colour arise from particular ratios of juxtaposition.

The theory of there being only seven colours, arranged in a one-dimensional line between white and black has some obvious flaws. If the goal of such as theory is to describe the relative positions and organisation of perceptual colours according to dissimilarity, we would expect such a mapping to adhere to at least three fundamental principles. First, colours that are close together in map space should be close in perceptual space, that is, they should exhibit a high degree of perceptual similarity. Second, neighbouring colours should be connected without the need to pass through colours which are perceptually distinct. Third, the mapping should be comprehensive, and accommodate either explicitly or implicitly all possible perceptual colours. Additionally, and moving beyond mere appearance, a comprehensive theory of colour will organise perceptual colours according to meaningful dimensions.

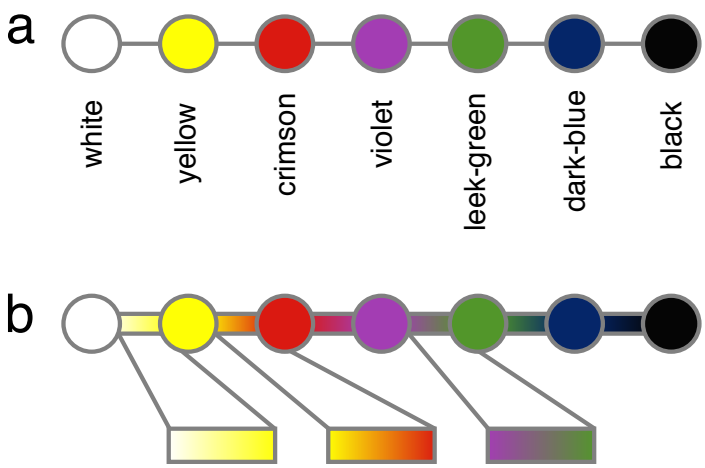


Figure 2.1: a) Aristotle's one-dimensional colour theory, in which the seven colours—white, yellow, crimson, violet, leek-green, deep-blue, and black—are ordered along a line. b) Interpolation shows that Aristotle's linear mapping violates fundamental principles of a perceptual colour space, as neighbouring violet and leek-green are perceptually very different from one another.

Aristotle's linear theory of colour organisation fails to uphold all three basic principles of perceptual organisation. First, colours that are perceptually similar, and hence close in perceptual space—yellow and green—are far apart in his linear mapping. Second, neighbouring colours are not directly interpolated without progressing through a third, perceptually distinct colour. This is illustrated in Figure 2.1.b, where connections between colours along the linear map are interpolated with a smooth gradient. Although some neighbours appear to show perceptual proximity (e.g. white and yellow, dark-blue and black), others do not. To connect yellow and red, it would seem necessary to pass through orange. This may not be that problematic, as 'orange' has only been identified as a basic colour term since the late sixteenth century, after Portuguese merchants began bringing large amounts of the fruit from India to Europe (Kastan and Farthing, 2018), and would previously be described as yellow-red. However, we also find violet situated next to leek-green, which by any account are perceptually highly dissimilar (and even oppositional (Hering, 1920)). To interpolate between the two, according to laws of colour mixture, requires passing through the achromatic point (grey), which Aristotle has already identified as a variant of black. Third, several types of colour are absent from Aristotle's theory, as it does not account for desaturated, pastel colours, or variations in lightness for a

given hue (there is a deep, or dark, blue, but there is no light blue, and it would not be clear where it should be positioned).

Of course, these problems arise because human colour perception is not one-dimensional, but three-dimensional[1]; while linear mappings are possible, they will inevitably suffer some or all of these drawbacks[2].

Finally, Aristotle's theory does not give a valid hypothesis to account for the (single) dimension of human colour vision. He claims directly that any colour may result from a mixture of white and black, directly contravening evidence from pigment mixing that this results in shades of grey. A generous reading would interpret his single dimension as luminance, with the corollary that light of any colour appearance would appear white if it were bright enough, and black if it were dim enough. Also in the *De sensu*, we find an observational description that fits with this theory:

> ...like that produced by the sun, which in itself appears white, but takes a crimson hue when beheld through a fog or a cloud of smoke...

Although the sun is not red or crimson when viewed through a turbid medium (unless at periods of sunset or sunrise, when longer wavelengths of light are refracted through the atmosphere to the observer, while shorter wavelengths fail to reach her), this passage provides a clearer insight into Aristotle's theory, and of which dimension he ay be attempting to describe.

Confusingly, much of the Aristotelian teaching on colour comes from another text, likely not authored by him, the *De coloribus* [*On Colours*]. In this text, far greater detail is given on how the mixing and juxtaposition of elements generates particular colours. Again it is stated that darkness is a privation of light, but black is now identified as the proper colour of elements in the process of transmutation. This

1 This is quite a simplification, stemming from the fact that most humans have three classes of cone photoreceptors. However, individuals with colour vision deficiency may possess only two, or one. Rod photoreceptors may also contribute to colour vision at mesopic light levels, and recently another type of cell—melanopsin-expressing intrinsically-photosensitive retinal ganglion cells—has been found to contribute to colour vision (Zele et al., 2018). However, in most cases three dimensions is sufficient for describing and representing the perception of uniform colour fields.

2 We should not be too critical of Aristotle, or adherents to his linear understanding of colour. Still today, many of the leading researchers in quantitative fields appear to support this view, by using 'rainbow' colour maps for figures showing one-dimensional data.

text is far more observational, concerned with combustion, birds' plumage, fruits, flowers, dyes, *et cetera*, but fails to cohere colorimetric data into a unifying framework or theory.

2.1.2 *Islamic innovations*

The linear mapping of seven colours remained highly influential and unchallenged through the classical period, until its reception by Islamic scholars during the period often referred to as the Golden Age. Some scholars simply repeated the Aristotelian theory that all colours were caused by different mixtures, in special proportions, of brightness and darkness, such as Ibn Rushd [Latinised as Averroes] (1126-1198) (Boyer, 1987). But others gradually expanded the one-dimensional understanding of colour into an implicitly two-dimensional space (Kirchner, 2013; Kuehni and Schwarz, 2008). Writing a century before Ibn Rushd, we find the first explicit departure from Aristotle in the writings of Ibn Sina's (980-1037) [Latinised as Avicenna]. In his *Kitab al-Shifa* [*Book of Healing*], published in 1027, Ibn Sina included a treatise on the soul, which was read in the Latin west as a commentary on Aristotle's *De anima* after its translation in twelfth-century Spain as *Liber de anima*. The fourth chapter of this text is on colour, in which Ibn Sina repeats Aristotle's assertion that colours are arranged between the two extrema of white and black, but that these extrema may be joined along more than one path (translation by Kuehni, (Kuehni and Schwarz, 2008), with direct translations of colour terms from the Dictionary of Medieval Latin from British Sources (DMLBS) added in bold):

> Moreover, if whiteness does not exist without light and blackness not in ways already discussed then whiteness and blackness cannot only be joined in one manner. A manifestation of this is the fact that white gradually passes to black by three paths. The first is via light yellow-green [DMLBS: **somewhat-pale**] (*subpallidus*) and its progression is pure: it will indeed be of pure progression, at first it progresses to light yellow-green,

> from there to yellow-green [DMLBS: **pale**] (*pallidus*), and continuing in this manner until black is obtained, because thus proceeding to its limit it does not veer from gradually stretching towards blackness, until it becomes pure black. There is also another path proceeding [from whiteness] toward light red [DMLBS: **somewhat-red**] (*subrubeus*) and from there to red [DMLBS: **red**] (*rubeus*), thereafter to black. The third path is the one going to blue-greenness [DMLBS: **green**] (*viriditas*) from there to indigo [DMLBS: **indigo**] (*indicus*), thereafter to blackness. And in these ways not all color diversity can exist, neither can they be the source of the diversity of [Aristotelian] median colors. A sort of black smoke mixed with fire will be red if blackness is abundant, or it will be orange (*citrinus*) if blackness is subdued and there is a superabundance of resplendent whiteness; if orange is mixed with blackness and splendor is not present greenness will occur; and altogether with blackness it will be more hidden and with whiteness more palpable When in the first case blackness is abundant the color will be dark [DMLBS: **dark**] (*fuscus*, the dark color of undefined hue), if blackness is abundant in the second the color will be that of leek leaves; on the other hand if it [greenness] becomes more and more it will be intense greenness, for which there is no name. But if whiteness is added to it will be grayish green; on on the other hand if blackness with a small amount of redness is mixed into grayish green the result will be indigo; but if redness is mixed to indigo carmine will result.

This passage is worth repeating in full, as it is not a simple case of expanding Aristotle's one-dimensional linear colour scheme between white and black into two-dimensions. Rather, Ibn Sina identifies three, disconnected routes along which one may pass from white to black: through *pallidus*, through *rubeus*, and through *indicus*. Here we already encounter a linguistic barrier, as the translation of colour terms featured in classical sources are often undetermined. *Pallidus* may variously be

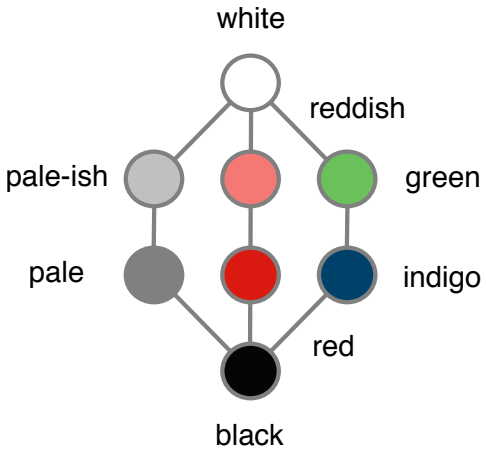


Figure 2.2: An interpretation of Ibn Sina's expansion of Aristotle's colour theory. He claims that there are three routes connecting white to black: one passing through pale (*pallidus*), one through red (*rubeus*), and one through both green (*viriditas*) and indigo (*indicus*).

translated as pale, pallid, colourless. This would suggest that, instead of the yellow-green interpretation decided upon by Kuehni and Schwarz, Ibn Sina is identifying an explicitly achromatic route from white to black, along with routes through reds (*rubeus*) and through greens and blues (*viriditas* and *indigo*, respectively).

Ibn Sina then supplements these colour routes with additional descriptions, or 'recipes', of colour mixtures, as he says these three routes in and of themselves are insufficient to account for all colours. He describes the following complex mixtures:

- Black smoke + fire (abundant blackness) = red
- Black smoke + fire (subdued blackness/abundant whiteness) = orange
- Orange + blackness (no splendor, with blackness) = hidden greenness
- Orange + blackness (no splendor, with whiteness) = palpable greenness
- Orange + blackness (no splendor, with blackness, blackness abundant) = dark
- Orange + blackness (no splendor, with whiteness, blackness abundant) = leek
- Green + green = intense greenness
- Green + whiteness = greyish green
- Greyish green + blackness + redness = indigo
- Indigo + redness = carmine.

Here already we begin to see a potential pitfall in constructing a colour theory, which is the sheer number of potential combinations of colours to mix and their results. Without an understanding of the abstract space that perceptual colour is embedded in, descriptions of this space can quickly become complicated and disorganised. This is more clearly seen in the colour theory of a later Islamic scholar, Nasir al-Din Tusi (1201-1274) (Kirchner and Bagheri, 2013). In a letter to one of his staff, Najm al-Din al-Katibi, Tusi writes that there are in fact five paths 'from which one gradually walks from white to black'. These are given as the following (translation by Kirchner and Bagheri):

> Regarding the production of colors from black and white there are numerous paths, from which one gradually walks from white to black. The path through yellow belongs there: First by the mixing of dense and fire, both in small amount, the straw-yellow is produced, then the lemon-yellow, then the saffron-yellow, then the orange-yellow, then the grenade-yellow, then in it the tendency towards black increases, according to the increase in the number of dense particles and the decrease of fire, until it becomes black. Another path goes through red. First it becomes rosy then like evening-red, then blood-colored, then purple, then violet, violet-colored. One path goes through green. It becomes pistachio-colored, then leek-colored, then verdigris-colored, then eggplant-colored, then naphta-colored. One path goes through blue. It becomes sky-blue, then Turkish blue, then lazur blue), then indigo blue, then like kohl. One path goes through turbidity/dirt. It becomes grey, then darkish/dirt-colored, then dark etc. These all occur according to the differences of particles in transparency, opacity (density), light and darkness. Now and then one sees a color together with another, and a different color is produced, such as green from yellow and blue, verdigris from green and white. There are infinitely many of such arrangements, and some are often found in small

> particles of plants and animals. Anyone who observes them is surprised by their number.

Tusi identifies five different paths connecting white and black, with an explicitly continuous space extending along them. There are 25 colour terms listed in his letter, mostly named after objects of nature such as foods, stones, and skies. He has clearly departed from the Aristotelian view that there are only seven basic colours, and has instead enlisted primary observation to compile a comprehensive inventory of perceptual colours, and how they are ordered in perceptual space. However, in so doing, his theory of colour, as well as Ibn Sina's, lacks engagement with the fundamental Aristotelian principles of perception, namely that sensory perception is about contrariety, and judgements along continua extending between extrema. Instead of more colour names, and more routes connecting them, it is an understanding of the space of perceptual colour as a mathematical structure—that is to say, the dimensions of contrariety—that is required to build on Aristotle's theory.

2.1.3 *The mathematics of sensory space in Grosseteste's De colore*

The first truly mathematical engagement with perceptual colour space is found in Robert Grosseteste's (c. 1175-1253) treatise on colour, *DC*. This text of under 400 words, written in the mid-1220s (Dinkova-Bruun et al., 2013; Smithson et al., 2012) is devoid of any colour terms other than white and black. He asserts that these form the two extrema of colour space; rather than routes connecting them through different, named colours, there are three qualities that may be changed, in different combinations, to navigate the space between them. He proceeds to enumerate these combinations, giving a mathematical justification for Aristotelian theory of there being seven colours. Instead of colour terms such as yellow, red and green (which Grosseteste used in other writings, such as Dictum 141), these seven 'colours' are routes extending from the extrema along different combinations of qualities. As there are three qualities that may be increased from black, where they are all minimal, or

decreased from white, where they are all maximal, the different combinations of these (given by the binomial coefficients $\binom{3}{1}$, $\binom{3}{2}$, and $\binom{3}{3}$, and painstakingly spelled out in linguistic terms by Grosseteste) sum to seven. While expressly trying to account for Aristotle's theory of there being seven colours, he has in fact advocated there being 16: seven routes proceed both from white, and from black, giving a total of 16 colours (white and black inclusive), rather than Aristotle's seven.

These qualities are, for the most part, not given a clear, observational basis. They are identified as three bipolar contrarieties, two pertaining to the light, and one to the medium in which the colour is seen. The light may be *multa* (copious) or *pauca* (scarce), and it may be *clara* (bright) or *obscura* (dim), while the medium may be *purum* or *impurum*. Of these terms, only *multa* is further elaborated on, connecting it to observation (Dinkova-Bruun et al., 2013) (translation by the Ordered Universe research group):

> I do not say that *multa* light is light diffused through a large expanse; rather, I say that *multa* light is gathered as if in a point when a concave mirror is positioned facing the sun and light falling over the whole surface of the mirror is reflected towards the center of the sphere of the mirror...

These three sets of bipolar qualities constitute dimensions, along which Grosseteste claims one may find any colour. The seven 'colours' from white, and the seven from black, are in fact idealised routes of travel within colour space. Instead of naming and locating colours, as Aristotle and Islamic scholars did before him, he is concerned with the space itself, and its mathematical properties. Here, we see a prime example of Grosseteste applying mathematical reasoning to the observable natural world. As with some of his other theories we do not see experimental science, where mathematical relationships are discovered by interrogating physical measurements, but rather an attempt to organise sense data within an abstract mathematical model.

2.1.4 *Grosseteste's De iride*

Grosseteste's treatise *DI* is perhaps his last scientific text of which we have record, dated to c. 1228. The text is far more extensive than the *DC*, opening with the claim that the rainbow can only be understood at the convergence of otherwise separate disciplines (translation by Sønnesyn):

> The inquiry concerning the rainbow falls within the remit both of the perspectivist and of the physicist.

We see that Grosseteste is a proponent of interdisciplinary science. It is not surprising that the rainbow, a phenomenon that ordinarily exists solely as an image (internal representation), while being undoubtedly grounded in physical conditions, would be a nexus for two branches of scientific inquiry.

Grosseteste identifies perspective as 'a science founded on visual figures', which he makes explicit are the mathematical lines and angles governing the directions of rays, emitted from both the eye (sight) and external bodies (light). This initially reads as a straight-forward case of Plato's *synaugeia* theory of vision, which requires both intromissive and extramissive visual rays. The victory of intromission over extramission, principally achieved by Ibn al-Haytham [Alhazen] (965-1040) in his masterful and exceptionally observational treatment of optics (Lindberg, 1967), is often heralded as a significant advancement in the history of science. But we should not be so quick to dismiss what may be painted as a plainly incorrect and superfluous theory. In combining the two, Plato sought to reckon the clearly apparent physical reality of luminosity—light-emitting and reflecting objects—with the subjective experience of 'looking' (Brett, 2014). In the *DI*, Grosseteste expands on this theory of vision, refuting the criticism that extramission is a baseless and unnecessary theoretical adjunct to an otherwise complete explanation of vision:

> Thus natural philosophers dealing with that which belongs to the natural and passive part of sight say that sight happens by intromission. But mathematicians and physicists who study what is beyond nature, dealing

> with the part of sight that is beyond nature and active, say that sight happens through extramission. This part of sight that happens through extramission Aristotle treats clearly in the last book on animals, saying "the inner eye sees remotely; for its movement is not broken up or consumed, but the power of sight goes out from it and passes straight to the things seen."

Grosseteste recognises that the natural mechanism of sight is explained by intromission, light radiating from external bodies to the eye. *Natural*, in a Platonic understanding, refers to the passive actions and behaviours of bodies according to their intrinsic qualities (natures), devoid of the agency or subjectivity of the soul, which is not passive but active. So, the physical, objective, and optical basis of sight is intromissive. But Grosseteste recognises that there is more to sight, which is determined not only by the external environment, but by the subjective observer, or agent. This is 'beyond nature', and instead an activity of the observer. Today, we can read this theory of extramission in the light of studies on visual attention and attention-based effects such as inattentional blindness, and more generally with regards to top-down processing in the visual pathway. These phenomena are not facilitated by physical, external substrates relating the observer to stimuli, and likewise Grosseteste is not claiming that a *physical* substance is emitted from the eye to that which is seen. Instead, there is something, not metaphysical, but psychological, at play, to which Grosseteste seems to be sensitive. To support his position, Grosseteste turns to Aristotle, a thinker who is often seen as a detractor of extramission, but here is correctly identified to be advocating a non-physical kind of extramissive theory in his writings on sensory perception and the *psyche*—that is, the soul, or perhaps more pertinently, the mind.

Moving on to the meteorological event of the rainbow, Grosseteste departs from the long-standing Aristotelian theory that its optical mechanism is explained by reflection, and instead places refraction at the centre of this theory. He hypothesises a complex geometric scene of transparent bodies—possibly droplets of rain—which

are gathered into the shape of an inverted cone. Sunlight passing through this conical body is gathered by refractive events and ultimately projected into the shape of a bow, as will be discussed at length in part ii.

2.2 SPEECH AND THE MEDIEVAL SENSORIUM

Part iii of this thesis concerns Grosseteste's theory of speech production and perception, described in his treatise *DGS*, and to a limited extent in another treatise of his, *DAL*. Speech was regarded as an act of power and perhaps beauty in the medieval world. Not only representational or symbolic, speech had an innate potency to effect change in the world—it was *declarative* (Woolgar, 2006). In fact more than change, it was the mechanism of divine creation. In Genesis, the world is created and ordered by declarative statements. In the Gospel of John, the *logos* [word] is identified and incarnated as Christ. In later medieval Mariology the Annunciation itself—the hearing of the words spoken by the angel Gabriel—is the mechanism of this incarnation, as the Virgin Mary conceives. On a day-to-day basis, speech was the messenger of the soul, and understood to be embodied in Mercury, the messenger god of the pagan pantheon, as found in Isidore of Seville's (c.560 - 636) encyclopaedic *Etymologiae* [*The Etymologies*] (Barney et al., 2006). Woolgar argues that as the medieval period went on, speech was increasingly identified as a sixth sense, of the mouth, and perhaps more so than taste (Woolgar, 2006).

However, speech as a perceptual phenomenon did not receive as much scientific attention in the classical world as colour. Nonetheless, certain theories of grammar, rhetoric, and phonetics, were suggested, and influenced Grosseteste's approach to this subject matter. The most influential of these was Isidore's *Etymologiae*, which he compiled during the first three decades of the seventh century. Within, Isidore identifies two categories of letter, the vowels and the consonants. Vowels are unimpeded, produced with a straightforward opening of the throat, and make a complete sound without an adjoining consonant. Consonants are produced by motions of the

articulators, that only produce sound when sounded together (*consonare*) with an adjoining vowel. We see this same description in the *DGS*, although Grosseteste also references the Roman grammarian Priscian (c.475? - c.525?), presumably with his eighteen-volume exposition of Latin grammar, *Institutiones grammaticae* [Grammatical Foundations], in mind.

Outside Europe, exciting innovations were again under way in the Islamic world, where a keen interest was taken in the physical anatomy of the vocal apparatus. Indeed, the earliest example of a phonetic diagram can be found from the late twelfth- or early thirteenth-century work of Al-Sakkākī (Hassan and Heselwood, 2011; Heselwood, 2013). Al-Sakkākī, a contemporary of Grosseteste, was a grammarian and rhetorician from Khwarazm in Persia whose 'systematic work of definitions and formulations became standard in the science of Arabic rhetoric' (Meisami and Starkey, 1998). Although this diagram did not take the shape of vowel diagrams that later came to dominate phonetic illustration, it marks the first precise location of articulators for human speech (Owens, 2013). A reproduction of the diagram can be seen in figure 2.3 (originally appearing in Bakalla, 1982, quoted by Hassan and Heselwood, 2011, and Owens, 2013). While Grosseteste's theory of speech is less empirically focused, it nonetheless illustrates the medieval period as one of growing interest in the mechanics of living, bodily systems.

2.3 SUMMARY

In this chapter, the medieval scientific theories of colour, the rainbow, and speech, were introduced, along with some discussion of their intellectual-historical background. The subsequent parts of this thesis will explore aspects of each topic in greater depth, and respond with fresh historical analysis, as well as experimental and computational investigations.

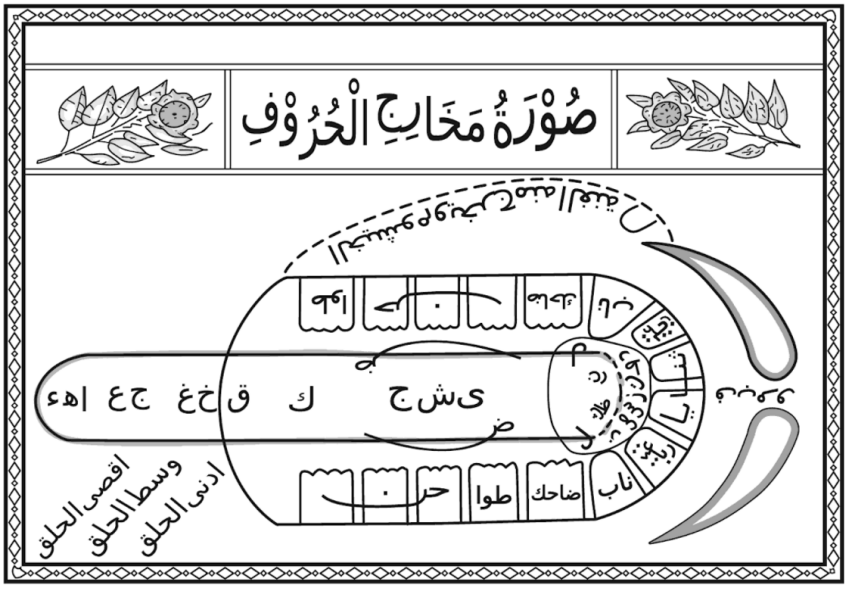


Figure 2.3: Vocal tract diagram from Al-Sakkākī's *The Key to the Sciences*, reproduced from Bakalla, 1982. The dashed line indicates the nasal passage, stemming from a nostril above the lip (shown on the right of the diagram).

3

PUBLIC ENGAGEMENT

3.1 THE ACCESSIBLE ACADEMY

Gone are the days when academics sought refuge in fortresses, fearful and ignorant of the townspeople with whom they shared a relationship marked by resentment and violence. Separated by language, dress, and physical space, medieval and early modern scholars saw little value or hope in sharing their knowledge with the general public, let alone including them into the academic process. Today, we are witnessing what some claim to be the gradual dissolution of educational bodies as defined and contained institutions, as the 'monastery' collides with the 'market' (Kirp, 2004). While this heralds new challenges and uncertainties for the academy, it also presents new opportunities to reconcile research communities with civil society. Indeed, it is now increasingly common for research funding to require some measurable degree of public engagement or 'impact', not least when it is awarded by public bodies.

The potential benefits to civil society are immediately clear, whether they be the cultural enrichment of the arts and humanities, or the translation of personal and environmental health matters into democratic policy enabled by scientific communication. What is perhaps less salient however, is how this relationship might benefit the academy. First, the limited case of 'public communication' necessitates the re-expression of complex ideas and findings without the specialist lexicon of a particular discipline. This is valuable—healthy, even—as jargon, while facilitating

everyday research, can occasionally conceal gaps in understanding or argument. To quote Ernest Rutherford (1871 - 1937), "An alleged scientific discovery has no merit unless it can be explained to a barmaid" (Whitrow, 1973)[1]. Second, and moving beyond the one-way traffic of communicating *to* the public, 'public participation' is an active and reciprocal dialogue *with* non-specialists. This typological distinction, discussed at length by Rowe and Frewer in the context of public policy and government (Rowe and Frewer, 2005), is also applicable to academic research, albeit with different objectives and measures of efficacy. While in a policy setting, efficiency may be measured in terms of maximising the transfer of information between different groups and its coalescence into a composite, for academic research we may value more the transfer of questions, associations, and perhaps enthusiasm. In this manner, public engagement confers not only a practical benefit to ongoing research, but is weaved into the creative context within which all research, whether in the sciences or the humanities, is ultimately rooted.

Lastly, the generational impact of public engagement should also be considered, as many events and methods of participation are accessible to children and young people. In this manner, it constitutes part of an extended curriculum, connecting early education with a tangible possibility of further education. In societies where opportunities to access higher and further education are still unequally distributed, this goes some way toward maximising the 'talent pool' from which academies can draw.

3.2 ACTIVITIES UNDERTAKEN DURING THE COURSE OF DOCTORAL RESEARCH

During the course of my doctoral research, I attended and contributed to four public festivals: *Being Human 2015*, *Being Human 2017*, *Cambridge e-Luminate 2018*, and *Napa Lighted 2018*. The material presented at these festivals included Robert Grosseteste's

1 While the intention behind this comment is clearly positive, it does also serve to underscore the importance of approaching public engagement without condescension or haughtiness.

scientific treatises, the history of colour theory, historical understandings of the rainbow, and material perception in medieval polychrome sculpture.

For *Being Human 2015* I participated at an event in Trevelyan College, Durham, called 'The Music of Light'. For this event, targeted to locals of Durham and surrounding towns, posters were displayed alongside hands-on demonstrations and workshops. Two posters were made in collaboration with Timothy Farrant, which presented Grosseteste's treatises *De luce* [*On Light*] (*DL*) and *De colore* [*On Colour*] (*DC*) in a visually engaging manner. For a hands-on demonstration, I showcased a re-enactment of Theodoric of Freiberg and Kamal al-Din al-Farisi's fourteenth-century experiments, which revealed the geometric-optical mechanism of the rainbow. By shining a narrow beam of light into a water-filled glass sphere, with the correct angle of entry, the caustics responsible for the primary and secondary bows can be seen, projected onto screens.

At *Being Human 2017*, which had the theme 'lost and found', I participated in the 'Lost Late' event at the Natural History and Pitt Rivers museums in Oxford. In addition to the Freiberg/Farisi experiment, I also demonstrated a physical model interpretation of Grosseteste's rainbow mechanism, described in his treatise *De iride* [*On the Rainbow*] (*DI*), and discussed further in part ii. In addition, I showcased a reconstruction of Ptolemy's (c. 100 - c. 170) refraction apparatus, which was used (or at least claimed to be used) in classical and medieval science to measure the angles of refraction for light passing between air, water, and glass. The event was targeted to locals of Oxford and surrounding towns, especially families and individuals still in primary or secondary education.

At *Cambridge e-Luminate 2018*, I delivered a public talk, 'Weaving a rainbow: the physical and perceptual construction of the rainbow's shape and colour', at the Cambridge Guildhall, which again presented material discussed in part ii. The event drew a primarily older crowd of approximately 20 people.

Finally, at *Napa Lighted 2018*, which had the theme 'beyond', I gave two public talks in Napa City, USA. The first, 'Beyond colour: historical and mathematical dimensions', gave an introduction to the history of colour theory, as it evolved through antiquity

into the modern era. Much of this material was presented in chapter 2. The second, 'Beyond materials: light, metals, and visual perception', presented research into the visual perception of materials used in medieval polychrome sculpture, which is discussed in part iv and part v.

While attending the 41st *European Conference on Visual Perception*, I showcased an immersive visual illusion with colleagues from the Oxford Perception Lab, at the *Un mare di illusioni* public exhibition night. The event was attended by researchers in the vision science community, artists, and many locals of Trieste, Italy. A publication describing the exhibit—the *neon fruit illusion*—is found in chapter 12.

As well as engaging with many young people, still in primary and secondary education, at public events, I have also been involved in the organisation and running of the Ordered Universe stream of the OxNet Summer School. Based at Pembroke College, this initiative seeks to widen access to higher education for students in the North East and areas of social deprivation. Students who are accepted onto the program are situated within a stream, typically defined by divisions such as humanities or sciences. The Ordered Universe stream is unique in delivering an interdisciplinary curriculum, spanning medieval history, translation, physics, and psychology. Currently in its third year of running, the program entails a series of seminars, delivered in Sunderland, an Easter school, and a week-long summer school located at Pembroke College, Oxford. Students have the opportunity to engage first-hand in interdisciplinary research, participating in collaborative readings of medieval texts, translation sessions, and creating a piece of written work that forms the basis for a tutorial. The material presented in chapter 1, as well as parts ii and iii, encapsulates much of what I have presented in collaborative readings, seminars, and tutorials within the OxNet program.

Part II

OPTICS AND THE RAINBOW

4

THE RAINBOW AS A HISTORICAL MOTIF

Where does the rainbow end,

in your soul or on the horizon?

—Pablo Neruda, 1991, trans. William O'Daly

PART ABSTRACT

The rainbow forms the focus of the research presented in this part. First, in chapter 4 the historical and scientific significance of rainbows is discussed. Following this, in chapter 5, Robert Grosseteste's theory of rainbow formation is explored with an interdisciplinary and experimental approach. It is the first theory in the medieval world to correctly recognise the central role of refraction in geometric optics, and makes a radical departure from the prevailing Aristotelian understanding that rainbows arise solely due to reflection events. With the work presented here, it is shown that his central claim about the geometric optical properties of transparent cones was correct, and hence that his theory was highly likely to have been informed by direct sensory observation, rather than a theoretical or inherited justification. Following this, in chapter 6 Grosseteste's cone-based scheme is found to be analogous to another variety of atmospheric dispersion phenomena, that of Parry's arc. This coincidence is discussed in terms of how it could be that Grosseteste, writing in the

thirteenth century, could have unwittingly hypothesised a valid optical mechanism for a phenomenon later found in nature.

4.1 WHAT IS SPECIAL ABOUT RAINBOWS?

The rainbow is a ubiquitous motif throughout both historical science and the history of science. There are intuitive reasons for this. Rainbows are a constant phenomenon throughout history—it is plausible that the rainbows we see on an inclement day are the same seen by Grosseteste in the thirteenth century, and by Aristotle before him. Whilst being constant, in the sense of an unchanging 'stimulus' or phenomenon upon which theory and science can evolve over centuries, they are also rare. Rainbows are not daily occurrences but rather special occasions, requiring a relatively narrow range of conditions to be met: namely that of rain, sunshine, and both the absolute and relative positions of the sun and an observer.

Rainbows are visibly special in two important ways—their shape, which is a perfect circle or arc thereof[1], and their colour appearance. Whilst the latter visual feature is perhaps more commonly how rainbows are primarily associated, it is worth considering how rarely in nature one finds a perfectly circular body. Indeed, with the exception of the sun and moon appearing as flat discs in the sky, rainbows may be the only other geometrically circular images found in nature[2]. It is therefore not surprising that theorising on the shape of the rainbow and geometric optics forms much of the content of Grosseteste's treatise *De iride* [*On the Rainbow*] (*DI*), before any mention of colour.

The colours of the rainbow have, naturally, also been discussed and theorised extensively throughout the history of science. It may be surprising, given the apparent

1 Rainbows are, in theory, circular. In practice, most are partially occluded by the horizon, leaving an arc or semicircle (at sunset or sunrise). However, when viewed from an aircraft or precipitous peak, the full circular extent may be seen.

2 This fact would have carried even greater significance in the ancient and medieval geocentric cosmos, which held that all celestial bodies—those situated in the sphere of the moon and above—were perfect and hence perfectly spherical. The rainbow then, with its perfect geometry and sublunary location, was a terrestrial manifestation of the divine.

consensus we have today[3], that the number of colours identified in the rainbow has varied considerably, ranging from as low as three (Aristotle, Book III of *Meteorologica* [*Meteorology*] (Aristotle, 1984c)), and four (Isidore (560 - 636), *De natura rerum* [*On the Nature of Things*] (Kendall, 2016)). In fact, Newton only settled on the number being seven later in life after being influenced by Pythagorean theories of music, having previously maintained there were only five colours (Fisher, 2015)[4]. If we turn to medieval manuscripts, we see a wide array of different colour combinations and orders in the rainbow. Nonetheless, they function, unambiguously, as representations of the same stimulus, due to the phenomenological uniqueness of the rainbow.

4.2 A BRIDGE BETWEEN DISCIPLINES

Grosseteste's treatment of the rainbow in his treatise *DI* was introduced in chapter 2. Already in the thirteenth century, Grosseteste identified that as an object of study, the rainbow required understanding in multiple disciplines, which today we see reflected in physics and vision science. Rainbows are a special kind of stimulus, because they only exist if there is an observer (eye, or camera) in the scene with the required geometric relationship to the sun and the rain. Thus, rainbows cannot be studied without studying the observer to some extent. We might also say that rainbows are not objects—they have no object permanence independent of observers, and are instead a class of event, much like a sunrise. Whether this event is framed in terms of the external world (a meteorological event) or the internal world of an observer (a sensory event), there is an unmistakable interdisciplinarity involved.

3 As if to stamp out dissent, many of us are taught from an young age that *Richard of York gave battle in vain*. The opposite approach can be seen in the parenting style of Guy Deutscher and Janie Steen, scholars of linguistics and English literature, respectively. Raising their daughter, Alma, they taught her basic colour terms without associating cloudless skies with blue. On some such days, Alma was asked what colour the sky was. Initially confused at the question, she eventually responded: white. Over time her answers changed to occasionally include blue, and eventually stabilised on blue (Howard, 2012).

4 Strictly speaking, Newton was enumerating the number of colours in the spectrum, as is produced by dispersing sunlight through a prism. It should be clearly stated that a rainbow is *not* a spectrum, and rainbow colours are mostly not spectral. This is due to the overlapping distributions of wavelengths, as spherical raindrops do not scatter rays to unique angles.

5

BOW-SHAPED CAUSTICS FROM CONICAL PRISMS: A THIRTEENTH-CENTURY ACCOUNT OF RAINBOW FORMATION FROM ROBERT GROSSETESTE'S *DE IRIDE*

Published in a special issue of *Applied Optics*: "Light and color in the open air", Vol. 56, No. 19, 2017.

AUTHORS

Joshua S. Harvey, Hannah E. Smithson, Clive R. Siviour, Giles E. M. Gasper, Sigbjørn O. Sønesyn, Brian K. Tanner, and Tom C. B. McLeish

ABSTRACT

The rainbow has been the subject of discussion across a variety of historical periods and cultures, and numerous optical explanations have been suggested. Here, we further explore the scientific treatise *De iride* [*On the Rainbow*] (*DI*) written by Robert Grosseteste in the thirteenth century. Attempting to account for the shape of the rainbow, Grosseteste bases his explanation on the optical properties of transparent cones, which he claims can give rise to arc-shaped projections through refraction. By stating that atmospheric phenomena are reducible to the geometric optics of a conical prism, the *De iride* lays out a coherent and testable hypothesis. Through both physical

Published in *Applied Optics*

experiment and physics-based simulation we present a novel characterisation of cone-light interactions, demonstrating that transparent cones do indeed give rise to bow-shaped caustics; a non-intuitive phenomenon that suggests Grosseteste's theory of the rainbow is likely to have been grounded in observation.

5.1 INTRODUCTION

In this paper we present an experimental approach to better understand historical discussion of a natural optical phenomenon. Rainbows are occasionally observed meteorological phenomena that have stimulated mythic and aesthetic thought in many human cultures. These optical events have also provoked scientific discussion on the nature of light, colour, geometry, clouds, and the atmosphere, leading to novel theory in both the physical and perceptual sciences.

The high medieval period saw advances in a diverse range of scientific fields, and perhaps chiefly that of theoretical and practical optics. A celebrated contributor to this advancement, Robert Grosseteste, explores the topic of the rainbow in his treatise *De iride* [*On the Rainbow*] (*DI*), written c. 1228. Within the short text, one of thirteen treatises written by Grosseteste on natural phenomena, an optical mechanism underpinning the rainbow is proposed. Previously, an interdisciplinary analysis has been applied to this treatise to explore its history (and likely dating), as well as its implications on colour theory in the context of Grosseteste's other writings (Smithson et al., 2014a). Here, we apply a similar research approach to the treatise's focus on how the shape of a bow is made to appear in the sky, opposite the sun, to an observer. Central to this explanation is the notion of a transparent meteorological body in the shape of a cone, and the refractive events that occur when light passes through it. Although Grosseteste does not explicitly offer an exemplar observation of cone-light interaction, his theory is reliant on what he claims is a universal property of nature; when illuminated in a certain fashion, transparent cones collect light into bow-like shapes.

Published in *Applied Optics*

This claim, predicting what would be described in modern terminology as caustics, is not obvious and cannot be arrived at through logical deduction, intuition, or simple mathematics. The directional propagation of a single ray of light through transparent media can be described classically by the laws of internal and external reflection, and Snell's law of refraction. If the refractive indices of media are known, and some trigonometric calculus is applied, the caustics of a sphere produced by parallel light may be modelled mathematically as the local maxima of distributed radiance. Apart from a sphere, whose every plane-section is circular, the light paths through three-dimensional objects are complex to compute, or even visualise. A successful recent example of this has been the simulation of rainbows from physically-accurate, non-spherical, raindrops (Sadeghi et al., 2012). But the geometric theory, optics, and mathematics required are far beyond the reach of a medieval scholar. Calculus was not formalised until the seventeenth century, and Willebrord Snellius—to whom Snell's law is attributed—was not born until the sixteenth century. While there is evidence that the law relating angles of refraction with angles of incidence was discovered earlier by the tenth-century Persian mathematician Ibn Sahl (Rashed, 1990), it is clear that Grosseteste was unaware of it as he put forward his own, incorrect, law (Eastwood, 1967).

With these barriers in mind, we suggest that the claim Grosseteste makes about cone-light interaction, and which underpins his mechanistic explanation of the rainbow, was rooted in observation. While we would not describe this as experimental science, it would have required astute observation of the material world, and an appreciation of the uniformity of nature across vastly different contexts; from the human scale of a conical glass vessel at the dinner table, to the meteorological scale of clouds and rainbows. Here we present a discussion of his theory, both its historical context and its content, and then explore his cone-light assertion further. The *DI* fails to arrive at the correct explanation for rainbow formation, but along the way it expresses novel and coherent ideas. Through both physical experiment with water-filled transparent cones, and physics-based rendering of caustics simulations, we show Grosseteste's

Published in *Applied Optics*

claim regarding cone-light interactions to be correct, and discuss its implications in the history of science.

5.2 EARLY HISTORY OF THE RAINBOW

The rainbow has been interpreted differently in a wide variety of human cultures. In Viking mythology it features as the bridge, Bifröst, between the realms of Gods and Humanity; for African and Australian peoples as the rainbow serpent - Kurreah for the Yualai of New South Wales; for Ancient Greeks, notably in the Iliad, as the messenger of the Gods (Lee and Fraser, 2001). Discussion of the rainbow both in terms of its colours and its shape occupied ancient and medieval authors considerably. To give an example of this and to situate Grosseteste's thesis in its historical context, the interconnection and distinction between the two aspects can be seen by the presentation of rainbows in the Bible. The Latin translation of the Bible, made by St Jerome in the later fourth century and the version that would have been known to Grosseteste, presents the rainbow shown by God to Noah as an arc in the sky with an intimate connection to clouds. The vision of heaven in the New Testament Apocalypse of St John (Revelations) refers, by contrast, to the rainbow by the Greek-derived word *iris*, the emphasis here placed on colour.

The most influential explanation of the rainbow as a natural phenomenon in the ancient world was that of Aristotle in Book Three of his Meteorology, which presents a geometric, rather than a physical, argument for the appearance and shape of the rainbow (Boyer, 1959). For Aristotle, the rainbow is produced by reflection off a cloud and consists of three colours, and the shape of the bow is related to the shape of the clouds onto which it is projected (Sayili, 1939). The Aristotelian theory remained unchallenged, and infrequently discussed in the West from Late Antiquity through to the High Middles Ages (c.1050-c.1250) (Lindberg, 1992). Isidore of Seville attributes the shape of the bow to clouds and sun, and then identifies four colours related to the four elements. For Bede, 'rainbows come to be when the sun illuminates clouds'.

Published in *Applied Optics*

Bede follows Isidore in his *De natura rerum* [On the Nature of Things], 'The Rainbow with its four colours is formed in the air from the directly opposed sun and the clouds' (Kendall and Wallis, 2010).

By contrast, scholars working in Arabic in the early Middle Ages, notably Ibn Sīnā (Latinised as Avicenna) (980 - 1037) and Ibn al-Haytham (Alhazen) (c. 965 - 1039), questioned the prevailing model, criticising the identification of rainbows only with clouds, but in neither case making much advance on Aristotle's position on the rainbow's shape (El-Bizri, 2016). Alhazen's optical theories would become influential in the West from the mid-thirteenth century, affecting discussion of the rainbow by Roger Bacon, Witelo and Theodoric of Freiberg (Lindberg, 1966).

Robert Grosseteste inherited a rather thin tradition of western scientific writing on the rainbow (Boyer, 1954). The new translations of Aristotle's Meteorology and Posterior Analytics offered the most comprehensive account of the formation and appearance of the rainbow available to him (Crombie, 1953; Eastwood, 1966). It is these texts that form the basis of his own, fresh, engagement with the phenomenon and the existing authoritative sources on it. Less clear has been the extent by which he drew on his own observations, a question that we tackle from a new angle in this work.

Published in *Applied Optics*

5.3 ROBERT GROSSETESTE'S THEORY OF THE RAINBOW

Our forthcoming, related publication will present a new edition of and commentary on the *DI*, and our present analysis of the text is based on the preparation of that edition and translation (Gasper et al., forthcoming). Grosseteste's treatise consists of two distinct halves. The first concentrates on the medieval science of 'perspective', or optics. Grosseteste asserts that the study of the rainbow is subordinate to the study of optics, that is, that the study and understanding of a complex phenomenon should be broken down into the simpler study of its underlying component parts. For thirteenth-century natural philosophy, this reads with a modern-sounding

scientific reductionism. Thus, the rainbow cannot simply be explained through its phenomenology, or observations alone, but requires an understanding of optical mechanisms.

Like many of his contemporaries, Grosseteste believed that sight is achieved by the emission of rays from the eye, though in the *DI* he insisted on the idea of *synaugeia* – Plato's proposition that rays from both the eye and the object were required in conjunction in order to facilitate sight (Lindberg, 1976). The lengthy discussion on the physics of refraction, which we do not explore in this paper, includes a theorisation of the law of refraction. Appealing to the simplicity of nature, as well as experimental evidence, Grosseteste states that the angle of refraction is equal to half the angle of incidence with respect to the normal. Earlier Islamic thinkers, of whose works Grosseteste was ignorant, had identified the correct law of refraction in the tenth century (Ibn Sahl) and, in the eleventh, recorded Ptolemy's second-century refraction experiment (Ibn al-Haytham) (El-Bizri, 2016; Smith, 2015).

In consideration of the rainbow, Grosseteste frames his account exclusively as the refraction of rays coming from the sun, finding agreement with modern optics. Accordingly, the formation of the rainbow is explained by the propagation of light rays through several transparent (*diaphanus*) media and the refraction of the rays at the boundaries between them. In particular:

> [T]he sun rays first are refracted at the interface of air and the cloud, and thereafter at the interface of the cloud and the drops, so that by these refractions the rays converge in the density of the drops, and are there once more refracted and spread out as though from a pyramidal cone. . . (Gasper et al., forthcoming)

An interpretation of Grosseteste's theoretical meteorology and optical mechanism of rainbow formation is shown in Figure 5.1, although it is important to state that no diagrams were drawn in the *DI* and Grosseteste's descriptions are ambiguous, due to both the language employed and the limitations of translation from the original Latin. The curved top of a cloud with drops descending from it to the ground forms

Published in *Applied Optics*

an inverted cone with a domed base. Light from the sun undergoes refraction at the air-cloud and cloud-drop interfaces, and subsequent refractive events amongst the drops themselves. Consequentially, light is gathered together within the cone. At the drop-air interface at the side of the cone, this gathered light is once again refracted, and spreads out:

> [I]n a shape similar to the curved surface of a round pyramid expanded opposite the sun. It therefore has the shape of an arch. (Gasper et al., forthcoming)

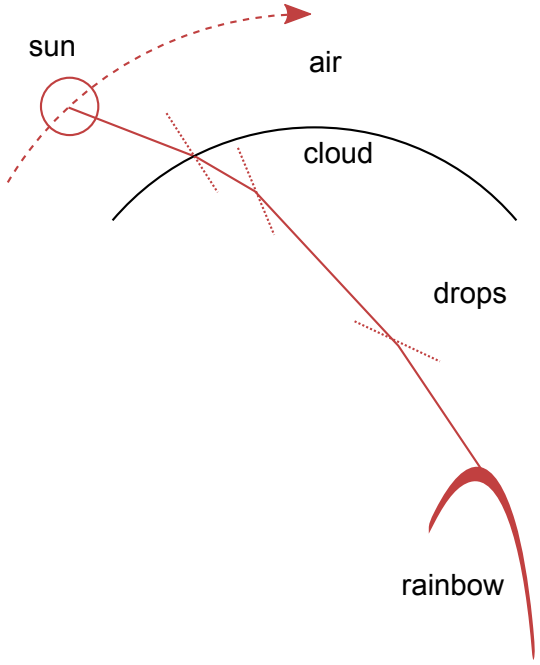


Figure 5.1: An interpretation of the geometric optics described in the *DI*. Rays from the sun undergo three refractive events: the interface between the air and the dome of the cloud, within the cloud's drops, and upon re-entering the air. These events collect the radiance into the form of a bow. In this depiction a single ray from the sun is shown.

Although the meteorological scene is complex with several refractive events resulting in the gathering of radiance into a bow (which is then seen as a projection, perhaps onto some other cloud), Grosseteste states that the overall transformation of the incoming light is analogous to that produced by a simple transparent pyramidal cone. While this is far simpler to consider than a scene with numerous transparent

Published in *Applied Optics*

media and associated refractive events at their boundaries, it is still a non-intuitive phenomenon as discussed earlier. Simply put, if asked what caustics are produced by a transparent cone when illuminated, most individuals will have no clear expectation, let alone base a theory of the formation of the rainbow on their conjecture. If transparent cones can produce such caustics, it is likely that Grosseteste knew this from observation.

We tested Grosseteste's theory of cone-produced caustics with both physical reconstructions of cone-light interactions and physics-based simulations. Observations of cone-light interactions accessible to Grosseteste, which could have provided a scientific justification for this theory, would have featured either divergent or collimated light. It is possible that Grosseteste had observed bow-shaped caustics from cones in an indoor setting, such as a candle-lit conical glass vessel, or in the open air, such as an icicle or glass vessel in sunshine.

5.4 HISTORICAL RE-ENACTMENT OF CONE-LIGHT INTERACTION

We hypothesised that the most likely setting for Grosseteste to observe the caustics from an illuminated transparent cone was that of a water-filled glass vessel on a candlelit dining table. At least since the Roman empire, we have evidence of cone-shaped drinking vessels and beakers throughout early European and Middle Eastern history (Harden, 1978). Two examples of conical beakers dated to the same period as Grosseteste are shown in Figure 5.2, taken from the website of the Corning Museum of Glass (Corning Museum of Glass, 2002).

To initially test our hypothesis, we filled a conical wine glass with water and illuminated it with a candle. The projected light collects into the shape of a bow, as depicted in Figure 5.2. While this may not meet the modern definition of a caustic as the envelope of light rays transformed by refraction or reflection, there is a clear collecting of rays (in keeping with the origin of the term), resembling other optical phenomena that are colloquially referred to as caustics. The formation of this shape,

Published in *Applied Optics*

and its striking display of colours, could well have provided an observation-based motivation for much of Grosseteste's *DI*.

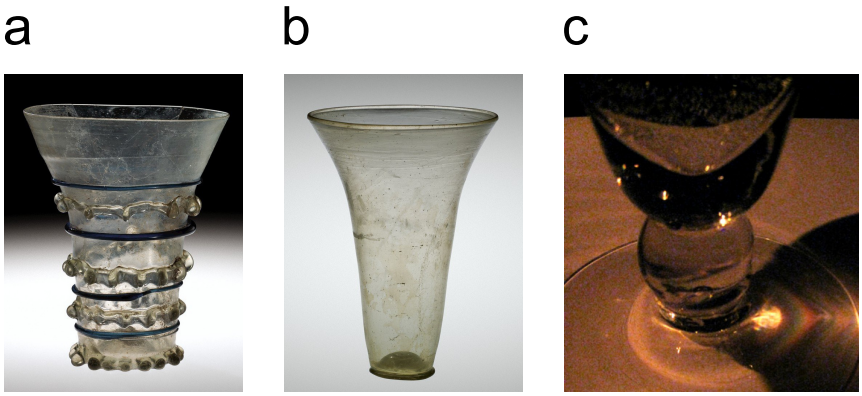


Figure 5.2: Cone-shaped glass beakers dated to the same period as Grosseteste from a) Central Europe and b) the Middle East. c) A modern conical wine glass filled with water, illuminated by a candle. A bow-shaped caustic can be seen to the right of the stem. The bow features a dispersive spectrum with the correctly-ordered colours for a rainbow.

To more precisely investigate cone-light interactions, we sourced a transparent conical vessel and lighting with properties to mimic light both from the sun and from an indoor light source such as a candle. A 1000 ml polycarbonate settling cone, with a base diameter of 110 mm, and an aperture (the angle formed in a cross-section by the cone's lateral sides) of 13.5°, was completely filled with water and illuminated by a xenon lamp at a constant distance of approximately 400 mm from the cone, with a varying angle of elevation, ϵ. The general plan of cone illumination is shown in Figure 5.3, which also serves as the basis for later simulations when cone aperture, α, can be varied. A large white piece of card was placed approximately 150 mm behind the cone to serve as a screen for the projected caustics, and the setup was photographed. The light source was used both with and without a collimating lens, to recreate possible observations made by sunlight (either outside in the open air or inside using a *camera obscura*), or candlelight, respectively. Figure 5.4 shows that for a range of ϵ values, a bow-shaped caustic is formed by both divergent and collimated illumination. Again, while this might not meet the definition of a caustic in the strictest sense (rays are not gathered into sharply defined lines or cusps), the projection appears similar to other optical phenomena, such as patterns of light seen

Published in *Applied Optics*

within an illuminated swimming pool, that are described as caustics, and we adopt the term throughout the following discussion.

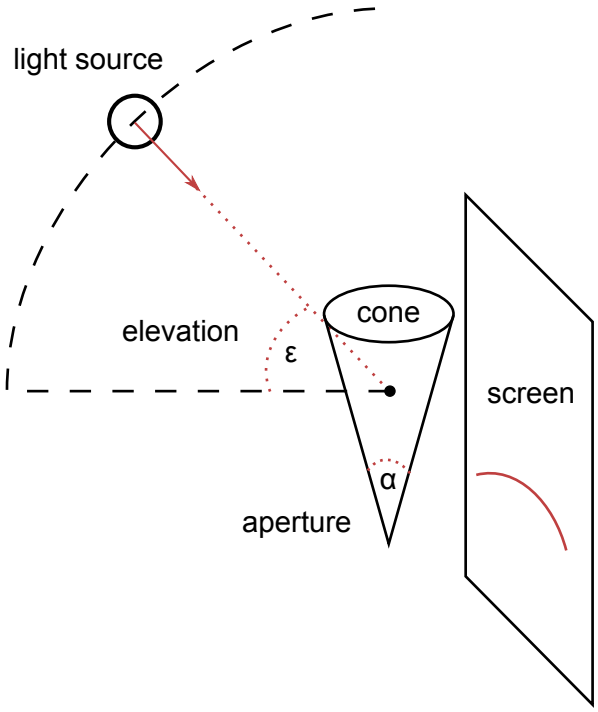


Figure 5.3: Parameters to describe the scene for both physical experiment and rendered simulations. The light source could emit either collimated light or divergent light.

It was evident that the colourful bow-shaped caustics shown in Figure 5.4 form from light entering the top of the cone, as small perturbations of the cone caused a rippling in the caustic. The bow was completely absent when the top of the cone was blocked with an opaque piece of plastic, as illustrated in Figure 5.5. This is concordant with Grosseteste's theory of the rainbow, where sunlight first enters the top of the cloud.

Although the focus of this paper is on the shape of caustics and the explanation in the *DI* for the shape of the rainbow, it is worth considering here a number of additional characteristics of natural rainbows that are at least partially reproduced in these caustics. As can be seen in the photographs of Figures 5.2 and 5.4, the bow-shaped caustics produced by transparent cones feature a strong display of colours, with red at the top of the bow as in a rainbow. Although Grosseteste is using cone-light interaction as an explanatory mechanism for the shape of the rainbow, it is

Published in *Applied Optics*

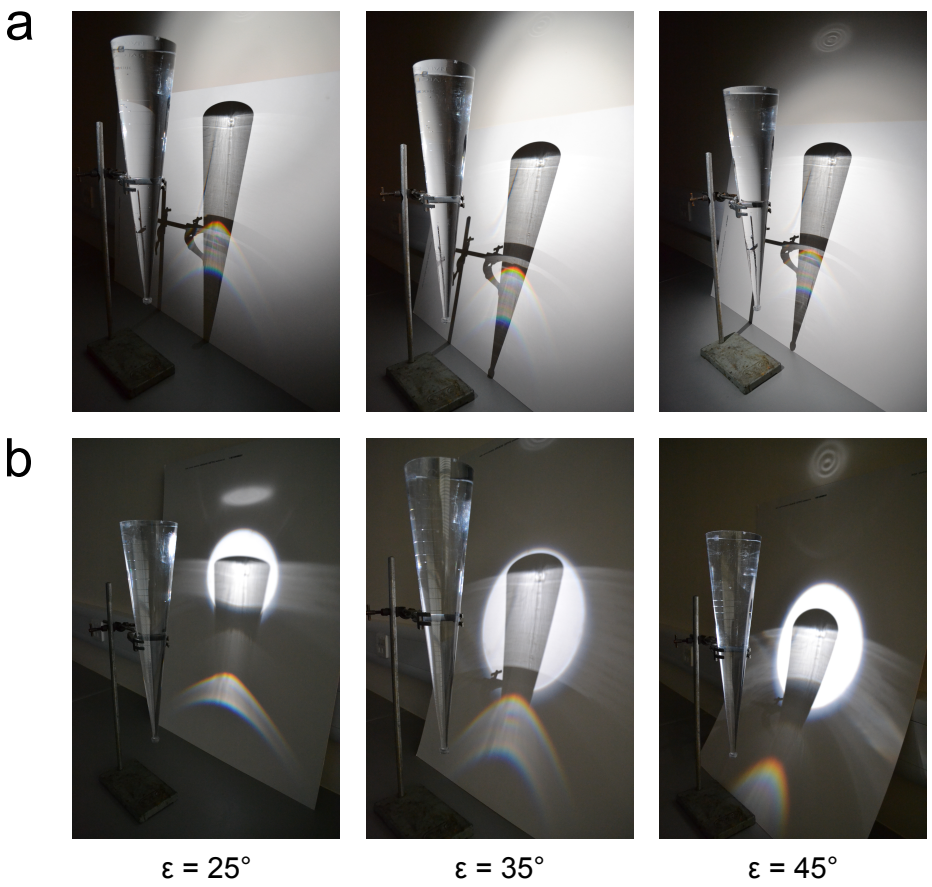


Figure 5.4: Physical experiment of cone-light interactions, for a cone of aperture $\alpha = 13.5°$, and elevation angles of $\epsilon = 25°$, $\epsilon = 35°$, and $\epsilon = 45°$. a) illuminated by divergent light. b) illuminated using a collimating lens.

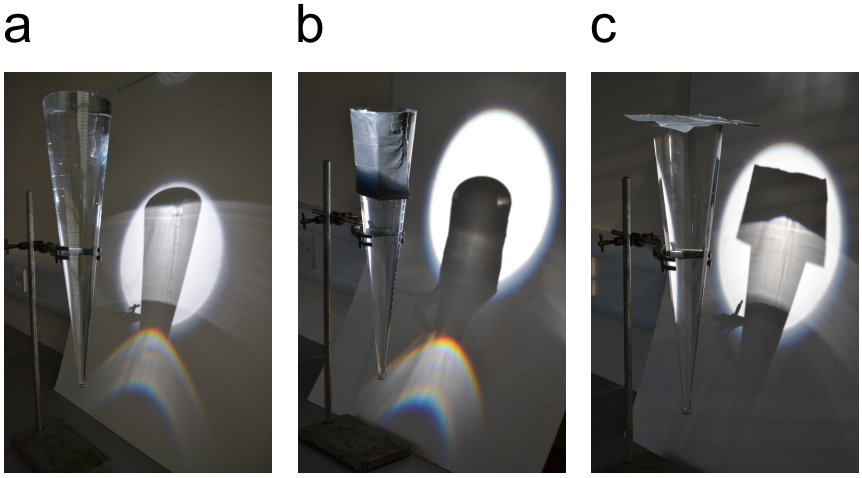


Figure 5.5: Light entering the top of the cone is responsible for the formation of the colourful bow seen projected onto the screen, collimated light, $\epsilon = 35°$. a) unobstructed illumination, b) blocking the side of the cone, c) blocking the top of the cone.

Published in *Applied Optics*

highly likely that the conceptual leap from glassware to meteorology was mediated by the observation of colour in both arcs. While these bows exhibit ordered variation in chromaticity, they are not the spectral colours that one sees from a beam of light exiting a dispersive prism. Rather, the colours appear desaturated to a varying degree, a feature of rainbows which has been discussed previously in connection to Grosseteste's theory of colour (Smithson, 2016; Smithson et al., 2014a). Other features, such as Alexander's dark band (the slightly darkened region of sky found between primary and secondary rainbows, first described by Alexander of Aphrodisias (fl. c. 200 BCE)), are also captured in the projection of many of these caustics, which may have contributed to the appeal of the cone-based theory.

5.5 PHYSICS-BASED RENDERING OF CONICAL PRISMS

To further explore cone-light interactions, physics-based rendering of simulated scenes was performed. As with the physical experiment, scenes were simulated to contain either an extended source of divergent light, such as a candle indoors, or a distant source of collimated light, such as sunlight in the open air or a *camera obscura*.

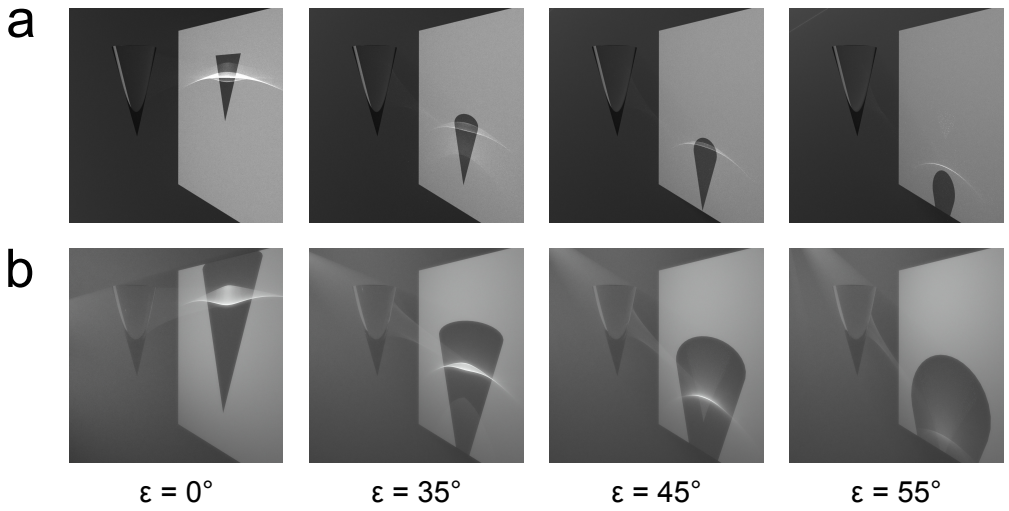


Figure 5.6: Images of caustics resulting from cone-light interactions simulated using the Luxrender engine in Blender. Upper panel: a simulation of the physical experiment photographed in Figure 5.4, featuring an extended source of divergent light. Lower panel: a scene lit by collimated light to simulate an observation in the open air illuminated by the sun.

Published in *Applied Optics*

Both scenes were constructed in the open-source program Blender to contain a light source, camera, transparent cone, and a screen onto which the cone's shadow and any resultant caustics would be projected (Blender Online Community, 2016). For scene descriptions, Blender-specific variables are printed in `monospace`. A degree of volumetric scattering was included in the `Exterior Volume` of the scene, so that rays could be visualised as they bunch together to form caustics. Further details for each scene can be found in the legend of Figure 5.6, which contains the corresponding rendered images. Although many rendering techniques do not readily yield caustics, simulations have previously been used to explore the lighting and caustics that would result from early Islamic glass lamps (Kider et al., 2009). For the work presented here, scenes were rendered with the Luxrender engine, a physics-based unbiased renderer that can be used to produce realistic images of caustics and other optical phenomena. While the study of the rainbow might naturally suggest an investigation into colour and dispersive phenomena, scenes were initially rendered without dispersion as we were testing Grosseteste's claims of how the shape of the arc is produced.

Cones with an aperture of 27° were constructed as `Homogeneous Volumes`, having a small amount of internal scattering, with a `Glass` surface, featuring an index of refraction (IoR) of 1.33 to model the optical properties of water. An opaque plane served as a screen for the projected shadow and caustics of the cone, positioned parallel to the cone's axis. Simulations were run with cones made with both flat and domed bases. While Grosseteste's meteorological description features a cone with a domed base, the outwardly-concave cloud, it is more likely that he observed caustics from cones with flat bases. Whether the bases of cones were flat or rounded did not have a strong impact on the resultant caustics, so for simplicity we have only presented the simulations featuring flat bases.

The first simulation used a small plane as an `Area Light`, emitting a spread of light over the scene from an extended source of light, to model a candle flame. The light object was positioned with locations and rotations corresponding to the desired ray angles and elevations. The second model simulated cone-light interaction in the open air by using the Luxrender `Laser` light source, which flooded the scene

Published in *Applied Optics*

with collimated light at a designated angle. It therefore modelled which optical phenomena would be observed from a transparent cone in collimated light, such as an icicle in sunlight.

5.5.1 *Rendering method*

For rendering, Luxrender's `Metropolis` Sampler was used in conjunction with the `Bidirectional` Rendering Mode. By using the Metropolis light transport algorithm to generate samples, and the bidirectional path tracer, the rate of convergence for rendering scenes with complex light-media interactions is greatly improved (Kelemen et al., 2002). This improved efficiency towards convergence is especially valuable for scenes featuring caustics. Rendering parameters were adjusted to improve the detail of caustics and reduce image noise, and the `Single` Volume Integrator was used in conjunction with `Noise-Aware Sampling`. Renders were tonemapped using the Reinhard kernel to preserve details in highly illuminated regions of the image (Reinhard et al., 2002). Rendering was halted when 99.99% of pixels passed a convergence test, producing images with a resolution of 1080 by 1080 pixels.

We note that a bidirectional path tracer, where samples are constructed through a combination of rays originating from the camera and rays from light sources, is essentially an implementation of Plato's *synaugeia* theory of vision, advocated by Grosseteste in the *DI*, discussed earlier.

5.6 RESULTS AND DISCUSSION

As illustrated by Figure 5.6, bow-shaped caustics are produced by both divergent and collimated light being projected through transparent cones for certain scene configurations. Luxrender simulations allowed for the construction of scenes with precise and modifiable geometry, such that a parameter space of cone-light interaction could be comprehensively explored. Two parameters were varied to investigate the

Published in *Applied Optics*

refractive optical interactions with transparent cones: cone aperture (the angle formed in a cross-section by the cone's lateral sides), and light source elevation (i.e. height and angle relative to cone, with a constant distance from the origin), as depicted in Figure 5.3. Narrow cones gave rise to arc-like caustics across a broad range of light elevations (for $\epsilon \approx 30 - 60°$), but as aperture increased to a more Martini-glass shape, arcs were less likely to form ($\epsilon \approx 35 - 45°$). For Luxrender simulations modelled with divergent light, we arrive at the parameter space illustrated in Figure 5.7.

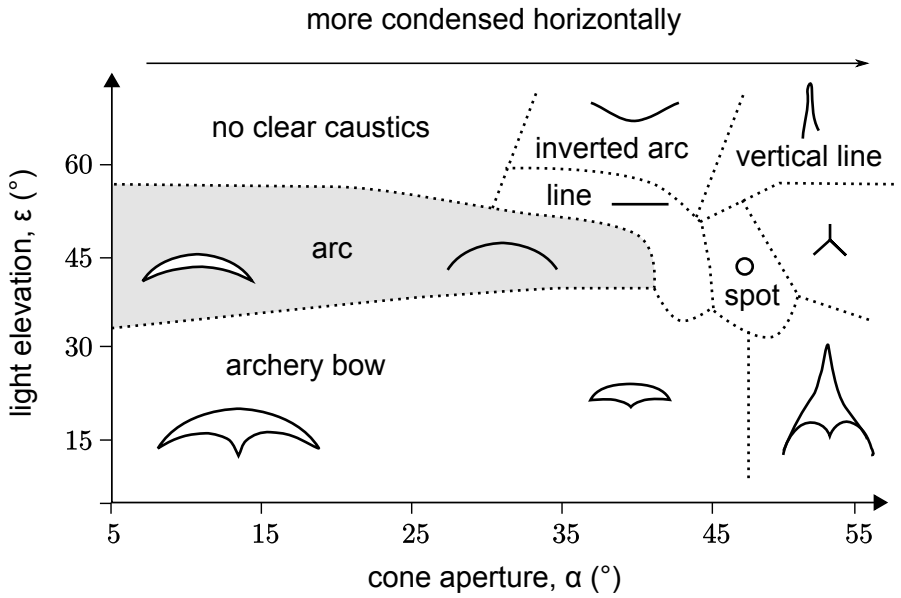


Figure 5.7: Parameter space of cone-produced caustics obtained from Luxrender simulations, for scenes containing a divergent light source. The region where arcs are produced is shaded gray for emphasis.

Not indicated in this parameter space is the proportion of light entering the top of the cone to that entering the side of the cone, although it logically follows that at low values of α and ϵ, light is predominantly passing through only the side. As values of α and ϵ increase, an increasing proportion of light enters the top of the cone, which according to physical experiment and the mechanism described in the *DI*, is the light responsible for the formation of bow-shaped caustics. To test if this was also the case for the bow-shaped light projections seen in the simulations, a scene was rendered containing a cone with its front face constructed with an opaque mesh. As seen in

Published in *Applied Optics*

Figure 5.8, it is also the case for this simulation that bow-shaped caustics arise from light entering the top of the cone.

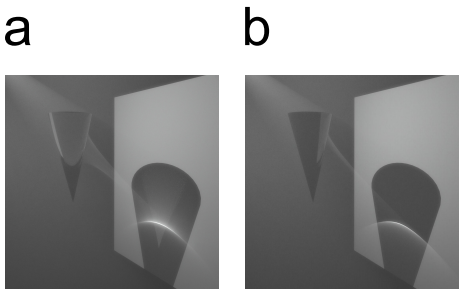


Figure 5.8: Light entering the top of the cone is collected into the bow-shaped caustic seen in the simulations. a) Divergent light of $\epsilon = 45°$. b) The same scene, but the portion of the cone's side facing the light source is opaque.

To simulate dispersive effects using Luxrender, twelve scenes were rendered containing cones of varying refractive indices, which corresponded to the effective IoR values for light at wavelengths equally spaced between 380 and 710 nm, as calculated for water by the three-term Selmeier equation:

$$n^2(\lambda) = 1 + \frac{B_1\lambda^2}{\lambda^2 - C_1} + \frac{B_2\lambda^2}{\lambda^2 - C_2} + \frac{B_3\lambda^2}{\lambda^2 - C_3}$$

where n is the IoR, λ is the wavelength, and B and C are experimentally determined Sellmeier coefficients (Daimon and Masumura, 2007).

The pixel intensities for each image were then plotted as a function of height through the render, as shown in Figure 5.9, which includes a composite image of the twelve renders recoloured according to wavelength. The caustic produced by the cone in the scene is dispersive, with light scattering to varying parts of the image as a function of wavelength.

During the physical experiment with the polycarbonate cone, another prominent bow-shaped caustic was observed to form at the same side of the cone that was illuminated, as seen in Figure 5.10. This presumably forms in the same way that transparent spheres (and in particular, raindrops forming a rainbow) produce caustics, namely by an internal reflection off the back face. However, Luxrender was unable to generate this caustic. To further explore this, the Matlab library Optometrika was

Published in Applied Optics

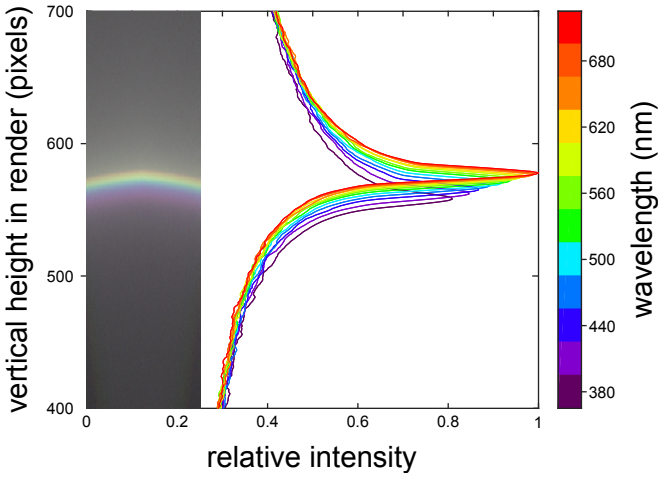


Figure 5.9: Simulated dispersive effects, after rendering images for twelve different wavelengths of light, $\epsilon = 45°$, $\alpha = 27°$. On the left is shown a composite image of twelve renders for light of wavelengths between 380 and 710 nm. The plot shows the distribution of pixel intensities for each wavelength through the center of the rendered image.

used to construct a ray-tracing scene similar to that of the physical experiment (Petrov, 2014, Retrieved January 2017). An aspherical lens was defined to approximate a transparent cone, and screens were placed both below the cone and behind, to capture any projected rays which had or had not been internally reflected, respectively. The more precise and customisable nature of the ray-tracing calculated by Optometrika provided more insight into how caustics may be created by cones. Light entering the side of the cone undergoes a minor transformation upon exiting the other side, with some bunching together horizontally but remaining parallel in the vertical axis. This is to be expected, as rays hitting the cone at the same radial angle, but different heights, are self-similar. However, rays entering the top of the cone (and therefore in agreement with Grosseteste's optical mechanism) do form strong caustics.

As shown in Figure 5.10, the internally-reflected caustic is well modelled by Optometrika, and the radiance of the scene forms a tight bow in front of the base of the cone. However, no rays seem able to make it through the back face of the cone, and the original caustic observed in both physical experiment and Luxrender simulation is absent. More fundamentally, it is necessary to address the clear differences between the results of the physical experiment shown in Figure 5.4, with the results from

Published in *Applied Optics*

simulations. While both the Luxrender and Optometrika simulations can accurately model important features of cone-light interactions, neither could produce the full range of caustics observed with the physical cone. The shortcomings of these simulations reflect the limitations of Grosseteste's own theory, which, whilst correctly describing some of the light projections from transparent cones, is silent regarding others.

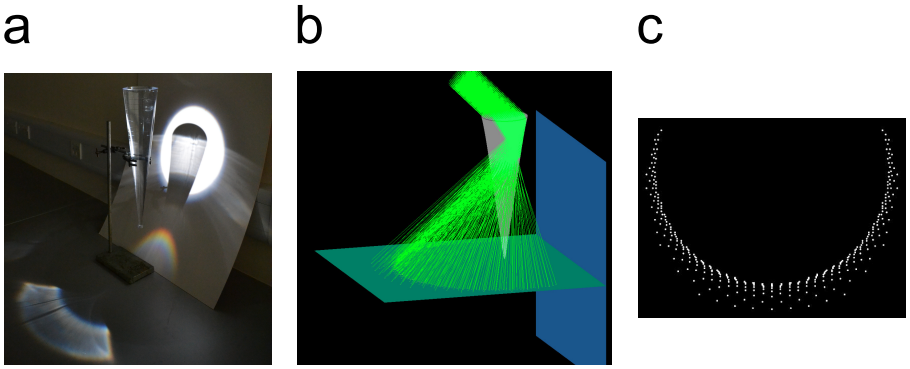


Figure 5.10: a) Photograph of physical experiment and b) Optometrika ray-tracing showing the bow-shaped caustic formed by internal reflection off the back face of the cone. c) The rays projected onto the horizontal screen in the Optometrika simulation. $\epsilon = 45°$.

Published in *Applied Optics*

5.7 CONCLUSION

While the history of scientific discussion on the rainbow has been written about extensively, we are still in the process of discovering and rediscovering contributions which were influential in the development of the field. Grosseteste has not normally been given prominence in the development of scientific understanding of the rainbow. In *Theories of Vision*, Lindberg dismissively writes:

> [M]uch has been made of Grosseteste's utilisation of refraction in explaining the rainbow, but the luster of that achievement quickly fades when we realise that his theory of the rainbow could not account for even the most basic phenomena and has remained largely unintelligible to the modern day. (Lindberg, 1976)

With this paper we have shown that, far from a lacklustre and unintelligible theory, Grosseteste's *DI* presents a novel, coherent, and pioneering theory of rainbow formation. By engaging with the text in an interdisciplinary way, we have uncovered features and implications that have been overlooked in previous discussions, and have found new evidence in support of Grosseteste's theory being grounded in observation of the material world. If this were the case, the *DI* contains possibly the first attempt at scientific explanation of the rainbow motivated by material world observation, situating Grosseteste among other pioneers such as Aristotle, Ibn al-Haytham, Kamal al-Din al-Farisi, Theodoric of Freiberg, and Descartes.

It is also interesting that while Grosseteste's theoretical mechanism of cone-light interaction can give rise to bow-shaped caustics as he claims, this is not universally applicable for all cones and all light angles. Rather, caustics form in the shape of a bow within a specific region of the parameter space mapped by varying cone aperture and light elevation. This too is somewhat addressed in the *DI*, where Grosseteste explores theoretically the relationship between the elevation of the sun and the extent to which a rainbow forms. It therefore captures something of the rarity of rainbows, which form in nature only when particular conditions are met.

While the aim of this work has been to explore Grosseteste's theoretical explanation for the shaping of the rainbow, doing so has revealed deep-running internal consistencies in his writings, particularly between the *DI* and his unique theory of colour, stated in his treatise titled *De colore* [*On Colour*] (*DC*). In this treatise, Grosseteste describes three pairs of opposing qualities, which together form a three-dimensional space containing all possible colours (Smithson et al., 2012). A difficulty in understanding the text, and analysing Grosseteste's colour theory in general, has been the precise translation of the words he uses for these qualities. One of these pairs, describing a quality of the light, is identified as *multa-pauca*. Grosseteste asserts in the *DC* that light is *multa* when it has been gathered together, such as by a burning glass. To a modern reader, in isolation of his other writings, this may seem to correlate simply with the intensity of light, and the question arises of how the gathering together of light might result in chromatic differences. Within the *DI*, Grosseteste states that

Published in *Applied Optics*

variations along the *multa-pauca* axis can account for the variations in colour seen within a single rainbow. If, as we suggest, Grosseteste had observed variations in colour arising from the gathering of light by transparent objects, this relationship between the term *multa* and variation in chromaticity would therefore seem to be an internal consistency in his understanding of colour.

Grosseteste's theory of rainbow formation is not correct, and does not advance our understanding of the rainbow as a natural phenomenon today. However, studying it reveals a movement toward a natural philosophy based on careful observation of the material world, the theoretical use of similarity in physics between different scales, and testable prediction. It would not be long until this emerging experimentalist approach discovered the correct mechanism of rainbow formation.

FUNDING

Arts and Humanities Research Council (AHRC) (AH/N001222/1); Andrew W. Mellon Foundation; Institute of Advanced Study, Durham University.

ACKNOWLEDGEMENT

We thank Nader El-Bizri for his correspondence and advice on Arabic science. The work presented here emerged from the collaborative Ordered Universe Project (http://ordered-universe.com/), which focuses on interdisciplinary readings of the scientific works of the remarkable English thinker Robert Grosseteste (c.1170–1253). The Ordered Universe Project is supported by the AHRC. Hannah Smithson held a Visiting Fellowship from the Institute of Advanced Study at Durham University from January to March 2017. In addition, Joshua Harvey holds a DPhil Studentship from the Andrew W. Mellon Foundation, hosted by the Oxford interdisciplinary research centre, the Oxford Research Centre in the Humanities (TORCH).

Published in *Applied Optics*

6

PARRY'S ARC AND GROSSETESTE'S *DE IRIDE*

A letter to be submitted to *Applied Optics*.

Authors

Joshua S. Harvey, Markus Selmke

Abstract

Parry's arc is a rare, though fascinating, atmospheric halo that may be seen above the sun under favourable conditions. A far more likely setting to encounter an analogous display is the dinner table, with light refracting through a conical glass filled with any clear liquid. Similarly, a rainbow may be observed if a cylindrical glass is illuminated appropriately. Here, we juxtapose two recent and independent publications to highlight a bigger picture we found connecting both: The intimate relationship between these two optical stages—the sky and the domestic table or the lab—may be a source of both inspiration and pitfalls when it comes to atmospheric optics, with Parry's arc being only the latest uncovered example.

Letter

When attempting to grasp the geometry at play in atmospheric optical phenomena—rainbows, halos and the like—it may be helpful to relate these mechanisms to analogies on a more familiar scale, such as glassware on a dining table, at least for the large subset of phenomena within the scope of geometrical optics. Of course, while dinner table caustics (in the loose sense of the term as projected patterns of light) may be easier to explain with a tangible experimental realisation at hand, one must not naively assume that they necessarily share in the mechanism of those seen in the sky. Most fundamentally there is a categorical difference between caustics seen projected from a single body onto a surface and those comprising numerous bodies, with each projecting a caustic, acting collectively to lighten or colour portions of the sky for an observer. Nonetheless, it has been established since the early fourteenth century that the mechanisms of both primary and secondary rainbows may be modelled with a single water-filled glass sphere, courtesy of experiments by both Kamal al-Din al-Farisi and Theodoric of Freiburg. Single-body caustics remain an important pedagogical tool in the classroom, and may be seen and appreciated frequently in everyday life. Here, we contrast two recent pieces of research that address the relationship between atmospheric optics and tabletop caustics, albeit from very different motivations.

For Selmke and Selmke's work (Selmke and Selmke, 2017), the inspiration came from a standard, simple classroom experiment. A cylindrical glass of water, with white light incident upon it through the top, gives rise to well defined arcs coloured by dispersion, often mistakenly likened to a rainbow caustic (the latter appearing for light entering through the lateral surface of the glass). Actually, this phenomenon has an analogous optical mechanism to the formation of the circumzenithal arc halo, depending on the angle of illumination. Moreover, a filled Martini glass—that is, a cone—was found to give analogous refractions to Parry's arc for light entering again through the cone's flat base, another colourful atmospheric phenomenon. This work

used a conventional experimental approach: observing a physical phenomenon, and from this observation building a mathematical model that can explain and predict the natural world. In this case, the focus was on the dining table stage.

For Harvey et al., the motivation for their work came from responding to Robert Grosseteste's thirteenth-century theory of rainbow formation, which posits that rays from the sun are bunched up into bows by pyramidal transparent bodies (clouds) in the heavens (Harvey et al., 2017b). That transparent cones do in fact give bow-shaped projections provides strong evidence that Grosseteste's theory of rainbow formation was not only the first using refraction, but was also the first based on observation, most probably of conical glassware illuminated on the domestic table stage. This research is therefore something of a mirror-image to Selmke and Selmke, 2017, starting from a model and working backwards to conclude whether or not direct observation had taken place. Most surprisingly, Grosseteste's proposed optical geometry of the rainbow is in fact analogous to the mechanism responsible for Parry's arc, as shown in Figure 6.1.

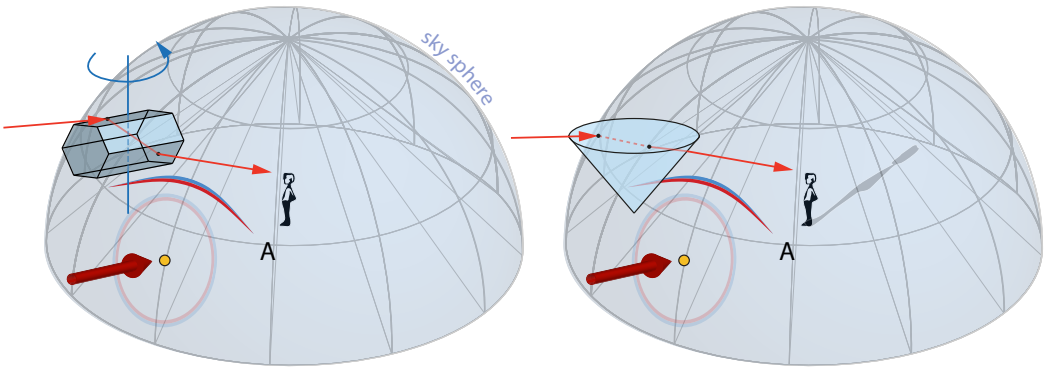


Figure 6.1: Refraction mechanism based on hexagonal columnar ice crystals responsible for the upper suncave Parry arc (A) (Greenler, 1980) and the equivalent refraction through a 60° apex angle cone (Grosseteste's conceived, and incorrect, mechanism for the rainbow, albeit with an unspecified apex angle).

Parry's arc is rare, and appears on the opposite side of the sky sphere from that of the rainbow (Parry's arc being concave towards the sun, the rainbow being a circle around the antisolar point) (Greenler, 1980). It would be wrong to imply that Grosseteste had knowingly found the mechanistic explanation for Parry's arc—he

almost certainly did not know of its existence, and it was not until 1820 that William Parry observed and depicted one. How then can we account for this coincidence?

Just as the classroom demonstration of the rainbow, using a cylindrical glass illuminated through the top air-water boundary, is in fact analogous to an altogether different atmospheric optical phenomenon (circumzenithal arcs), so too is Grosseteste's scheme. These coincidences, while perhaps unexpected, arise from the consideration of single-body optics with simple, rotationally symmetric geometries (cylinders and cones) that are effectively replicated in nature by hexagonal ice crystals in random orientations suspended in the atmosphere. Yet another mistake of this category is found in Huygens' rather successful, though physically incorrect, halo theory based on partially opaque ice cylinders (Tape and Moilanen, 2006). Both Grosseteste's cone, and the classroom cylinder, illustrate a principle of optics and more generally of science: a convergent result is not necessarily correct. This may initially provoke pessimism, or even cynicism, casting a shadow of doubt over scientific progress. But seen in another light, the complexity of the natural world, and our developing understanding through history, may inspire curiosity, research, and wonder.

ACKNOWLEDGEMENTS

The research into Robert Grosseteste's theory of the rainbow was carried out during workshops organised by the Ordered Universe Project, which is supported by the Arts and Humanities Research Council (AHRC) under grant AH/N001222/1. JSH holds a DPhil Studentship from the Andrew W. Mellon Foundation, hosted by the Oxford interdisciplinary research centre the Oxford Research Centre in the Humanities (TORCH).

Part III

PHONETICS AND SPEECH PERCEPTION

7

SPEECH PRODUCTION AND PERCEPTION

PART ABSTRACT

In this part, Grosseteste's theory of speech production and perception, as laid out in his treatise *De generatione sonorum* [*On the Generation of Sounds*] (*DGS*), is explored. First, the medieval *liberal arts* are introduced, and the categories of grammar and rhetoric are discussed with regards to the emerging science of phonetics. Chapter 8 follows, presenting a published dialogue between psychology and theology from an extended conference proceedings published in *Bulletin de Filosophie Medievale*. Then, in chapter 9, the interdisciplinary research into the *DGS* is presented, in a paper that explores the geometry, psychology, and phonetics of Grosseteste's treatise.

7.1 THE LIBERAL ARTS AND THE MEDIEVAL CURRICULUM

Medieval institutes of higher learning followed a strictly-defined curriculum, the *septem artes liberales* [seven liberal arts]. 'Art', in this sense, refers to the studies and practices by which the fallen, corrupted states of human thinking and behaviour are elevated to a heavenly perfection. The seven liberal arts were clustered into two groupings: the lower, literary, division—the *trivium*—of grammar, logic, and rhetoric, and the upper, numerical, division—the *quadrivium*—of arithmetic, geometry, music, and astronomy.

Grosseteste's views on the medieval curriculum are expressed in his treatise *De artibus liberalibus* [*On the Liberal Arts*] (*DAL*). The *DAL* is considered the earliest of Grosseteste's treatises, placed in the final decade of the twelfth century. He begins the work by explaining how the arts function as a mechanism for the correction of human error, through the lens of the Latin terms *aspecuts* and *affectus*. These terms refer to two inseparable yet distinct pieces of the soul (or perhaps, streams of perceptual-cognitive processing), the initial grasping of the mind's sight (*aspectus*) and the embracing or rejection of what is grasped, by the mind's desire (*affectus*). An excerpt from a published conference report, where the *aspectus-affectus* dynamic is discussed in relation to ideas in modern neuroscience, is found in chapter 8.

8

ASPECTUS AND *AFFECTUS*

An excerpt taken from a conference report published in *Bulletin de Philosophie Médiévale*[1].

AUTHORS

Timothy Farrant, Joshua S. Harvey[2]

ASPECTUS AND *AFFECTUS* AND SENSORY PERCEPTION

There are striking symmetries that become apparent when considering the different stages of sensory perception—from data acquisition through the bodily senses, through to perceptual experience—alongside the dialogue of *aspectus* and *affectus* and its considerations of the parts of the soul and their interrelations. A key motivator for dissections of *aspectus* and *affectus* was the desire to shape and perfect *aspectus* through the altering of *affectus*; an illustrative example being the zoological observations of Neckam which pervaded his *De naturis rerum* [On the Nature of Things]. Although the fabulous descriptions of animals and their features would not have been directly observed by readers of Neckam's didactic text, acquiring

1 The remainder of the report includes discussions of other conference presentations, which are not of direct relevance to this thesis.

2 The ideas and content presented here are my own, with the exception of specific references to the works of Hugh of Fouilloy, Hugh of St. Victor and Alexander Neckam.

Published in *Bulletin de Filosophie Medievale*

knowledge of them through study could still have a positive, moralising influence on the *aspectus* when such animals were encountered or meditated upon. For it is the condition of the *affectus* which may be responsible for error in the *aspectus*. With such an emphasis on the didactic function of twelfth-century zoological descriptions, and their transformative effects upon the soul, it is unsurprising that empirical accuracy almost always became a secondary concern. As Hugh of Fouilloy stressed to his readers at the commencement of his *Aviarium*, a textual description alone becomes a valid representation from which moral teaching and learning is discovered:

> Nec tantum volui columbam formando pingere sed etiam dictando describere, ut per scripturam demonstrem picturam; velut cui non placuerit simplicitas picturae, placeat saltem moralitas scripturae.

Likewise, faculties of sensory perception are profoundly shaped by low-level mechanisms that process sensory data, many of which are acquired and tuned during early development. For the case of vision (although these principles equally apply to other faculties, for example speech perception), the image formed on the retina is meaningless unless statistical features can be extracted from the sensation of sight. These range from the most fundamental features of intensity and colour variations, such as edges, lines and corners, up through to shapes of more complex and useful spatial information, which may then be identified as objects, faces, *et cetera*. We might describe this as a microscale scheme of the *aspectus-affectus* relationship, compared to the macroscale relationship at work in higher, abstract cognitive processes such as zoology. This comparative arrangement is summarised in Figure 8.1.

A number of interesting parallels may then be drawn. First, for both microscale and macroscale cases, faults in *affectus* may be to blame for an error-prone *aspectus*. For Neckam, without a grounding in theology and moral knowledge the soul will fall short of perfection in its mental processing – a rule coined by Hugh of St. Victor in the simple phrase: *mores ornant scientiam*. For considerations of the microscale, we need only look to cognitive disorders such as prosopagnosia, or 'face blindness'. Despite

Published in *Bulletin de Filosophie Medievale*

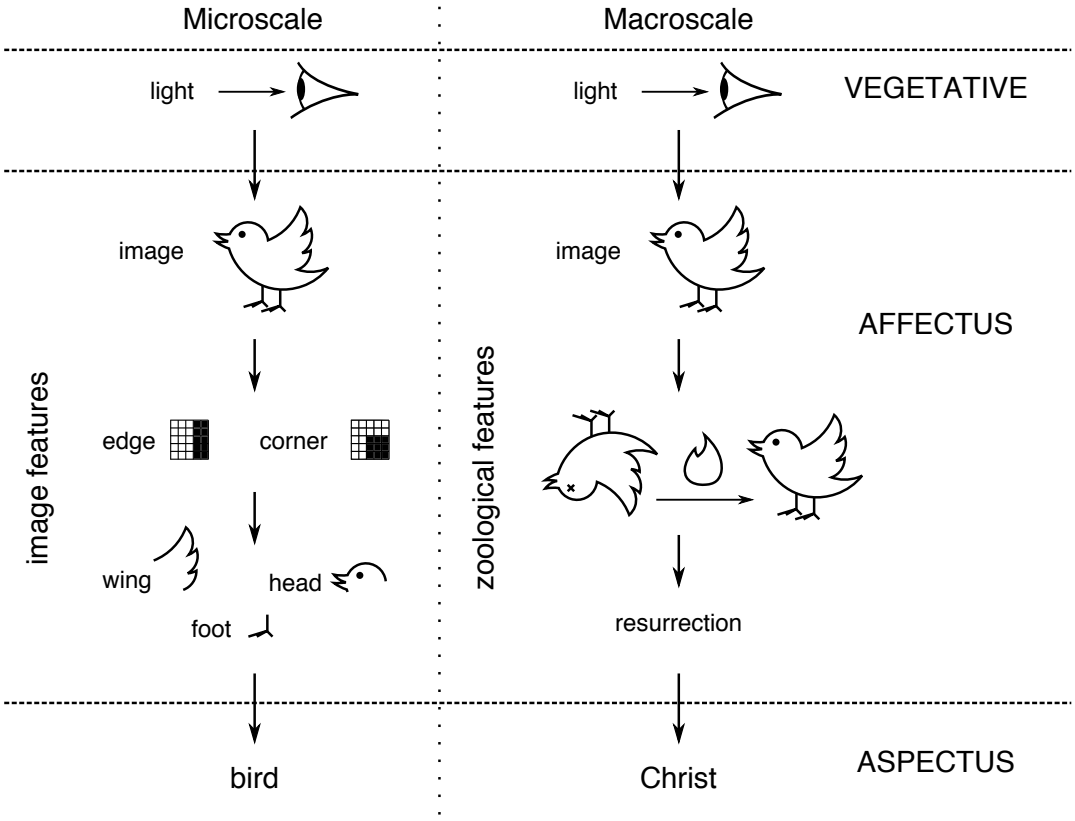


Figure 8.1: A graphical comparison of the 'microscale' and 'macroscale' *aspectus-affectus* relationships at play in sensory perception and theological learning, respectively.

the visual system being otherwise unaffected, individuals with the condition are unable to recognise the identity of even the most familiar faces, including their own. This may be caused by injury to regions of the brain shown, through neuroimaging, to be structural correlates for face perception (such as the fusiform face area, occipital face area, and the posterior superior temporal sulcus).

Second, the interactions of the *aspectus* and *affectus* can take place without an objective stimulus external to the soul or mind. For the microscale, both memory and imagination are plainly able to bring about internal cognitive states corresponding to those when a stimulus is directly encountered. Interestingly, however, there is evidence from neuroimaging that memory is not simply reducible to a retrieval process, but is additionally constructive. For example, during retrieval of visual memories of particular shapes, there is evidence that cortical regions associated with processing feature-specific information are reactivated. As said before for the

Published in *Bulletin de Filosophie Mediewale*

writings of Neckam, imagination or meditation focused on a moralised zoological feature may be equally corrective to the *affectus* as a direct encounter with the animal in nature. What mattered most was not the grounding of the textual description in reality, but the transformative power of the moral, theological truth it articulated.

Third, we may observe critically that both microscale and macroscale schemes of perception and cognition are engineered toward some purpose or goal, and contingent upon them. Human perceptual potential has been developed to confer evolutionary advantage, primarily as a species and secondarily as individuals. While this has endowed us with a broad range of capabilities, we additionally have limitations on our senses and cognition to which other animals may not be confined. Neckam's end goal is the perfection of the soul in knowledge of and harmony with God and creation. Indeed, as the hugely influential Hugh of St. Victor stressed in his *Didascalicon*, a love of Divine Wisdom lay at the heart of all scholastic pursuits, whether they were logical or ornithological. When seen in such a light, medieval zoological writings that do not feature direct observation of the natural world cannot be dismissed as false or simply unscientific to a modern reader; this would be akin to chastising an individual for being 'blind' to ultraviolet radiation. Though this position was obvious to the medieval intellectual, the undeniable presence of recent historiography which dismisses the fabulous, in favour of the scientifically empirical, continues to rear its anachronistic head.

Finally, it can be noted that neither the microscale nor the macroscale frameworks of understanding considered here are engineered toward direct comprehension (or, indeed, experience) of an objective reality external to the mind, instead being geared towards biological survival and theological knowledge, respectively. One immediate consequence of this is the potential for disjuncture between what might be reasonably understood about an external stimulus, and the mind's direct perception of it. While this is clearly the case in disorders of the mind and states of hallucination where sensory frameworks are further altered, an illustrative case at the microscale is that of simple optical illusions, which might provoke perceptions of motion, colour, or geometry known to be in contradiction to the properties of the stimulus as measured

Published in *Bulletin de Filosophie Medievale*

external to the mind. Similarly, a student of Neckam's zoological writings might experience consternation upon witnessing the burning of a bird said to be a phoenix, which nonetheless did not rise from the ashes. This is especially so since Neckam's *De naturis rerum* included an additional chapter on the phoenix, *Item de phoenice*, stressing the typological resurrection of the bird from its own ashes (comparable to the regeneration of the eagle, which was also made popular by Neckam, Gerald of Wales and Hugh of Fouilloy). In both cases, particular states of the *affectus* have been conditioned without directly encountering the external stimulus. Upon this encounter, subsequent modification of prior beliefs is required; not everything that appears to move does in fact move (for the case of optical illusions), that was not a phoenix, or phoenixes do not rise from the ashes, in which case it is less likely that Christ was resurrected (for the case of the burning of the bird).

Published in *Bulletin de Filosophie Medievale*

9

A CATEGORICAL THEORY OF VOWELS FROM *DE GENERATIONE SONORUM*

Published in *The Journal of the Acoustical Society of America*.

AUTHORS

Joshua S. Harvey, Hannah E. Smithson, Clive R. Siviour, Giles E. M. Gasper, Sigbjørn O. Sønesyn, Tom C. B. McLeish, and David M. Howard

ABSTRACT

In this historical paper we examine a pioneering theory of speech production and perception from the thirteenth century. Robert Grosseteste (c. 1175—1253) was a celebrated medieval thinker, who developed an impressive corpus of treatises on the natural world. Here we look at his treatise on sound and phonetics, *De generatione sonorum* [*On the Generation of Sounds*] (*DGS*). Through interdisciplinary analysis of the text, we find a theory of vowel production and perception that is notably mathematical, with a formulation of vowel space rooted in combinatorics. Specifically, Grosseteste constructs a categorical space comprising three fundamental types of movements pertaining to the vocal apparatus: linear, circular, and dilational-constrictional; these correspond to similarity transformations of translation, rotation,

Published in *The Journal of the Acoustical Society of America*

and uniform scaling, respectively. That Grosseteste's space is categorical, and low-dimensional, is remarkable vis-a-vis current theories of phoneme perception. As well as his description of vowel space, Grosseteste also sets out a hypothetical framework of multisensory integration, uniting the production, perception, and representation in writing of vowels with a set of geometric figures associated with 'mental images'. This has clear resonances with contemporary studies of motor facilitation during speech perception and audiovisual speech. We additionally provide an experimental foray, illustrating the coherence of mathematical and scientific thinking underpinning this early theory.

9.1 INTRODUCTION

This paper explores and responds to a historical theory pertaining to the psychology and physiology of speech. This theory was developed in the early thirteenth-century, but within it may be found many of the same considerations as those of modern neuroscience—the nature of mental representations, the relationship between those representations and external stimuli, and correspondences between the sensory faculties. Examining this theory, from such a contrasting intellectual context to our own, raises questions of the role of experimentation, observation, and modelling, and what constitutes permissible evidence for supporting or rejecting hypotheses.

Robert Grosseteste (c.1175–1253) was a celebrated medieval thinker, who, as well as writing on philosophy and theology, developed an impressive corpus of treatises on the natural world. Here, we analyse one of these treatises—his text on sound and phonetics: *De generatione sonorum* [*On the Generation of Sounds*] (*DGS*). The *DGS* was probably written in the first decade of the thirteenth century, several centuries before the apparent 'scientific revolution' in Early Modern Europe. It was a formative period, however, for the development of European scientific thought, during which the reception of Greek natural philosophy, enabled by their transmission, translations, and commentary from Arabic and Greek into Latin, prompted new conceptual frameworks

Published in *The Journal of the Acoustical Society of America*

for the consideration of natural phenomena (Burnett, 1996; Burnett et al., 2013; De Leemans, 2010). For modern science, reading medieval works presents several significant challenges, starting not least with that of editions and translations. This analysis of the *DGS* has only been possible through interdisciplinary collaboration between science and humanities scholars, resulting in the compilation of a new critical edition and translation of the text[1] (Gasper et al., in press 2019)[2].

Previous interdisciplinary research has already explored other scientific treatises written by Grosseteste: the *De colore* [*On Colour*] (*DC*) (Smithson et al., 2012), *De iride* [*On the Rainbow*] (*DI*) (Harvey et al., 2017b; Smithson et al., 2014a), and *De luce* [*On Light*] (*DL*) (Bower et al., 2014). In the *DC*, Grosseteste develops a pioneering application of mathematics to psychology. Within the space of approximately 400 words, he claims that colour occupies a continuous, three-dimensional space, contrary to the prevailing one-dimensional theory of the time (Smithson et al., 2012). It is surprising to find this theory articulated six centuries before three-colour printing techniques were established (Le Blon, 1720) and trichromacy was formulated by Thomas Young (Young Thomas, 1802). In the *DGS*, the treatise we explore and respond to in this paper, Grosseteste attempts a similarly mathematical, combinatorial abstraction for phonetics—specifically for vowels—as he attempts for colour. Several features of how he goes about doing this are of interest to the modern reader. Whereas Grosseteste's colour space is explicitly continuous, the vowel space described in the *DGS* is explicitly categorical. Underpinning his theory is a multimodal framework identifying correspondences between the mental representation of vowels, their physical production, their perception, and their external representation as letter shapes. Within this framework, the correspondences between speech perception, letter perception, and shape perception, have particular modern resonances in audiovisual speech and involvement of the motor system during speech perception. In the second half of this paper we present an experimental interpretation of the text, using artificial

1 This new critical edition and translation incorporates three manuscripts not known by the last editor, Baur, 1912.

2 The critical edition and English translation of the *DGS* which form chapter 11 of this volume was provided by S. O. Sønnesyn.

Published in *The Journal of the Acoustical Society of America*

vowel synthesis and psychophysics to test the claims of correspondence between abstract, geometric acoustic chamber shapes and vowel perception.

Before presenting a detailed discussion of the *DGS*, a question that might first be addressed is why one ought to concern themselves with medieval science. Modern neuroscience is already at an interdisciplinary juncture between psychology, physiology, biology and mathematics; why should matters be further complicated with the inclusion of medieval history and Latin? An answer may be found in the sheer wealth of scientific theory and observation that was amassed during this period, which largely remains untapped. The history of science is highly non-linear, despite its frequently linear presentation, leaving worthwhile questions and suggestions unresolved in every historical age (Chang, 2017). Psychological phenomena such as the perception of speech are not new, and have been prompting rational discourse throughout many historical and geographical cultures. By engaging with these theories today, we may find unexpected agreement with, or perspectives that are strikingly different to, our own. In either case, we stand to gain much from the exercise.

9.2 ROBERT GROSSETESTE'S *DE GENERATIONE SONORUM*

The *DGS* begins with a physical description of vibrational mechanics: a sounding body is such that when struck, its smallest parts move away from, return towards, and overshoot their natural places, with vibrations occurring as a result. This is to be expected from the given title of the treatise. However, only a quarter of the way through the text there is a change of focus, as Grosseteste presents a case study of a particular sounding body, that is, the production of human speech:

> And since there is no such movement continuously in beings that have a soul, such movement cannot come from a vegetative soul, but from a sentient motive force and in a voluntary movement, which by necessity is preceded by the making of a mental image or by apprehension. Therefore,

Published in *The Journal of the Acoustical Society of America*

> a sound formed by a primary motive force in which there is an ability to form mental images is a voice.

The remainder of the treatise is an attempt to characterise those 'mental images' that initiate the voice, and the relationships between mental representations of origination, the physical gestures of the vocal tract, the acoustic qualities of vowels, and the movements of the hand that draw out letters to represent speech sounds[3]. Immediately following on from the above passage, Grosseteste demarcates the difference between an intelligible and an unintelligible speech sound:

> But the actualising shaping itself of the vocal instruments and the shaping of the movement of breaths able to move the vocal instruments gives to a certain voice its kind and perfection; to a certain other voice, however, such shaping does not give perfection. The voice, therefore, to which the aforementioned shaping gives outward appearance and perfection, will be [called] a lettered voice. And the voice that is completed by a single shape will be a letter. The voice that is completed by several shapes will be composed of letters.

Here, Grosseteste establishes a direct relationship between the shapes—or as they may be understood, *figures*—of mental images, vocal tract shapes, and the movements of the breath during speech. These three figures, when perfected, give rise to a 'lettered voice', i.e. an acoustic output of intelligible speech. Grosseteste does not yet describe these figures geometrically, though that will come in the next section of the treatise. It is interesting to note the particular emphasis on the natures of certain voices due to the 'actualising shaping itself of the vocal instruments'; any voice is preceded by a mental image, but the intelligibility of that voice additionally depends on the speaker's ability to precisely execute the required motor programs. Or, to further unpack this notion, the acquisition of speech requires first the presence

3 Although Grosseteste only refers explicitly to the shapes of letters, when read in the broader context of the *DGS* and its focus on motion, a valid interpretation is that the movements of handwriting played a central role in his thinking.

Published in *The Journal of the Acoustical Society of America*

of mental representations for speech sounds (it is unclear whether Grosseteste is of the opinion these are innate or acquired), and second the learning of distinct motor programs encoding muscular coordinations for the production of these speech sounds. While Grosseteste does not explicitly describe this in terms of language acquisition, and the development from an imperfect to perfected voice, it is heavily implied when understood in the broader medieval context of discussions on the liberal arts. The seven liberal arts—and in this case the first art, that of grammar—provide a means whereby the fallen and corruptible things of the world may be refined and perfected through study and practice. In this case, the notion of a 'perfect' or completed voice is related to the art, and study, of grammar, and the acquisition of vocal tract coordinations that give rise to a 'lettered' voice, i.e. intelligible speech[4].

In isolation, it may seem from this passage that Grosseteste understands that both diaphragmatic breath control ('shaping of the movement of breaths') and muscular coordination of articulators ('shaping of the vocal instruments') are required to produce intelligible speech sounds. However, he later makes clear that he is instead claiming a direct identity between control of the vocal apparatus and the resultant movements of the ('motive') breath, and it is these motive breath shapings that determine the 'outward appearance and perfection' of a voice. Writing six hundred years before Fourier and modern notions of frequency, resonance, and spectral analysis, this provided a sensible hypothesis for the causal relationship between the shape of the vocal tract and the acoustic qualities of the generated sound.

Grosseteste then moves beyond the production of speech (the shaping of the vocal instruments and motive breaths) and its perception (its outward appearance) to the visual representation of speech in writing, and in doing so provides further discussion on the nature of these fundamental geometric figures:

> The voice's capacity for being written down, therefore, is nothing other than this same shaping of the vocal instruments and of the breaths by which the letter is generated internally. It may therefore be represented by

4 A comprehensive discussion of the seven liberal arts, and Grosseteste's treatise on them, may be found in Gasper et al., 2019.

Published in *The Journal of the Acoustical Society of America*

> a visible shape similar to the shape of its generation. It is clear, moreover, that, since art imitates nature and nature always acts in the best possible way, and art does similarly when not in error; however, representation by exterior shapes assimilated to interior will be better than [representation done] otherwise: to write is, according to the art of grammar, to represent interior shapes by means of exterior shapes similar to these same interior shapes.

Here Grosseteste is guided by two Aristotelian principles: first, that 'art imitates nature', or mimesis; and second, that nature always acts in the best possible way. There is clear indication of his reading Aristotle's *De anima* [On the Soul] (Aristotle, 1984a), although Grosseteste does not reference Aristotle directly, as he does in some other scientific works[5]. These principles motivate one of the most central and clearly articulated claims of the treatise: the capacity for speech to be written lies in the visual representation of shapes similar to the geometric figures (mental, gestural, and of the 'motive breaths') at play during speech production which is summarised in Figure 9.1. This claim that 'representation by exterior shapes assimilated to interior will be better than otherwise' is particularly interesting, and has strong resonances with recently resurfacing theories of non-arbitrary representation, or 'iconicity' (Schmidtke et al., 2014; Turoman and Styles, 2017).

For many languages today, including modern English, such a direct relationship between speech-sound (phoneme) and written letter (grapheme) would be impossible; individual letters have diverse pronunciations in differing lexical contexts, themselves quite different to the letter name. As an example of phonological inconsistency, while an English speaker with received pronunciation today may read the letter 'O' as a diphthong /əʊ/, it could be similarly pronounced as /əʊ/ in 'go', but also as /u/ in 'do', /ʌ/ in 'tonne', /ʊ/ in 'woman' and even /ɪ/ in 'women'. This complication was not known to Grosseteste, who saw a mostly direct and consistent grapheme-phoneme relationship in the languages it is likely that he knew (Middle English,

5 A deeper dive into the intellectual influences for Grosseteste's *DGS* can be found in (Sønnesyn and Gasper, 2019).

Published in *The Journal of the Acoustical Society of America*

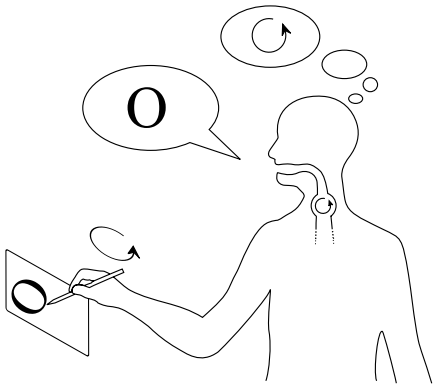


Figure 9.1: A diagrammatic depiction of one of the claims in Grosseteste's *DGS*. Grosseteste claims that the capacity for speech to be written lies in the visual representation of shapes similar to the geometric figures (mental, gestural, and of the 'motive breaths') at play during the production of speech. Because 'art imitates nature', the representational potential of letter shapes is maximized when those letters display geometric features common to the geometric figures at play in a vowel's production.

Latin, and French). Any exceptions, such as variations in regional accents, could be accounted for as being 'accidental'.

The treatise then gives a special consideration of vowels, for which Grosseteste provides a comprehensive study of his hypothesised geometric figures.

> The whole sound of the vowel and of any part of the vowel are the same as each other. It is necessary, therefore, for it to be generated by a movement the parts of which are the same as the whole. But there are seven movements in which the parts are the same as the whole: straight movement, circular movement, dilation and constriction—these last two do not differ except as straight movement forwards and backwards—, circular movement over a centre in a straight movement and a circular movement over a centre in a circular movement, and likewise dilating and constricting movement over a centre in a straight movement and over a centre in a circular movement.

In fact, this is a combinatorial system related to that described in the *DC*: three simple elements are combined in various ways to give rise to a full set including complex combinations, except that for this scheme only two simple elements may be

Published in *The Journal of the Acoustical Society of America*

combined rather than all three. It is also different in that, rather than being defined by independent dimensions as in the case of the bipolar qualities of colours, only some of the simple elements may be combined, and one—circular movement—may be self-combined. The choice of three simple movements may not appear such an obvious choice, and it may be even more puzzling why only one of the three may be self-combined. Grosseteste states clearly that this is the comprehensive list of movements 'in which the parts are the same as the whole'. We may rephrase this description as one of time-invariant functions on position.

One way of interpreting the scheme that seems to resolve these confusions is by viewing the three classes of simple movements as geometric linear transformations. In which case, these movements correspond perfectly to the allowed operations for Euclidean similarity transformations: straight movement for translation, circular movement for rotation, and dilational movement (and constrictional) as uniform scaling. Matrix notation provides a convenient and efficient way of describing these transformations; while Grosseteste would not have had this notation at his disposal, imagining these movements *per se* is not contingent on any particular form of mathematical description. Expressed as two-dimensional transformation matrices of translation, rotation, and scaling—$\mathbf{A_t}$, $\mathbf{A_r}$, and $\mathbf{A_s}$, respectively—these three simple geometric transformations are given as:

$$\text{Translation} : \mathbf{A_t} = \begin{bmatrix} 1 & 0 & t \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}; \quad \text{Rotation} : \mathbf{A_r} = \begin{bmatrix} \cos(t) & \sin(t) & 0 \\ -\sin(t) & \cos(t) & 0 \\ 0 & 0 & 1 \end{bmatrix};$$

$$\text{Scaling} : \mathbf{A_s} = \begin{bmatrix} t & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Published in *The Journal of the Acoustical Society of America*

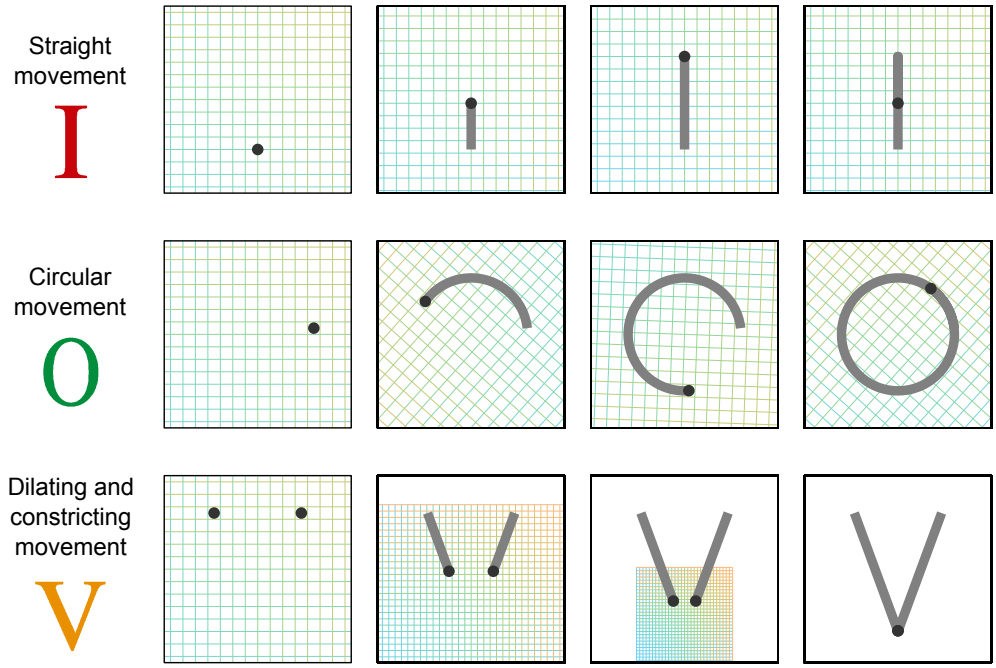


Figure 9.2: The simple, self-similar geometric movements that Grosseteste describes as the basis for vowel categorisation. We have interpreted his categories of simple movements—straight movement, circular movement, and dilating and constricting movement—as the three fundamental classes of linear geometric transformation: translation, rotation, and uniform scaling. Points (shown in black) embedded in planes undergoing these transformations trace out movements that agree well with Grosseteste's descriptions of simple movements, shown in grey. Videos are provided in the online version of this paper.

Using this interpretive scheme, the geometric figures which Grosseteste describes naturally arise by the consideration of points in Euclidean space experiencing these transformations. These simple and combined movements may be visualised in Figure 9.2 and Figure 9.3, respectively.

This interpretation also accounts for why straight movement does not give rise to a distinct movement when self-combined. The result of two translation transformations, $\mathbf{A}_t^2\mathbf{A}_t^1$, is simply another (different) translation, $\mathbf{A}_t^3$, even when the coordinate systems of the two transformations are allowed to differ. The same can be said for two consecutive or simultaneous operations of scaling, or of dilational-constrictional movement. Circular movements can, however, be self-combined to give a new class of self-similar movement, as in Figure 9.4. The combination of circular movement over another circular movement strongly connotes the epicyclic approach employed

Published in *The Journal of the Acoustical Society of America*

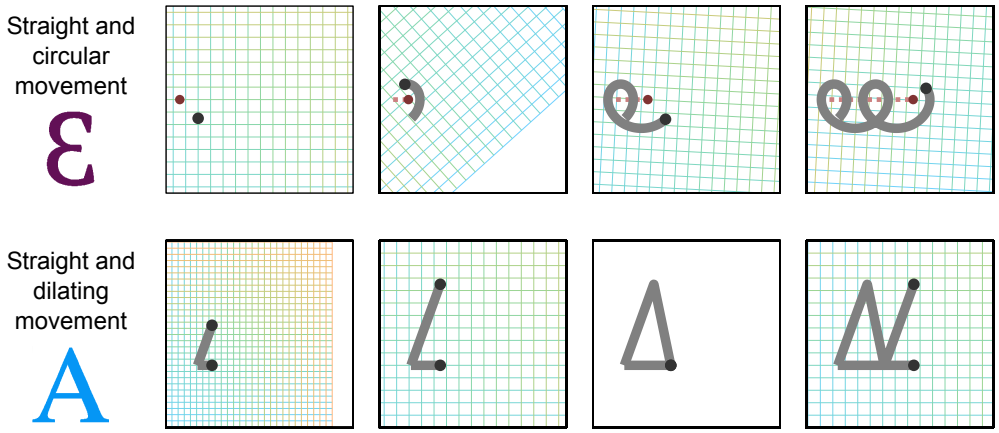


Figure 9.3: The combined movements that give rise to vowels in Grosseteste's model of phonetics. For the combination of straight and circular movement, the translating origin of rotation is indicated by a red dot. For the combination of straight movement with dilating and constricting movement, two dots repeatedly expand from, and collapse to, a single point that itself undergoes translation. Circular movement, or rotation, can be self-combined mathematically, as shown in Figure 9.4, but Grosseteste discounts it for vowel production as overly complex for the speaker. Videos are provided in the online version of this paper.

in classical and medieval astronomy, which comprises highly organised structures of rotating, nested spheres. In this case it is clear that an additional rotational transformation is applied to the space experiencing the first rotational transformation, but the centre of this rotation is at a point offset from the origin, itself experiencing rotation. In terms of linear transformations, the two transformations operate on the same space but under different systems of coordinates[6]; in the case of epicyclic astronomy, these spaces are volumes bounded by the spheres. What first appears as an arbitrary selection of movements, in fact constitutes the complete scheme of self-similar, geometric similarity transformations of the two-dimensional plane, such that points in this plane trace out movements. However, to limit the number of vowels from seven to five ('A', 'E', 'I', 'O' and 'U'), Grosseteste discounts complex movements over a point itself tracing a circular movement—circular movements and dilational-constrictional movements over a centre already experiencing circular movement are infeasibly difficult:

6 Equally, a single coordinate system could be maintained, with translation operations (and their inverse) effectively providing a second origin of rotation, i.e. $\mathbf{A}_\mathbf{t}^{-1}\mathbf{A}_\mathbf{r}^2\mathbf{A}_\mathbf{t}^1\mathbf{A}_\mathbf{r}^1$.

Published in *The Journal of the Acoustical Society of America*

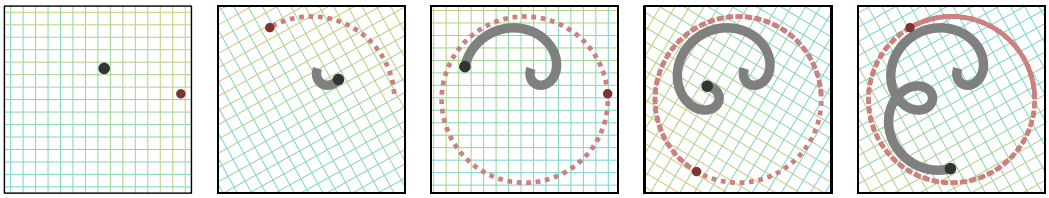


Figure 9.4: Grosseteste describes a self-combination of circular movement, which he discounts as too complex for use in speech. This movement strongly evokes the mathematical constructions of epicycles in medieval astronomy. Here, the rotating origin of rotation is indicated with a red dot. Videos are provided in the online version of this paper.

> On account of these seven movements the ancient Greeks posited seven vowels. But the abovementioned two movements over a centre in circular movement, granted that they are possible in imagination, are nevertheless difficult in reality. For this reason, there only remain five movements that are possible or easy to produce.

He then gives an in-depth geometric description of the remaining five self-similar movements, and how they generate the letters that represent their corresponding vowels:

> It is therefore clear that in a straight movement of the motive breathings through the vocal tract an 'I' is shaped. But this straight movement is not a single continuous movement – for then the lack of interruption would not cause a vibration – but is very frequently coming and going. A circular movement over a centre makes the shape 'O'. A circular movement over a centre [moved] in straight movement subtends a chord by the movement of the centre, and, by the movement of any point of the circumference, describes an arc over the chord and thus makes the shape 'E'. A constricting and dilating movement, on the other hand, makes the figure 'V', that is, two lines running together in a centre. And a dilating and constricting movement over a centre moved straight in a straight movement subtends the base of a triangle. And any point, when there is dilation, because it is moved by a double movement, describes one side of the triangle from the base to the top, and when there is constriction, it

Published in *The Journal of the Acoustical Society of America*

describes the remaining side from the top to the base, and thus it makes the figure 'A'.

As shown in Figures 9.2 and 9.3, these descriptions align well with a linear transformation interpretation of movement schemes. All five of the figures that Grosseteste traces out in words can indeed be traced out by points or combinations of points embedded in the plane experiencing the simple or combined similarity transformations of translation, rotation, and or uniform scaling.

As made clear by these descriptions, the abstract figures that correspond to phonemes (and, on account of the art of grammar imitating nature, graphemes) are not static geometric shapes, but rather categories of movement, which are ascribed to the vocal tract during speech. Therefore, for Grosseteste the perception of a speech sound, whether in hearing speech or in reading, is intrinsically connected with vocal gestures, and the 'mental images' that encode their associated motor programs. This multisensory framework readily lends itself to current discussions of the motor theory of speech (Galantucci et al., 2006), and involvement of the motor cortex in speech perception.

Eight centuries after Grosseteste was writing, we now have experimental evidence from brain imaging and transcranial stimulation that his intuitions were solid. Involvement of the motor system was established fifteen years ago in response to visual and auditory speech perception (Watkins et al., 2003), and soon after, that specific motor circuits in the precentral gyrus are recruited to facilitate phoneme identification—serving as 'speech-sound-specific neuronal substrates' shared across the sensory and motor processes (Pulvermuller et al., 2006). Motor cortex involvement has been found to be beneficial for speech perception under noisy conditions (Meister et al., 2007), and possibly under normal listening conditions (Sato et al., 2009) (although possibly not (D'Ausilio et al., 2012)). Of particular relevance to Grosseteste's theory, Möttönen and Watkins found direct evidence for motor representations playing a complementary role in the categorisation of speech sounds when they are found along continua (Möttönen and Watkins, 2009). As they point out, the mapping

Published in *The Journal of the Acoustical Society of America*

of highly variable acoustic signals onto discrete motor representations could support the intelligibility of speech in challenging environments. Even more intriguingly, Tian and Poeppel proposed a common sequential estimation mechanism underpinning both the quasi-perceptual experience of articulator movement and the corresponding auditory percept of speech mental imagery (Tian and Poeppel, 2012). They claim that the experimental evidence from both task demands and stimulus properties demonstrates the top-down role the motor system is playing in this type of mental imagery. In which case, Grosseteste's claim that the mental imagery of speech is in fact a mental representation not of sound, but of motion (albeit of a simple, geometric nature), was remarkably apt.

In the light of these recent investigations, we can again consider Grosseteste's approach to understanding speech. Acoustic signals show enormous variety, and to the thirteenth-century researcher writing before the advent of spectral analysis, this would have proved impossible to organise. Confronted with the curse of dimensionality, Grosseteste limits his study of sound to that of speech—a subset of natural sounds that the human auditory system can reliably organise, doing so in a categorical manner. Aristotelian principles, the scientific paradigm of the day, provide the methodological approach, with the movements of the hand during writing perhaps constituting a permissible form of evidence for understanding the mental and anatomical origins of speech, and its perception. That speech sounds differ due to differences in movement category sits well with what Grosseteste understands about the vibrational mechanics of sound; sound is the perception of a special class of movements made by physical bodies, either when struck (the sounding body) or when formed by a primary motive force capable of forming mental images (the voice).

Published in *The Journal of the Acoustical Society of America*

9.3 A PSYCHOPHYSICAL INTERPRETATION OF THE TEXT

The claims in the *DGS* are bold, and may read today as 'unscientific', lacking any evidential basis. But before dismissing these claims out of hand, it is worth considering exactly what evidence would have been available at the time to a shrewd observer. The morphology of the vocal tract would largely have been unknown, although from the end of the twelfth century very good diagrams of the vocal tract and its articulators were being produced in the Arabic-speaking world (Bakalla, 1982). These would not have been accessible to Grosseteste, and we can reasonably say that any data he had regarding vocal tract morphology would have come from his own direct experience of vision and proprioception. As has been remarked by others, the resemblance of the 'O' letter shape and the pronounced rounding of the lips when producing the /ɔ/ phoneme may suggest a non-arbitrary grapheme-phoneme relationship (Allott, 2000), and could have been a motivating factor for the theory as a whole.

To experimentally determine whether Grosseteste's theory could have been constructed in a way commensurable with the available evidence, we created a set of synthetic vowels, using physical models of vocal tracts. These models were designed to incorporate the geometric figures Grosseteste identified at the front of the mouth end of the tract. This is, categorically, not to refute or accept the theory expounded in the *DGS*; we have ample data on the morphology of the vocal tract, and nowhere does it feature idealised geometric shapes as described in the *DGS*. However, in this manner we are able to evaluate whether Grosseteste's theory would have been consistent with the observational data available to him—the visual and proprioceptive measurements of the mouth and lips. The question is, therefore, not whether the theory is correct, but the following: can we construct acoustic chambers that incorporate Grosseteste's ideal geometric figures at the 'mouth end' (the end furthest form the acoustic source), and yet are perceived as the five vowels in question? We tested

Published in *The Journal of the Acoustical Society of America*

this using established methodologies of phonetics and speech perception, namely, spectral analysis, and both multidimensional scaling and classification experiments.

9.3.1 *Stimuli*

Synthetic vowels were produced by plate-type model vocal tracts, constructed to resemble the five geometric figures Grosseteste describes at the mouth end. This is a one-dimensional model developed by Arai *et al.* (Arai et al., 2001), comprising 75 mm wide acrylic squares, each 10 mm thick, with central holes of different diameters. The plates are clamped together in a specified order, leaving a central cavity of varying size down the length of the tract. A rubber coupler allows the introduction of an electrolarynx to acoustically stimulate the model at the laryngeal end, which produces a falling pitch excitation in the male range from 100 Hz to 60 Hz lasting around two seconds. Adjustments were made to the laryngeal end of the models such that the output best approximated the associated phoneme. The resultant plates are shown in Figure 9.5, which also includes an overlay in red of the region made to resemble the geometric shape for each vowel. The acoustic outputs of these vocal tract models were then analysed acoustically (formant analysis) and perceptually (two psychophysical listening tests), to evaluate how successfully the synthetic speech-sounds approximate natural vowels.

9.3.2 *Formant analysis*

Spectrograms for each sample were generated with a Hamming window of 20 ms, as shown in Figure 9.6, Upper Panel. The Lower Panel shows smoothed spectral slices calculated as the mean of each spectrogram across time. Difference between these synthesised stimuli and natural vowels are the shape of the acrylic plates vs the speaker's vocal tract—which is our primary interest—and the acoustic excitation (electrolarynx vs a speaker's larynx). The electrolarynx for the Arai tubes provides

Published in *The Journal of the Acoustical Society of America*

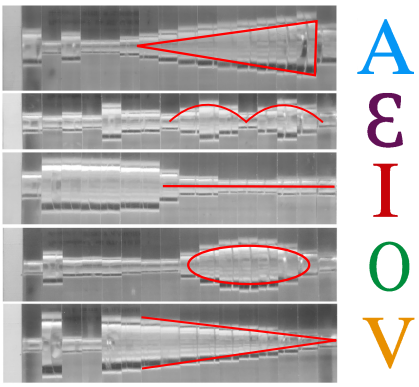


Figure 9.5: The configurations of the plate-type vocal tract model (VMT-10) of Arai *et al.* (Arai et al., 2001) used to synthesise the five samples corresponding to Grosseteste's geometric figure associations for each of the five vowel letters, with the mouth-end on the right. From top to bottom: A, $\mathcal{E}$, I, O and V. The models are overlaid with the geometric shapes inferred from Grosseteste's descriptions.

a signal that has a constant spectrum whereas the output from the vibrating vocal folds of the speaker vary as a function of the airflow loading owing to the shape of the vowel being uttered, sub-glottal lung air pressure through breath control and the nature of the voice quality being employed and any pitch variation.

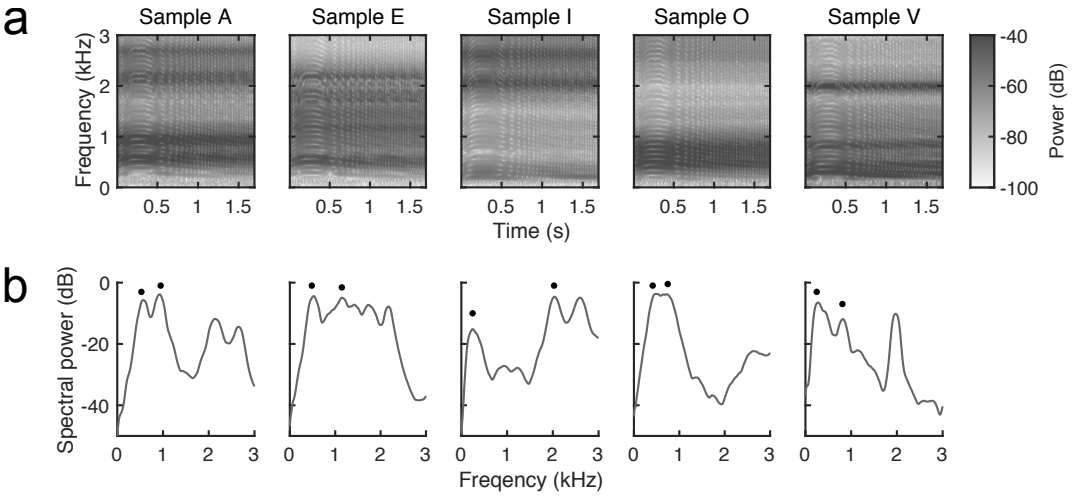


Figure 9.6: (a) Spectrograms produced from each of the five synthesised samples. (b) Spectral slices given by the mean of each spectrogram across time for each sample, from which the frequencies of the first two formant peaks,F_1 and F_2, were taken (indicated by black dots).

The horizontal dark bands in the spectrograms show formants (peaks in spectral power) that result from filtering the input acoustic excitation of the electrolarynx by the passive acoustic resonances of the chambers. The primary acoustic features of

Published in *The Journal of the Acoustical Society of America*

vowels are the locations in frequency space of their two lowest-frequency formants, F_1 and F_2. When, for different vowels, F_1 is plotted on the ordinate and F_2 is plotted on the abscissa, the vowel quadrilateral results, and different vowels plot in well-separated regions of this acoustic space (Catford, 2001 p. 161). A vowel quadrilateral for the synthetic vowels produced via the plate-type model is shown in Figure 9.7. This plot confirms that the acoustic properties of the synthetic samples are broadly consistent with the patterns of formants of natural vowels documented in the prior literature, with all samples falling within the quadrilateral. Additionally, the samples locate to disparate regions of the quadrilateral, suggesting they may be perceived as separable vowels.

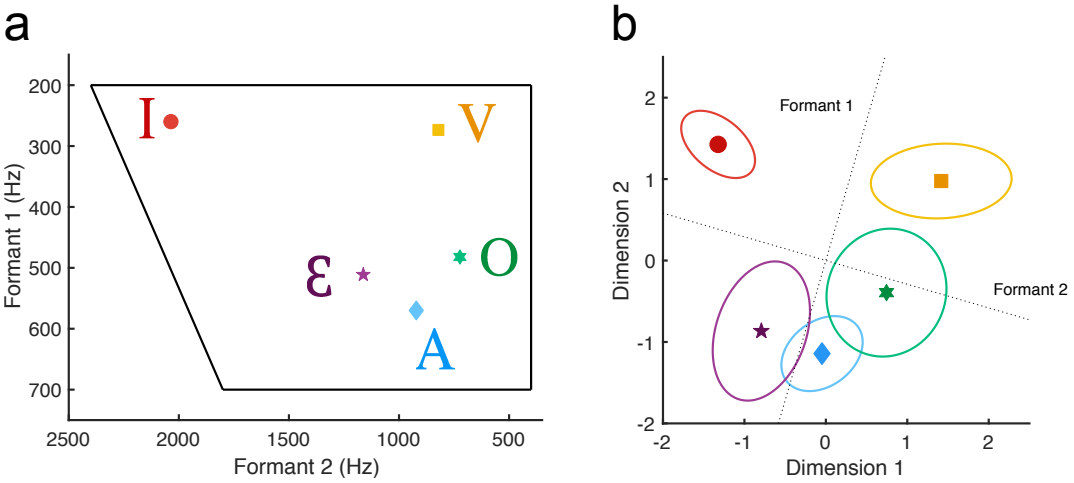


Figure 9.7: (a) Acoustic map of the recorded synthetic vowels based on their measured first and second formant (F_1 and F_2) frequencies. The quadrilateral indicates the area within which discernible vowels are expected from previous literature (Catford, 2001). Blue diamond = sample A, purple pentagram = sample $\mathcal{E}$, red circle = sample I, green hexagram = sample O, orange square = sample V. (b) Scatter plot of the multidimensional scaling analysis for the perception of the same five synthetic vowels. Mappings were averaged across participants after Procrustes realignment. The mean locations for each sample are shown, with ellipses representing 1 *SD* of bivariate normal distributions fitted to the data. Interpretative axes were obtained by Procrustes analysis with the data from (a), and plotted as dotted lines.

Critical to the success of vowel production is whether or not the vowels are discriminable and identifiable, that is whether or not they can be easily differentiated and transmit the intended vowel to the listener, regardless of how non-overlapping their formant locations may be in frequency space. These qualities were evaluated in an experimental program. First, distances in perceptual space between the stimuli

Published in *The Journal of the Acoustical Society of America*

were obtained by asking participants to rate inter-stimuli dissimilarity for all possible pairings. A multidimensional scaling analysis was performed on the distances, which could be mapped to a two-dimensional projection with minimal stress, in order to establish if the five synthetic sounds occupy discernibly different regions in perceptual space. A vowel classification experiment was then carried out to assess vowel identity and its consistency both within and between individuals.

Vowels and their pronunciations have evolved considerably since the time of Grosseteste, and it goes without saying that we were unable to run experiments with participants with a medieval language background. However, it is reasonable to expect that the mechanisms of vowel perception have broadly remained constant to the modern era, although some finer elements of speech perception vary as a result of differing cultural and language contexts (Kuhl et al., 1997). For this reason, we selected participants from a range of language backgrounds.

9.3.3 *Multidimensional scaling experiment*

In the first psychophysical experiment the five stimuli were presented to both native and non-native English speakers to obtain dissimilarity scores. The *.wav* files (sampling rate 44,100 Hz, 16 bit, monophonic) were all normalised to 0 dB relative to full scale and limited to a duration of 1.70 s in Audacity, to be played through a pair of Sennheiser HD201 Closed Dynamic Stereo headphones. The experiment was built using the open-source Matlab function set Psychtoolbox (Brainard, 1997), and run using the same laptop and headphones in quiet conditions. 20 participants took part in the experiment (12 female, 8 male, mean age 25 years). Participants were asked for their country of origin (13 UK, 1 USA, 2 India, 2 Bulgaria, 1 Germany, 1 Poland), if they were native or non-native English speakers and if non-native what their native language was (16 native English speakers [13 monolingual UK, 1 monolingual USA, 2 bilingual in English and Hindi], 4 non-native [2 Bulgarian, 1 German, 1 Polish]).

Published in *The Journal of the Acoustical Society of America*

Participants registered their informed consent with a button press, before testing began.

Participants were first played each of the five stimuli once for familiarity. Pairs of recordings were then presented separated by a 300 ms pause, and participants registered their perceived dissimilarity via a keyboard, from 0 (identical) to 7 (very dissimilar). For stimuli $i, j = 1, \ldots, 5$, all possible pairs were presented once in a random order, for both (i, j) and (j, i) sequences, to give a dissimilarity response matrix. From this, a symmetric matrix was constructed for each participant by taking means of (i, j) and (j, i) values. For six of the participants a single set of dissimilarity judgments was collected, while 14 went through the experiment twice, and their symmetric matrices were averaged across repeats.

Kruskal's non-metric multidimensional scaling analysis (MDS) (Kruskal, 1964) was performed on the symmetrical matrices to approximate the relative locations in perceptual space of the samples for individual participants. Once Euclidean coordinates were obtained from MDS analysis, these were plotted to inspect their agreement with the formant plots of the samples. Visual inspection of the mappings showed a clear correspondence between the first dimension of scaling and F_2, and the second dimension of scaling and F_1, for the majority of participants, which was later formally analysed as described below. This agrees with previous studies that find human vowel discrimination primarily tracks the frequency position of F_2, which corresponds to perceived vowel advancement, and secondarily tracks the frequency position of F_1, corresponding to perceived vowel height (Sinnott et al., 1997). There were four exceptions for this agreement; notably, these data sets were from the four non-native English speaking participants. Further inspection showed that these data agreed with F_2 and F_1 when plotted in the first and third dimension from the MDS, respectively, and hence these mappings were taken forward in the analysis.

Data sets then underwent Procrustes analysis, which permitted similarity transformations of the mappings (uniform scaling, orthogonal rotation, translation and reflection) in order to give the best concordance across participants while maintaining relative perceptual distances within mappings (Sibson, 1978). Once realigned, data

Published in The Journal of the Acoustical Society of America

sets were analysed to extract the statistics for each stimulus as located in perceptual space by participants. Figure 9.7.b shows the mean positions for each stimulus, plotted as solid symbols. Ellipses show one standard deviation of the bivariate distribution of each vowel within the two dimensions of scaling. Sample O gave rise to the most spread compared to the other vowels, indicating that participants differed most in where to locate it in their perceptual space, relative to the other vowels. This is likely related to the strong degree of variation present in open back vowel pronunciations across dialects of English.

Procrustes analysis was also performed between the realigned perceptual space data and the acoustics-based vowel quadrilateral generated from formant data, in order to obtain axes for interpretation of the MDS analysis, labelled as 'Formant 1' and 'Formant 2'. The distribution of relative perceptual locations for the five synthetic samples (Figure 9.7 b) show a clear agreement with their placing in the F_2/F_1 frequency space (Figure 9.7 a), primarily with the samples occupying separate (i.e. discriminable) regions in perceptual space, albeit with some overlap between participants.

Monte Carlo simulations were carried out to evaluate the likelihood of stimuli being mapped to distinct regions due to chance, and consistently with the same relative orientation. From 26 simulations, only 20 generated data that could be mapped by MDS. After Procrustes analysis of these 20 mappings, none gave rise to a distinct region for any of the stimuli (i.e. non-overlapping regions bound by one standard deviation of stimuli mean position), and all stimuli regions had an area above 5 scaling space units^2, compared to a mean of 1.2 scaling space units^2 for participant-generated data. For all mappings, shown in Figure 9.9 in the Supplementary Material, the relative orientation of vowels were different. A more extensive simulation was carried out to generate 100 mappings, whose ellipses had a mean of 7 scaling space units^2, shown in Figure 9.10. We therefore conclude that the results of mapping the participant data, with stimuli occupying separable regions and a relative orientation in agreement with the acoustic analysis, are not owing to chance.

Published in The Journal of the Acoustical Society of America

9.3.4 *Vowel classification experiment*

Fourteen of the participants (ten native English speakers; four non-native English speakers) also completed a second test, to obtain vowel classifications for the stimuli. Participants were asked to listen to the recordings with headphones and assign them labels which best agreed with their percepts. Participants were not expected to be familiar with IPA notation, instead selecting one of the following options: "'ah' as in spa", "'eh' as in get", "'ee' as in beat", "'o' as in cot", or "'oo' as in zoo"; corresponding to /ɑ, ɛ, i, ɔ, u/, respectively. These options are also summarised in Table 9.1 in the Supplementary Material. Each stimulus appeared in a familiarisation phase once in this order, followed by a test phase in which they were presented a further four times in a randomised order.

Responses from the familiarisation phase were not included in the analysis, as participants had not heard all of the vowels at that time. The data from individual participants did not show any correlation between classification confusions and being a native/non-native English speaker, which is not surprising given the coarseness of the classification system. Figure 9.8 shows the distributions of responses for each stimulus, with pie charts for each stimulus being centered at the stimulus' position in acoustic space as calculated above. The data are also given in Table 9.3 in the Supplementary Material.

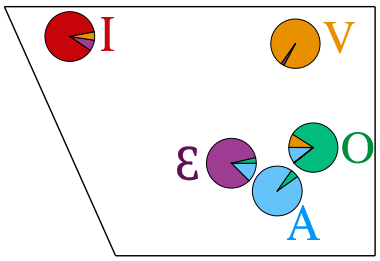


Figure 9.8: Classifications obtained for each of the five samples from the second listening test. The pie charts for each sample, showing participants' classifications, are centered at the samples' locations when mapped in acoustic space, as shown in Figure 9.7 a. Responses are indicated by color: "'ah' as in spa" (/ɑ/) in blue, "'eh' as in get" (/ɛ/) in purple, "'ee' as in beat" (/i/) in red, "'o' as in cot" (/ɔ/) in green, and "'oo' as in zoo" (/u/) in orange.

Published in The Journal of the Acoustical Society of America

9.3.5 *Results: MDS and classification experiments*

Listening to isolated vowels is not a common activity in daily life, and listening to isolated vowels without having any reference to the speaker is also unusual. In addition, these stimuli are clearly non-human in origin given the identical electrolarynx acoustic input in each case. Some confusion is therefore inevitable. As may be expected, the synthetic vowel with the broadest spread of placement in perceptual space (indicated by its ellipse in Figure 9.7 b having the greatest area) was also the least reliably classified sound, Sample O, which received 80.4% correct classifications and 10.7% and 8.9% misclassifications as 'ah' and as 'oo' respectively. The greatest source of misclassification was the assigning of both Sample $\mathcal{E}$ and Sample O as 'ah' (12.5% and 10.7% respectively). The perceptual space generated by MDS analysis and the acoustic space from formant data both show Sample E and Sample O located in close proximity to Sample A, which itself was classified as 'ah' with high agreement. Indeed, on the perceptual map these are the only two instances of overlapping standard deviations from the samples' means. It can be said with confidence that the samples are perceived, imperfectly, as vowels, spanning a large proportion of vowel perceptual space.

As well as the samples being consistently classified by participants, these classifications were overwhelmingly in accordance with the mapping specified in the *DGS*, according to which the vocal tract models were constructed, when these five vowel letters are related to phonemes, as given in Table 9.1 in the Supplementary Material. Of course, we cannot be sure that Grosseteste would have had these same phonetic sounds in mind (namely 'A' mapped to /ɑ/, '$\mathcal{E}$' mapped to /ɛ/, 'I' mapped to /i/, 'O' mapped to /ɔ/, and 'V' mapped to /u/). The classification task did not test for exact identity between stimuli and labels; participants were asked to select the closest match from the five options given rather than provide their own labels. However, it is worth stating that as there are 120 possible permutations of mapping five labels to five stimuli ($P(5) = 5! = 120$), it would be unlikely to observe this specific mapping

Published in *The Journal of the Acoustical Society of America*

by chance alone across numerous participants. We can therefore conclude that the shapes Grosseteste specified for shaping the vocal tract during vowel production are compatible with their related phonemes when present in the mouth end of the vocal tract (or other acoustic chamber), in a five-vowel system.

9.4 DISCUSSION

While sometimes described as a scientist, and undoubtedly instrumental in the conception of the scientific experimental method (Crombie, 1953), we must be careful when reading Grosseteste's treatises not to impute any sense of experimental or even observational basis for his theories, however elegant the logical or mathematical arguments found therein. Recent interdisciplinary research has found that the origin of such theories, though they may be wrong within the context of current scientific understanding, may still best be explained as resulting from direct observation, such as for his novel theory of rainbow formation (Harvey et al., 2017b). However others, though they may have been correct, are unlikely to have had a direct observational basis, such as his three-dimensional theory of colour space as expressed in the *DC* (Smithson et al., 2012). These works remain remarkable achievements, and the desire to mathematicise the mental or material world was a fundamental evolution for intellectual history in the medieval and early modern era.

In his treatise on sound, Grosseteste is applying a similar mathematical framework of combinatorics as his theory of colour, but to vowels. There are, however, some interesting differences between the two. In the *DC*, Grosseteste is clear that colour space is continuous, as he describes the infinite 'diminutions' between the extrema of the space. That he constructs the parameter space to reflect established intuitions about space and distance is therefore quite sensible; colours are connected along routes, which may be traversed by increasing or decreasing one, two, or all three of the space's parameters. This particular feature of the theory we can presume was likely based on direct observation, and the subtle and continuous variations

Published in *The Journal of the Acoustical Society of America*

in colours seen in the world and, explicitly, in rainbows. In the *DGS*, Grosseteste again constructs a generative scheme to account for the variety within a perceptual phenomenon, but it is this time categorical and discrete, accounting for the varieties of vowels and their external representational forms, letters.

The scheme is defined by what he says are the three types of simple, self-similar movements: linear, circular, and dilational-constrictional. These simple movements may be combined, but only a subset yield novel categories of movement: combining linear with circular, linear with dilational-constrictional, circular with circular, and circular with dilational-constrictional. These descriptions of movement are readily interpreted as the three types of geometric similarity transformations—translation, rotation, and uniform scaling (with reflection being equivalent to rotation through a higher dimension)—though it should be noted that no diagrams are found in extant manuscripts, and this is just one possible interpretive scheme[7]. The treatise can be read as one primarily about types of movement, and relies heavily on the false premise that sounds of different qualities are discriminable based on the category of vibrational movement, rather than the spectral filtering achieved by differently-shaped acoustic chambers with varying resonant frequencies, and other language-specific factors. Although this theory is mistaken about the underlying source of vowel timbre, Grosseteste nevertheless constructs an elegant theory that attempts to account for the categorical nature of vowel perception, and the representation of vowels as letters.

Reading this text today prompts us to examine what may constitute permissible evidence in science. For Grosseteste, the shapes of letters could serve as the primary evidence for his claims regarding the shape of the vocal tract, and the forms of mental representations of vowels; within the medieval paradigms of Aristotelian mimesis and the liberal arts, this was a scientifically orthodox and justifiable use of observations to infer properties of the natural world. Although we do not share these paradigms as modern scientists, we share in the methodological framework of setting our own standards for permissible evidence; in many cases such sources of evidence

7 Other schemes of interpretation may be found in a forthcoming volume, which contains the critical edition, translation, and interdisciplinary analyses of the *DGS* (Harvey et al., 2019c).

Published in *The Journal of the Acoustical Society of America*

are far-removed from the phenomenon we attempt to study. A generous reading of the *DGS* could be that Grosseteste is engaged in modelling; do abstract movement categories offer a viable framework for the robust, categorical representation and perception of speech sounds, despite their continuous variety and noisy instances? Although our models of speech processing have matured in their awareness of acoustics and physiology (Hickok and Poeppel, 2007; McClelland and Elman, 1986; Stevens, 2002), they share the underlying goal of understanding how speech signals are processed and represented.

The *DGS* does make strong claims about the morphology of the vocal tract during vowel production, which are clearly incorrect in asserting the presence of geometric shapes. However, we have shown, through artificial vowel synthesis and the methods of spectral analysis and psychophysical testing of vowel perception, that these geometric shapes can in fact be incorporated at the mouth end of acoustic chambers that give rise to discriminable vowel sounds. This is plausibly due to degree of freedom present in the remainder of the acoustic chamber, i.e. the laryngeal and pharyngeal cavity, and the many-to-one property of acoustic chambers and their spectral output (Atal et al., 1978), meaning that unique speech sounds may have multimodal or highly nonlinear mappings in articulator space (Qin and Carreira-Perpiñán, 2009). In the thirteenth-century Grosseteste would only have had visual and proprioceptive measurements of the lips, teeth and tongue, so any requirements of the rest of the vocal tract for vowel production could not have impacted his theory.

How influential the *DGS* was on the developing field of phonetics is difficult to say. Roger Bacon, a student of Grosseteste's who praised his mathematical approach to understanding nature, describes similar notions of relating the number of vowels in languages to the number of fundamental classes of movements in his text on Greek Grammar (Nolan and Hirsch, 1902). However, he seems to criticise these theories as falling outside the scope of the 'pure grammarian', instead they should be left to the disciplines of metaphysics and of music (Roser-Catach, 1997). Specifically, he is engaging with the content of the *Tractatus de grammatica*. Circulating at the time, the

Published in *The Journal of the Acoustical Society of America*

anonymous *Tractatus* was widely attributed to Aristotle, but Bacon shows this to be unjustified, and the treatise was later sometimes ascribed to Grosseteste.

Readers familiar with Hangul, the native Korean alphabet devised by King Sejong the Great (1397-1450) in the fifteenth century, may find similarities between Grosseteste's theory of non-arbitrary letter shapes and the apparent similarity between Hangul consonant forms and their corresponding places of articulation (Kim-Renaud, 1997). However, we have no record of a reception of Grosseteste's work in east Asia, and any direct connection seems improbable. Moreover, while the articulatory basis of the Hangul alphabet is often stated as matter of fact, and has been written about since only a few years after Hangul was devised (such as in *Hwunmin cengum haylyey* [Explanations and Examples of the Correct Sounds for the Instruction of the People], published in 1446), there are competing theories. It seems equally likely that Hangul consonants were instead influenced by or modelled on the Mongol 'Phags-pa alphabet, itself derived from Tibetan, as suggested by Keith Whinnom (Hope, 1957). It could therefore be the case that in Hangul and its reception we find a thesis parallel to claims made in the *DGS*: the notion of glyph iconicity being used as a kind of pedagogical or philosophical device to explain their forms.

Theories attempting to draw direct relationships between the shaping of articulators and the shapes of letters surfaced again in the seventeenth century, with Franciscus Mercurius van Helmont claiming that intrinsic to the Hebrew alphabet was found a phonetic guide to its pronunciation (Helmont, 1667), and Bishop John Wilkins attempting to construct a visual alphabet of speech sound diagrams (John Wilkins, 1668). In neither case is there an explicit connection to the *DGS*. Such theories relating letter shapes to vocal tract shapes paved the way for the speaking machine of Wolfgang von Kempelen in 1780, and, later, the set of 'visible speech' symbols by Alexander Melville Bell (Bell, 1867; Dudley and Tarnoczy, 1950).

Lastly, an essay published in 1772 by Charles Davy makes near identical claims regarding the representations of the vocal tract in the letter shapes of vowels (Davy, 1772) (p84-87), but again, any connection to Grosseteste's theory is not made explicit and may be entirely accidental. It should also be noted that Davy's text was not written

Published in *The Journal of the Acoustical Society of America*

as a serious scientific endeavour but as an amusing romp through classical trivia, with Davy himself writing: "The Editor will not undertake to defend it: as a *whimsical* conjecture, it may still afford some entertainment. *Better* reasons might perhaps be offered in its favour than what appear at present", before stating his belief that the Greeks' visual representation of the vocal tract in letter shapes is what enabled their literary success. It may simply be the case that such theories were best appreciated as a form of intellectual entertainment, rather than serious scientific endeavour. Now, with the advent of recent studies into glyph iconcity (Schmidtke et al., 2014; Turoman and Styles, 2017), theories of non-arbitrary representation of letter shapes are again being considered, albeit from a more nuanced and experimental standpoint.

9.5 CONCLUSION

In the treatise *De generatione sonorum* [On the Generation of Sounds], Robert Grosseteste attempts a mathematicisation of the perceptual space of vowels. With this paper we show that the treatise formulates vowels—their production, perception, and representations both mental and in writing—into a coherent framework of geometric figures, which are combinatorially generated from basic types of movement. Although clearly incorrect in his understanding of vocal acoustics, and ignorant of the supporting physiology, Grosseteste shows remarkable insight in his approach to explaining why vowels are categorical in nature, and how auditory, visual, and motor faculties play complementary roles in speech perception. His theory touches on principles highly relevant to contemporary neuroscience, namely the nature of mental representations and their relationship to external stimuli, and the integration of different sensory faculties. Finally, aspects of Grosseteste's theory of speech can be expressed in a scientific, falsifiable manner, which we show here to have been potentially commensurable with the sensory data available at the time.

Published in *The Journal of the Acoustical Society of America*

ACKNOWLEDGEMENTS

This work was supported by the Arts and Humanities Research Council (AHRC) under grant AH/N001222/1. JSH holds a DPhil studentship from the Andrew W. Mellon Foundation, hosted by the Oxford interdisciplinary research centre the Oxford Research Centre in the Humanities (TORCH). HES was supported by a Visiting Fellowship from the Institute of Advanced Study at Durham University. The work presented here emerged from the collaborative Ordered Universe Project (http://ordered-universe.com/), which focuses on interdisciplinary readings of the scientific works of Robert Grosseteste (c. 1170–1253).

The authors would like to thank all the participants of the collaborative workshops during which the *DGS* was discussed [2nd-3rd October 2014, "13th Century Science in a Multi-Disciplinary Perspective", Pembroke College, Oxford (funded by the Mahfouz Foundation); 8th-10th April 2015 "Knowing and Speaking: On the Generation of Sounds and On the Liberal Arts", Bishop Grosseteste University, Lincoln; 25th-28th November 2015, "On the Liberal Arts and On the Generation of Sounds: Robert Grosseteste's Early Treatises and Their Reception", Durham University]. We thank John Coleman for his advice on RP, Middle English, and Latin pronunciation, the use of IPA symbols, and the origins of the Hangul alphabet. We appreciate Brian Tanner's contributions to the discussion of sound and the movements of the vocal tract, Neil Lewis's suggestions on issues of translation and interpretation of the medieval texts, and Cecilia Panti's discussions on the text. We also thank the two anonymous reviewers for their comments on an earlier edition of this manuscript. Finally, the authors thank all the participants who took part in the listening tests.

Published in *The Journal of the Acoustical Society of America*

9.A SUPPLEMENTARY MATERIAL

Letter shape	Phoneme	Example
A	/ɑ/	'ah' as in 'part'
$\mathcal{E}$	/ɛ/	'eh' as in 'get'
I	/i/	'ee' as in 'beat'
O	/ɔ/	'o' as in 'cot'
V	/u/	'oo' as in 'zoo'

Table 9.1: Our interpretation of phonemes from the vowel letters Grosseteste uses in the *DGS*. The third column also shows the options given to participants in the classification listening test.

	larynx	.	.	.	.	.	.	.	.	.	.	.	.	.	.	lips
A	22	8	18	8	8	12	16	20	24	26	28	30	32	34	38	24
E	12	8	12	8	22	14	14	10	16	24	18	10	16	24	18	10
I	16	32	32	32	32	30	30	20	12	12	8	8	8	10	10	10
O	8	20	12	12	12	10	8	8	16	24	30	32	30	24	16	10
V	8	32	10	8	30	28	26	24	22	20	18	16	14	12	10	8

Table 9.2: Diameters (in mm) of the employed plate-type model of Arai *et al.* (Arai et al., 2001) used to create the tracts shown in Figure 9.5 and to synthesise the five speech sounds (Sample A, Sample E, Sample I, Sample O, Sample V) based on Grosseteste's five movement types.

	'ah' as in 'part'	'eh' as in 'get'	'ee' as in 'beat'	'o' as in 'cot'	'oo' as in 'zoo'
A	64	0	0	6	0
E	8	59	0	3	0
I	0	8	59	0	3
O	7	1	0	57	5
V	0	1	0	1	68

Table 9.3: Results from the classification experiment ($N = 14$). Each participant classified each sample five times, choosing from the five possible responses in the top row of the table.

Published in *The Journal of the Acoustical Society of America*

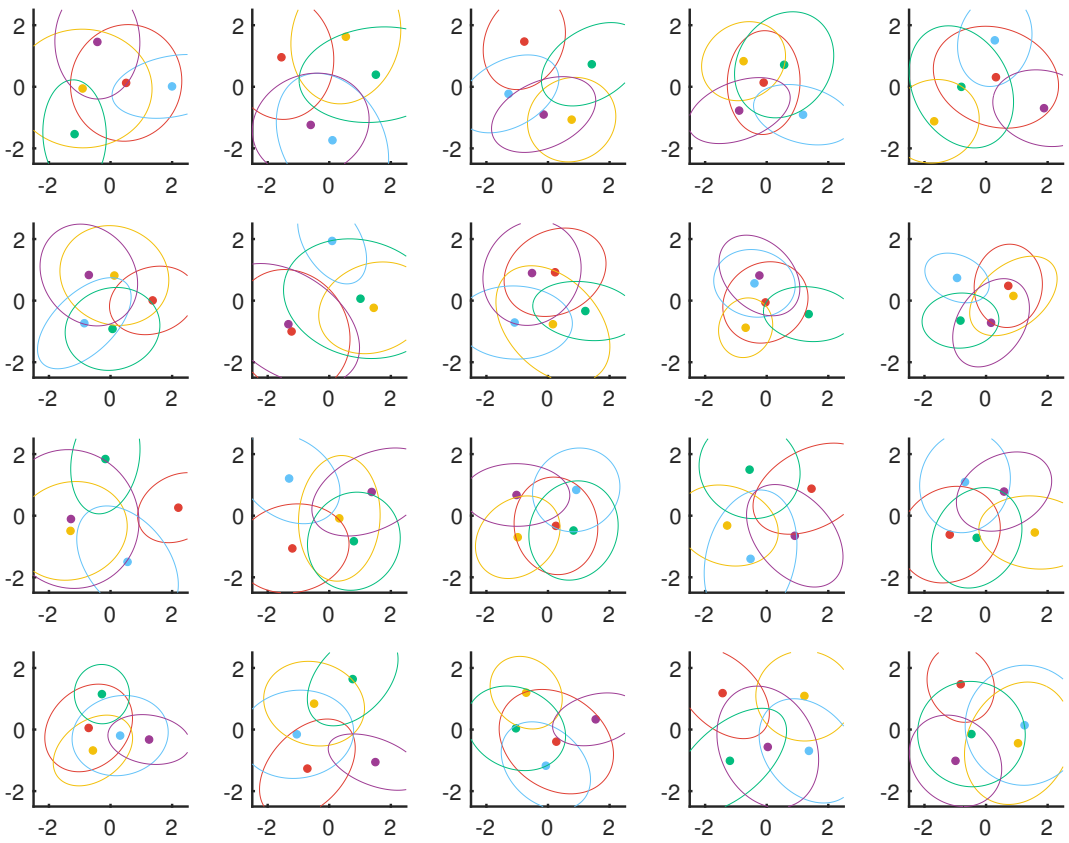


Figure 9.9: 20 examples of Monte Carlo simulations that generated data sets for which an MDS mapping was possible. No simulation produced dissimilarity data that when mapped featured a distinct area for a stimulus, as bound by one standard deviation from its mean position (indicated by ellipses).

Published in *The Journal of the Acoustical Society of America*

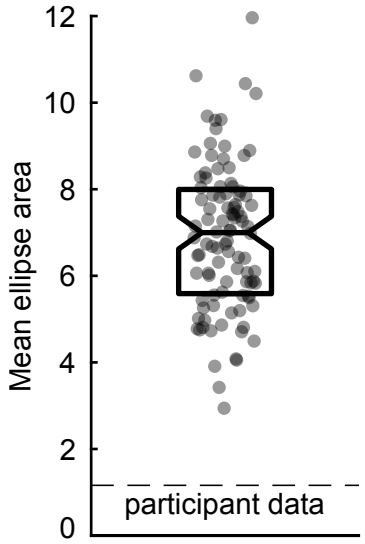


Figure 9.10: The results of 100 Monte Carlo simulations of the MDS experiment. The mean ellipse areas from each simulation (which comprised 20 randomised participant data sets) are shown. The box plot indicates the mean and quartiles of the distribution, with a 95% confidence interval on the mean shown as a notch. The mean of the participant data set is indicated by a dashed line.

Published in *The Journal of the Acoustical Society of America*

Part IV

COLOUR AND MATERIALS

10

VISUAL PERCEPTION OF MATERIAL QUALITIES

The very mountains and stones have a richer gloss,
and lovelier transparency,
and intensity of hue.

—Aldous Huxley, 1961, quoting Plato's *Phædo*

PART ABSTRACT

As previously discussed in part i, scientific—that is, structured and ordered—understandings of the human colour sense have evolved considerably throughout history. This evolution required radical departures from established dogma, as in the case of the Islamic thinkers, as well as a complete rethinking of how colour itself should be formulated, with Grosseteste making an early case for a continuous space with a defined dimensionality. In this part, I will discuss colour perception in the context of optics—the physical properties of illuminants, surfaces, and their interactions, and material perception—how the (normally) trichromatic visual system might infer material qualities from the spatial and chromatic distributions of colour signals. I will provide an overview of light transport, which governs how objects, light sources, and their relative geometries to an observer determines the retinal image. Chapter 11 presents a framework for visualising metamers, and the relationship between physical spectral distributions and perceptual colour, expanding the dimensionality

of colour beyond trichromacy. Chapter 12 explores the *neon fruit illusion*, which was showcased at a public exhibition of visual illusions, and exploits the physical properties of illuminants and surfaces, and their interactions with the properties of the cone photoreceptors, to impressive effect. Lastly, chapter 13 introduces an 'imitation gold' material used in medieval polychrome sculpture. The physical and optical qualities of the material is explored with physical measurement and computer graphics simulations, and related to its perceptual qualities.

10.1 HUMAN COLOUR VISION AND ITS STIMULI

In chapter 1, the classical and medieval theories of colour were discussed, culminating with Robert Grosseteste's thirteenth-century formulation of colour as a three-dimensional space. Grosseteste was later vindicated as Young-Helmholtz trichromacy became the accepted model of normal human colour vision. Possessing three types of sensors—the photosensitive cone cell classes of the retina—incoming spectral distributions of light are reduced to a three-dimensional signal, the relative excitations of each sensor type (the two-dimensional chromaticity plane) as well as their summed activity (the luminance axis).

There are two immediate limitations to human colour vision. First, there exist visual stimuli with distinct spectral distributions that give rise to the equivalent physiological responses, and hence phenomenal colour experiences, for trichromats. These spectrally different, though perceptually indistinguishable, colours—*metamers*—are discussed further in chapter 11. Second, we are sensitive not to surface colours, but the appearance of those surfaces under a given illuminant spectrum. If the goal of human colour vision is to differentiate, identify, and recognise objects in the world, this problem is fundamentally ill-posed. Nonetheless, we ordinarily possess the remarkable ability to stably perceive object colour under changing illumination, termed *colour constancy* (Foster, 2011). Chapter 12 describes the *neon fruit illusion*, which exploits a failure of colour constancy under narrowband illumination.

However, the visual world is not comprised of flatly arranged coloured surfaces[1], and surfaces interact with light with greater complexity than simply absorption and diffuse reflectance. As such, the appearances of surfaces are determined by reflectance functions not only in terms of their spectral distribution, but also their spatial distribution. If colour constancy is the ability to characterise a surface's reflectance function with respect to its spectral distribution, in a world of three-dimensional material objects, the ideal visual system could additionally characterise reflectance functions with respect to all possible angles of incidence and reflectance—the bi-directional reflectance distribution function (BRDF) of the surface.

With these two considerations in mind, we can expand our dimensional conception of colour as it is found in the natural world, just as Islamic thinkers first expanded Aristotle's one-dimensional theory of colour into two dimensions, and Grosseteste further described colour as a three-dimensional space, as discussed in chapter 2. These additional dimensions have direct bearings on perceptual mechanisms of colour vision: the scene illuminant on colour constancy, and the spatial distribution of surface colour on material perception.

10.2 THE VISUAL PERCEPTION OF MATERIAL QUALITIES

The material properties of an object dictate whether it will shatter, bounce or deform on impact, be slippery or easy to grip, feel dry or wet, and be warm or cool to the touch. In turn, they govern how we interact with such objects and their surfaces. They additionally determine the way in which a material interacts with light, providing a basis for material properties to be perceived through the visual system. One of the central questions at the heart of material perception is how the human visual system is capable of so accurately, and rapidly, inferring material qualities from images. This remains an impressive feat; rarely in life do we mistake one material for another, and many imitations whether physical or virtual are unconvincing. To answer this

1 This simplistic visual stimulus, referred to as Mondrian patterns after the Dutch abstract painter, was used in early investigations into colour perception (for example see Arend and Reeves, 1986).

question we must characterise the relationship between the physical world of optics and materials, our mathematical models and approximations of the physical world, and the image features available to the visual system.

There is a rich variety of optical mechanisms governing material-light interactions. It is implausible that the visual system is calculating any form of inverse optics in material perception—such computations would be intractably complex and expensive. Still, a solid understanding of the physics of material-light interactions paves the way for the creation of physically-plausible stimuli, as well as the potential for relating perceptual cues to physical states of the world.

10.3 THE PASSAGE OF LIGHT THROUGH OPTICAL MEDIA

Often the focus on colour and material appearance is on surface properties, and 'volumetric' effects are perhaps seen as supplementary. However, in order to understand surface properties, it is first helpful to address the simple case of light travelling through a medium of given material properties. This will serve us well when thinking about surfaces and the optical properties of boundaries between media, as the majority of surfaces in the natural world are not strictly reflective, but permit light to enter into the object to some degree.

Any optical medium can be described in terms of two physical processes: absorption, where the spectral power of light is attenuated and its energy lost as heat, and scattering, where the direction of light propagation is changed. Of these, it is absorption that bears the most influence on a medium's colour as a result of unequal spectral attenuation in the visible range, while scattering primarily determines the opacity of the medium, but can in some cases also affect the wavelength composition of light (and hence its chromaticity). Both absorption and scattering have the same

effect on intensity of light, which is attenuated during passage through a medium as described by Lambert's laws:

$$I_T(z) = I_0 \cdot \exp(-a_A \cdot z), \tag{10.1}$$

$$I_T(z) = I_0 \cdot \exp(-a_S \cdot z), \tag{10.2}$$

These equations relate the transmitted intensity of light, I_T, to the initial intensity, I_0, the path length, z, and the absorption and scattering coefficients of the medium, a_A and a_S[2], respectively.

More completely, we can describe not only the intensity of a light ray, but the function of an electromagnetic wave with respect to space, time and the the absorbance of a medium, A. Absorbance should not be confused with absorption—the ratio of radiant flux absorbed to radiant flux received of a surface—and instead is a measure of transmittance, T, a material's ratio of radiant flux transmitted, Φ_e^t, to radiant flux received, Φ_e^i (see Eq. 10.3). An object with a low absorbance may still absorb much of the light shining on it, such as a thick piece of darkly coloured glass, while an object of high absorbance may absorb close to no light, as in the case of a clean metal surface.

$$T = \frac{\Phi_e^t}{\Phi_e^i}, \quad A = \log_{10} \frac{\Phi_e^i}{\Phi_e^t} = -\log_{10} T \tag{10.3}$$

Materials exist on a continuum of absorbance, through which different optical phenomena will dominate appearance. With no spectral attenuation ($A = 0$, $T = 1$), materials are perfectly transparent. Low-absorbance materials will appear translucent, and may be modelled by subsurface scattering and diffuse reflectance. Materials with higher absorbance can be modelled without subsurface scattering, and be described by diffuse reflectance alone. Finally, for materials with a very high absorbance which do not allow incident light to penetrate into the material, any reflectance will be

2 More completely, light can also be scattered into the propagating ray (in-scattering). This will most commonly result in a desaturation of colours as seen in foggy conditions.

primarily specular, although microstructure irregularities of the surface (roughness) may still contribute a diffuse-like appearance.

Attenuation of light through a material can be modelled by introducing an imaginary component to its refractive index. The complex refractive index, $\underline{n}$, then describes both the phase velocity of light through the medium, relative to a vacuum, in its real component n, as well as its extinction coefficient in its imaginary component, $i\kappa$ (occasionally as $-i\kappa$). We therefore describe an electromagnetic wave, $\mathbf{E}$, propagating in a direction z as follows:

for a vacuum,
$$\mathbf{E}(z,t) = e^{i(kz-\omega t)} \quad \text{where} \quad k = \frac{2\pi}{\lambda} \tag{10.4}$$

for a medium,
$$\underline{n} = n + i\kappa \tag{10.5}$$
$$\underline{k} = \frac{2\pi \underline{n}}{\lambda} = \frac{2\pi(n+i\kappa)}{\lambda} = k_r + k_i \tag{10.6}$$
$$\mathbf{E}(z,t) = e^{i(\underline{k}z-\omega t)} = e^{-k_i z} e^{i(k_r z-\omega t)} \tag{10.7}$$

Where ω is the angular frequency, k is the wave number, λ is the wavelength, k_r and k_i are the real and imaginary component of the complex wave number, $\underline{k}$, respectively. These equations have been used to plot examples of real and complex refractive indices in Figure 10.1.

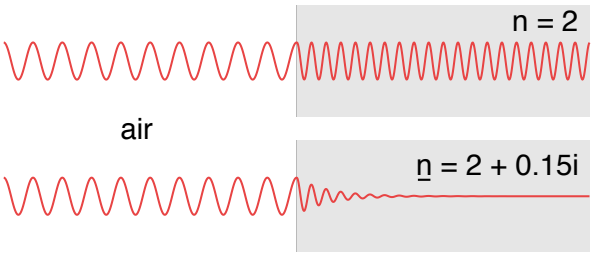


Figure 10.1: A simple model of light passing between two media. In the upper case, the medium has a real refractive index, $n = 2$, and the apparent wavelength of the wave halves. In the lower case, the medium has a complex refractive index, $\underline{n} = 2 + 0.15i$, evident by both the decrease in apparent wavelength, and the reduction of the wave's amplitude over distance (or time).

10.4 SURFACE REFLECTANCE PROPERTIES AND METALLIC MATERIALS

When light impinges on a boundary between different optical media it will undergo Fresnel reflection, sending it back through the same medium, and refraction, as it enters the new medium and propagates through to its skin-depth. In each case, scattering may also be caused by the geometry of the boundary. Surface reflectance properties are therefore best understood as describing the optical events occurring at the boundary between two optical media. Metals, and more broadly what could be described as metallic materials, present a distinct category of optical properties. While most other materials permit light to penetrate into the surface to some degree, allowing it to transmit and *diffuse* throughout a material, metallic objects are effectively opaque at their surface, and hence their reflectance functions can be accurately approximated by a sole specular component, without a diffuse or subsurface scattering component. This distinction is usually not due to surface roughness or microstructure geometry—for most white diffuse materials, no amount of polishing will produce a mirror, as light will still penetrate through and diffuse before reflecting off the surface.

For materials of high absorbance, light will be either absorbed or reflected at the surface. The appearance of these, generally metallic, surfaces therefore depends on the geometry of the object at both a macro- and micro-scale, as well as how much light in the visible spectrum is absorbed (if any). When polished, metal surfaces have a relatively flat surface microstructure geometry which gives rise to sharp specular reflection, and correspondingly a wider distribution of both highlights and lowlights, compared to diffuse materials, as shown in Figure 10.2. Increasing irregularity will effectively blur the reflected image of the object's environment, resulting in a more 'diffuse' appearance, although even very worn, rough metallic surfaces are noticeably more specular than low absorbance materials.

The specular component for most materials reflect an unaltered image of the illuminant, as this component is by definition the light that has not been transmitted

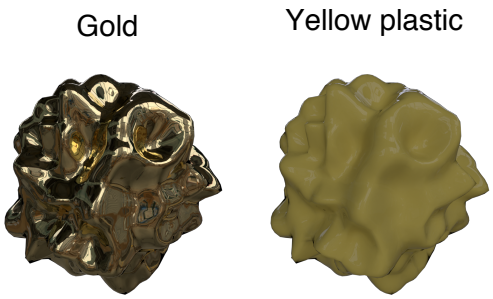


Figure 10.2: A defining characteristic of metallic appearance is an exclusively specular reflectance. The specular component of metallic (conductor) surfaces can affect the chromaticity of the reflected illumination. In comparison, glossy plastic (dielectric) surfaces retain the chromaticity of the illuminant in their specular reflectance component, but their overall appearance is dominated by a diffuse (albedo) reflectance.

into the material, reflecting directly off the surface. Unlike the specular components for most other materials (e.g. glossy plastic), metals can impose a strong colouration seen in the reflected image of their environment, as a consequence of their quantum-electrical properties. While dielectric (insulator) materials specularly reflect light of all wavelengths equally from their surface, coloured metals, such as gold, absorb light of certain wavelengths in the visible range corresponding to the excitation of the metal's electrons, regardless of the angle of incidence. For most elemental metals, the energy of excitation lies in the ultraviolet range; as no visible light is absorbed there is no effect on the spectral distribution (and hence chromaticity) of the reflected light. Instead, the metals appear silver, or mirror-like if highly polished.

The two metals commonly encountered that exhibit specular colouration are gold and copper. For the particular case of gold, the 5d orbital is elevated while the 6s orbital is lowered due to relativistic effects. This results in light towards the blue end of the visible spectrum being absorbed, as it has sufficient energy for the 5d-6s electron transition of 2.3 eV (Pyykko and Desclaux, 1979). For copper, the 3d-4s transition of 2.1 eV is responsible for the absorption of light through to the green part of the spectrum, giving the metal a reddish-orange chromaticity. The resultant chromatic specular appearances of these materials is illustrated in figure 10.3.

While specularity is such an important feature for metallic materials and their reflectance, there is some confusion within the literature surrounding metallicity,

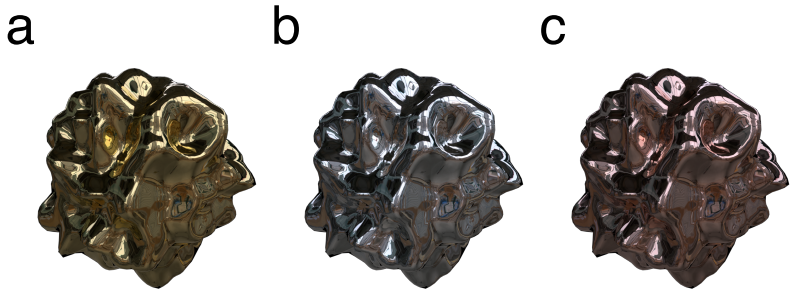


Figure 10.3: Renders using predefined Mitsuba materials for a) gold, b) silver, and c) copper.

gloss, and the reflective properties of metals. It would therefore be helpful to distinguish between 'glossy' specular components, which arise due to Fresnel reflection as a function of refractive index, and do not absorb any light in the visible range, and metallic specular components, which are uniformly present over a metal's surface regardless of angle of incidence and may have colouration. The appearance of metal cannot be rendered by simply adding, increasing, or maximising the specular component of a material, unless the diffuse component is achromatic, as illustrated in Figure 10.4. Adding an unnaturally high specular component to a diffuse material may instead produce a 'pearlescent' or painted-metal effect, depending on the underlying diffuse chromaticity.

There is some indication that the names of metals such as 'gold' and 'silver' may serve as categorical colour terms, or 'basic colours' (Matsumoto et al., 2016; Okazawa et al., 2011; Tanaka and Horiuchi, 2017). This has been summarised as $metal = color + gloss$ in numerous cases. However, this is false, as illustrated in Figure 10.4, and one cannot render a convincing image of a gold object by increasing the specularity of a yellow-orange object. The confusion comes from an incomplete consideration of 'gloss' or specular reflectance, as these studies exclusively consider highlights of a similar chromaticity to the underlying diffuse colour, resulting from either coloured illuminants (Tanaka and Horiuchi, 2017), coloured specular reflectance (Matsumoto et al., 2016), or stimuli post-processed to uniform chromaticity (Okazawa et al., 2011).

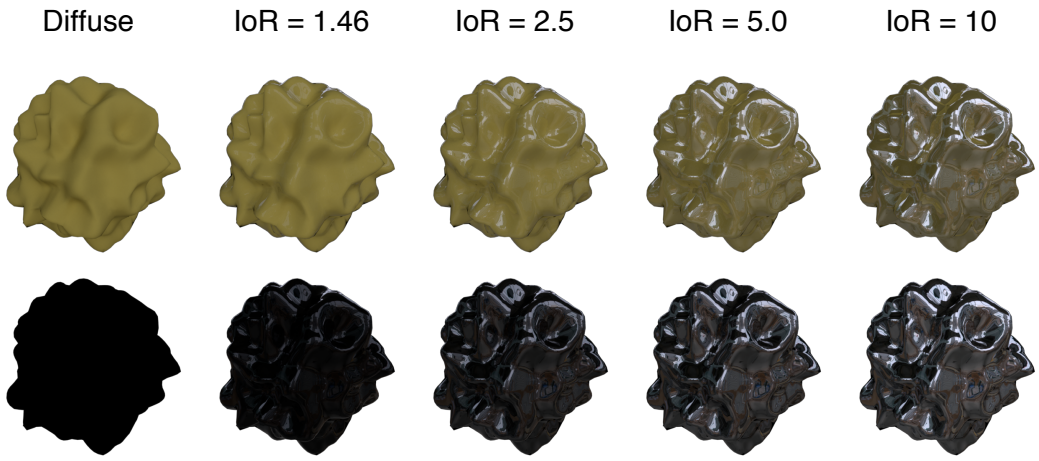


Figure 10.4: Metallic appearance is not simply extreme gloss. Upper Panel: A rendered material with a diffuse component does not approach a metallic appearance, however unrealistically high the specular component is set, unless the specular reflectance function incorporates a colour term that matches the diffuse term. As the specular component increases, more of the environment illuminant is seen. Lower Panel: A rendered material with an achromatic diffuse component does resemble a metallic appearance when a high enough specular component is added.

Just as binocular disparity serves as an important cue for the perception of gloss, it also plays a central role in the perception of shiny metallic materials. Many metallic objects exhibit a lustrous sheen that cannot be recreated with individual images such as photographs or renders, as it is a consequence of binocular lustre.

10.5 SIMULATING LIGHT-MATERIAL INTERACTIONS

The optical interactions of materials and surfaces can be simulated for a given scene. In computer graphics, this process is called *light transport*. Depending on their material, surfaces will be assigned a BRDF, which can be used to numerically relate the spectra of light impinging on, and reflecting off, the surface at all angles. Rendering engines such as Mitsuba, used in chapter 13 and again in part v, are optimised to calculate light transport for a variety of material BRDFs, including conductors, dielectrics, and specific classes such as glass and plastic. Given a three-dimensional scene, containing objects of defined geometries (meshes) and light sources, we are able to render an image that may be formed at any point in the scene, simulating light transport. This

is achieved in a hyperspectral fashion, giving rise to, for example, a thirty-channel hyperspectral image, that may then be converted to a three-channel image for viewing on a display.

10.6 SUMMARY

In this chapter, the physical optics of materials were discussed, in the context of the human colour sense. The three-dimensional understanding of colour first described by Grosseteste was expanded with relevance to mechanisms of human colour perception: the role of illuminants in colour constancy, and the role of spatial distributions of surface colour in material perception.

11

FROM SPECTRA TO PERCEPTUAL COLOUR: VISUALISATION TOOLS FOR THE DIMENSIONAL REDUCTION ACHIEVED BY THE HUMAN COLOUR SENSE

This is an extended version of a short paper originally published in *Eurographics 2018*, EG 2018 Posters collection.

AUTHORS

Joshua S. Harvey, Clive R. Siviour, and Hannah E. Smithson

ABSTRACT

Radiant spectra, defined as unique combinations of photon populations whose wavelengths lie in the visible range, occupy an infinite-dimensional real vector space. Any number of photon populations from the continuous spectrum of monochromatic wavelengths may be present to any non-negative amount. For normal vision, this space collapses to three dimensions at the retina, with any radiant spectrum eliciting just three sensor values corresponding to the excitations of the three classes of cone photoreceptor cells. While there are many mappings and visualisations of three-dimensional perceptual colour space, attempts to visualise the relationship between infinite-dimensional radiant spectral space and perceptual colour space are

Published in *Eurographics 2018*

lacking. We present a visualisation framework to illustrate this mathematical relation, using animation and transparency to map multiple radiant spectra to locations in perceptual colour space, revealing how the perceptual colour solid can be understood as intersecting surfaces and volumes. This framework provides a clear and intuitive illustration of colour metamerism.

11.1 INTRODUCTION

Radiant spectra occupy an infinite-dimensional real vector space, H_{colour}[1], and may be represented by countable infinite column vectors $\boldsymbol{s}$. At the retina, this space collapses to three dimensions of sensor values, that is, the photon catches and resultant excitations of the three classes of photosensitive cones (Koenderink, 2010). The cone spectral responses, L, M and S, of cones with response function row vectors $\boldsymbol{r}$ are then given by:

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} \boldsymbol{r}_L \\ \boldsymbol{r}_M \\ \boldsymbol{r}_S \end{bmatrix} \boldsymbol{s}$$

Where $\boldsymbol{r}$ is an N-dimensional row vector containing the response function for a given cone.

While the three-dimensionality of perceptual colour space is a familiar and intuitive concept, the same cannot be said for H_{colour}. We are accustomed to conceptualising spaces of three dimensions or fewer, living in a spatially three-dimensional world. Perceptual colour space is, by nature, relatable to the perceptual experience of colour, which may be used to navigate the space such as when selecting colours with a

1 In some literature this infinite-dimensional vector space of radiant spectra is identified as a Hilbert space (Drösler, 2013; Lenz, 2002). Certainly, the space does have an orthonormal basis (wavelengths) and is complete over the non-negative real numbers. However, negative radiant spectra do not exist in reality (though the filtering of light through absorbing media may be modelled as interaction of positive and negative spectra), and as such the space of radiant spectra is better described as an infinitely dimensional Hausdorff space (Koenderink, 2013).

Published in Eurographics 2018

hue-saturation-value (HSV) or red-green-blue (RGB) colour-picker. Any projection of radiant spectral space must violate our intuitions of either space or colour: namely that objects should not occupy the same space, or that colours should be separated according to their perceptual dissimilarity. The visualisation framework proposed here maintains the distance relations of colour dissimilarity, using transparency and animation to illustrate the intersecting surfaces and volumes of radiant spectra projected throughout the perceptual colour solid.

11.2 TWO-WAVELENGTH RADIANT SPECTRAL SURFACES

We begin by considering radiant spectral sets comprising just two wavelength populations, w_1 and w_2. Each colour set, C, is defined by the intensities of each population, $C\{w_1, w_2\}$, giving rise to a surface extended between the monochromatic axes $C\{w_1\}$ and $C\{w_2\}$. Figure 11.1 shows two colour sets, $C\{610, 540\}$ and $C\{460, 540\}$.

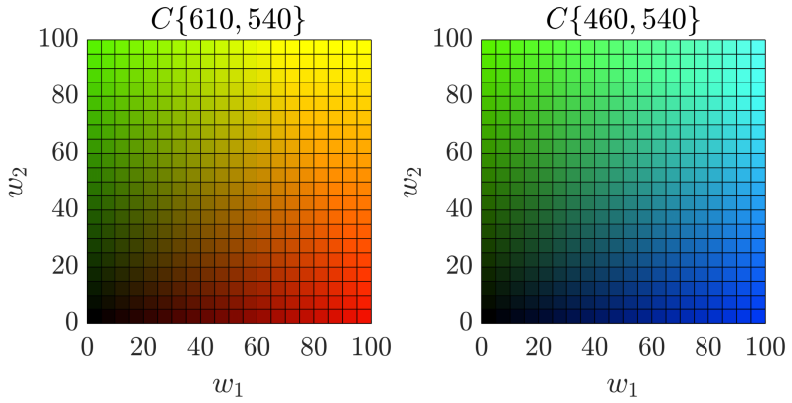


Figure 11.1: Projection of the $C\{610, 540\}$ and $C\{460, 540\}$ two-wavelength radiant spectral sets onto the $[w_1, w_2]$ plane.

Spectral sets of different two-wavelength populations may be projected simultaneously, by mapping resultant perceptual colours according to hue along the third dimension, as shown in Figure 11.2 with the two surfaces from Figure 11.1. They intersect at the w_2 axis, where both radiant spectra are monochromatic and $C(w_1) = 0$.

Published in Eurographics 2018

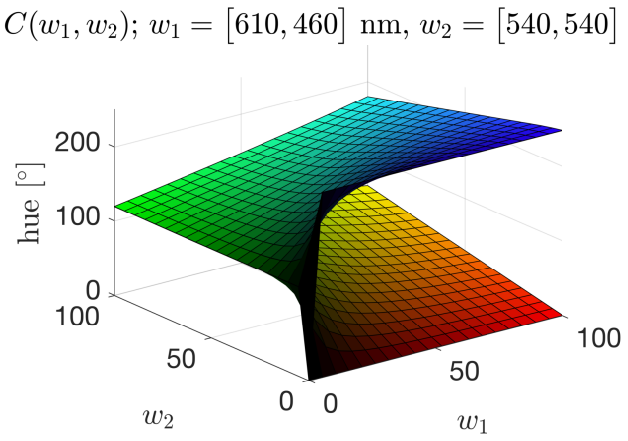


Figure 11.2: Projection of two $C(w_1, w_2)$ surfaces onto a $[w_1, w_2, \text{hue}]$ mapping. $C(w_1, w_2)$; $w_1 = [610, 460]$ and $w_2 = [540, 540]$. The surfaces intersect at the monochromatic axis $C(w_1) = 0$.

When $C(w_1, w_2)$ surfaces are projected according to their wavelength composition and resultant hues, surfaces may be discontinuous for two reasons. First, mapping hue to a linear axis will give rise to a discontinuity at hue = 0°, due to the circular nature of the parameter (0°=360°). Second, the requirement for colours to be mapped spatially according to perceptual similarity is violated due to the unaccounted parameter of saturation, which will also give rise to discontinuities when saturation = 0. Examples of these discontinuities are shown in Figure 11.3.

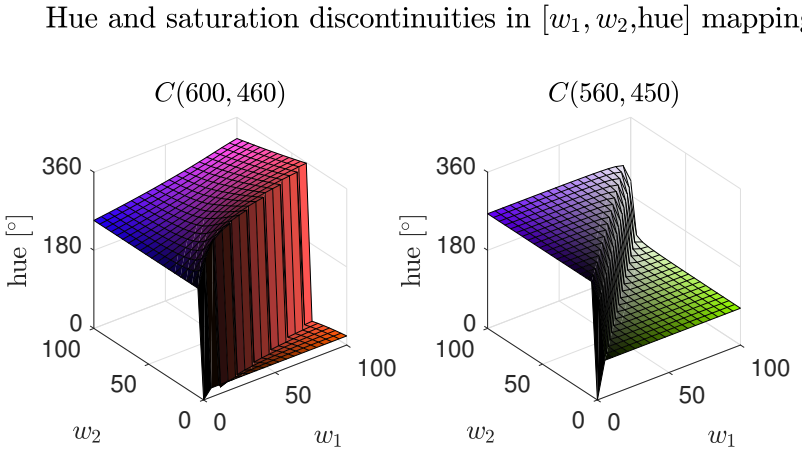


Figure 11.3: Example discontinuities due to hue (left) and saturation (right) in the $[w_1, w_2, \text{hue}]$ mapping.

Projection to a fully three-dimensional colour space will result in continuous $C\{w_1, w_2\}$ surfaces, as shown in 11.4. These spaces may be geared towards intuition

Published in Eurographics 2018

(HSV), practical considerations (RGB), or representing perceptual space (La*b*). Figure 11.5 shows several $C\{w_1, w_2\}$ colour sets projected to a HSV space, using the spherical coordinate system (r,θ,ϕ) such that r=value, θ=hue, and ϕ=saturation. The summation of primaries rather than proportional mixing makes surfaces and primaries more readily identifiable.

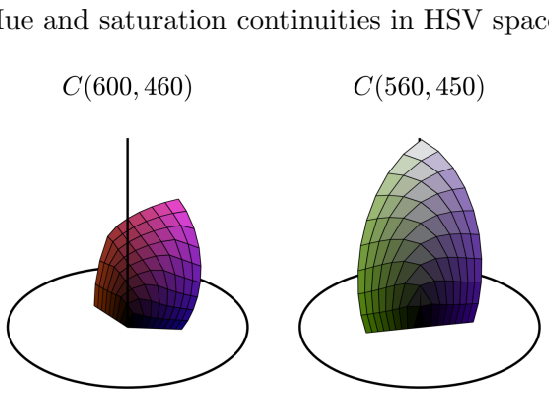


Figure 11.4: A fully three-dimensional space such as HSV projects $C(w_1, w_2)$ spectral sets as continuous surfaces.

Published in Eurographics 2018

11.3 POLYCHROMATIC PRIMARIES

As well as the simple case where both w_1 and w_2 are monochromatic, radiant spectral sets may be generated by combining polychromatic spectra. Figure 11.6 illustrates the colour set generated by combining the D65 standard daylight spectrum (**D65**) with a monochromatic 610 nm light.

11.4 VISUALISING ANOMALOUS TRICHROMACY

Anomalous trichromacy is a type of colour vision deficiency affecting 5-6% of the Caucasian male population (Sharpe et al., 1999); colour perception in such individuals is still three-dimensional, but with an altered spectral sensitivity of the L or M cones, giving a reduced wavelength separation between the L and M sensitivities. While this

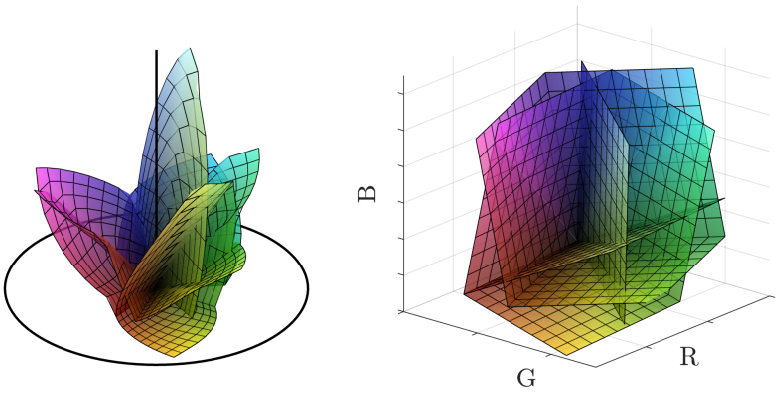


Figure 11.5: Left: $C\{w_1, w_2\}$ spectral sets projected to HSV space. Black circle = plane of saturated hues; central vertical line = grey axis. Right: the colour sets projected to RGB space. Sets were selected to span the spaces. In both projections, metamers are clearly identified by the intersections of two or more surfaces.

results in impaired discrimination overall, some radiant spectra that are metameric to normal observers may be discriminable to the anomalous trichromat. Colour vision deficiencies have been simulated (Brettel et al., 1997), but visualisation tools for the relationship between radiant spectra and an individual's sensor space are lacking. Figure 11.7 shows the projection of two $C\{w_1, w_2\}$ spectral sets to normal- and anomalous trichromacy-defined RGB spaces. Matrices show inter-surface inter-point distances (darker = shorter distance), with distances below a threshold coloured

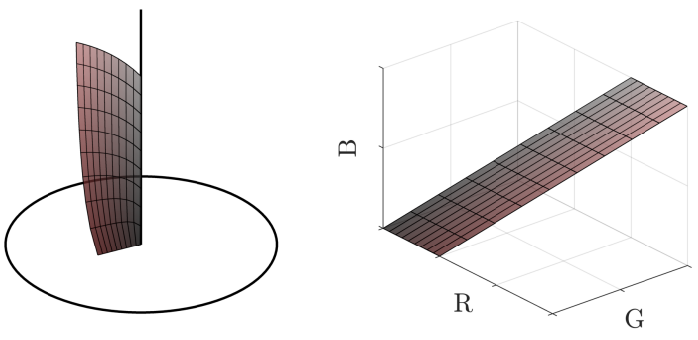


Figure 11.6: The surface generated by the spectral set $C\{\mathbf{D65}, 610\}$.

Published in *Eurographics 2018*

red, corresponding to metamers (i.e. difference signals overwhelmed by neural noise). While there are more metamers in the anomalous space (2.09%, vs 1.50% for the normal space), some spectra that are metameric in the normal space are now discriminable. This is further illustrated by a difference matrix.

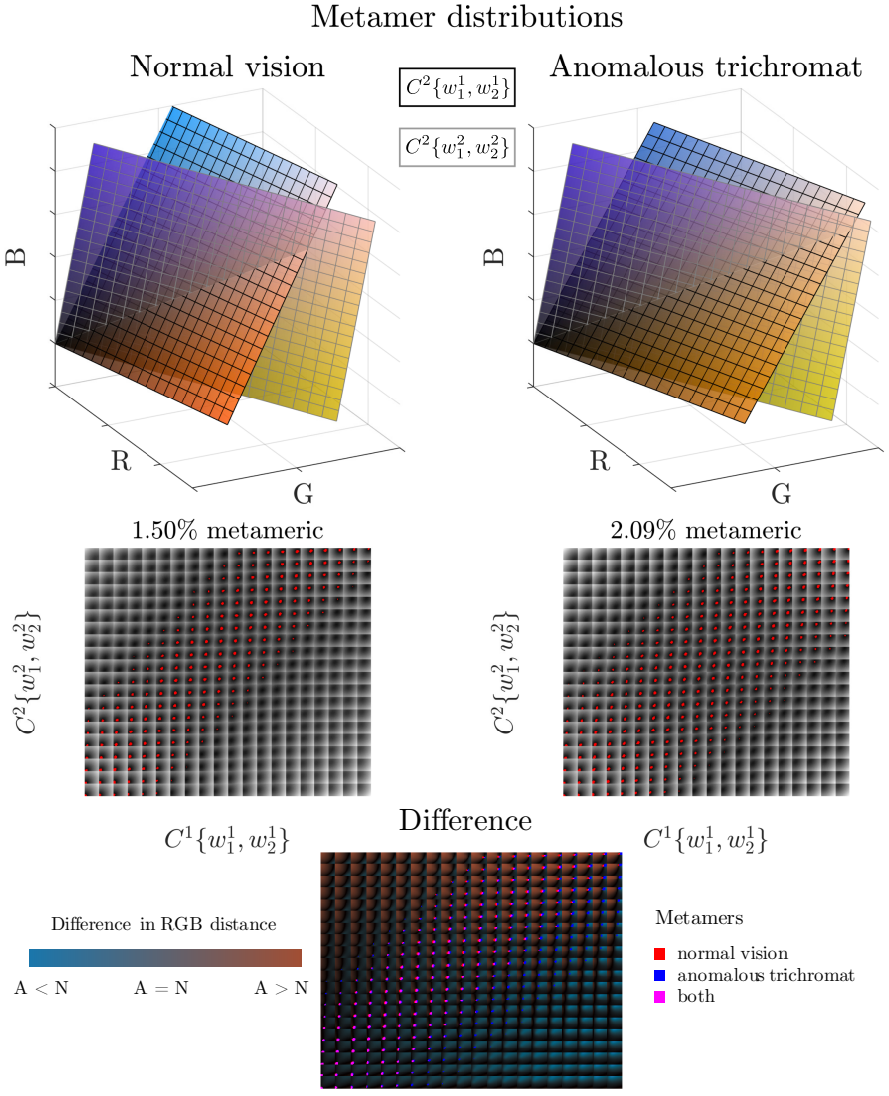


Figure 11.7: Two $C\{w_1, w_2\}$ spectral sets projected to RGB spaces defined by normal vision and anomalous trichromacy. While there is a greater number of metamers in the anomalous space (coloured red in the matrices), some colours that are metameric in the normal space have inter-point distances above threshold.

Published in Eurographics 2018

11.5 THREE-WAVELENGTH RADIANT SPECTRAL SET VOLUMES

Just as we may project $C\{w_1, w_2\}$ radiant spectra to surfaces, so too can we project three-primary $C\{w_1, w_2, w_3\}$ spectra to volumes. The resultant volume would be an opaque solid, but instead may be regularly sampled for combinations of primaries, as in Figure 11.8. The addition of a fourth primary may be visualised with animation.

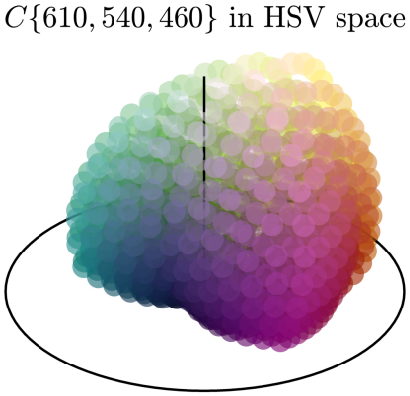


Figure 11.8: The three-primary radiant spectral set $C\{610, 540, 460\}$.

11.6 CONCLUSION

In this chapter, the three-dimensional understanding of colour (as first theorised by Robert Grosseteste in the thirteenth century) was expanded to consider the infinite-dimensionality of radiant spectra, and their relationship to the normal and anomalous-trichromat sensor spaces of human colour perception, with the aid of a visualisation framework. The following chapter will illustrate this relationship further with an illusion that exploits the properties of illuminant spectra, surface reflectances, and human colour vision.

12

THE NEON FRUIT ILLUSION

Published in *Perception*, Vol. 48, No. 3, 2019.

AUTHORS

Joshua S. Harvey, Takuma Morimoto, and Manuel Spitschan

ABSTRACT

At this year's European Conference on Visual Perception we debuted a novel colour science demonstration—and visual illusion—for the *Un mare di illusioni* exhibition. Under carefully curated lighting conditions, cycling through different illuminant spectra, certain fruits and vegetables appear to glow and dim in an unchanging environment. Encouraged by the positive reactions it received, and the numerous and specific questions from conference delegates, we here describe what this illusion is, why we believe it may work, and how this particular low-cost setup may be assembled and demonstrated for the amazement of your friends, students, and members of the public.

Published in *Perception*



Figure 12.1: A sequence of frames taken from a video of the neon fruit illusion.

12.1 INTRODUCTION

Human colour vision is remarkably stable across a range of conditions. One of the mechanisms underpinning this stability is colour constancy, when objects do not change their colour appearance despite changes in environmental lighting, even when the objects' reflected spectra may change considerably. This is a significant achievement of the human visual system, as colour constancy is a mathematically ill-posed and under-constrained problem. Our novel *neon fruit illusion* capitalises on this, exploiting the physical conditions of lighting changes when colour constancy inevitably breaks down.

The illusion relies on both the dimensionality reduction from spectral to trichromatic information, and the distinct reflectance properties of different objects. Two illuminant spectra may themselves be metameric, yielding the same amount of cone excitations and therefore causing no noticeable change in lighting. Depending on their respective reflectance functions, objects in a given scene may also yield the same amount of cone excitations under the two illuminants, therefore not exhibiting a change in luminance or chromaticity. The *neon fruit illusion* occurs when a largely non-noticeable modulation in illumination causes a pronounced change in the appearance of *some* objects in the scene, while others remain constant. A reasonable inference from this stimulus is therefore that the object itself is emitting light, or glowing.

In fact, the illuminants need not be nominally metameric, as long as the perceptual difference between them is very small relative to the change in 'glowing' objects. An effective and affordable setup illustrates this effect using a commercially available four-primary LED bulb. If the scene is flooded with its relatively narrowband green channel, adding small amounts of the red channel will not noticeably affect the

Published in *Perception*

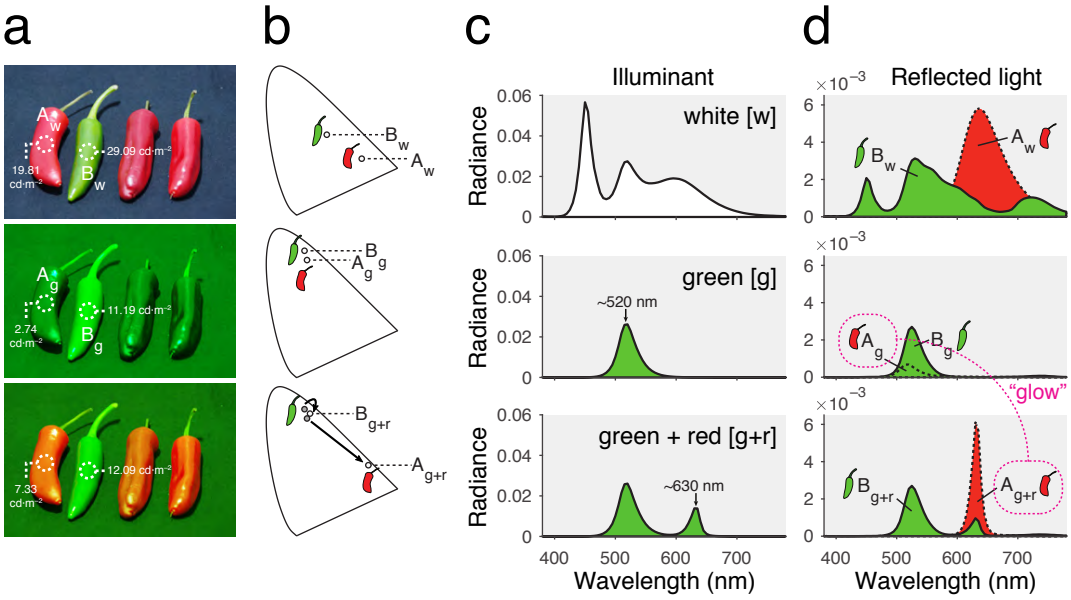


Figure 12.2: a) Typical stimuli, in this case an assortment of red and green peppers, under white, green, and green + red LED illumination. The change between the latter two illuminants causes a pronounced increase in the luminance of the red peppers, while the green peppers show little change. b) The chromaticity of the red peppers also shows a substantial shift. c) and d) The spectra of different illuminant conditions, and the measured reflected spectra of the peppers shown in (A), respectively. Radiance units are mW/cm²/sr/nm.

chromaticity of the light, or the colour appearance of most objects (if they are white, yellow, green, blue etc. under white illumination). However, reddish objects that were previously 'in the dark' will exhibit a strong change in luminance, transitioning from dark grey/black to brightening red. Spectral measurements from this setup are shown in Figure 12.2. Observers will typically ascribe this colour change not to the lighting, but to the objects themselves, either due to glow or colour change. This is not an impossibility—the objects could indeed have light sources buried within them; it therefore meets the criteria of an illusion suggested by (Todorovic, 2018).

These manipulations do not merely cause changes in the apparent surface colour, but can evoke a strong sensation of glow (where an object appears to emit light) (Katz, 2014). This phenomenon has been investigated in past studies with relatively simple stimuli (Bonato and Cataliotti, 2000; Bonato and Gilchrist, 1999; Speigle and Brainard, 1996; Uchikawa et al., 2001; Yamauchi and Uchikawa, 2000, 2005). It is

Published in Perception

worth noting that such a sensation can occur even for natural objects, despite our priors that peppers should not be emitting light.

Interestingly, when shown as still images such as in Figure 12.2, the impression of glow weakens or may be entirely absent. It may therefore be the dynamic manipulation of spectral illumination that evokes strong sensations of glow in this case, and temporal modulation may be a fruitful avenue of research for further investigations into perceived glow.

We debuted such a setup at the 41st *European Conference on Visual Perception* evening exhibition *Un mare di illusioni*. Sourcing a range of fresh fruits and vegetable from the local DESPAR supermarket in Trieste, we obtained fruits and vegetables with a broad range of natural reflectances. We set up shop in our own private room within the dock warehouse Magazzino 42, which gave us excellent control of the lighting. The fruits and vegetables were placed on a piece of standard black-out cloth on a table, and illuminated by four LiFX bulbs (LiFX A60 LED Light) under software control (IFTTT running on an Android smartphone). The resultant exhibit is shown in frames of a video recorded on a mobile phone (without any post-processing) in the *Abstract*, and displayed in Figure 12.3 as a video-slice diagram.

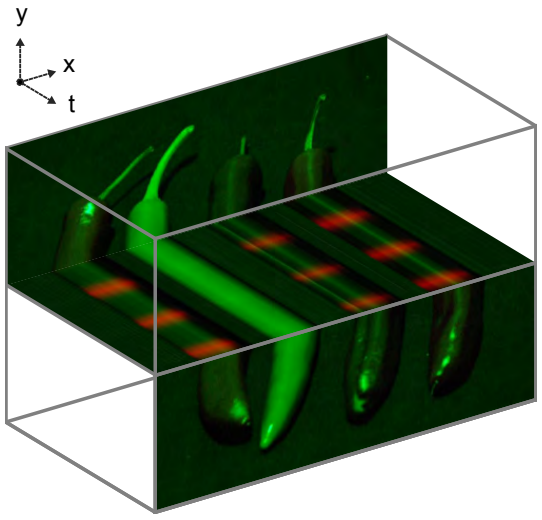


Figure 12.3: A slice-through view of a neon fruit illusion video. Certain regions of the image feature a pronounced colour change from dark grey to red through time, while the rest of the image remains constant.

Published in *Perception*

If you would like to prepare this illusion yourself, we present instructions below in recipe format.

12.2 INGREDIENTS

- 1 or more controllable multicolour LED bulbs (e.g. LiFX A60 LED Light);
- 1 black table cloth;
- Fresh (or fake) fruits and vegetables of a wide range of colours.

12.3 RECIPE

- Take the produce, and arrange to your liking on the black cloth.
- Immerse the fruits and vegetables in light emitted from the bulb's green channel.
- Add a dash of the red channel to taste, until the desired mixture is obtained and some of the fruits and vegetables are nicely reddened. A little goes a long way, and too much will overpower the dish.
- Remove the red channel component, subsequently adding and removing in a gradual manner.
- Serve immediately and enjoy. Alternatively, the dish may be videoed with a standard camera and enjoyed another time.

For best results, use LEDs of narrowband spectra and ensure the recipe is prepared in a very dark environment. It is worth noting that the illusion is easy to set up and allows many people to experience it at the same time. For maximum surprise, consider using a set of spray-painted plastic bananas, and revealing their true colours with a white illuminant at the end. *Buon appetito!*

In case fresh fruit and vegetables, or controllable multi-primary LEDs, are unavailable, the neon fruit illusion may also be simulated. This would not be possible with standard three-channel photographs, but a good approximation can be achieved

Published in *Perception*

using a recently published hyperspectral database of fruits and vegetables (Ennis et al., 2018). The results of a lime and a red pepper, illuminated by a narrowband green LED with and without the minor contribution of a narrowband red LED, are shown in Figure 12.4. Although the database does not contain actual reflectance measurements, these were estimated from hyperspectral image regions taken to contain flat matte reflectance, dividing out the illuminant spectrum.

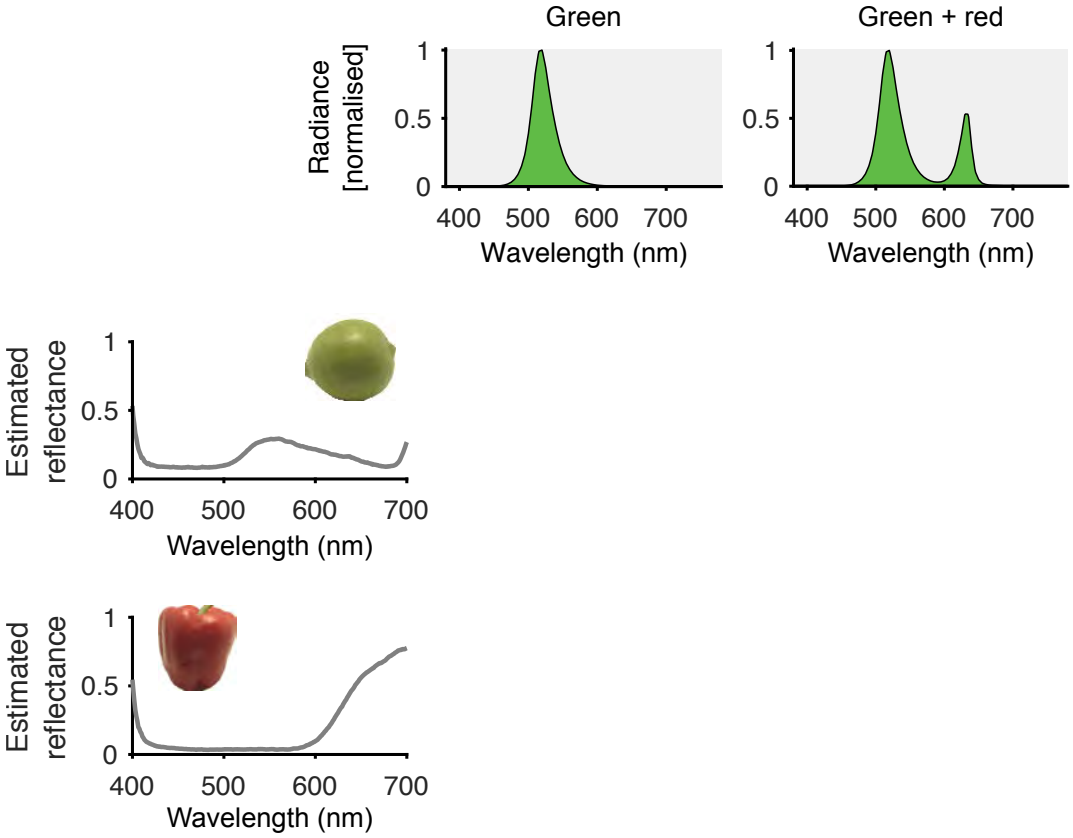


Figure 12.4: A simulation of the neon fruit illusion, using hyperspectral images of a lime and a red pepper from a database of hyperspectral images of fruits and vegetables (Ennis et al., 2018). We estimated reflectance of the lime and red pepper by dividing the spectrum of reflected light reported in the database by a measurement of the illuminant spectrum at a pixel location free from obvious specular highlights. We then rendered the lime and the repper under the illuminants given in Figure 12.2c using the software provided by Ennis et al., 2018.

Whether consumed fresh, preserved as a video, or simulated from hyperspectral images, the neon fruit illusion provides an engaging and surprising demonstration of colour science.

Published in *Perception*

12.4 ACKNOWLEDGEMENTS

The authors wish to thank the ECVP 2018 organising committee for allowing us to showcase the illusion; everyone who attended the *Un mare di illusioni* exhibition for a fun and stimulating evening; Cécile Gal for her discerning eye in fruit and vegetable selection; and Robert Ennis for providing the illuminant spectrum. JSH and TM received W.D. Wright Travel Awards from the Colour Group (GB) to attend ECVP 2018. JSH is funded by the Andrew W. Mellon Foundation and Clarendon Foundation through the Oxford interdisciplinary research centre, the Oxford Research Centre in the Humanities (TORCH). TM's DPhil studentship is funded by the Aso Scholarship, and awards from the Sasakawa Fund, the Kikawada Foundation and the Japanese Student Services Organization. MS received the AVA Tom Troscianko Travel Award to attend ECVP 2018 and is supported by a Wellcome Trust Sir Henry Wellcome Postdoctoral Fellowship (204686/Z/16/Z) and a Junior Research Fellowship from Linacre College, University of Oxford.

13

MEDIEVAL VISUAL EFFECTS IN POLYCHROME SCULPTURE

13.1 INTRODUCTION

Far from dark or drab, the medieval period offered an assortment of visual delights within the privileged context of religious practice. While everyday life featured a muted palette, devoid of the synthetic dyes and shiny plastics that have come to dominate the modern visual landscape, churches offered an exposure to the arts to all members of society. Paintings, stained glass, and sculpture greeted the medieval observer with a gamut of hues and material qualities not encountered outside religious experience, providing both a treat for the sensorium and narrative representations for a preliterate society. Nowhere is this better typified than in *polychrome*. While *polychromy* is a more limited term referring only to the presence of multiple paint layers (from Greek *poly-* 'many' + *chroma* 'colour'), polychrome refers to the widespread medieval artistic practice of applying diverse materials including chalk ground, paint, gilding, relief, and ornaments to a single substrate, such as a wooden sculpture or panel (Nadolny, 2000).

Within this artistic tradition, there is a recurrent combining of metal-leaf applied below an overcoat of transparent, colourful glaze. When a colourless metal, such as tin or silver, is set below a yellow glaze, this gives rise to a bright golden material finish. Indeed, it could be glibly referred to as an 'imitation gold' material. However,

a closer examination of both the material, and its use in medieval aesthetic practice, invites a more nuanced appreciation.

This chapter presents an overview of this material, including its application and cultural significance in European church art in the medieval era, in the light of recent scholarship (Kollandsrud, 2018; Pereira-Pardo et al., 2019). The physical composition of the material is then discussed, and how it might give rise to particular optical phenomenon. These characteristics are further evaluated with the use of both physically-based rendering computer graphics simulations, and physical measurement of samples produced according to historically accurate methods. Lastly, perceptual qualities are discussed, with a particular focus on how polychrome materials might affect image saliency and visual attention when present on sculptures.

13.2 MEDIEVAL METALLICITY AND GILDED POLYCHROME

While metallicity was not such a common feature of the medieval visual landscape as it is today, observers frequently encountered a variety of metallic materials and surface finishes, some of which fell out of use in later historical periods. From antiquity, and climaxing in the Islamic Golden age and Western medieval period, glazed ceramics and glassware featured a so-called 'metallic lustre' (Caiger-Smith, 1985; Pradell, 2016). This lustre is caused by nanocomposite layers of copper or silver, dispersed within the matrix of the glaze or glass, the physics of which has attracted significant investigation (Bobin et al., 2003; Borgia et al., 2004; Molera et al., 2007; Pérez-Arantegui et al., 2001; Pradell et al., 2005). These layers of only a few hundreds of nanometres thickness—the first reproducible nanostructured thin metallic film made by humans (Pérez-Arantegui et al., 2001)—give rise to striking iridescent appearances, thanks to the quantum confined optical response of metallic nanoparticles (i.e. surface plasmon resonance).

The optical and structural properties of these ceramics have been investigated with a range of techniques, including scanning electron microscopy (Borgia et al., 2004;

Pérez-Arantegui et al., 2001), energy-dispersive X-ray spectroscopy (Bobin et al., 2003; Pérez-Arantegui et al., 2001), X-ray photoelectron spectroscopy (Pérez-Arantegui et al., 2001), transmission electron microscopy (Borgia et al., 2004; Pérez-Arantegui et al., 2001; Pradell et al., 2005), spectroradiometry Bobin et al., 2003, UV spectroscopy (Borgia et al., 2004; Molera et al., 2007), X-ray fluorescence analysis (Borgia et al., 2004; Pradell et al., 2005) and X-ray diffraction (Molera et al., 2007; Pradell et al., 2005). Little focus has fallen on their perceptual qualities, despite their physical characteristics locating them at an interesting juncture between metals and ceramics. The implications that these characteristics may have for perceptions of colour and material are prefigured by a twelfth-century Persian treatise on making coloured and enamelled glass, contained within the *Bayan al-Sana'at* [The Explanation of Crafts]. Within this text the writer, alchemist Hobaysh Teflisi, repeatedly associates the desired colour of the glassware with metals (translation by Holakooei (Holakooei, 2016)):

> To make a white glassware like silver... To make a yellow glassware like gold...

These passages are particularly interesting given recent studies into colour naming and metals (Komatsu et al., 2013; Okazawa et al., 2011; Yang et al., 2013), as well as the perceptual overlap between the appearances of metallic and transparent stimuli (Tamura et al., 2018).

Gold adorned medieval churches and luxury objects in a variety of forms. In its solid state, as jewellery and coins, but more commonly applied as gold leaf. Gold is the most malleable of metals, and medieval artisans exploited this material quality, placing small amounts obtained from streams (*placer*, or alluvial gold), from the Rhine, Tiber, Po, Rhone, and Garonne rivers, between two pieces of leather and hammering to give gold leaf (Benton, 2009). Alternatively coins were used, with Florins—24-carat gold Florentine coins—being the most highly prized (Berzock, 2019; Buccolieri et al., 2009). The gold leaf was then used in manuscript illumination, the gilding of wooden panels, sculptures, and incorporated into the tesserae (tiles) of

mosaics, the latter most notably featuring in Byzantine art (Neri et al., 2017; Neri and Verità, 2013; Neri et al., 2016).

In medieval polychrome objects we find another example of lustrous, metallic appearance, which cannot be discretely identified with the common categories of material models: conductors and dielectrics[1]. It is a composite material, comprising an achromatic metal base layer, either tin or silver leaf, set beneath a transparent, colourful glaze. When the glaze has a yellowish chromaticity, such as pine tar resin, the resultant appearance is similar to that of gold. Recent studies have resulted in historical recreations of this material, produced according to medieval tar burning methods[2]. Pine tar extraction is a lengthy and involved process. It requires a large amount of dry, pitch-rich pine roots to be placed in an inverted cauldron, resting on a large piece of slanted slate. Grooves are incised into the stone, allowing the resin to pool as it releases over the course of about an hour, while a small fire burns on the top of the cauldron. Over the course of the process flammable terpenes are released, posing a not insignificant risk if the fire is not separated by a physical barrier from the 'collection zone'.

13.3 OPTICAL CHARACTERISTICS

The 'imitation gold' material consists of just two layers, yet this simple composition gives rise to a complex set of optical interactions, shown graphically in Figure 13.1.

To quantify the effects of these optical interactions, the material was explored in two ways. First, a collaboration with researchers at the Cultural History Museum, University of Oslo, enabled us to produce physical samples of a flat, square geometry according to historical methods. These were then used to obtain spectroradiometric

1 These categories, frequently used in computer graphics rendering models, come from the electrical properties of materials. Ordinarily, metals are categorised as 'conductors', while all other materials are 'dielectrics'. Because light-material interactions are principally determined by their electro-chemistry (and secondarily by their geometric structure), this is normally a reasonable classification.

2 Recently, the Madonna from Biri church, Norway, was reproduced according to these methods. The Tjærebrenning (tar burning) report from the workshop at Bekkbråten, in 2012, describes this process from pine tar extraction through to application on the sculpture.

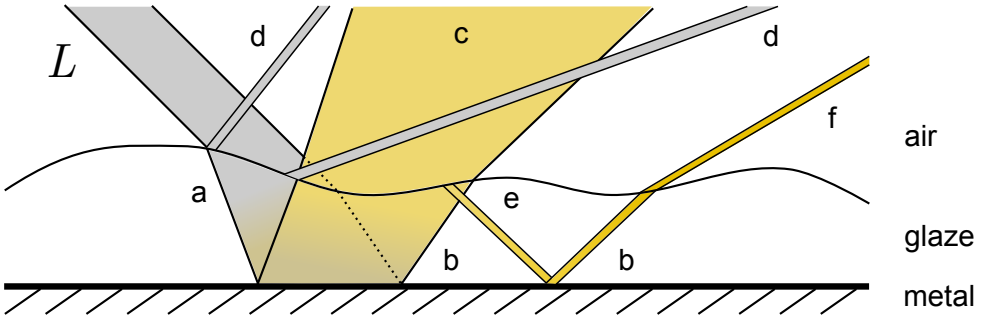


Figure 13.1: A graphical overview of the optical interactions at play in the medieval polychrome 'imitation gold' material. The incoming illumination, *L*, is subject to lensing (refraction) upon entering the uneven surface of the transparent glaze. a) As it travels through the glaze, it takes on increasingly deep colouration, as shorter wavelengths of light are absorbed. b) Upon striking the metal base layer, it is specularly reflected. c) The majority of incoming light then leaves the material, as a yellowed, geometrically distorted image of the environmental illumination. d) A small fraction of light will specularly reflect off the air-glaze boundary, before it penetrates into the material, giving rise to achromatic highlights, and binocular glitter. e) Another small fraction of light will be internally reflected instead of exiting the material, which will then be subject to further colouration, giving rise to chromatic highlights and glitter (f).

measurements, to give an initial descriptor of material appearance. Second, hyperspectral, physically-based computer graphics rendering was performed, to simulate similar samples. This allowed a more exact characterisation of reflectance properties.

13.3.1 *Spectroradiometry of physical samples*

Five flat, square samples of the 'imitation gold' polychrome material were produced by Kaja Kollandsrud in Oslo, Norway, and sent to Oxford for physical measurements. Using a broadband white LED lamp for illumination, the spectrum of light reflected off the samples was measured for different angles of reflectance (with equivalent angles of incidence), using a SpectroCAL MkII spectroradiometer (Cambridge Research Systems). Three different samples were measured, and the means for each angle of incidence/reflectance were taken. Spectra were normalised using a diffuse, white standard with a known reflectance spectrum. The measured spectra are shown in Figure 13.2. The chromaticity of each spectrum was then calculated using the

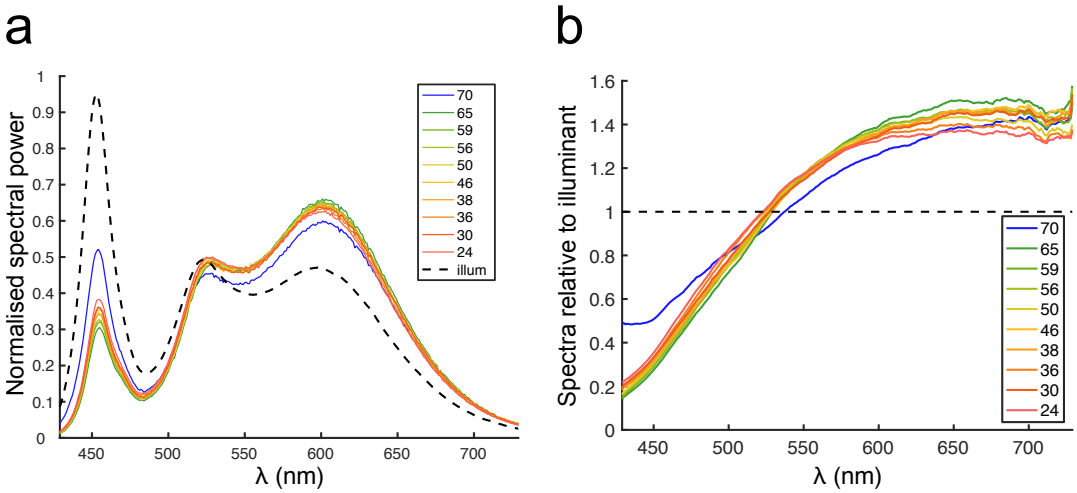


Figure 13.2: Spectroradiometric measurements of physical samples of the medieval polychrome 'imitation gold' material. a) Mean spectra across repeated measurements for different angles of incidence/reflectance of the flat samples. Spectra are normalised such that each spectrum has the same total power across all wavelengths. b) Spectra normalised by the illuminant.

standard observer 2° cone fundamentals. These chromaticities are plotted onto the u'v' isoluminant plane[3] in Figure 13.3.

It is clear that as angles of incidence and reflectance increase from 0° (i.e. normal to the sample plane), the yellow chromaticity of the reflected light begins to deepen, moving further away from the achromatic illuminant. However, at 70°, there is a sudden shift back towards the achromatic illuminant, indicating a weakening of yellow chromaticity. Presumably, this is because at grazing angles, incident light is more likely to specularly reflect off the air-glaze boundary, before penetrating into the material, resulting in an achromatic 'highlight'.

13.3.2 *Physically-based rendering simulations*

To further investigate the reflectance properties of this composite material, the square samples were simulated using physically-based computer graphics rendering. This

3 Although this is a rough linear approximation for perceptual colour space, given that human colour perception is non-Euclidean, it is nonetheless recommended by the CIE for some applications (Ohno and Blattner, 2014). So, distances in the u'v' chromaticity plane are roughly equivalent to perceptual dissimilarity, although this approximation worsens with increasing distance from the Planckian locus.

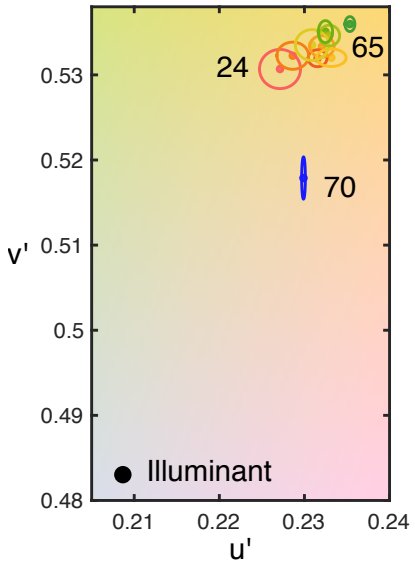


Figure 13.3: Chromaticity variations of the 'imitation gold' physical samples as a function of angle of incidence/reflectance. Angles of reflectance/incidence are colour-coded the same as in Figure 13.2, with some indicated on the plot. For each angle, the mean chromaticity of three measurements is plotted, with the surrounding ellipse indicating the standard deviation for three measurements.

approach endeavours to accurately simulate how light interacts with materials to form images as a function of scene-observer geometries. Scenes were constructed in the open-source 3D illustration and animation software Blender (Blender Online Community, 2016), and hyperspectral ray-tracing was computed using the rendering engine Mitsuba (Jakob, 2010). The scene was configured to contain a flat square mesh assigned with a silver `conductor`[4] Mitsuba material, set beneath a `homogeneous medium` which absorbs in the blue region of the spectrum. The absorbance of the medium was adjusted for similarity to the physical samples. The surface of the medium was subject to a smooth, random displacement field to give the slight unevenness present in the physical samples, and assigned an index of refraction (IoR) = 1.4.

Two sets of scenes were rendered for two illumination conditions: one set with ambient, environmental illumination from a spherical light probe, and the other with an additional direct light source, emitting equal energy white (EEW). Images were taken at 5° intervals rotating around the object, giving angles of reflectance

4 Rendering parameters are printed in monospace.

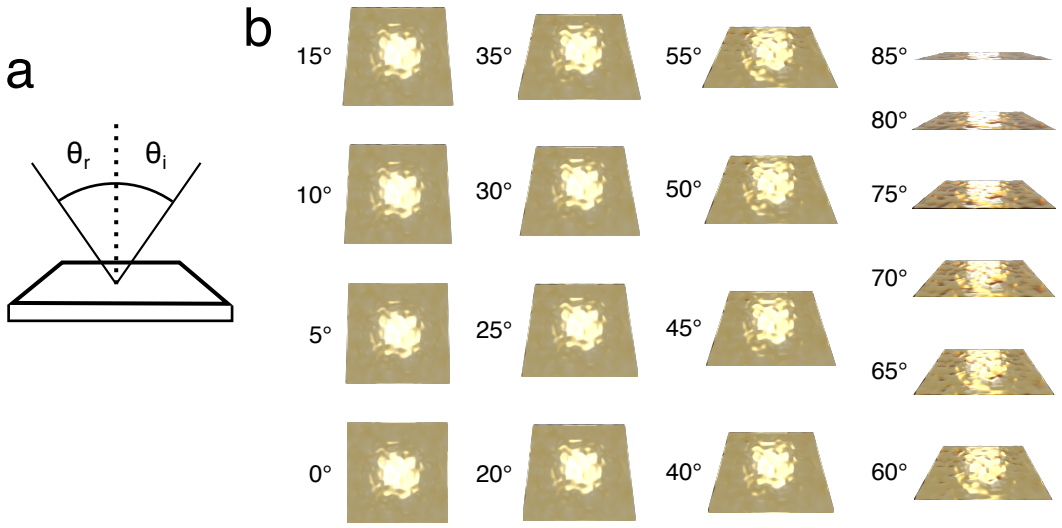


Figure 13.4: a) Measurements and simulations had a scene geometry such that the angle of incidence, θ_i, was equal to the angle of reflectance, θ_r. b) The results of simulating the polychrome 'imitation gold' material with physically-based computer graphics rendering, for different angles of reflectance. Results are shown for the 'direct' lighting condition, which included an EEW light source in addition to the environment light probe.

(and incidence, for the 'direct' lighting condition) between 0 and 85°. Rendering was performed with the `extended volumetric path tracer`, to 256 samples. The renders are shown in Figure 13.4.

As with the physical spectroradiometric measurements, the mean chromaticity for each reflectance angle was then calculated, again by converting the 31-dimensional hyperspectral data to u'v' values. The resultant chromaticity plots are shown in Figure 13.5 for both the 'ambient' and 'direct' lighting conditions.

The results of this analysis agree with the findings from the physical measurements, namely that the reflectance function of the material is complex—as the angle of reflectance increases, the chromaticity of light reflected off the object deepens in yellowness. This is likely due to the reflected light having to take an increasingly longer path through the glaze, attenuating more short-wavelength (blue) light. But, at a critical angle of around 70° (depending on the IoR of the glaze), this abruptly changes, as the reflected light becomes less yellow, and its chromaticity is increasingly like that of the illuminant. This is due to reflection from the glaze-air boundary at glancing angles.

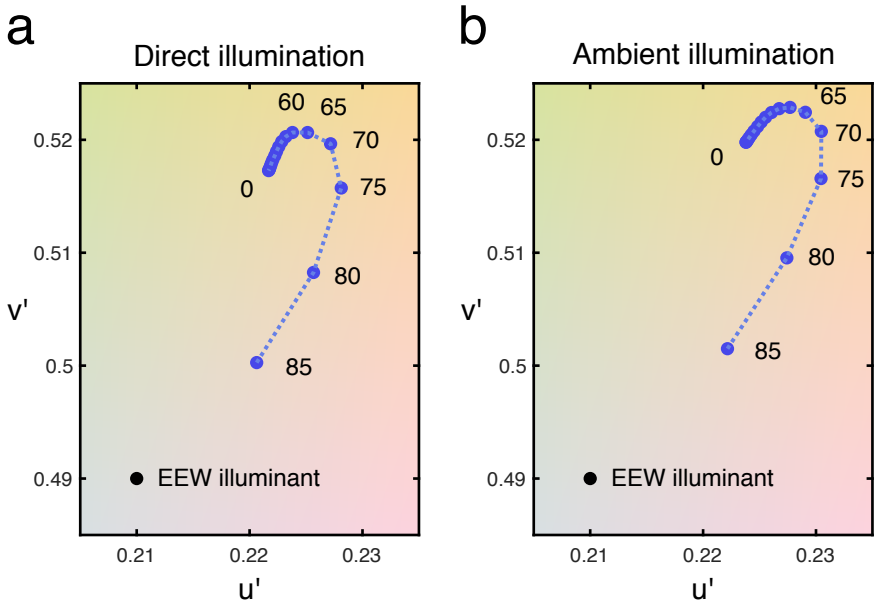


Figure 13.5: The mean chromaticities of images obtained by simulating the 'imitation gold' material. The flat square simulated sample was illuminated under a) direct illumination, for which an EEW was included in the scene, opposite the camera at an equal angle, and b) exclusively ambient illumination.

Further evidence in support of this comes from inspecting the colour distributions of pixels in the renders, which reveals a distinct bimodal distribution at glancing angles. As we are primarily interested in yellowness, we can limit our scope of inquiry to the blue channel of images in sRGB space. If we fit distributions using kernel density estimation, we see that higher angles of reflectance (above 70°) have a bimodal distribution, shown in Figure 13.6.a.

Lastly, we can also use this approach to look for changes in luminance as a function of reflectance angle, computed as a weighted intensity of the sRGB channels. The mean luminance of simulated images of the 'imitation gold' material, rendered under both 'direct' and 'ambient' lighting conditions, is plotted in Figure 13.6.b as a function of reflectance angle. An ideal mirror, or metallic surface, would have a close to flat luminance profile with respect to reflectance angle, but instead we see strong changes in both the direct and ambient conditions. This suggests that a three-dimensional object clad in this material may retain some shape-from-shading cues, despite its metallic appearance.

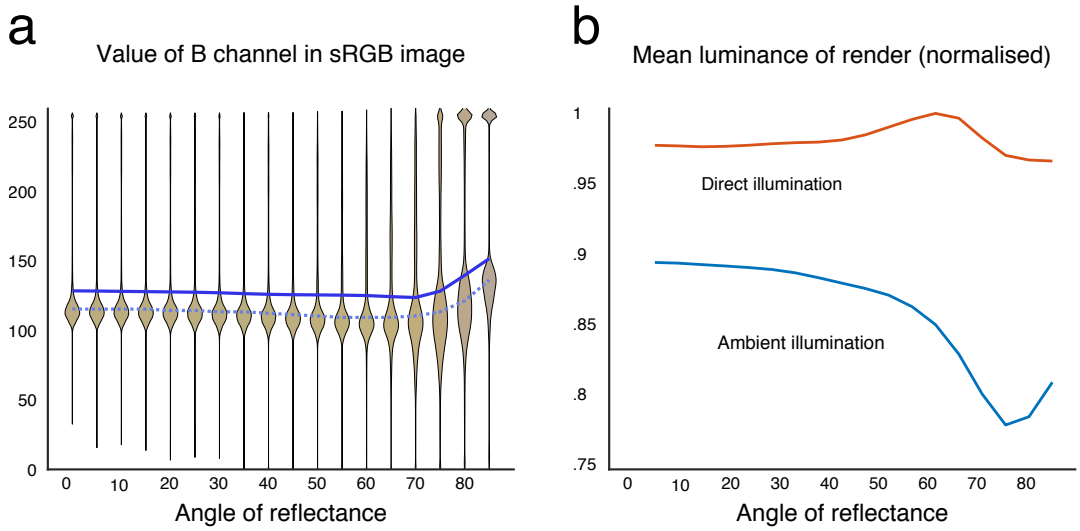


Figure 13.6: Chromaticity and luminance as a function of reflectance angle in simulated images of the 'imitation gold' material. a)The distributions of pixel colours in simulated images of the 'imitation gold' material, indicated by the value of the blue channel in sRGB images. As angle of reflectance increases from 0, the blue channel depletes further, until around 70°, when the mean (solid blue line) and median (dotted blue line) start to increase. A bimodal distribution can be seen above 70°. Each violin plot is coloured according to the mean chromaticity of the rendered image. b) Mean luminance of 'imitation gold' material simulations as a function of reflectance angle. Both the 'direct' and 'ambient' illumination conditions are shown. Data have been normalised by the brightest mean.

13.4 PERCEPTUAL QUALITIES

From the physical measurements and simulations performed here, there is evidence that the 'imitation gold' medieval polychrome material exhibits reflectance properties starkly different to that of real gold. Real gold has a reflectance distribution function that is constant with respect to angles of incidence and reflectance, analogous to an ideal mirror convolved with a filter of yellow or yellow-orange chromaticity (depending on its chemical composition and purity). However, initial analyses show that the polychrome material exhibits a surprising and complex reflectance distribution function, with variations of chromaticity as well as luminance, as a function of reflectance angle. An intriguing question to ask is whether medieval artisans would have been aware of these differences in optical qualities, whether they contribute to distinct perceptual qualities when objects clad in these materials

are situated in their environments, and if so, whether these aesthetic qualities might be driving material choices in medieval sculpture.

To investigate potential effects on object perception, physically-based rendering was performed to simulate how statues would appear when clad in the 'imitation gold' material, and how this might differ to objects decorated with real gold. First, spheres were rendered in order to inspect luminance profiles, as shown in Figure 13.7. Four materials were rendered: smooth gold, slightly rough gold (`ggx` distribution of 0.06), the 'imitation gold' material (silver with a `ggx` distribution of 0.06, beneath a slightly uneven homogeneous medium with an IoR of 1.4), and a yellow diffuse material for reference. The scene contained low-level constant environment illumination, as well as a bright achromatic lamp. The luminance profiles, taken through the central horizontal of the spheres, are informative in illustrating how much shape-from-shading information is available in the image. The flattest of these, for the smooth gold object, gives essentially no information, while the yellow diffuse object shows the greatest variation. The slightly rough gold object has a luminance profile with a central peak, corresponding to a the specular lobe of the emitter, but no variation throughout across the other regions of the object. The imitation gold profile, though noisier due to the irregularity of the coating, has additional decreases in luminance at the edges of the profile, providing a subtle cue to three-dimensional shape.

13.4.1 *Material simulations of sculptures*

To see if different materials resulted in distinct looking behaviours for objects, sculptures were simulated using physically-based rendering. A three-dimensional model of the sculpture 'Penelope' was used, taken from a collection of high-resolution laser-scans created by Oliver Laric (Laric, 2016). The model was rendered with the three material conditions used in the previous analysis: smooth gold, rough gold, and the polychrome 'imitation gold' material, using Mitsuba physically-based rendering materials. The results of the rendering simulations are presented in the top panel of

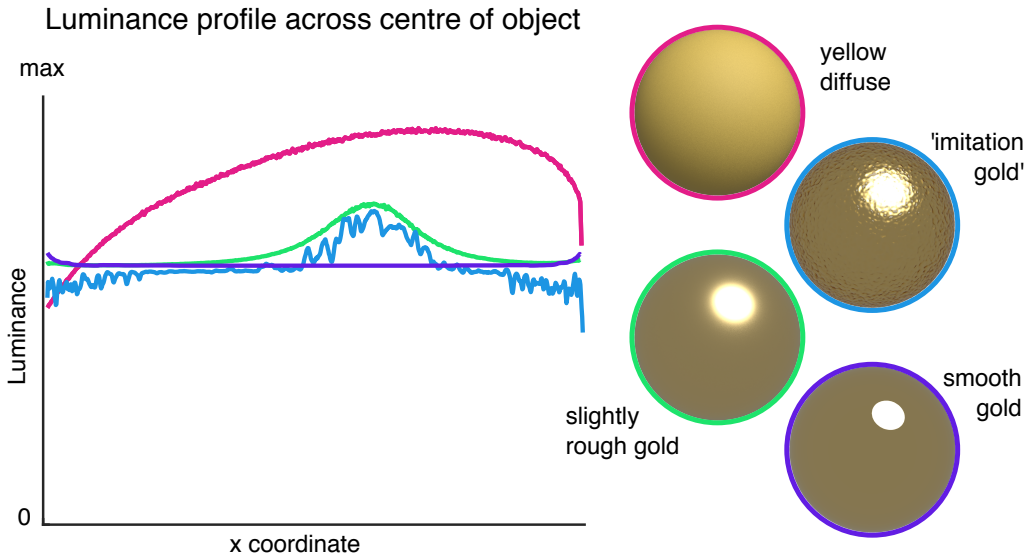


Figure 13.7: Shape-from-shading cues in medieval polychrome 'imiation gold'. Four different materials—yellow diffuse, smooth gold, slightly rough gold, and polychrome 'imitation gold' are rendered as spheres, and the luminance profile through the central horizontal of the image is plotted. The polychrome material was configured such that its silver base had the same roughness setting as the 'slightly rough gold' material.

Figure 13.8. An initial inspection of the images suggested that the 'imitation gold' condition was better at making visible the three-dimensional features of the sculpture, and importantly, at communicating the semantic features (e.g. face, pose). To test this, fixation prediction analysis was performed with the DeepGaze II network, the state-of-the-art computational approach for predicting where in images observers are likely to direct their gaze (Kümmerer et al., 2016, 2017)[5]. These predictions are shown in the lower panel of Figure 13.8, underneath their corresponding images.

It should be noted that DeepGaze II was trained on human observer data sets that contained stimuli of natural images, including objects and in particular human faces, so it would be expected to place a high priority on faces and other semantic features present in images as well as merely low-level geometric features[6]. As can

5 It should be acknowledged that this is a simulation of observer viewing behaviour in relation to a simulation of objects' material properties. Ideally, we would have real objects with real observers. However, the creation of real objects was unaffordable. A psychophysics experiment with real observers will be carried out in the near future.

6 Of course, the model was not trained with any images directly similar those analysed here. So, although it might be a reasonable expectation that face-like stimuli attract more fixation prediction density, it is somewhat impressive that the model performs in this manner with inputs that are so distinctly different to its training data.

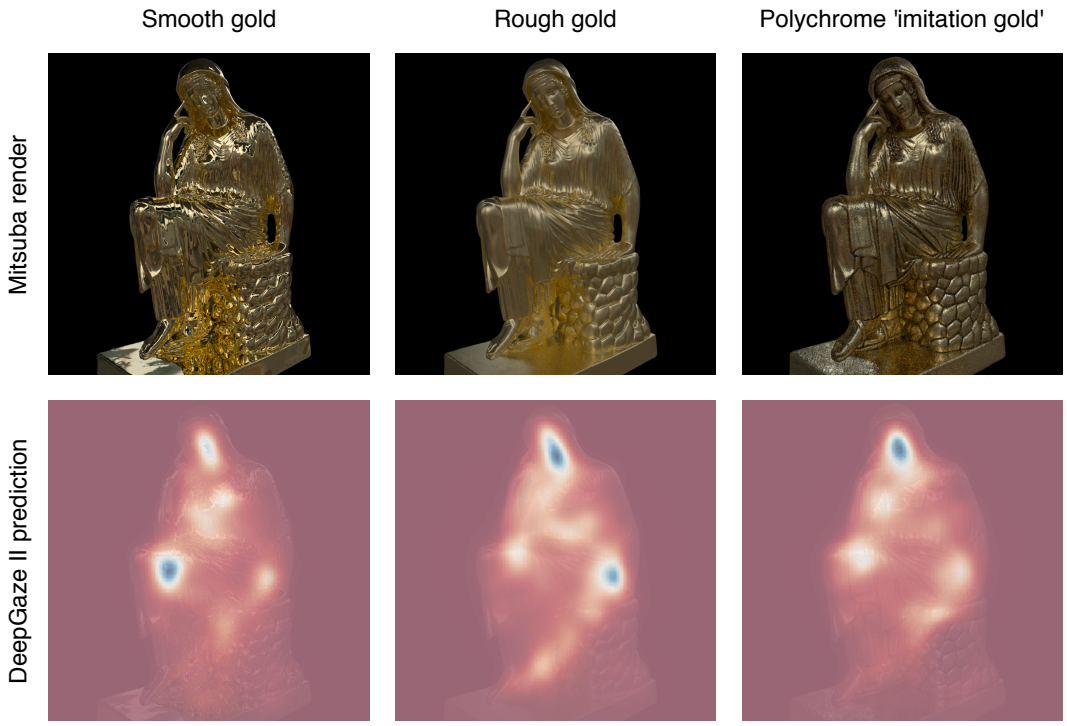


Figure 13.8: Top panel: renders of sculptures clad in smooth gold, rough gold, and polychrome 'imitation gold'. Lower panel: prediction density distributions generated by the DeepGaze II network, shown as a heatmap overlaid on the image.

be seen in the generated 'saliency maps', the smooth gold condition appears to give rise to spurious features, due to the sharply defined highlights and lowlights (spatial variations in the specular reflection giving rise to regions of high contrast). The rough gold condition has a more widely-spread distribution of fixational prediction, with more over the face (a key semantic feature), while also over the rest of the object. The polychrome 'imitation gold' condition, however, shows high density of fixation prediction *only* over the face. To test if this is consistent over viewing position, ten different perspectives were rendered, rotating round the object. Each corresponding saliency map was segmented with an ellipse into regions of the face and the rest of the object. Summary statistics for the renders of each material condition are shown in Figure 13.9.

As shown on the plots in Figure 13.9, there are statistically significant effects on how DeepGaze II distributes fixational prediction as a function of material, as evaluated by a one-way analysis of variance (ANOVA) and multiple comparisons. One-way

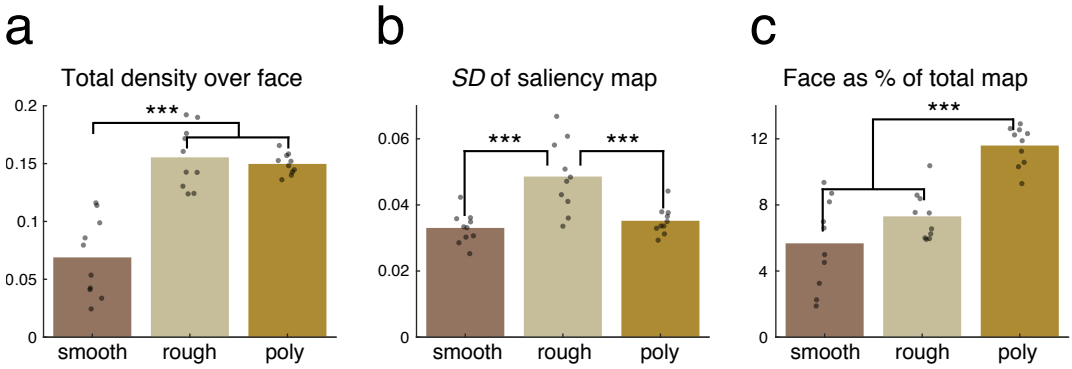


Figure 13.9: Summary statistics of DeepGaze II saliency predictions, as a function of material: smooth gold (smooth), rough gold (rough) and polychrome imitation gold (poly). The mean of each statistic across ten viewpoints is plotted as a bar chart, with each viewpoint represented in a scatter. a) The total density of prediction over the face is far lower for the smooth gold material than either the rough gold or polychrome 'imitation gold'. b) The standard deviation of the saliency maps, which describes how spread out or spatially constrained the distribution of fixation prediction is, is far higher for the rough gold condition than either the smooth gold or 'imitation gold'. This indicates that there are fewer features that are strongly salient, and stand out against the rest of the object. c) As a percentage of the total saliency map, the 'imitation gold' condition has a far higher level of fixation prediction over the face of the sculpture, the primary semantic feature, than either the smooth or rough gold-clad sculptures. ***$p < 0.001$

ANOVAs found a statistically significant effect of material category (smooth gold, rough gold, imitation gold) on each of the three summary statistics—density over the face, standard deviation of the density map, and density of the face proportional to the map. Multiple comparisons were then computed, using Tukey's range test to control for the Type I error rate. The face of the sculpture appears to be more salient to the model when the object is rendered under rough gold and 'imitation gold' material conditions, than the smooth gold condition ($p < 0.001$) This can also be seen directly in the renders, as the features of the face are not clearly portrayed amidst the sharp, bright specular reflections of polished gold. The saliency maps for smooth gold and 'imitation gold' are more tightly constrained than that of rough gold ($p < 0.001$ in both cases), as measured by the standard deviation of the density distributions over the images. These two effects in combination result in a significantly higher percentage of total fixation prediction density falling over the face in the 'imitation gold' condition, than either the smooth or rough real gold conditions ($p < 0.001$).

13.5 EMBODIED LIGHT OR FOOL'S GOLD?

As we have seen from direct physical measurements, as well as physically-based computer graphics simulations of both simple samples and three-dimensional objects, the 'imitation gold' material used in medieval polychrome has some surprising optical and visual characteristics that distinguish it from real gold. Its reflectance distribution is notably non-linear, exhibiting changes in chromaticity and luminance that are absent for pure gold. It also gives rise to distinct visual qualities, such as binocular glitter and lustre, which likely play with and subvert learned (or innate?) expectations or models for material appearance.

It is difficult to make strong claims about the decision-making processes of medieval artisans. Presently, we know little about the relative economic costs and benefits of using one material over another—the historical data seem to indicate that both gold and silver prices (and their ratios) fluctuated chaotically throughout the twelfth and thirteenth centuries, with strong local effects (Munro, 1983). Moreover, there are documented cases of medieval artists using more expensive materials (such as ultramarine), even when they gave virtually indistinguishable results to cheaper, more available alternatives (such as indigotin), so we should not assume that cost-cutting was a priority (Lautié et al., 2005). That said, there seems to be compelling evidence that medieval artisans applied the 'imiation gold' material in their polychrome sculpture specifically for its visual qualities, rather than more practical considerations. Perhaps the most persuasive evidence lies in the examples of polychrome sculpture that contain (real) gold set beneath transparent glazes (Plahter, 2006). In any case, the 'imitation gold' material discussed here, and the broader categories of metallic and coated metal surfaces, present unique visual qualities worthy of further investigation.

Part V

METALLICITY

14

PSYCHOPHYSICAL SCALING METHODS FOR SUPRATHRESHOLD MEASUREMENTS

What is now proved was once, only imagin'd.

—William Blake, 1790

PART ABSTRACT

In this part, investigations are made into the percept of metallicity[1]—the visual quality of an object that indicates to an observer it is made of metal. Specifically, the 'imitation gold' material introduced in chapter 13 will serve as the basis for a psychophysical study into the perception of metallic surfaces, now ubiquitous in our 'visual diet'. First, the quantitative methods of maximum likelihood difference scaling (MLDS), and maximum likelihood conjoint measurement (MLCM), will be introduced, in the historical context of measuring sensory events as functions of physical attributes. These methods will be used in chapter 15 to quantify the relative, and possibly interacting, contributions from both a metal base layer and its transparent coating to the percept of metallicity. Image statistics will be analysed using both Fourier decomposition and steerable pyramids, with the goal of relating material qualities, image statistics, and visual features. Finally, in 16 a spatio-chromatic statistic

1 Readers familiar with the field of astrophysics will no doubt associate this term with the fraction of a star that is constituted by neither Hydrogen nor Helium, but heavier elements, which may or may not be true metals. That is not the intended meaning here.

approach to material quality is explored, through the use of three-dimensional colour distribution warping to confer a gold-like appearance on images of objects.

14.1 DIFFERENCE SCALING

To address questions pertaining to material perception in a rigorous manner, we first need a mathematical approach to studying supra-threshold perception. We are looking to map distances in physical parameter space (e.g. metal roughness) to perceptual dissimilarity space; we are interested in the dimensionality of these spaces, as well as how those dimensions exhibit independence or interactions. The use of notation can be helpful in formulating these relationships. We say that a physical quantity, ϕ_i, has a corresponding perceptual attribute, $\psi(\phi_i)$ or ψ_i. For the most basic type of supra-threshold quantitative psychophysics, we aim to map a physical continuum, $\mathbf{\Phi} = (\phi_1, \phi_2, ..., \phi_n)$ to a perceptual continuum, $\mathbf{\Psi} = (\psi_1, \psi_2, ..., \psi_n)$, in a process now commonly referred to as *difference scaling*.

Early psychophysical investigations sought to map subjective sensory continua, such as perceptual loudness, by the stepwise cumulations of just noticeable differences (JNDs)—a psychophysical measurement formulated by Ernst Heinrich Weber (1795 - 1878) and Gustave Fechner (1801 - 1887). The JND, essentially a measurement of confusability, was an attractive 'unit' for psychometric studies, as even naive observers gave reliable, concordant responses. A serious issue arose, however, as such scales disagreed considerably with those obtained through fractionation, wherein observers are required to estimate the fractional relationships between stimuli, and direct magnitude estimation (Luce and Edwards, 1958). Indeed, it became clear that scales arising from JND summation were, reliably, nonlinearly related to ratio scales of subjective magnitude: a stimulus 100 JNDs above threshold is not twice as great as a stimulus 50 JNDs above threshold, but more than twice as great (Stevens, 1936, 1957). Stevens argued, emphatically, that it was wrong to take the JND—a measure of confusability—as a proper unit of sensation. Given Weber's law, $\Delta S = kS$, which

states that the JND of a stimulus, ΔS, will be proportional to its magnitude, S, we would expect perceptual scales to grow exponentially with JNDs.

The JND is defined as the smallest physical increment in a stimulus that is correctly detected an arbitrary proportion of the time, such as 50% of presentations. To express this metric in terms of a observer's decision, δ, we may write the following:

$$\delta = \psi(\phi_i + JND) - \psi(\phi_i) = 0.5 \tag{14.1}$$

There is another problem with this measure. Supposing that an observer was 'trigger happy', and decided to respond 'yes' on every trial to whether or not a change in stimulus presentation was perceived. According to the JND as it is simplistically defined, even a physical increment of zero would be correctly detected 100% of the time. This measure fails to address response bias and 'false alarms' (responding 'yes' in the absence of the change), which can only be measured with a comprehensive formulation of signal detection theory. So, instead of a JND specifying a decision for a single stimulus increment, what we require is a measure of observer decision making for a continuous range of the stimulus parameter space, accounting not only for correct detection but for correct *response*.

An early attempt to move beyond the rigidity of the JND ('misses' and 'hits') was carried out by Barbara Sakitt. By asking observers to respond with a rating between 0 and 6, in response to stimuli near the absolute threshold for vision, Sakitt sought to unify threshold measurements with signal detection theory. In doing so, she showed that some observers (i.e., herself) were capable of detecting the visual sensation of a single quantum (photon) of light when under the most favourable conditions (Sakitt, 1972). Sakitt found that ratings could be directly related to cumulative Poisson distributions, and that this allowed for the mathematical modelling of action potentials from both signal and noise events. In fact, a rating scale is not necessary, as binary response data ('unseen', 0, vs 'seen', 1) can still be fit to a continuous parameter space through the use of a *link function*, which enables the application of ordinary linear regression to models other than normal distributions. In this

manner, the *psychometric function*—the function relating a physical stimulus scale to the probability distribution of observer responses—can be fit using a generalised linear model (GLM).

More recently, a method of fitting observer responses to sensory continua, which explicitly models stochasticity (i.e. noise, or uncertainty in response, and possibly also in stimuli) has been developed: maximum likelihood difference scaling (MLDS) (Maloney and Yang, 2003). MLDS obtains a signal-detection-theoretic model of how a unidimensional physical stimulus space is mapped to a perceptual space, in terms of both the relative spacings of internal responses to stimuli and the uncertainty, or 'spread', of their distributions, as shown graphically in Figure 14.1. For t trials, an observer is shown two pairs of stimuli, (i, j) and (k, l) and simply asked to report which pair had the most intra-pair difference. Each decision can be modelled as the comparison of four perceptual scale values, perturbed by Gaussian noise, $\epsilon \sim \mathcal{N}(\mu, \sigma^2)$:

$$\delta_i = (\psi_i - \psi_j) - (\psi_k - \psi_l) + \epsilon \tag{14.2}$$

In Matlab, the implementation of MLDS is achieved by the optimisation of an objective function, shown in Listing 14.1. This takes as its input the trial indices and response data, and calculates the log likelihood of obtaining these responses for a given perceptual scale. The goal of optimisation is to obtain perceptual scale values that maximise the likelihood of the data, which for convenience we frame as minimising the negative log likelihood.

The objective function first applies a sigmoid (logistic) transform to perceptual scale values, $\mathbf{\Psi}$, ensuring that they lie in the range of 0 and 1, and an exponential transformation of σ, ensuring that it is positive[2]. For these transformed values, a set of prospective perceptual differences, δ, are computed, as a function of the trials used in the experiment. From these, a likelihood is computed given the observer's

2 In fact, these transformations are not strictly required, as Matlab functions such as `fmincon` readily allow for lower and upper bound constraints to be placed on optimisation solutions.

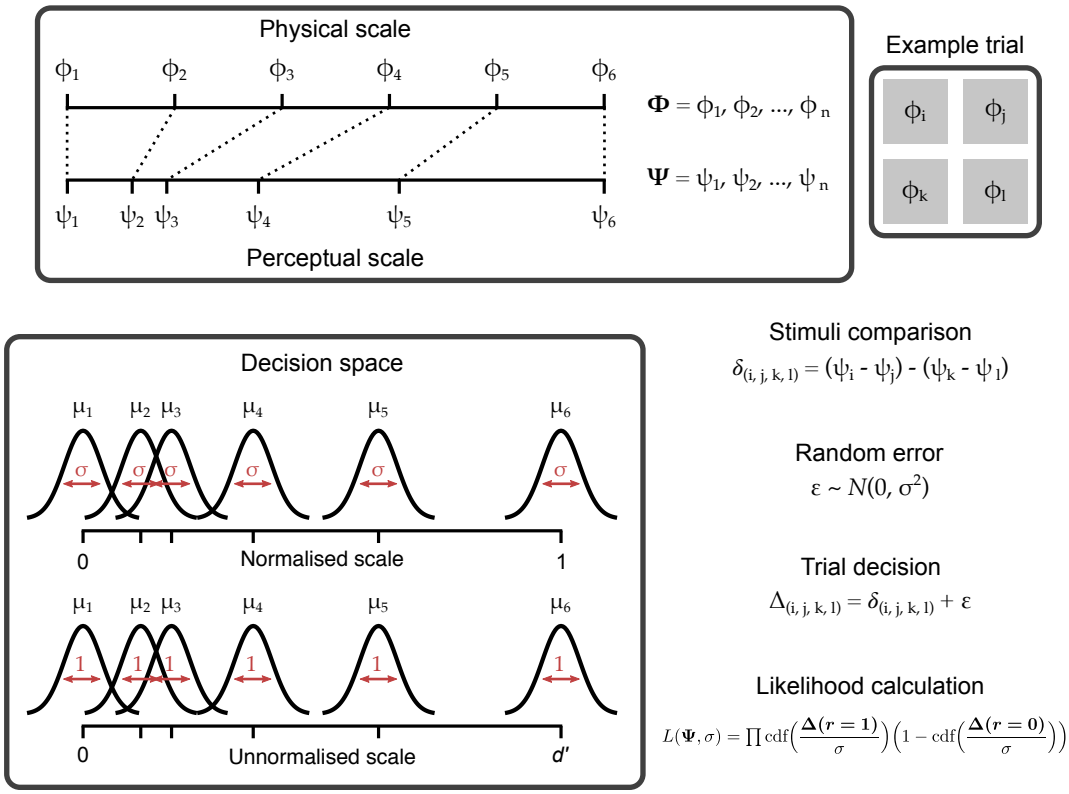


Figure 14.1: A graphical overview of MLDS. The underlying model assumes that a unidimensional physical parameter of scale $\mathbf{\Phi} = \phi_1, \phi_2, ..., \phi_n$ is related to an observer's perceptual scale $\mathbf{\Psi} = \psi_1, \psi_2, ..., \psi_n$. In a decision space, such as is hypothesised to account for behaviour during trials in an experiment, each stimulus evokes a response perturbed by random error, modelled as drawing a sample from a Gaussian of mean $\mu_i = \psi_i$ and variance σ, when the perceptual scale is normalised such that $\mathbf{\Psi} \in [0, 1]$. The likelihood function of a perceptual scale, given a pattern of binary responses ($r \in \{0, 1\}$) is computed by equation shown in bottom right, using the cdf of the standard normal distribution as a link function. Alternatively, the decision space can be transformed such that each Gaussian has a variance of 1, and $\mathbf{\Psi} \in [0, d']$.

responses, using the cumulative distribution function (cdf) of the standard normal distribution as a link function, generated by the Matlab function `normcdf`. Given `stim`, a tx4 matrix that defines the scale indices for stimuli quadruples, and `resp`, a binary vector of length t containing responses (0 for the first pair having greater intra-pair difference, 1 for the second pair), the vector `ps`, containing $[\mathbf{\Psi}, \sigma]$, can be optimised by a global optimisation function such as `fconmin`. The perceptual scale obtained can be expressed in terms of d', the sensitivity index defined as $(\mu_1 - \mu_2)/\sigma$, or on a normalised scale between 0 and 1, with σ given as a separate parameter. These two representations of difference scales are shown in Figure 14.1.

Listing 14.1: objMLDS.m

```
function lL = objMLDS(stim,resp,ps)
sigma = ps(end);
psi = ps(2:end-2);
psit = 1./(1+exp(-psi)); % sigmoid transformation of psi scale
psit = [0 psit 1]; % fix start and end of scale to 0 and 1
esig = exp(sigma); % exponential transformation of sigma
del = psit(stim) * [1 -1 -1 1]'/esig; % calculate del function
lL = - ( sum( log(normcdf(del(resp(:,1)==1)))) +...
        sum( log(1-normcdf(del(resp(:,1)==0)))) ); % log likelihood
end
```

14.2 CONJOINT MEASUREMENT ANALYSIS

Once the theoretical and mathematical details of difference scaling are understood, it can be seen that the approach is readily generalisable to stimuli of more than one physical and perceptual dimension. In other words, the two or more physical quantities in the stimuli, $\mathbf{\Phi}^a$ and $\mathbf{\Phi}^b$, may relate to their own corresponding perceptual evaluations, e.g. $\mathbf{\Psi}^a$ and $\mathbf{\Psi}^b$ with elements $\psi_i^a(\phi_i)$ and $\psi_i^b(\phi_i)$. At first glance, this is the same playing field as seen in Part iii, where we parameterise the space of human vowel perception as being two-dimensional, and solve the relative locations of stimuli with an experimental approach of dissimilarity measurement and multidimensional scaling analysis (MDS) analysis. MDS analysis assumes that the dimensions of perceptual relevance in stimuli are independent—the presence or absence of one does not contaminate judgements of the other. However, if we are interested in the interaction between different dimensions of perception, this raises a fundamental question of whether such interactions can even be approached mathematically: we cannot perform mathematical (concatenation) operations on unrelated perceptual magnitudes such as colour and roundness, so how could we hope to model the perceptual judgement of the freshness of an orange, which we presume would depend on both?

This puzzle, among others, caused a great deal of concern and debate for the psychological sciences in the early twentieth century, evident in the reports of 'the Committee appointed to consider and report upon the possibility of Quantitative Estimates of Sensory Events' (a.k.a. the Ferguson Committee), published under the aegis of the British Association for the Advancement of Science between 1932 and 1940 (Michell, 1999). Within these reports, we encounter perspectives that now seem frankly irrelevant to the community of psychological science, such as the competing statements of John Guild ("Are Sensation Intensities Measurable?" Answer: No, primarily due to the inadequacy of the JND as a unit of sensation, and not merely sensitivity.) and James Drever ("The Quantitative Relationship between Physical Stimulus and Sensory Event", which responds to Guild with the simple assertion that all measurement is relative comparison, and such comparisons are both possible for the sensations occurring in perceptual experience, and relatable to variations in physical quantities) (Ferguson et al., 1938). Nonetheless, the final report of the Ferguson Committee, published in 1940, concluded that because psychological attributes were not capable of sustaining concatenation operations (i.e. the lack of additive relations between sensory intensities), such attributes were not continuous quantities, and hence psychophysical methods did not constitute scientific measurements (Ferguson et al., 1940; Michell, 1997).

A resolution came two decades later, with the theory of *conjoint measurement*. Developed within the increasingly cross-pollinating fields of mathematics, economics, and psychology, conjoint-measurement theory provided mathematical proof that for situations where attributes are not known to be formal quantities in and of themselves, they may still be expressed as continuous quantities via their (measured) relation to another attribute (Krantz and Tversky, 1971; Luce and Tukey, 1964; Tversky, 1967). As such, it provides a way of evaluating the hypothesis that perceptual dimensions are independent and, if it must be rejected, of quantifying the nature of inter-dependence between the two or more perceptual dimensions of the stimuli. It bears strong similarities with difference scaling[3], in that relative locations in stimulus space

3 In fact, strictly speaking difference scaling is an example of conjoint measurement.

are obtained, in order to model the mathematical relationship between physical parameters and sensory continua. Only, rather than a line, in this case the space is of two or higher dimensionality.

As with difference scaling, models may be fit to experimental data by maximising the likelihood function of scales given a pattern of binary responses to stimuli, with the observer selecting one of a pair of stimuli for each trial that she weights more highly for a given 'sensation' (Ho et al., 2008). To solve MLCM, we again use the cdf of the standard normal distribution as a link function. As with MLDS, implementation in Matlab is possible by optimising over an objective function which calculates the likelihood of a given conjoint analysis model, shown in Listing 14.2.

Listing 14.2: objMLCM.m

```
function lL = objMLCM(stim,resp,ps,model)
pn = max(stim,[],'all'); % get number of levels in parameter space
ps(1) = 0; % fix to 0
if strcmp(model,'add') % for additive model
        ps(2:end,2:end) =  ps(2:end,1) + ps(1,2:end);
elseif strcmp(model,'ind') % for independent model
        ps = repmat(ps(:,1),[1,pn]);
end % otherwise assume saturated model
del = ps(sub2ind([pn pn],stim(:,4),stim(:,2))) -... % calculate del
        ps(sub2ind([pn pn],stim(:,3),stim(:,1)));
lL = - ( sum( log(normcdf(del(resp==1)))) +...
        sum( log(1-normcdf(del(resp==0))) ) ); % log likelihood
```

For this function, `ps` is an $m \times n$ matrix, $\mathbf{\Psi}$, where m and n are equal to the resolution of the parameter space's two dimensions. The difference function for each trial is computed as `del` using the level of each parameter, ψ_a and ψ_b, to index $\mathbf{\Psi}$. The likelihood is then calculated from these difference functions given the response pattern of the observer, as shown in Figure 14.2.

As shown in Figure 14.2, there are three candidate models, through which we can fit a pattern of observer responses to trail conditions, depending on if and how the two dimensions of the stimuli contribute to the perceptual magnitude in question. The

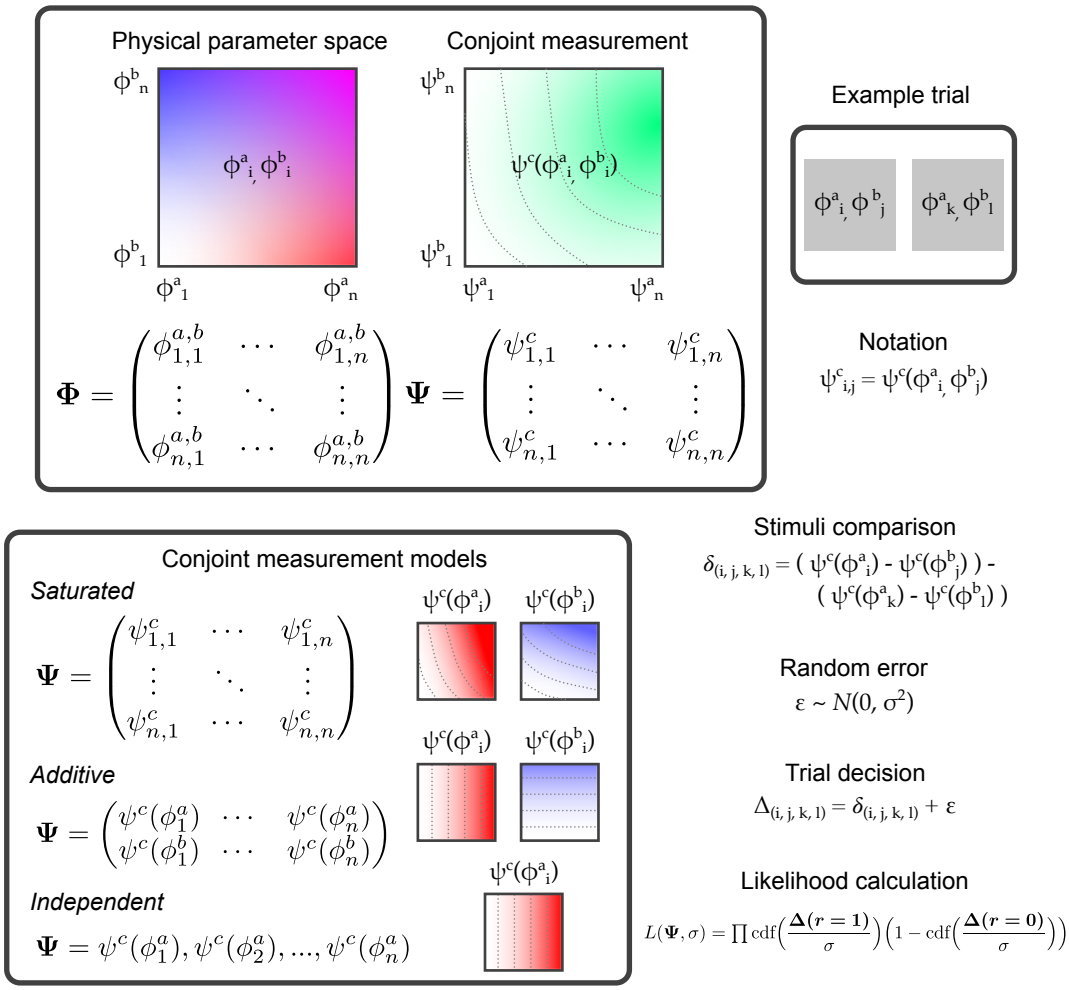


Figure 14.2: A graphical overview of MLCM. The underlying models relate a two- (or higher) dimensional physical parameter space, $\mathbf{\Phi}$, shown as an $m \times m$ matrix, to an observer's conjoint measurement space $\mathbf{\Psi}$. Depending on the model, $\mathbf{\Psi}$ may be a fully two-dimensional matrix (in the *saturated* case), a $2 \times m$ matrix (in the *additive* case), or a $1 \times m$ vector (in the *independent* case). As with MLDS, a decision space is hypothesised, such that each stimulus evokes a response perturbed by random error, modelled as drawing a sample from a Gaussian of mean $\mu_i = \psi_i$ and variance σ. The likelihood function of a conjoint measurement model, given a pattern of binary responses ($r \in \{0, 1\}$) is computed by equation shown in bottom right, using the cdf of the standard normal distribution as a link function, as with MLDS. Again, the decision space can be transformed such that each Gaussian has a variance of 1, and $\mathbf{\Psi} \in [0, d']$.

simplest case is that of the *independent* model, where only one dimension has a direct bearing on the task—perhaps observers are only concerned with colour saturation when appraising the freshness of oranges, and are insensitive to deformations of shape. In this case, which is identical to difference scaling as discussed above, we have only as many parameters as the resolution of our single dimension of interest minus

one (as the first term is anchored at 0), corresponding to the means of Gaussians with $\sigma = 1$, spaced in units of d'[4], such that $\mathbf{\Psi} = [0, \psi_2^a, ..., \psi_n^a]$. If observer judgements are affected by roundness as well colour saturation, such perceived freshness is a linear function of the two dimensions—it may be described by an *additive* model. The additive model is defined by twice the number of terms as the parameter space resolution minus two, as the perceptual contribution from the first level of both physical dimensions is anchored to 0, giving $\mathbf{\Psi} = [0, \psi_2^a, ..., \psi_n^a; 0, \psi_2^b, ..., \psi_n^b]$. Finally, it could be that not only does roundness factor in to observer judgements of freshness, but that it does so *as a function* of the other physical dimension, colour saturation, that varies in the stimuli. In this case, $\mathbf{\Psi}$ is a fully two-dimensional $n \times n$ matrix, with the first element equal to 0, and all others representing the means, in d', of Gaussians with $\sigma = 1$, corresponding to perceived sensation—in this case orange freshness—magnitudes of each location in the stimulus space.

The three categories of conjoint measurement observer model—independent, additive, and saturated—form a nested hypothesis, that is, a set of related, ordered hypotheses. This is because the independent model is in fact a special case of the additive model, which in turn is a special case of the saturated model. All three may be expressed as fully two-dimensional $n \times n$ matrices, except that in the additive and independent models, certain terms are fixed to be equivalent, shown in Figure 14.3 for a stimulus space of $n = 4$.

Each level of the nested hypothesis is a model of differing complexity that describes the data, $\boldsymbol{X}$. These models can be characterised by their likelihoods, which are a function of the model parameters given the observations, $\mathcal{L}(\mathbf{\Psi}; \boldsymbol{X})$. If two models have equivalent likelihood maxima, we accept the more parsimonious hypothesis, which has fewer (or a greater number of fixed) parameters. For example, if a saturated conjoint measurement model has the same likelihood, $\mathcal{L}_s$, as the additive model, $\mathcal{L}_a$, we would accept the latter, as the inclusion of additional free parameters in the model did not better describe the data. In practice, we determine if maximal likelihoods are

4 Alternatively, for representation on a normalised scale we have terms specifying each stimulus increment other than the first and last in the series (anchored at 0 and 1, respectively), and a term specifying the spread of the Gaussians.

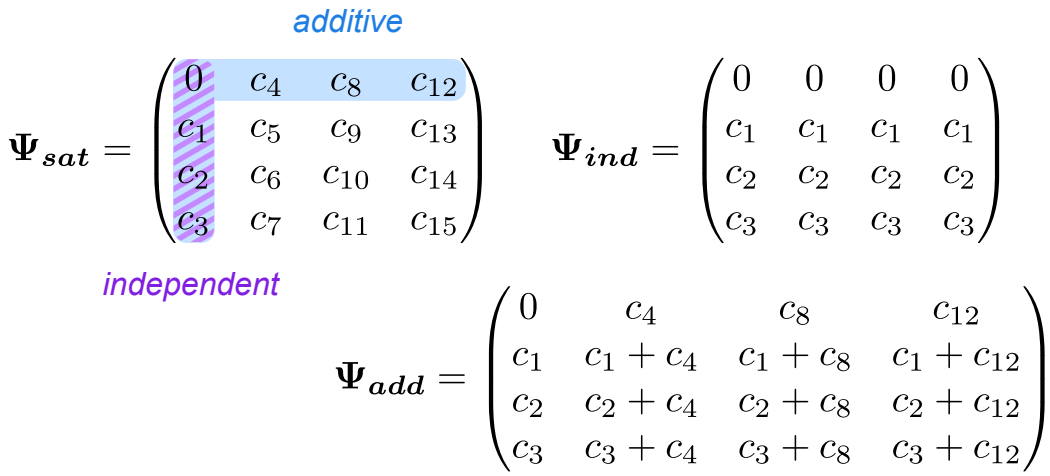


Figure 14.3: The independent, additive, and saturated models of conjoint measurement as a nested hypothesis. The $\boldsymbol{\Psi}$ matrix of coefficients is shown for each of the three models of conjoint measurement, for a 4×4 two-dimensional parameter space of stimuli. For the saturated model matrix, the coefficients that are required for the additive and independent models are indicated in blue and purple shading, respectively. The saturated model requires 15 free parameters, the additive six, and the independent only three.

significantly improved by additional model complexity by performing a likelihood ratio test. Given two models, their likelihood ratio is computed as $\lambda_{a/s} = \mathcal{L}_s / \mathcal{L}_a$, and it follows that $-2 \log \lambda$ asymptotically has a χ^2 distribution with degrees of freedom (DoF) equal to the difference in free parameters (Wilks, 1938). The two models can therefore be evaluated for statistically significant difference using a p-test, easily performed in Matlab using the function `lratiotest`, which takes as its arguments $\mathcal{L}_s$, $\mathcal{L}_a$, and the difference in DoF of the two models, $24 - 8 = 16$, and returns a Boolean decision of accepting or rejecting the more complex model and the p-value.

With this approach, we have a mathematically and statistically sound methodology to characterise how sensory attributes relate to multidimensional physical stimuli, and to test specific hypotheses. These methods will now be applied in chapter 15, to quantitatively evaluate the visual perception of metallicity.

15

MATERIAL QUALITIES, IMAGE STATISTICS, AND VISUAL FEATURES DETERMINING THE PERCEPTION OF METALLICITY

ABSTRACT

How the visual system judges the material composition of objects, accurately and robustly, remains for the most part an open question. We present experimental data and modelling that sheds light on what drives the percept of metallicity—the quality of appearance that suggests an object is made from metal. With computer-graphics stimuli, hyperspectrally rendered in a physically plausible manner, we use a conjoint-measurement experimental paradigm to decouple the contributions made by qualities both of a metal object and its transparent coating to the perception of metallicity. We find that under viewing conditions of a standard dynamic range, such as during an overcast day, coating qualities can strongly detract from a metallic appearance. However, when stimuli are viewed in high dynamic range viewing conditions, perceptions of metallicity are more robust. Spatio-chromatic features and luminance distributions of stimuli were analysed, and computational models were built to extract image correlates of physical parameters. These models replicated the key findings of the observer data, and suggest that metallicity is not the same as estimating metal roughness, but more akin to global contrast.

15.1 INTRODUCTION

The twenty-first-century visual landscape is both littered and adorned with the unmistakable flash of metallic objects. Across architecture, fashion, food, transport, and technology, we are presented with more metallic surfaces than at any other time in history. Unlike other visual features, such as the reddening of ripening fruit, or the gloss of wettened surfaces, the human visual system did not evolve in an ecosystem of metallic stimuli. Nonetheless, we are highly accurate in our judgements of metallicity from vision alone.

Material perception may at first glance appear a simple and closed deductive problem; the material composition of objects will determine their surface reflectance properties, and hence the image received on the retina as a function of illumination. However, the visual system exhibits an impressive robustness across a myriad of undetermined parameters, both of surfaces themselves, and their surrounding lighting environment. Metallicity, the visual quality of an object suggesting it is made of metal, is a prime example of this. The surfaces of most metallic objects return a solely specular reflection to observers, and their appearance is therefore highly contingent on both neighbouring objects and illuminants in their environment. At the same time, the physical roughness and oxide-layer depths of metallic surfaces can result in appearances ranging anywhere from an immaculately polished mirror to the matte sheen of an anodized laptop. Throughout this range, and across viewing environments, we accurately judge such objects as metallic, rather than other shiny alternatives such as porcelain or glass (Fleming et al., 2013; Sharan et al., 2014; Tamura et al., 2018; Vangorp et al., 2017). At the same time, the recognition of metal can prove particularly problematic for computer vision systems (Bell et al., 2015; Liu et al., 2010).

A sizeable economic value is placed on the visual appearance of metallic surfaces in sectors such as the automotive and jewellery industries, and the literature on metallic appearance reflects this. There are numerous studies quantifying and perceptually

evaluating the visual qualities of metallic surfaces and coatings, looking at qualities such as colour, gloss (Christie, 1969), scatter, brilliance, and lustre (McCamy, 1996), as well as 'visual texture' such as glitter, glint, and sparkle (Kirchner et al., 2007; McCamy, 1998). Recently, more attention is being placed on how physical qualities might give rise to a metallic appearance (Todd and Norman, 2018; Toscani et al., 2018). In the present study, we investigate how physical parameters defining coated metallic objects interact and contribute to perceptions of metallicity. We also interpret psychophysical data through image analysis and the outputs of computational models, in order to connect the optical properties of materials, image statistics, and visual features.

15.1.1 *Metallicity vs glossiness*

Metallicity as a visual property of materials has received little attention when compared with glossiness, and it is worth contrasting the two. Gloss, and its associated characteristics of highlights, sheen, haze, and texture distinction, is present on the surfaces of materials when incident light is reflected at the air-material boundary, giving a 'specular' reflectance component, before otherwise penetrating the material and giving rise to the object's 'diffuse' reflectance, or albedo. Glossiness therefore affects not just the surface colour of an object, but also the spatio-chromatic statistics of its corresponding image, with different glossy objects sharing related chromatic signatures of the illuminant despite having different surface colours. Within a natural scene, illuminants will typically be weak in chroma, giving rise to glossy highlights that are both an increase in lightness and a decrease in saturation on an object's surface.

Metallicity can be misconstrued as an extreme form of gloss, or, in the case of coloured metals such as gold, as a simple combination of colour plus gloss (Komatsu et al., 2013). However, there are notable differences between the two. First, a typical metallic surface lacks a 'diffuse' reflectance; all incident light is reflected at the

air-material boundary, due to the electro-physical properties of metals, and there is therefore no distinction between highlight and surface colour. There are some notable exceptions, particularly with regards to certain patinated metals, which due to weathering or chemical treatment exhibit a dulled or even completely matte finish, and these will be discussed later. Second, the proportion of incident light that is reflected specularly is far higher for metals than for glossy materials. This can result in far brighter regions on the surface, and in a perceptible reflected image of the full environment, rather than the array of highlights—illuminants divorced from their environment—frequently seen on smooth, glossy dielectric objects. Third, a metallic surface specularly reflects equally at all angles of incidence, whereas dielectric surfaces are subject to Fresnel reflectance, and specularly reflect to a greater extent at grazing angles. Whether human observers would be sensitive to this difference is unclear.

15.1.2 *Approach*

15.1.2.1 *Physically-based computer graphics stimuli*

For our stimuli we have used physically-based rendering, which permits precise parametrisation of stimulus spaces (which would be problematic with real objects) while staying within a physically plausible range. Specifically, we model objects with a silver base-layer underneath a colourful, transparent coating. This configuration may seem arbitrary, but is in fact a common material composition; with the majority of metals being colourless, foils and surfaces that would have an otherwise plain silver appearance are frequently coated with a clear, coloured layer. A common example may be found in the wrappers for assorted confectionery. It can also be seen as a simplified model of automotive paints, which typically comprise metal flakes dispersed amongst pigments in a transparent medium. There are several advantages to using this stimulus. First, it gives rise to both achromatic and chromatic specular reflections, from light that has both reflected off the coating-air boundary and the

silver base layer, respectively. Therefore, by varying the roughness of the silver, we can populate a parameter space of objects having both metallic and plastic appearances. Second, controlling both the roughness of the reflective metal and the evenness of a transparent body around it allows great control over different image properties. As shown in Figure 15.1, increasing metal roughness essentially introduces blurring, as the reflected image of the environment undergoes increasing scattering. Increasing the bumpiness of the coating has an effect very similar to applying a local disarray image transform, where pixels are subject to a random displacement field of a given scale and extent. The two axes of our parameter space are therefore analogous to an 'Eidolon factory' (Koenderink et al., 2017). Third, every image in the stimulus set is a natural image, albeit a synthetic one (as opposed to an artificial image). There are sound arguments to avoid the use of natural images in visual neuroscience, due to their complexity and unwieldiness (Rust and Movshon, 2005). However, manually programmed synthetic stimuli (e.g. sinusoidal gratings, plaids etc.) can only be useful insofar as the 'recipes' for particular percepts (or neuronal activations) are known; for material perception it is paramount that stimuli can appear as convincing examples of real life materials, and using synthetic natural images, obtained with physically-based rendering (Pharr et al., 2016), allows experimenters to investigate high-level visual percepts with precision and control. There is neurophysiological evidence from both humans and nonhuman primates that material qualities and judgements may be represented in higher cortical areas, particularly in medial regions of the ventral extrastriate cortex, such as the inferior extrastriate area (Newman et al., 2005), collateral sulcus (Cant et al., 2009; Cavina-Pratesi et al., 2010), and FG/CoS, the region extending from the fusiform gyrus into the collateral sulcus (Hiramatsu et al., 2011). Gloss-selective neurons have also been found in the inferior temporal cortex of macaque monkey (Nishio et al., 2012). As Rust and Movshon say themselves, "for neurons with complex properties whose circuitry is unknown (such as those in higher cortical areas), these methods [exploratory experiments with natural stimuli] may be the best or even the only way to begin". For metallic objects, this necessitates

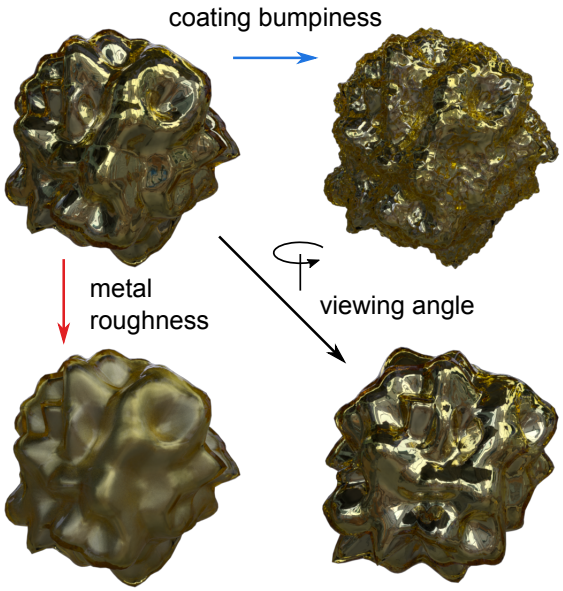


Figure 15.1: The parameter space used in experiments. Stimuli were physically-based rendered computer graphics of a composite coated metal object. Roughness of the metal base, and bumpiness of the coating, vary from zero to a plausible upper limit. The objects could be viewed from eight possible rotations around the z-axis.

their embedding within a (natural) environment, even if they are to be displayed without any background, as in this study.

15.1.2.2 *Conjoint measurement analysis*

As we are interested in the contributions to metallicity from properties of both the metal base and the transparent coating, we have framed this as a question of conjoint measurement. Methods of conjoint measurement analysis, and in particular the fitting of models through maximum likelihood estimation, have been developed for some time in psychophysics (Ho et al., 2008; Krantz and Tversky, 1971). Maximum likelihood conjoint measurement (MLCM) has been applied to colour (Rogers et al., 2016), the watercolour effect (Gerardin et al., 2014), as well as material judgements (Chadwick et al., 2018; Hansmann-Roth and Mamassian, 2017; Ho et al., 2008). The basis of this approach is to find the relative contributions of two (or more) independent physical parameters to a single perceptual judgement. In our case, we are interested in the effect of metal roughness, R, and coating bumpiness, B, on the

percept of metallicity, M. When an observer is given the choice of two images, one of metal roughness and coating bumpiness levels (i, j), the other of (k, l), and must decide which she finds more metallic, we can model the decision variable, Δ as:

$$\Delta(i, j, k, l) = \psi^M(r_i + b_j) - \psi^M(r_k + b_l) + \epsilon \tag{15.1}$$

where ψ^M is a function that relates the physical levels of a stimulus— for example of ith-level metal roughness, r_i, and the jth-level of coating bumpiness, b_j—the to the percept of metallicity. The decision variable is perturbed by Gaussian noise, $\epsilon \sim \mathcal{N}(0, \sigma^2)$.

These latent functions can be modelled with a set of nested hypotheses, with each hypothesis formulated as a generalized linear model. The most restricted model is the *independent model*, where only one physical parameter influences the perceptual judgement, and the observer's decision process is uncontaminated by other physical parameters. For example, if the percept of metallicity were only a function of metal roughness, the decision variable could then be evaluated as:

$$\Delta(i, j, k, l) = \psi^{M:R}(r_i) - \psi^{M:R}(r_k) + \epsilon \tag{15.2}$$

where $\psi^{M:R}(r_i)$ is the additive contribution of metal roughness to metallicity, computed for the ith-level of that parameter.

If such contamination is present, the next level of the model with increasing complexity supposes a linear mixing of the two physical parameters, and is said to be *additive*.

$$\Delta(i, j, k, l) = \left(\psi^{M:R}(r_i) + \psi^{M:B}(b_j)\right) - \left(\psi^{M:R}(r_k) + \psi^{M:B}(b_l)\right) + \epsilon \tag{15.3}$$

If this mixing is not linear, but the contribution of one physical parameter to the perceptual judgement depends on the level of the other, a fully comprehensive—or *saturated*—model is required to best account for the variance of the data. In this case, the functions relating physical levels of metal roughness to the percept of

metallicity depend on how bumpy the coating is, $\psi^{M:R(b_j)}$, and vice-a-versa. The decision variable is then given by:

$$\Delta(i,j,k,l) = \left(\psi^{M:R(b_j)}(r_i) + \psi^{M:B(r_i)}(b_j)\right) - \left(\psi^{M:R(b_l)}(r_k) + \psi^{M:B(r_k)}(b_l)\right) + \epsilon \tag{15.4}$$

With a set of collected responses from observers, these models can be fit via optimization, to maximize the likelihood function of parameter estimates. This accounts for the stochasticity of observers, and allows us to compare across levels of the nested hypothesis by evaluating their goodness of fit using the log likelihood. If a more complex model results in a significantly better fit (as found by a chi-squared approximation), it is preferred over the simpler model. Ultimately we obtain values of d', the sensitivity index (Green and Swets, 1966), which gives a measure of how discriminable the effect of changing physical parameters of the stimuli is on the percept of metallicity.

In order to determine the spacing of stimuli, we initially performed a maximum likelihood difference scaling (MLDS) of the parameter spaces (Maloney and Yang, 2003). This indirect method has advantages over more traditional appearance-based (suprathreshold) methods such as Thurstonian scales and just-noticeable difference estimation, and may arrive at more accurate estimates (Wiebel et al., 2017). MLDS has been used to evaluate colour perception (Brown et al., 2011; Radonjić et al., 2015), depth perception (Aguilar et al., 2017), image quality (Charrier et al., 2012, 2007), emotion (Junge and Reisenzein, 2013) and gloss perception (Obein et al., 2004). By obtaining approximately equivalent spacings throughout the parameter space, we safeguard against nonlinearities and prevent one dimension from dominating the other during the conjoint measurement task.

15.1.2.3 *Color spaces and dynamic range*

As well as observer performance in psychophysical tasks, we are interested to see if image features can be extracted that correlate with performance data, and

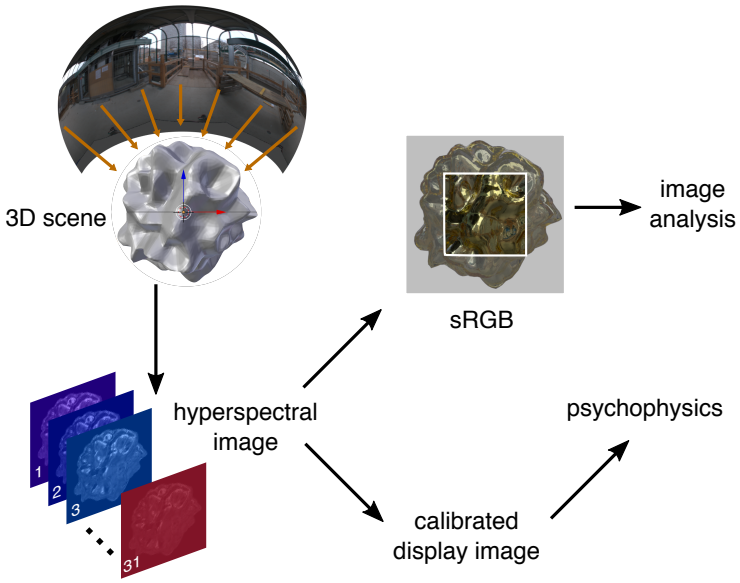


Figure 15.2: An overview of the pipeline for the research presented here. Scenes are modelled in Blender, and hyperspectral images are computed with physically-based rendering. From these hyperspectral images (n = 31), three-primary images are produced for either psychophysical testing on a calibrated display, or statistical analysis.

incorporated into models that show the same performance. For image analysis and modelling, we convert hyperspectral renders to sRGB space. This is a standardised colour space best suited for general displays and printing, and is also used for images that appear in the figures throughout this paper. An overview of the pipeline, including rendering, psychophysical experiments, and image analysis, is shown in Figure 15.2.

Due to their 'conductor' reflectance properties, many metallic objects are prone to exhibiting extremely high luminance values in their corresponding images when under certain lighting conditions. Standard displays, such as CRT displays and LCDs (liquid crystal displays) are unable to faithfully present such stimuli due to an inadequate dynamic range—the luminance gamut of the display and the gradations accessible within that gamut. For this reason, we ran experiments both under standard dynamic range (SDR) conditions on a calibrated CRT display, and under high dynamic range (HDR) conditions on an LCD/DLP projector display (Seetzen et al., 2004; Wanat et al., 2012). As we rendered the same three-dimensional

models and materials, with different environmental illuminations, this is analogous to viewing the same objects in both standard and high dynamic range lighting environments.

Ethics

This study was approved by the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford, in accordance with the Declaration of Helsinki.

15.2 EXPERIMENT 1: PERCEPTUAL SCALING OF PHYSICAL PARAMETERS

15.2.1 *Method*

Modelling

Stimuli consisted of computer-generated images, the primitive of which was a metal sphere randomly deformed to contain smoothly curving projections, enveloped within a coating. This shape exhibits several required properties. It has a wide array of curvatures, including regions of both highly positive and negative curvature, as well as numerous saddles, which ought to aid in visual judgements of material qualities (Vangorp et al., 2007). This being so, the shape has a distinct solidity, with flatter regions of the structure preventing it being perceived as a smoothly flowing or rippling liquid. In combination, the flatter and more curvaceous regions of the shape give rise to specular reflections of the environment across the gamut of those seen day-to-day in metallic surfaces, such as cars, cutlery, and cookware. While it has rotationally invariant statistical features—the height and frequency of projections—these are distributed randomly, ensuring that when viewed from different angles observers must make a material judgment, rather than a direct image comparison. Direct image comparisons (i.e. comparing the same region of

the environment reflected in two or more objects) are also hindered by the high frequency of random projections, which scramble the locations of the environment's features.

Stimuli were modelled using Blender 2.77, open-source 3D computer graphics software, in a similar manner to a recent study on metal and glass visual perception (Tamura et al., 2018). First, the object was created as an icosphere with two subdivisions. Then, the mesh is subdivided, with the number of 'cuts' set to two, 'fractal'—a measure of random deviation in the mesh—set to 10, and 'along normal' set to one, ensuring projections lead directly out from the center. The 'subdivision surface' modifier is then applied with two subdivisions, and smooth shading was specified. The coating is defined as the same shape scaled up by a factor of 1.05. This has a further two subdivisions applied, giving a resolution that permits smooth bumpiness. The bumpiness is defined with a 'displace' modifier, configured with a Stucci texture of size 0.05 and turbulence 1.00. The 'strength' parameter of the modifier varies between 0 and 0.05. As the name may suggest, the 'Stucci' procedural texture provides a random displacement field that evokes a decorative process, as opposed to naturally occurring or unrealistic options such as Voronoi patterns or simple spheres.

Rendering

Physically-based rendering was carried out with Mitsuba, configured hyperspectrally with 31 10-nm wavelength bins between 395 and 705 nm. Rendering parameters (materials, rendering engine, integrator etc.) were specified with RenderToolbox where possible (Heasly et al., 2014). The object was assigned a 'rough conductor' BRDF with the spectral reflectance of silver. Roughness was varied between 0 and 0.15, using the 'ggx' microfacet distribution model, as it has been found to give more physically accurate results over the standard Beckmann distribution (Walter et al., 2007). The coating was assigned a 'dielectric' BRDF with a refractive index of 1.4, a plausible value for a varnish or glaze, and an interior 'homogeneous medium' with absorbance defined as 0, 0.05, 0.60 in RGB (which is then interpolated to hyperspectral

data by Mitsuba), at a 'scale' of 20. Eight renders of each coated object were obtained by rotating at 45° intervals around its vertical axis.

Objects were illuminated by the 'Overcast day/building site' environment light probe made by Bernhard Vogel (http://dativ.at/lightprobes/), which provides a realistic environment that does not require tone mapping or compression of the luminance space for presentation on a standard display. Path tracing was computed with the 'extended volumetric path tracer' with an infinite maximum path length. The 'hide emitters' function created images of objects in a black void. The 'low discrepancy sampler' was used; a sample count of 256 was found to give images of sufficiently low noise. For the MLDS stimuli sets, 176 images were rendered over 17 hours on a Macbook Pro (early 2015).

Rendering outputs were 31-dimensional 512 x 512 pixel hyperspectral images. These were transformed to LMS values based on Stockman and Sharpe 2° cone fundamentals before conversion to RGB for display on a calibrated cathode ray tube (CRT) display. The final images were 512 x 512 single precision floating point matrices.

Participants

Five practised observers were recruited from the Department of Experimental Psychology at the University of Oxford. All participants had normal colour vision (as determined using an HRR test) and normal or corrected-to-normal visual acuity. Participant age ranged from 23 to 28 years.

Procedure

The experiment was carried out at the University of Oxford. Participants viewed all stimuli in a dark room on a CRT monitor (NEC, FP2141SB, 21 inches, 1600 x 1200 pixels) controlled with ViSaGe MkII (Cambridge Research Systems), which allows 14-bit intensity resolution for each phosphor. Gamma correction was performed with a ColorCAL MkII colorimeter (Cambridge Research Systems) and spectral calibration was performed with a SpectroCAL MkII spectroradiometer (Cambridge Research

Systems). Viewing distance was maintained with a chin rest positioned 92 cm from the CRT monitor. Participants viewed the screen binocularly.

On each trial, participants were presented with a quadruple (i.e. two pairs of stimuli simultaneously presented) and asked to indicate for which pair the material composition of the objects appeared to have a greater within-pair difference. Trials comprising variations in metal roughness, and variations in coating bumpiness, were interleaved. In either case, the values of the varying parameter of the four objects on each trial were drawn from eleven possible values, such that all four objects had different values. Additionally, each image in the quadruple was viewed at a rotational angle drawn without replacement from the eight 45° intervals, hindering direct image comparison. Participants viewed all 310 non-overlapping quadruples, for both metal roughness and surface bumpiness parameters, in a randomised order. Participants were given 3 s to respond to each trial and entered their response by pressing either up or down on a response box, selecting either the upper pair or lower pair in the quadruple.

15.2.2 *Results*

The estimated perceptual scales of roughness and bumpiness for five participants are shown in Figure 15.3, with 95% confidence limits obtained with the bootstrap method. Each participant's data were described by a monotonically increasing perceptual scale, which was concave with respect to the ends of the scale. In other words, all participants found greater perceptual differences at lower levels of roughness and bumpiness, with differences weakening, and in some cases saturating, as the physical parameters increased.

The construction of the conjoint measurement parameter space required a perceptually even spacing for both of the physical parameters to be varied. As there was good agreement across participants, data were aggregated across individuals for both roughness and bumpiness difference ratings and a representative perceptual scale

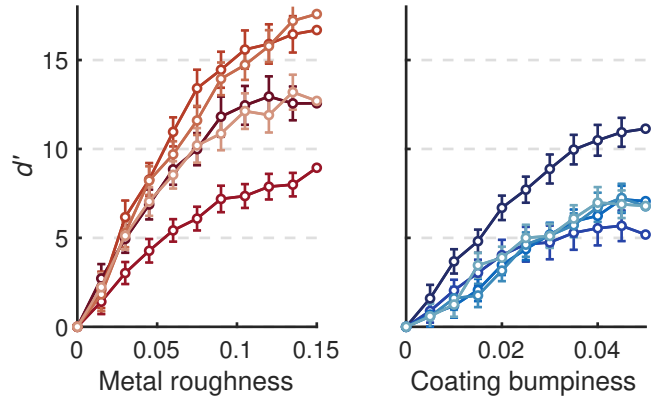


Figure 15.3: Perceptual scale estimates for individual participants for roughness (left) and bumpiness (right). Values are shown in d', and it is clear that some observers show greater sensitivity to the parameters than others. Error bars show 95% confidence intervals obtained via bootstrap.

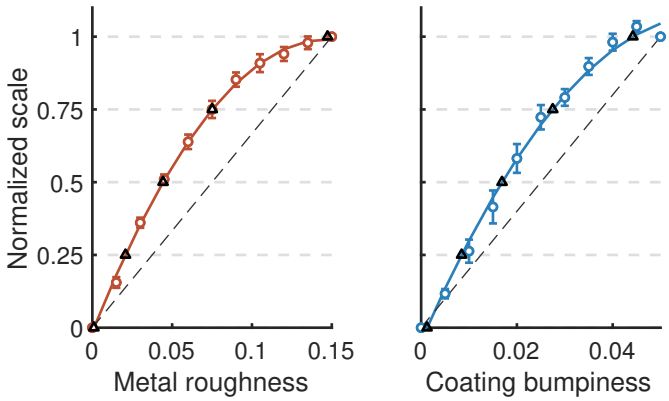


Figure 15.4: Mean perceptual scale estimates from aggregate data for roughness (left) and bumpiness (right), plotted on normalized scales between 0 and 1. A uniform scale is indicated by a dashed line. Curves are second-order polynomial fits to means across participants. Error bars show $\pm SEM$ ($N = 5$).

was estimated, shown in Figure 15.4. Four perceptually-equal intervals in roughness and bumpiness were obtained by fitting a polynomial to the respective aggregate perceptual scales, and evaluating at the quintiles of the function.

15.3 EXPERIMENT 2: CONJOINT MEASUREMENT OF ROUGHNESS AND COATING BUMPINESS

15.3.1 *Method*

Stimuli

Stimuli were modelled and rendered in the same manner as for the difference scaling experiment, varying the physical parameters of metal roughness and bumpiness of the coating according to the perceptually equal spacing obtained from the difference scaling experiment. The five-by-five parameter space was modelled with the same eight viewing angles described for the first experiment, giving a total stimuli set of 200 images, rendered over 10 hours.

Participants

Participants were the same as took part in Experiment 1, with the exception of one participant change.

Procedure

The experiment was carried out in a similar manner as for Experiment 1. Each trial consisted of simultaneous presentation of a pair of stimuli, with participants asked to indicated which object appeared most likely to be made of metal. In each trial, the values of the varying parameter of the two objects were drawn at random, without replacement, from the five possible values. Additionally, each image in the pair was viewed at a rotational angle drawn without replacement from the eight 45° intervals, hindering direct image comparison. Participants viewed all 325 pairs, in a randomised order, four times, totalling 1300 trials. Participants were given 2 s to respond to each trial and entered their response by pressing either left or right on a response box, selecting either left or right in the pair.

15.3.2 *Results*

The data from each participant were fit to the three nested models of conjoint measurement: independent, additive, and saturated. For every participant the independent model, where only one parameter varying in the stimuli has an effect on the decision-making of the observer, was rejected following an ANOVA comparison with the additive model ($p < 0.001$), confirming that both metallic roughness and coating bumpiness of the objects affected perceived metallicity. This is a somewhat surprising finding; even when coatings profoundly affect the appearance of objects, we might expect observers to exhibit a material constancy for the underlying surface. Instead, we see that for four out of five participants (P1, P2, P3, and P5), increasing the bumpiness of the coating has a deleterious effect on perceived metallicity. As would be expected, decreasing the surface roughness of the metal caused increased perceptions of metallicity for most participants in concordance with a previous study (Todd and Norman, 2018). However, one participant (P5) showed the opposite effect, suggesting that judgements of metallicity may be heterogeneous in the population. When asked if that participant could give a reason for rating rougher objects as more likely to be made of metal, they gave the following response: "I feel that the blurred specular pattern makes it more metallic. I imagine it would be because the surface of metals in real life are somewhat rough."

A direct question that comes out of these results is therefore at which level the contamination from coating bumpiness occurs. It could well be, as suggested by Todd and Norman, that surface roughness and metallicity are essentially the same appearance judgement. In which case, would coating bumpiness also contaminate perceptions of roughness? Alternatively, metallicity could be a separate judgement to both roughness and coating bumpiness, although for most participants these judgements seem to be highly correlated for the former, and anti-correlated for the latter.

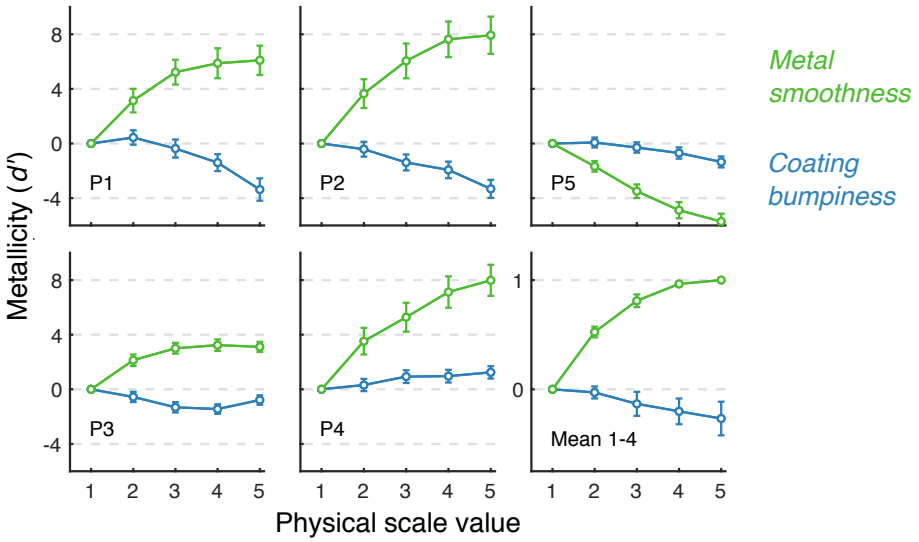


Figure 15.5: Conjoint measurement estimates for individual participants fit to the 'additive' model, for the SDR condition. Plots show the contributions of metal smoothness (green, the inverse of roughness) and coating bumpiness (blue) to discriminations of metallicity. Error bars for individual participants show 95% confidence intervals obtained via bootstrap. The lower right-hand plot shows the mean of normalized estimates for P1-4, with error bars showing $\pm SEM$ ($N = 4$).

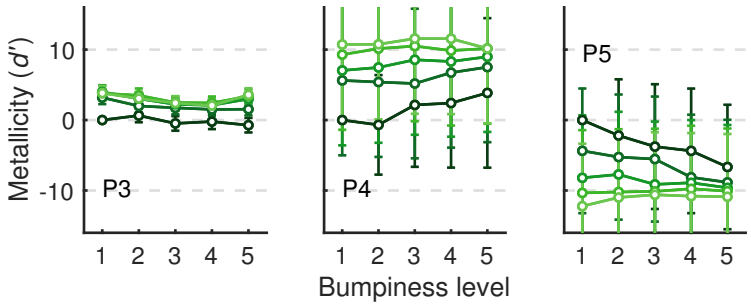


Figure 15.6: Three of the data sets from the conjoint measurement were best fit to the saturated model for the SDR condition. Coating bumpiness variation is given on the abscissa, while increasing metal smoothness is indicated by brightening green colour. Error bars show 95% confidence intervals obtained via bootstrap.

15.4 IMAGE ANALYSIS AND MODELLING

Little is known about the visual features and statistics of images that drive the perception of metallicity. From a physics standpoint, metals owe their appearance to a uniform specular reflection over their entire surface, and their exceptionally low absorptivity in the visible region of the spectrum (with the exception of chromatic metals such as gold). Recent work has related metallic appearance to rendering parameters such as surface roughness and the quality of illumination (Todd and Norman, 2018). While this does inform on which parameters of a scene give rise to metallic objects, it does not explain why they look the way they do to observers. To address this question, we developed models in order to simulate the experiments described above, testing to see if low-level image statistics could explain some of the general findings. Although we are primarily concerned with image correlates for metallicity, developing models for discriminating metal roughness and coating bumpiness—the two physical parameters we varied in the conjoint measurement analysis—will allow us to evaluate whether these percepts are independent.

15.4.1 *Spatial frequency analysis*

The relationship between image statistics and material appearance has received much attention in recent years, and the use of global statistics, such as intensity histograms, has proved controversial (Anderson and Kim, 2009; Kim et al., 2016; Motoyoshi et al., 2007; Sawayama and Nishida, 2018; Sharan et al., 2008; Wiebel et al., 2015). The role of spatial frequency analysis has been implicated for texture appearance (Harvey and Gervais, 1978) and in particular has proved useful in accounting for fabric appearance (Giesel and Zaidi, 2013); here we adopt a similar approach. An immediate evaluation of whether spatial frequency distributions can account for material appearance can be made by swapping the Fourier amplitude spectra or phase responses of images of different materials (Oppenheim and Lim, 1981; Piotrowski and Campbell, 1982),

as shown in Figure 15.7. It is clear form these amplitude-phase swaps that metal roughness is primarily determined by amplitude spectra, while coating bumpiness appears to be determined by both amplitude and phase spectra.

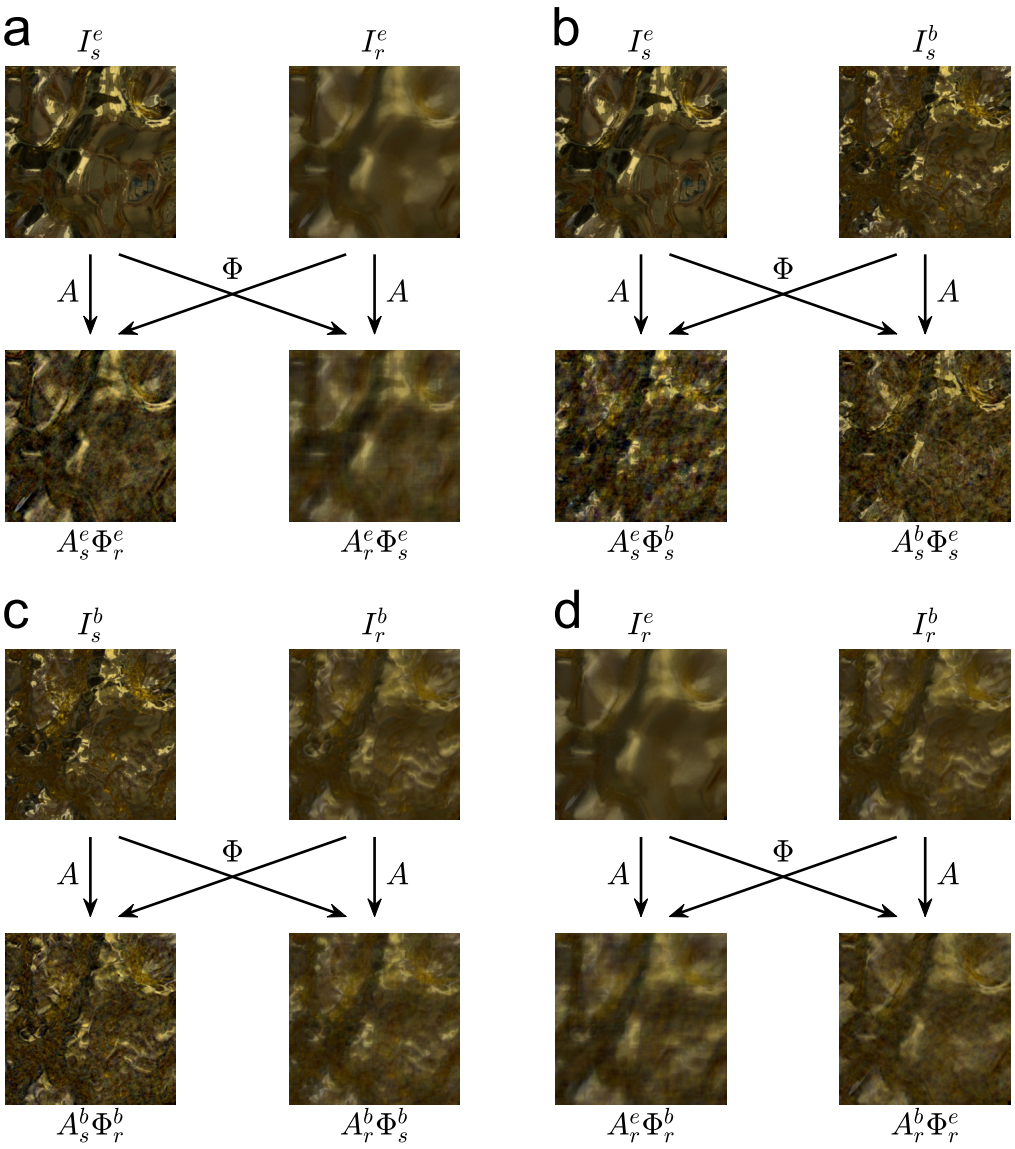


Figure 15.7: Swapping the amplitude spectra (A) and phase spectra (Φ) of the stimuli shows that metal roughness is predominantly determined by amplitude, as shown in plots a and c, while coating bumpiness has a significant phase component, shown in plots b and d. Key: I = original image, A = amplitude spectrum, Φ = phase spectrum, s = smooth (low metal roughness), r = rough (high metal roughness), e = even (low coating bumpiness), b = bumpy (high coating bumpiness)

To quantify these effects, power spectra were calculated for the stimulus set. Images were cropped to exclude any background or object edges, as shown in Figures 15.8

and 15.9, and a window was applied to prevent edge effects when computing the Fourier transform. Spatially two-dimensional images give rise to two-dimensional spectra in frequency space; these were interpolated from x- and y-axis frequencies onto a one-dimensional radial (r-) frequency axis, equivalent to averaging frequency spectra by orientation as per convention (Tolhurst et al., 1992). As with studies into the statistics of natural images, power spectra showed a slope of approximately $1/f^2$ when plotted on log-log axes (Cohen et al., 1975; Ruderman and Bialek, 1994; Tolhurst et al., 1992). Increasing metal roughness showed a strong effect on this slope, attenuating high frequencies in the image. This was well conserved across the eight viewpoints of the object, as shown in Figure 15.8 which also shows the mean spectral powers over the high-end of the spectrum (between 23 and 53 cycles per image, or cpi, as given in previous studies (Giesel and Zaidi, 2013)). Coating bumpiness also affects the power spectra of images; the divots and bumps of the coating create fine details in the images where previously image regions were smooth, in a similar manner to that of a random local disarray field. This results in low-frequency components being slightly attenuated, as the coating disrupts the repetitive curvatures of the objects, seen in Figure 15.10. However, there is a poor concordance of spectra across the different viewpoints of the object, shown in Figure 15.9, and these changes do not hold for different levels of roughness (the fully two-dimensional parameter space of conjoint measurement). Perhaps, surprisingly, increasing coating bumpiness results in only a very slight increase in power for high-frequency components, which is only apparent when spectra are heavily smoothed and normalized by the zero-bumpiness spectrum.

15.4.2 *Steerable pyramid analysis*

Although global spatial frequency features can sometimes provide good metrics to evaluate texture features, in particular when images are of a single object and viewpoint, and other material features are kept constant, they perform poorly outside

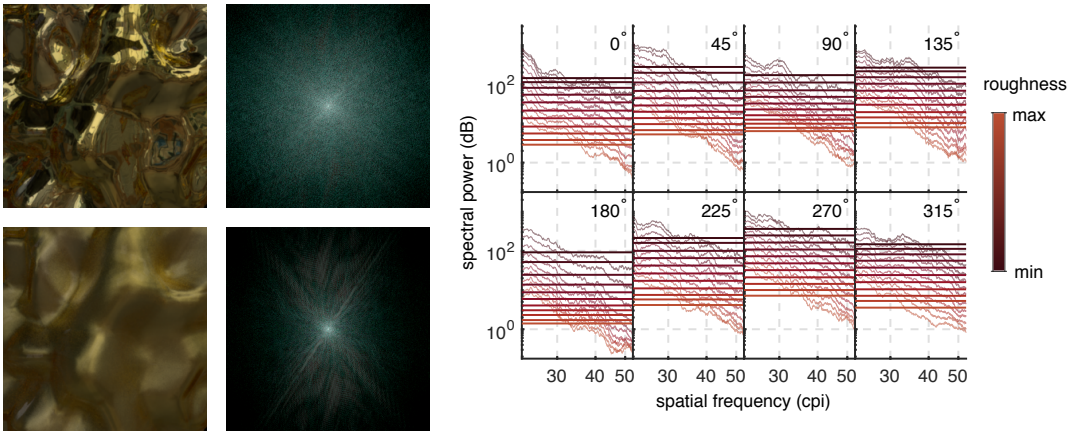


Figure 15.8: Roughness of metal has a direct effect on the spatial-frequency content of the stimuli. Example images of the stimuli are shown on the left, alongside their Fourier spectrum, with the smoothest stimulus on the top and the roughest below. Shown on the right are plots of power spectra varying with roughness (thin lines) and their mean over the range of 23-53 cpi (thick bars), with increasing roughness indicated by brightening red colour, for the different viewpoints of the objects. Across viewpoints, there is good agreement between the means of the power spectrum over this frequency range. Specifically, the rank-order of the means is invariant across viewpoint, and the range of means is similar. Still, this is not a direct 1-to-1 mapping from roughness to mean, as viewpoint does introduce some variation.

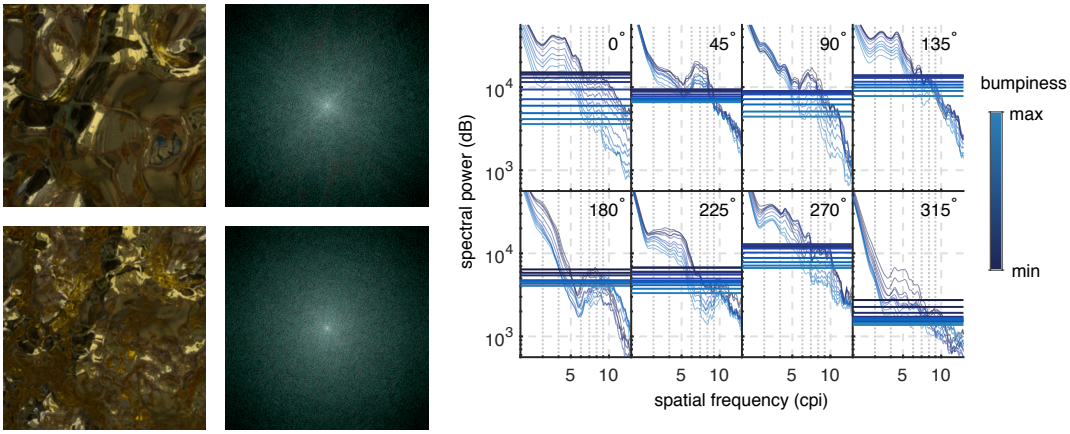


Figure 15.9: Bumpiness of coating also has an effect on the spatial-frequency content of stimuli. As bumpiness increases for individual viewing angles, the means of the spectrum over the range of 2-15 cpi generally decrease. However, as compared with roughness, this is a less direct relationship showing less consistency across viewing angles, with the rank-order of means occasionally changing, and the rand of means varying considerably across viewing angles.

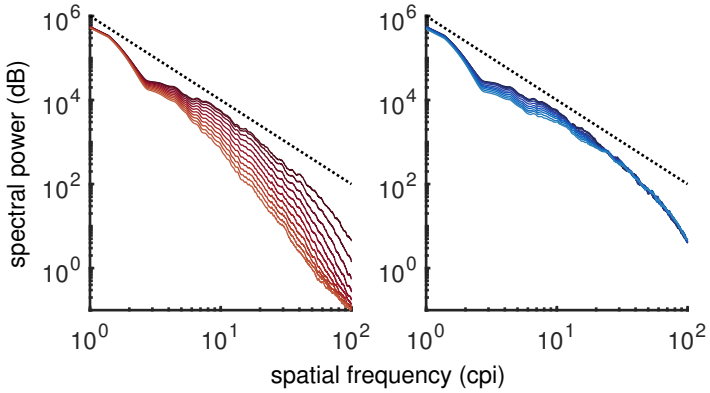


Figure 15.10: Amplitude spectra for the roughness and bumpiness stimuli used in the MLDS experiments, interpolated onto a one-dimensional scale of radial frequency. Spectra have been averaged across all eight viewing angles of the objects. Left: increasing roughness, indicated by red. Right: increasing coating bumpiness, indicated by blue. Black dotted lines indicate a gradient of $1/f^2$ (the intercept is arbitrary).

limited cases. Another image processing approach, which is more robust against changes of these nature, is to use steerable pyramids (Freeman and Adelson, 1991; Portilla and Simoncelli, 2000; Simoncelli and Freeman, 1995). Steerable pyramids are constructed by computing the outputs of oriented linear (Gabor-like) filters, at different 'scales' of an image. Algorithmically this is achieved by iteratively blurring and downsampling the image, taking filter outputs at each level, as well as marginal statistics over the images and residuals such as variance, mean, and kurtosis. At a basic level, this method can be used to map local orientations throughout an image, obtaining an 'orientation field' (Fleming et al., 2011), and is sensitive to both local and global image features. Because autocorrelations and correlations across scales and orientations are also calculated, the resultant catalogue of parameters captures many of the spatially invariant statistics and features of textures, most directly demonstrated by the *de novo* synthesis or extrapolation of visibly similar textures. Although this catalogue may contain several hundred parameters, they are retained in a comprehensible format and are useful for analysis. While steerable pyramids are no longer the state-of-the-art in texture synthesis, having been supplanted by convolutional neural networks featuring many thousands or millions of parameters (Gatys et al., 2016), they remain useful to vision science in their ease of implementation,

interpretability (i.e. they are not a 'black box'), and similarity in principle to the processing ongoing in the human visual system.

The synthesis-end of steerable pyramid analysis provides an alternative to phase-swapping (shown above in Figure 15.7). New images may be synthesised starting with either noise or another image, such that the steerable pyramid parameters match a desired template. Performing this synthesis with an image taken from the centre of the parameter space as the base image allows us to evaluate if the steerable pyramid parameter catalogue captures the necessary image statistics to describe (and synthesise) the wider parameter space, and to probe how different components of the parameter catalogue contribute to each parameter's associated percept. As shown in Figure 15.11, which contains the central image of the parameter space in the middle, and four synthetic images around it, steerable pyramids can sufficiently capture the properties of the parameter space for both metal roughness and coating bumpiness.

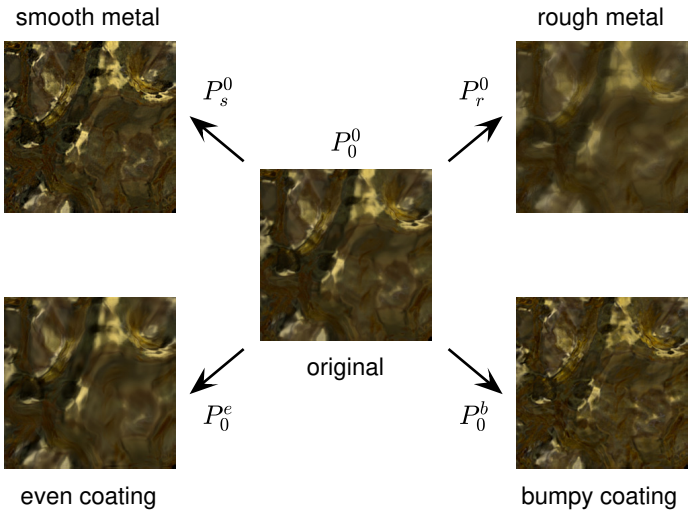


Figure 15.11: Steerable pyramid synthesis, starting from an image at the center of the parameter space (P_0^0). The synthesis process attempts to match the pyramid parameters of that image with the parameters of another pyramid, in this case from four different sides of the parameter space, reaching smooth metal (P_s^0), rough metal (P_r^0), even coatings (P_0^e), bumpy coatings (P_0^b).

Images in the stimulus set were analysed with a steerable pyramid of four layers and four orientations. The outputs (or 'activations') or the linear filters, averaged over orientations, and across all viewing angles of the objects, are given as heatmaps in Figure 15.12. All values presented are logarithmically transformed, as per the typical

responses of neurons in the early visual pathway, although this did not effect results. As expected (and in concordance with the results of spectral analysis), increasing roughness correlated with a decrease in the magnitude of response by the oriented filters at early levels, corresponding to the detection of fine details in the image, for all viewing angles. Increasing bumpiness correlated with a clear increase in these responses at the first two levels of the pyramid and the high-pass residual. This corresponds with the fine details that coating bumpiness introduces, which are poorly described by a global spatial-frequency (Fourier) analysis. Outputs of the filters for the fourth level of the pyramid reduced with increasing bumpiness, corresponding to the weakening of low-frequency components, which is also seen in Figure 15.9. Bumpiness also markedly affected phase statistics within the steerable pyramid, which are more informative than the global phase spectrum given by Fourier analysis. Specifically, increasing coating bumpiness correlated with decreases in the autocorrelation of the real parts (ACR) of the pyramid at early scales of the pyramid, possibly because otherwise smooth and continuous edges in the image were 'scrambled' by the refraction of light through the variable coating, an image transformation similar to local disarray (discussed at length in previous literature (Koenderink et al., 2017)). These correlations are given in Figure 15.13. It is clear from these results that while the magnitudes of both fine (high spatial frequency) features and broad (low spatial frequency) features are directly affected by coating bumpiness, these measures are not necessarily comparable across different viewing angles of the object, giving rise to fairly low aggregate R^2 values, despite individual correlations being very high (reflected in the mean of R^2 values for single viewing angles, $\overline{R^2}$). A visual system employing these metrics alone would therefore exhibit limited viewpoint invariance, and poor material constancy—the same issue that arises from basing judgements purely on global spatial-frequency content, as in Figure 15.9. However, by combining the activations of linear filters across different scales, namely by taking the sum of Levels 1 and 2, and computing its ratio to Level 4, a metric is obtained that shows strong correlation with coating bumpiness both within and across viewing angles, shown in Figure 15.14.

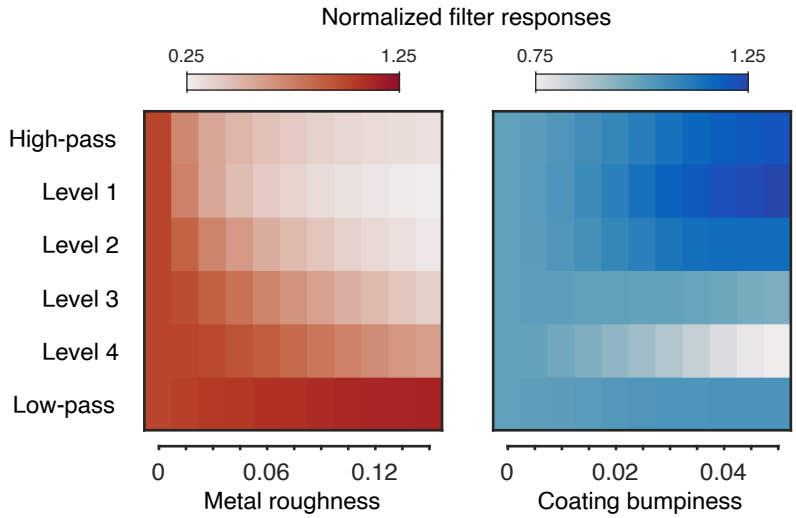


Figure 15.12: Heatmaps showing the magnitudes of filter responses summed over orientations, for each level (or 'scale') of the steerable pyramid analysis. Values of each level have been normalized to zero roughness (left heatmap) or zero bumpiness (right heatmap), to show the relative changes throughout the parameter space, averaged across all viewing angles.

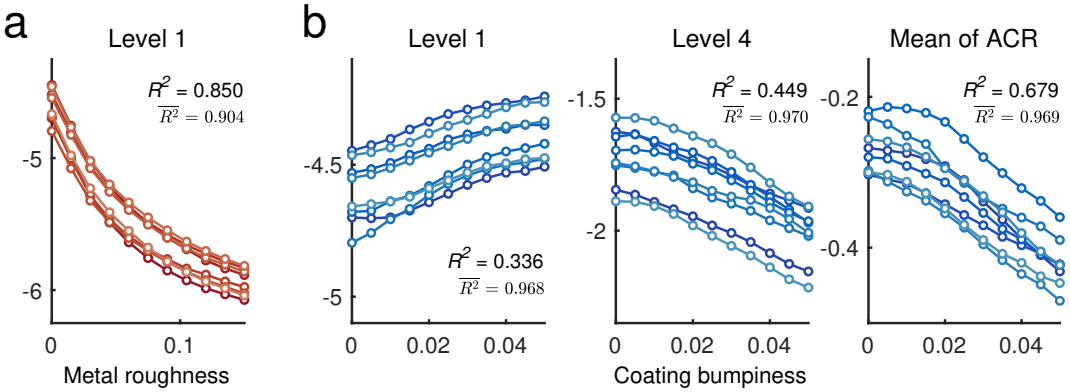


Figure 15.13: a) Increasing metal roughness correlates with decreasing magnitude responses of the oriented filters for early levels of the steerable pyramid (Level 1 is shown here). b) Increasing coating bumpiness correlates both with changes in the magnitude responses of oriented filters, and with a decrease in the autocorrelation of the real part for early levels (the high-pass residual is shown here). For each plot, lines show data from a single viewing angle of the objects.

15.4.3 *Modeling metallicity*

We then sought to model the results of the conjoint measurement task assessing metallicity of the stimuli. In keeping with the literature on gloss, which suggests that glossiness in natural images may be inferred by global luminance distributions, we

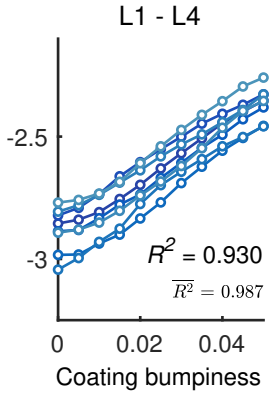


Figure 15.14: Taking the ratio of the linear filter outputs at Level 1 and Level 4 of the steerable pyramid strongly correlates with coating bumpiness, even across viewing angles. A visual system estimating bumpiness would therefore exhibit good material constancy and viewpoint invariance, while individual steerable pyramid level outputs (shown in Figure 15.13), and global spatial-frequency analysis (shown in Figure 15.9), would give poor viewpoint invariance.

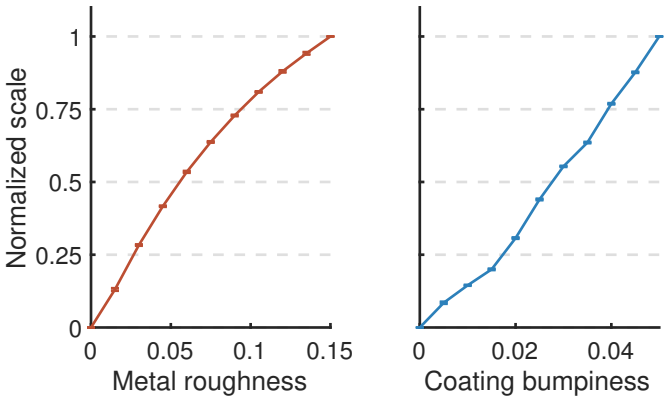


Figure 15.15: Simulation results of the MLDS experiment, using models that estimate metal roughness and coating bumpiness. Left: Results of a model estimating roughness according to the activations of Level 2 linear filters in a steerable pyramid. Right: Results of a model estimating coating bumpiness by the ratio of Level 4 filters to the sum of Level 1 and Level 2 filters. Error bars show 95% confidence intervals obtained via bootstrap; the robustness of the models results in small errors, despite the simulated experiment requiring comparison across viewpoints.

compared a range of models that used different estimators of global image contrast to model task performance, the results of which are found in Figure 15.16. In total eight models were designed, using global contrast parameters summarised in Table 15.1. Models estimating metallicity according to dynamic range, either simply taking the ratio of the brightest and dimmest pixel (DR), or a more robust average of the brightest and dimmest 10 pixels (rDR), showed an increase in metallicity judgement with increasing smoothness of the metal base, as seen in observer data. However,

these models were not affected by coating bumpiness, while all of the observer data sets were best fit by additive models rather than independent. Models using the skew of the luminance distribution of images (SK), or Michelson contrast (MC), the ratio of the difference between the brightest and dimmest pixels and their sum, gave similar results.

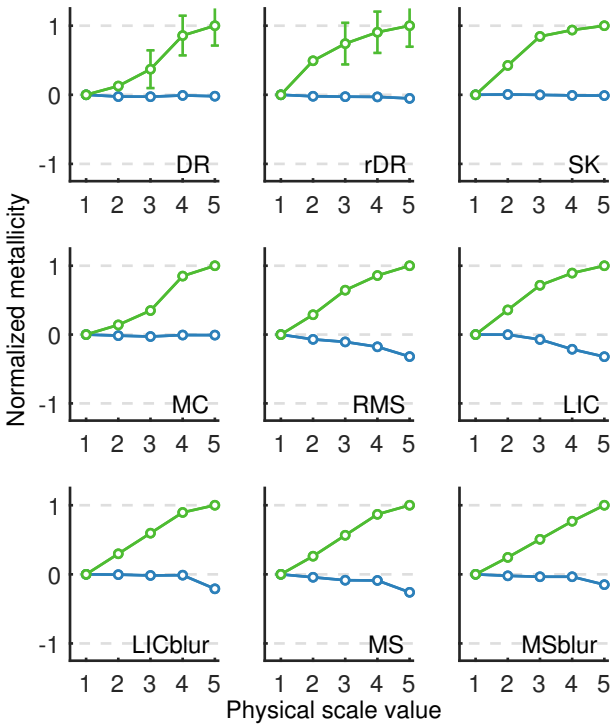


Figure 15.16: Conjoint measurement results from the models using different estimators of global contrast, shown in Table 15.1. Estimated contributions to metallicity discrimination are shown for metal smoothness (green) and coating bumpiness (blue). Four of the models are not significantly affected by coating bumpiness: dynamic range (DR), robust dynamic range (rDR), skewness (SK), and Michelson contrast (MC). The remaining five models, root mean square (RMS), local image contrast (LIC), local image contrast calculated against a blurred image reference (LICblur), Meese and Summers contrast (MS), and Meese and Summers contrast calculated against a blurred image reference (MSblur), give broadly similar results, with some agreement with the mean additive fit from observers shown in Figure 15.5. Error bars show 95% confidence intervals obtained via bootstrap.

A model using root mean square (RMS) as a global contrast measure gives rise to results highly concordant with those from observers, which is not surprising given its established record as a reliable indicator of image contrast (Bex and Makous, 2002). However, as Bex and Makous remark, a global metric such as RMS is difficult to relate to channel models of the human visual system, and is unlikely to be computed

Abbrev.	Name	Equation
DR	Dynamic range	$\frac{L_{max}}{L_{min}}$
RDR	Robust dynamic range	$\frac{1/n \sum_{i=1}^{n} L_{max(i)}}{1/n \sum_{i=1}^{n} L_{min(i)}}$
SK	Skewness	$\frac{1}{N} \sum_{i=1}^{N} \frac{(L_i - \mu)^3}{\sigma^3}$
MC	Michelson contrast	$\frac{L_{max} - L_{min}}{L_{max} + L_{min}}$
RMS	Root mean square	$\sqrt{\frac{1}{N} \sum_{i=1}^{N} (L_i - \mu)^2}$
LIC	Local image contrast	$\sum_{i=1}^{N} \lvert L_i - \mu \rvert$
LICblur	Local image contrast (blurred)	$\sum_{i=1}^{N} \lvert L_i - B_i \rvert$
MS	Meese and Summers	$\sum_{i=1}^{N} \frac{\lvert L_i - \mu \rvert^{2.4}}{1 + \lvert L_i - \mu \rvert^2}$
MSblur	Meese and Summers (blurred)	$\sum_{i=1}^{N} \frac{\lvert L_i - B_i \rvert^{2.4}}{1 + \lvert L_i - B_i \rvert^2}$

Table 15.1: A summary of the global contrast estimators used in modelling conjoint measurement assessment of metallicity. N indicates the total sample number (i.e. the number of pixels in the image, while n indicates the number of a sub-population. B is a highly blurred copy of the original luminance image, L.

by observers. A method of estimating local image contrast (LIC), by simply averaging differences between each pixel and the image mean (for the region containing the object) gave good results, with metallicity increasing monotonically with metal smoothness, and decreasing gradually with coating bumpiness. If differences were computed not against the image mean, but rather a local mean (the same pixel in a heavily blurred image), the model (LICblur) also gave similar results to participant data. We also tried the Meese and Summers method for estimating global contrast with gain control (Meese and Summers, 2007), which has recently been found to more accurately predict observer performance in contrast estimation tasks (Meese et al., 2017). Models employing this estimator, using either the mean of the image or

pixels in a blurred image to calculate local contrast, MS and MSblur, respectively, both performed similarly to the observer average.

Despite the success of many of these models in recreating the observer results from the conjoint analysis experiment, none is plausibly representative of the processing that may be achieved by the visual cortex and its physiology. RMS, a frequently used metric for evaluating image contrast, gives the best fit to experimental data, but the operations of calculating local luminance and squaring have no evidential basis in the visual system. Meese *et al* developed an estimator that captures observer performance better than RMS, but it too requires local luminance measurement, which they suggest could be achieved by "a convolution kernel that is, or approximates, a non-oriented delta function from the weighted sum of the oriented basis filters (or wavelets)". This seems unlikely, and even if it were the case, summation over such a kernel's outputs would simply integrate luminance, not contrast, requiring another step to compute the variance or some other summary statistic of these values (i.e., the visual system would be calculating differences, in order to calculate absolute levels, in order to calculate differences). Instead, we considered a model that computes global contrast directly from the outputs of local linear filters, which are combined across different spatial scales in a nonlinear, gain-controlled fashion. With this approach, the percepts of metal roughness, coating bumpiness, and metallicity may all be framed as functions of the outputs and parameters of steerable pyramids, shown diagrammatically in Figure 15.17. This is in keeping with an established, earlier 'generalized multiple-channel model of visual detection and perception' (Harvey and Gervais, 1978; Schnitzler, 1976a,b).

One question is whether a simple computational model, taking as inputs the linear filter responses across different spatial scales, can simulate the performance of observers during the conjoint measurement of metallicity. To address this, a computational model was built that took the responses of just two different levels of the steerable pyramid—Level 1 (corresponding to fine details in the image) and Level 4 (corresponding to coarse image features). The model took the weighted outputs of these levels and compared them. As shown in Figure 15.18, the model has only

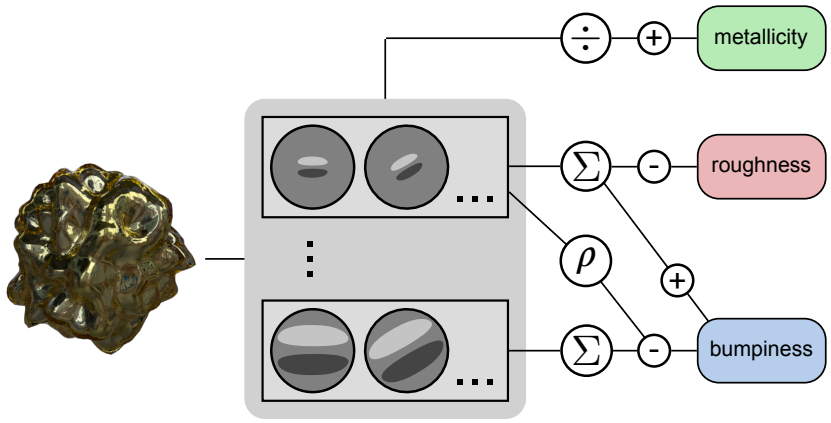


Figure 15.17: The models of roughness, bumpiness and metallicity using only the outputs of linear filters. Metal roughness may be modelled by inverse correlation with the summation of small-scale linear filters. Coating bumpiness may be modelled by positive correlation with the same small-scale linear filters, and negative correlation with their autocorrelations, and with large-scale linear filters. Metallicity may be computed as a weighted, gain-controlled combination of the filter bank.

two free parameters—the weightings (which may take either positive or negative values) for each level. As Figure 15.19 shows, this two-parameter model can give very similar results to observer data, when the weightings of each level are fit via global optimisation to each observer's conjoint measurement model. The same figure also shows how a model performs that is using only the outputs of Level 1, that is, a model that is essentially estimating roughness differences as a proxy for metallicity. This model performs similarly to P4 and the inverse of P5 (ie. P4 is equating smoothness with metallicity, while P5 is equating roughness with metallicity), while the conjoint measurement models of other observers agree more with a Level 4-only model. This is also apparent by examining the optimised weights for the computational models fit to each observer, shown in table 15.2, as for both P4 and P5 the optimal weighting for Level 1 has a higher magnitude than Level 4.

15.5 EXPERIMENTS WITH A HIGH DYNAMIC RANGE DISPLAY

Metallicity may be related to perceptual judgements of glossiness—although the two should not be confused—and the dynamic range of stimuli has been found to limit

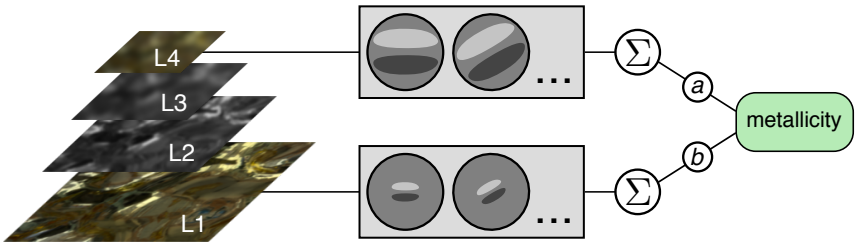


Figure 15.18: The computational model of metallicity, taking a weighted comparison of the outputs of Level 1 and Level 4 of the steerable pyramid decomposition. The pyramid is shown on the left, with higher levels resulting from applying a Gaussian blur to preceding levels and downsampling the spatial resolution by a factor of 2 (not to scale). The same linear oriented filters operate on all levels, such that lower levels generate responses owing to an image's fine details, and higher levels generate responses owing to coarse features. The model has only two free parameters, a and b, the weightings of responses at Level 1 (fine details) and Level 4 (coarse features).

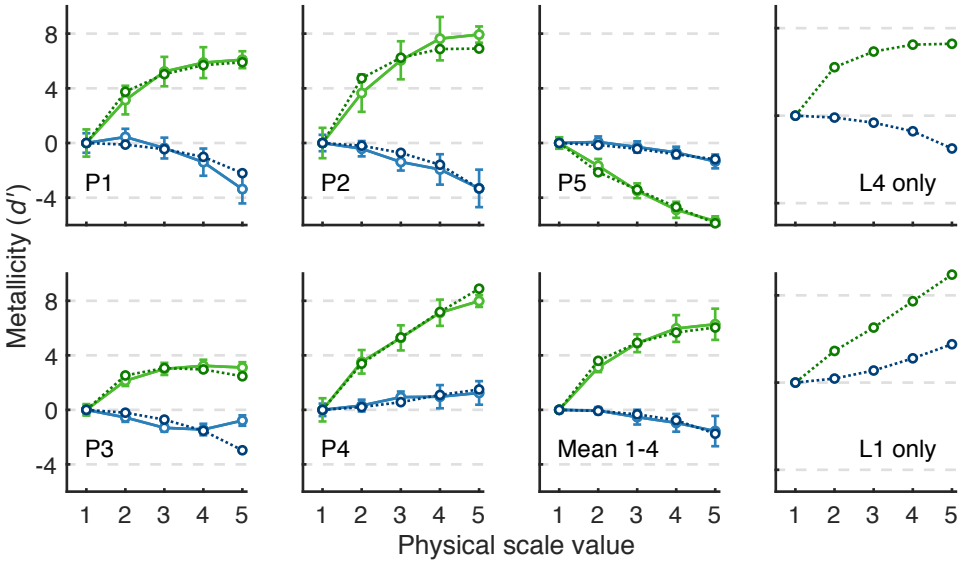


Figure 15.19: A simple computational model of metallicity that compares the weighted outputs of Level 1 and Level 4 oriented filters accounts for the observer data. Conjoint measurement contributions are shown for metal smoothness (green) and coating bumpiness (blue). Also shown are the conjoint measurement contributions for computational models using only a single level of the pyramid. Level 4, corresponding to coarse image features, at the top right, Level 1, corresponding to fine image features, at the bottom right. Error bars for participant data show $\pm SEM$ ($N = 4$) and for observer model results show 95% confidence intervals obtained via bootstrap.

perceptions of glossiness (Adams et al., 2018). From the results of the experiments and modelling, it was likely that global image contrast, and the visual system's estimation of contrast in the reflected image of an object, could be a key determining factor for metallicity. We therefore carried out the same experiments but under

Observer	L1	L4
P1	0.488	12.909
P2	-0.166	17.316
P3	-1.832	11.514
P4	5.527	4.988
P5	-3.852	-2.691
Mean 1-4	1.0041	11.6819

Table 15.2: The optimal weightings of Level 1 and Level 4, for a computational model of metallicity fit to observer data.

an HDR condition. As we did not change the physical parameters of the objects themselves for rendering, this was akin to viewing the same objects in a different lighting environment. It may therefore be expected that, due to the robustness of the visual system and its mechanisms of constancy, we would not see any change in appearance judgements made by observers.

15.5.1 *Method*

Stimuli

For this condition, the physical parameters for stimuli (metal roughness and coating bumpiness) were kept the same, but they were rendered using a light probe with a significantly higher range of luminance values than in the SDR condition, the 'Kiara interior' probe made by Greg Zaal (https://hdrihaven.com/). This environment has an unclipped luminance range captured over 12 EVs (i.e. a wide range of exposure values, $2^{12} : 1$). The luminance distributions of the two environments, used for both the SDR and HDR conditions, are shown in Figure 15.20. The environment used to render stimuli for the HDR condition has a population of pixels centred at around 700 times the median value of the light probe, while the environment used for rendering the SDR condition's stimuli has a population of brightest pixels centred at only around 25 times the median of the light probe.

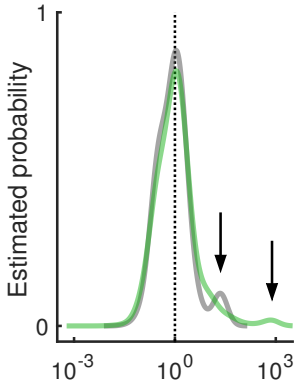


Figure 15.20: Distribution of luminance values in the light probes used for both the SDR (grey line) and HDR (green line) rendering conditions, normalized by their median values and fit with a kernel density estimator. Note the logarithmic x-axis. Light sources will be represented by only a small number of pixels. The HDR condition light probe contains a population of pixels centred at around 700 times the median value, while the SDR condition light probe's brightest pixels are centred at around 25 times the median (both indicated by arrows).

Participants

For the high dynamic range condition, five observers were recruited from both the Oxford Department of Experimental Psychology and the Computer Laboratory at the University of Cambridge.

Ethics

All participants gave their informed consent prior to testing.

Procedure

Experiments for the HDR condition took place at the University of Cambridge. Stimuli were viewed in a dark room on an HDR display, comprising a DLP projector and an LCD display, capable of reaching a maximum luminance of 15,000 cd/m^2, with a black level of below 0.001 cd/m^2. Spatio-temporal dithering was used to achieve a 10-bit colour channel resolution on both DLP and LCD, with a spatial resolution of 2048 x 1536. Viewing distance was maintained with a chin rest positioned 92 cm from the LCD panel.

15.5.2 *Results*

15.5.2.1 *Experiment 1: HDR Perceptual Scaling*

Results of the perceptual scaling for the HDR condition are shown in Figure 15.21, and agree well with the SDR condition. An exception was found in the estimated difference scales obtained for the roughness condition; for some participants, estimated scales showed very little perceptual difference between the conditions of no roughness and the smallest degree of roughness. This could be due to glare effectively blurring the least rough condition for the high luminance levels at which the stimuli were displayed. A logistic function was therefore a better fit of the aggregate data than the polynomial fit from the SDR experiment, as shown in Figure 15.22. Importantly, the perceptually-even intervals obtained from fitting the aggregate SDR difference scale estimates still gave significant perceptual difference in the scales estimated from the HDR condition, validating that the same parameter space of stimuli could be used for the HDR condition of the conjoint measurement experiment.

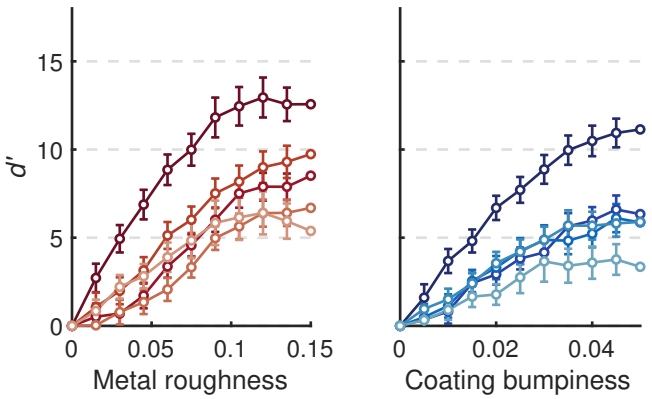


Figure 15.21: Perceptual scale estimates for individual participants for metal roughness (red, left) and coating bumpiness (blue, right) in the HDR condition. Values are given in d'.

15.5.2.2 *Experiment 2: HDR Conjoint analysis of metallicity judgments*

As for the SDR condition, none of the participants gave responses that could be sufficiently described with an independent model of conjoint measurement. Again,

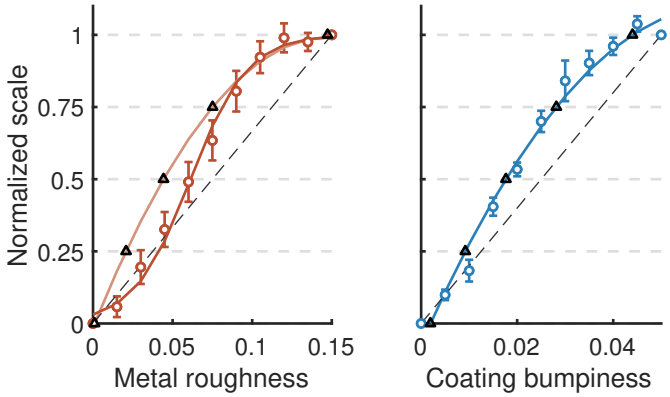


Figure 15.22: Mean perceptual scale estimates from aggregate data for metal roughness (red, left) and coating bumpiness (blue, right) in the HDR condition. A uniform scale is indicated by a dashed line. The second-order polynomial fits obtained from the SDR condition are replotted for roughness and bumpiness (in orange and blue, respectively). For roughness scaling, a best-fitting logistic function is also shown (red). Error bars show $\pm SEM$ ($N = 5$)

there was a subset of participants whose responses could be described with an additive model, and there was a subset who required a saturated model. Model fits to the data varied widely between observers, with one data set best fit by a model with a very minor deleterious contribution from metal smoothness to metallicity, and another showing a strong inverse relationship—P5, the same participant whose data show a negative contribution of metal smoothness to perceived metallicity in the SDR condition. Additionally, all observers reported after the experiment that the task was difficult, and that the majority of stimuli looked convincingly made of metal.

Two observers (P5 and P7) gave responses that were significantly better fit by saturated models, rather than additive. These are shown in Figure 15.24. For both saturated models, increasing the bumpiness of the coating for the most metallic-appearing roughness level (i.e. the objects with the roughest metal for P5, and the smoothest metal for P7) causes a decrease in perceived metallicity, while increasing coating bumpiness for objects with the least metallic roughness (those with the smoothest metal for P5, and the roughest metal for P7) results in an increase in perceived metallicity.

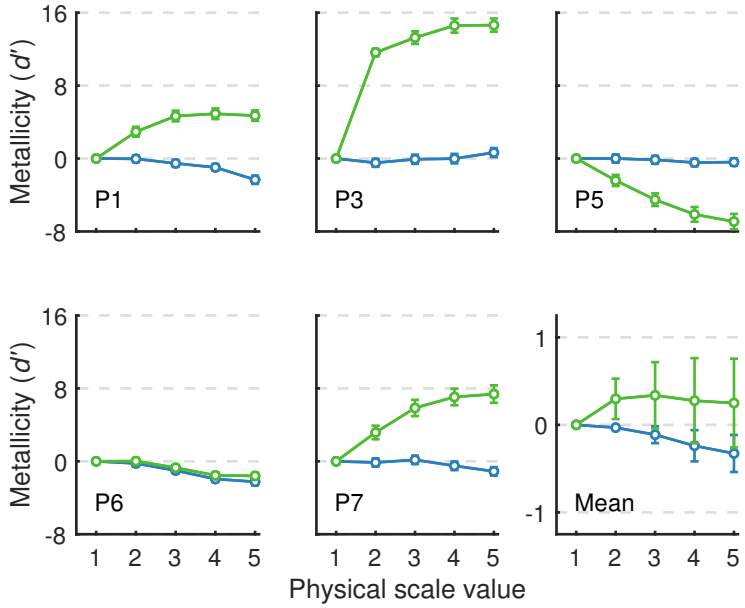


Figure 15.23: Conjoint measurement estimates for individual participants fit to the 'additive' model, for the HDR condition. Contributions to metallicity discriminations are shown for metal smoothness (green) and coating bumpiness (blue). Error bars for individual participants show 95% confidence intervals obtained via bootstrap. For the aggregate data in the lower right-hand plot, error bars show $\pm SEM$ ($N = 5$).

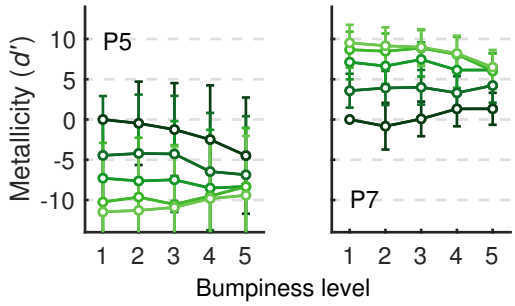


Figure 15.24: Conjoint measurement estimates for individual participants fit to the 'saturated' model, for the HDR condition. Increasing coating bumpiness is shown along the abscissa, while increasing metal smoothness is indicated by brightening green colour. Error bars show 95% confidence intervals obtained via bootstrap.

15.6 DISCUSSION

Investigating material perception presents some particular challenges to visual neuroscientists and psychophysicists. Stimuli require both verisimilitude and versatility; they must be convincing representations of the material percept in question, and they should have properties that can be predictably and independently varied. Our stimuli,

of physically-based rendered computer graphics, meets these requirements. The model of a coated metal object, with varying metal roughness and coating bumpiness, allows for a continuous parameter space extending across both metallic and plastic appearances, and every location in this parameter space is a (synthetic) natural image. Additionally, we have shown that these parameters predictably determine image statistics in the stimuli, through a steerable pyramid analysis, which are then directly relatable to the percept of metallicity.

For standard dynamic range viewing, participants mostly find objects with smoother metal more metallic, confirming a previous finding for ambient lighting conditions (Todd and Norman, 2018). However, one participant (P5) showed the opposite effect, raising questions over the homogeneity of the population of observers for metallicity judgements, and material perception in general. Indeed, it is reasonable to assume that visual material judgements are heavily influenced by each observer's typical environment, in this case whether they are more likely to see matte metallic surfaces such as anodized aluminium electronics, or brushed metal appliances etc., rather than highly-polished and mirror-finish objects. If the former is the case, 'shiny' objects' appearances may be more frequently associated to plastic-coated objects, or—and particularly if seen without a background for context—to transparent glass objects. As objects in our stimuli were presented in a void, one initial question the perceptual system may have to answer is whether they are reflecting light or transmitting light—in other words, mirrors or windows? For a perfectly polished metal surface, which reflects a mirror image of a given environment, there exists a corresponding transparent object of another environment, which gives rise to the same image for the observer. Tamura et al. recently investigated this ambiguity in material perception, and found that without motion cues from dynamic stimulus presentations, observers classify mirror objects as glass between 8 and 40% of the time, depending on the object geometry (Tamura et al., 2018). This study only considered perfectly smooth surfaces, but we can speculate as to how surface roughness might improve or disrupt classifications. When light reflects off a rough metal surface, microscale geometric inhomogeneities have the effect of

scattering a fraction of the light at angles other than the angle of reflection, the result being a blurring of the image. For rough transparent objects, light is scattered at both entry and exit points. Rather than simply blurring the transmitted image, this also tends to flatten and desaturate it, giving a 'frosted' or translucent appearance, as light from illuminants may be distributed to any region of the object after the initial scrambling at entry. This effect is shown in Figure 15.25, where silver and glass objects are rendered with different levels of physical roughness. For silver and glass objects whose roughness gives a similar blurring of the reflected environment, glass objects show a more tightly constrained luminance distribution, lacking the extreme highlights and shadows present over the surface of the metal object. We would therefore expect that observers who confuse smooth metal for glass might be less prone to this error for objects whose surfaces are not perfectly smooth.

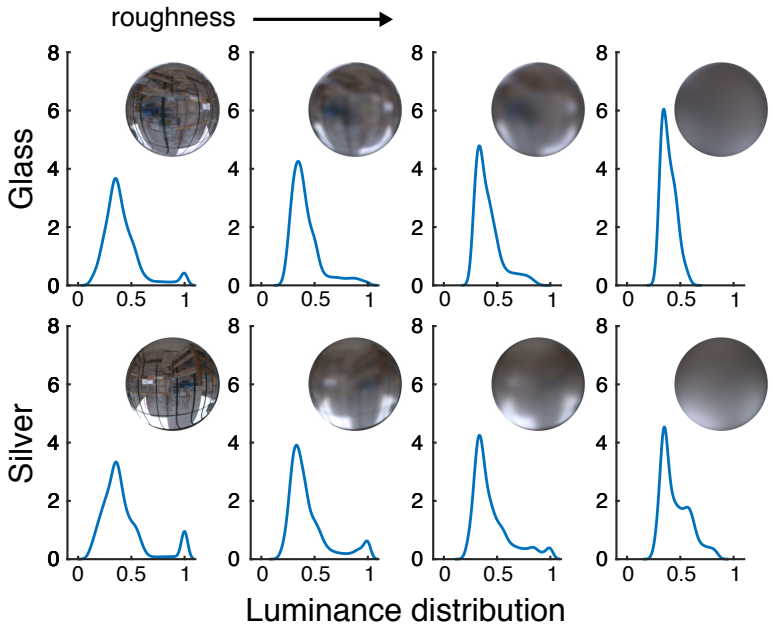


Figure 15.25: Silver and glass objects, rendered with different physical roughness. The objects have a similar luminance profile when completely smooth, making it hard to discern which material is which. However, increasing roughness has a more pronounced effect on the glass object than the silver object, as the luminance distribution is more tightly constrained, despite a similar apparent level of image blurring.

As increasing metal roughness tends to decrease the strength of metallic appearance in standard display conditions, this may initially suggest that observers are simply

estimating the degree of articulation within the reflected image of the object's environment, and relating that to metallicity. Rougher metals blur specular reflections, attenuating high spatial-frequency components as seen in Fourier power spectra. However, the negative contributions to metallicity of coating bumpiness, which actually increases the degree of articulation over the object's surface, suggests a different mechanism. Two possibilities remain when the experimental results are considered in the light of the modelling we have undertaken. It could be that fine (high spatial-frequency) details and edges are less important to observers than broad (low spatial-frequency) details. Coating bumpiness increases fine edges while disrupting smoother, gradual components, and this is particularly the case when the reflected environment image is already blurred (in the rougher regions of the parameter space). Or, it could be that observers are selectively relating fine details to metallicity only when they are seen as part of the natural image of the environment, whose edges are normally continuous and extended. Although coating bumpiness increases the strength of fine details on the object's surface, these show a reduction in continuity (i.e. correlations between neighbouring oriented filters of the same orientations). It is possible that observers are sensitive to this, and that the resultant irregularly wrinkled appearance does not contribute to perceptions of metallicity in many observers, and perhaps may even disrupt it.

The modelling we have undertaken additionally reinforces the work of others[1] in illustrating some important limitations for using Fourier analysis (and particularly amplitude spectra) to probe the human visual system. Visual stimuli in the real world are not composed of infinitely-extending sinusoidal frequencies, and in fact are rarely periodic at all in the strict sense that a global Fourier analysis is best tuned to deal with. Instead, our visual system is primarily concerned with higher-order image structures such as edges, and local differences in luminance and colour. These

1 See Westheimer, 2001 for a review of the historical origins of approaching vision with a Fourier perspective, and how vision science reckoned with mounting evidence of local organisation in the visual pathway. This can be readily seen in use of Gabors (Gaussian-enveloped sinusoids) rather than sinusoidal gratings as stimuli for psychophysics. It can also be seen in the development of computational tools and their application to vision, from Fourier analysis to Gaussian, Laplacian, wavelet and—eventually—steerable pyramids (Adelson et al., 1984; Beyerer et al., 2016).

local features may be completely described by various frequency components and their phase alignment, but the phase spectrum of a natural image is generally not informative for analysis (Franz and Schölkopf, 2005). Disrupting the phase spectra of natural images results in decreased edge contrast (Wichmann et al., 2006); in the limiting case, where the phase spectrum is scrambled or consists of random noise, the resultant image is pink noise—devoid of any salient edges, corners, or features that we would identify with perceptual sharpness—despite having the same amplitude spectrum (and high frequency components) as the original image. Steerable pyramid analysis of pink noise does show a reduction in the oriented filter outputs of early levels, as compared to the original image, which agrees with the perceptual loss of edges and fine details.

For this study, we have performed experiments using both a standard dynamic range (SDR) condition, and a high dynamic range (HDR) condition, using an HDR display. We find that when the same 'objects' (simulated material conditions) are viewed under HDR conditions, even those that observers were confident did not look metallic under the SDR conditions are now unmistakably metallic. This was reported by all observers, who said that due to this, the MLCM task (of selecting the most metallic object) was very difficult. Importantly, observers could still discriminate between different levels of metallic roughness and coating bumpiness, as found in the HDR MLDS experiment. Due to this, it is likely that observers adopted a strategy to make judgements for the MLCM task, besides simply estimating which object was more likely to be made of metal, and we therefore cannot place too much stock in the resultant observer models. Computational models that used global image statistics to predict metallicity gave results concordant with the informal reports of observers, namely that even the roughest and bumpiest objects still appeared highly metallic, as shown in Figure 15.26.

In this study, stimuli have consisted of objects devoid of any background, although it is almost certainly the case that in everyday vision, observers are heavily influenced by environmental context in their appraisal of metallic surfaces. In such cases, the relative luminance of surfaces with respect to environmental illumination is likely

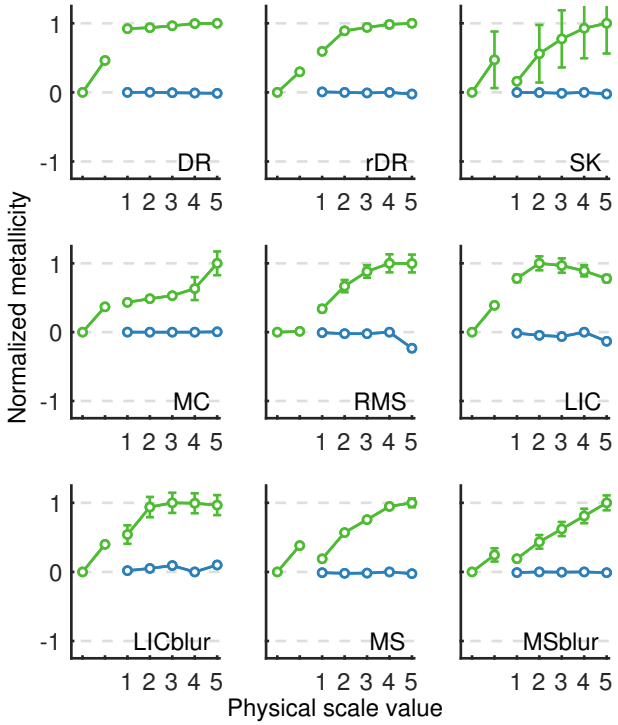


Figure 15.26: HDR conjoint measurement simulation results, using different models to estimate metallicity. Each plot shows the relative scaling of contributions to metallicity from levels of metal smoothness (shown in green) and coating bumpiness (shown in blue). The first two steps of metal smoothness indicate the roughest, and smoothest, images from the SDR condition. For the majority of models, the roughest object in the HDR scene is judged as more metallic than the smoothest object in the SDR scene.

one of the most reliable potential cues, as typical metallic surfaces reflect all of the incident light falling on them. This may be particularly critical for the perception of very rough, even matte, metal surfaces such as anodized aluminium, or patinated metals. The surfaces of such materials cannot be said to be devoid of a 'diffuse' appearance, but rather fall along a continuum of matte and mirror appearance. It is therefore also possible that there is not one single visual feature that confers the appearance of metallicity, but rather several categories e.g. polished, matte, painted, rusted etc.

15.7 CONCLUSION

Metallicity remains an intriguing special case of material perception, and one with high significance for several commercial and technical sectors. The work presented here indicates that the material judgement of metallicity does not correspond simply with metal roughnes, but that other factors such as coatings, lighting and viewing conditions, and individual differences play a role. Moreover, we have shown that while global image statistics are unable to account for observer decisions in a conjoint measurement task, a simple model incorporating low-level visual features—as represented in the activations of oriented filters at different spatial scales—was able to simulate the results for all individual observers.

16

THREE-DIMENSIONAL COLOUR DISTRIBUTION WARPING FOR GOLD APPEARANCE TRANSFER

16.1 INTRODUCTION

In chapter 15, it was found that global contrast across an object's surface could be a critical image heuristic for metallicity, and the visual classification of an object as made of metal. In fact, there has been ongoing debate over the use and relevance of global image statistics for the perception of material qualities, particularly gloss(Anderson and Kim, 2009; Motoyoshi et al., 2007; Wiebel et al., 2015). In this chapter, I argue that this debate can be reconciled by limiting considerations to the subset of natural images, for which global statistics are necessarily spatially allocated in a deterministic way. The statistical operation of distribution warping provides a means of arbitrarily manipulating global statistics without violating the spatial organisation of picture elements, and higher-order statistics, of natural images. Operating on three-dimensional colour distributions is therefore a simple yet powerful approach in graphics and image processing. I will show that a combination of image segmentation and three-dimensional colour distribution warping can be used to transfer a gold-like material appearance to objects in both real and rendered images. I also present counter-examples where this approach fails to produce a convincing or visually attractive output, and offer some discussion of why this may be the case.

16.2 HISTOGRAMS AS STATISTICAL CORRELATES FOR MATERIAL APPEARANCE

Global image statistics do not, in and of themselves, dictate the material appearance of images. This may be plainly illustrated by scrambling the pixels of an image; when spatial information is lost the resultant image is noise, and global image statistics such as the mean, skew, and kurtosis of luminance distributions are relatively meaningless, as in Figure 16.1. Nevertheless, a considerable amount of debate has resulted within the vision community over the value, importance, and perceptual relevance of global (spatially-invariant) image statistics when it comes to estimating material qualities such as gloss and surface finish (Anderson and Kim, 2009; Motoyoshi et al., 2007; Wiebel et al., 2015). This debate, and the apparent otiosity of global statistics illustrated in Figure 16.1, may be reconciled by limiting the discussion to *natural images*. This limitation is justified: if the visual system were in some way using image statistics for estimations of material quality, these relations could only be learnt from natural images. Natural images, and their processing in human visual cortex, are a central feature of vision research and their properties have been well studied (Field, 1987; Nascimento et al., 2016; Ruderman and Bialek, 1994; Simoncelli and Olshausen, 2001; Stephens et al., 2013; Thomson and Foster, 1997). Within natural images, global image statistics are distributed spatially and deterministically according to the scene, comprising light sources and objects (surfaces), as a function of the geometric relations between them and the materials that determine light transport. It is therefore unsurprising that scrambling or rearranging elements of a natural image is sufficient to violate relationships between global statistics and visual percepts of material qualities.

Distribution warping—a generalisation of the common practice of distribution normalisation or equalisation—provides an approach to modify the global statistics of an image or region without perturbing the spatial allocation of these statistics. In classical one-dimensional distribution normalisation (often in the narrow case of 'histogram normalisation'), an image is modified such that it has an even distribution

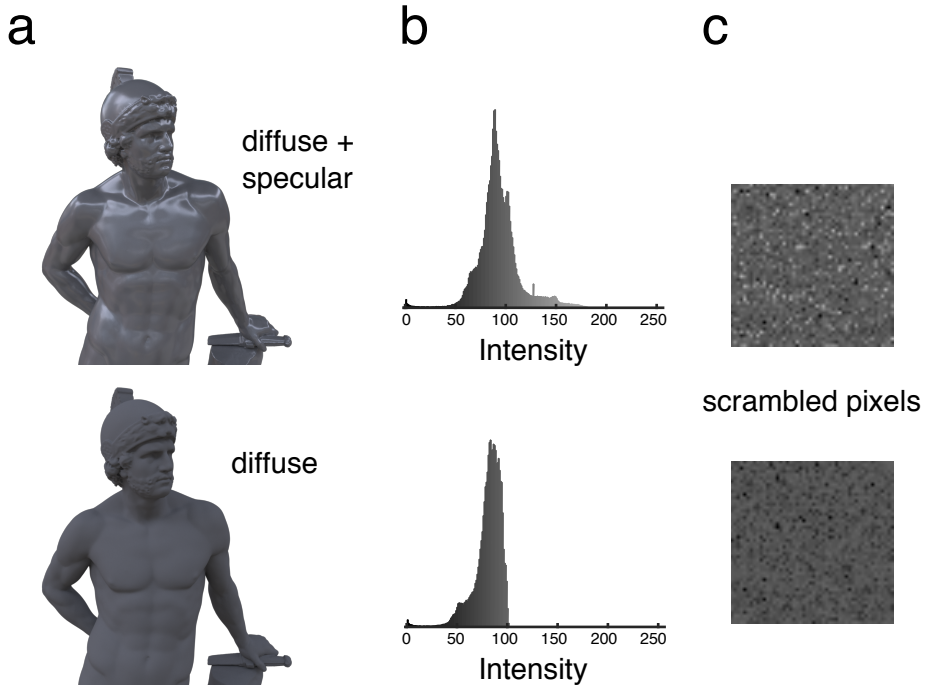


Figure 16.1: One-dimensional luminance histograms of an object rendered with a matte and shiny material appearance. a) The same statue of Mars rendered with a diffuse (bottom) and diffuse + specular (top) reflectance function. b) There is a clear difference in the luminance histograms of the two images. However, specific spatial organisation of these distributions is necessary to give impressions of 'mate' or 'gloss'. c) Scrambling the pixels of the statues results in neither a matte nor a glossy appearance, though the results share the same chromatic statistics of the original images.

of luminance values, i.e. every luminance value is equally likely to feature in the image. This typically allows details to be seen in otherwise low-contrast and perceptually 'flat' images, while—crucially—preserving the ranked order of pixel values, that is, the comparatively brighter pixels remain as such. Warping also preserves the ranked order of pixel values, but rather than flattening the distribution, fits it to any given shape. The mathematical origins of optimising this operation are found in transportation theory, first formulated by Gaspard Monge in his 1781 essay, *Mémoire sur la théorie des déblais et des remblais* [*Thesis on the theory of holes and piles from digging*] (Monge, 1781), and more recently as the earth mover's distance (EMD). The EMD presents a true metric of the similarity between multidimensional distributions, providing they have the same overall mass, and was indeed first used as a corollary for perceptual similarity between images in terms of colour and texture distributions

(Rubner and Tomasi, 1998; Rubner et al., 1998, 2000). For colour distributions, the 'ground distance'—the distance between basic elements in the distribution—was established by psychophysics and represented in (approximately) perceptually uniform colour spaces such as CIEu'v' (CIE, 2014).

The EMD provides a solid metric for comparison, but does not immediately offer a method of warping—the transfer of one distribution to another in an optimal manner. Intuitively, luminance distribution warping is optimised by minimising the ground distance without changing the rank order of pixel values. A rough approximation of this, employed in rudimentary image processing, is to operate on a histogram, shifting the contents of entire bins up or down so that they best match the expected capacities. Generalising to probability distributions of higher dimensions, implementation becomes more taxing, both conceptually and computationally. For three-dimensional colour, one cannot necessarily apply transformations to dimensions (i.e. the RGB colour channels) independently, as their distributions are often correlated in a perceptually salient way that should be preserved. This is illustrated in Figure 16.2. Converting first to a decorrelated colour space such as $l\alpha\beta$ (Ruderman et al., 1998) results in significant improvement, as used in an initial colour transfer approach that adjusted the means and variances of three decorrelated distributions (Reinhard et al., 2001). The authors suggest that accounting for 'higher moments such as skew and kurtosis' of the distributions, i.e. allowing non-linear transformations, would improve results further, but do not suggest tackling the problem in a fully three-dimensional manner.

16.3 THREE-DIMENSIONAL COLOUR HISTOGRAM TRANSFORMATIONS

For full three-dimensional distribution warping—specifically colour transfer—an algorithmic approach was developed by Pitié and Kokaram (Pitie and Kokaram, 2007), which is implemented here. Rather than simply matching the means and variances of decorrelated colour channels across images, this algorithm approximates a complete

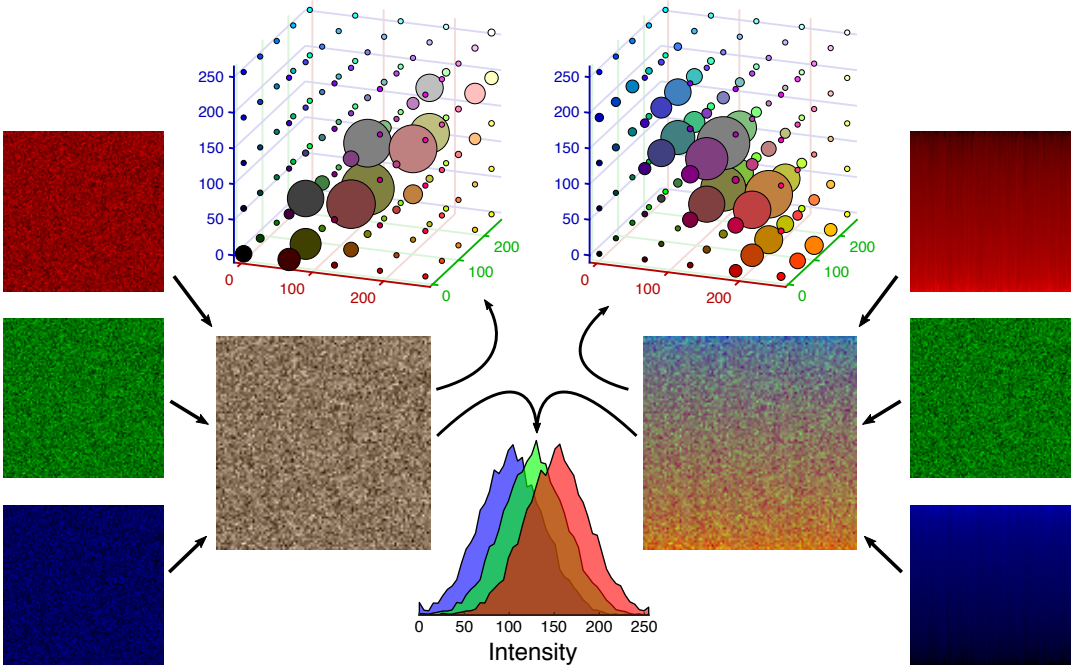


Figure 16.2: Single-channel RGB histograms fail to describe perceptual colour distributions, as the spatial distributions of the channels are correlated in a perceptually salient manner. On the left and right of the figure, noisy RGB images are generated from Gaussian distributions for each channel. For the right-side image, the red and blue channels have been sorted vertically, with red decreasing and blue increasing up the image. The resultant colour images have remarkably different chromatic features, while having the same single-channel RGB histograms, shown bottom-centre. The fully three-dimensional histograms, shown at the top of the figure, faithfully reflect the differing chromatic distributions of the two images.

matching of fully three-dimensional distributions. The transformed images therefore retain the spatial organisation of natural scenes, allowing the editing of global image statistics that correlate with perceptual judgements of material appearance. The image elements are still spatially organised into objects of distinct colours, with the shading of these objects still determined by their geometric relation to light sources and the sensor and the material properties of their surfaces.

The algorithmic approach to transfer objects a gold-like material appearance is presented as a flowchart in Figure 16.3. It requires two input images, a source containing a gold object, in this case a photograph of an Academy Award, and a target containing the object to be gilded, in this case an apple. First, the gold source image is segmented, and the pixels identified with the golden object are used to compute a

three-dimensional histogram. The target image is likewise segmented, and pixels identified with the object to be transformed are reshaped into a square—this will permit de-graining. The three-dimensional colour histogram of the source square is then reshaped to approximate that of the target, and undergoes de-graining. Finally, the source square is remapped onto the original image, which now shows a golden apple.

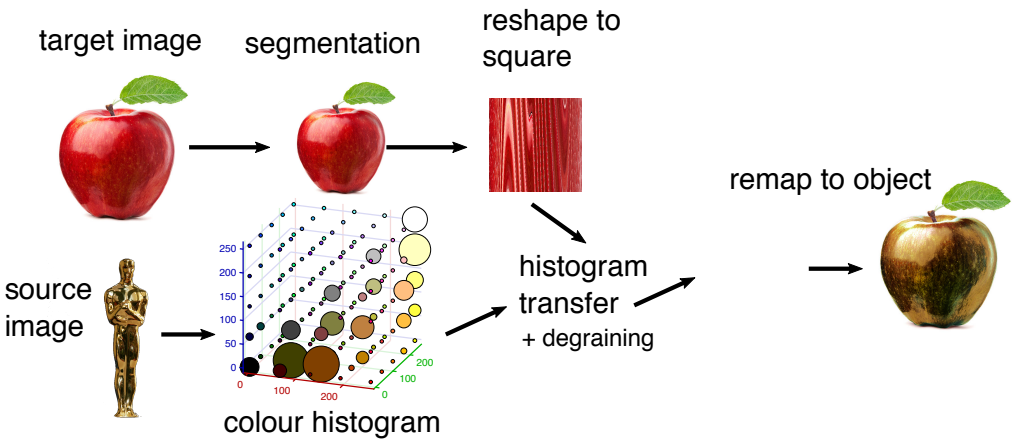


Figure 16.3: The algorithm for transferring a gold-like appearance to objects in images, using three-dimensional colour histogram warping.

16.4 RESULTS OF THE GOLD APPEARANCE TRANSFER ALGORITHM

Figure 16.4 shows the results of applying the three-dimensional colour histogram warping algorithm to the images presented in Figure 16.1. It is evident that the approach fails for base images of matte (diffuse) appearance. In these cases, what little chromatic variation there is over the surface of the object is primarily a result of the object's geometry, and not the articulation (spatial complexity) of the object's environmental illumination. As such, the algorithm simply increases the shape-from shading cues of the object. However, when a specular reflection is present in the base image, this enables the chromatic distribution that is being transferred to distribute not only according to shape-based shading, but also with some correspondence to

the articulation of the environment. The result is a fairly good approximation of the object if it were made of gold (the 'ground truth' is also shown in Figure 16.4).

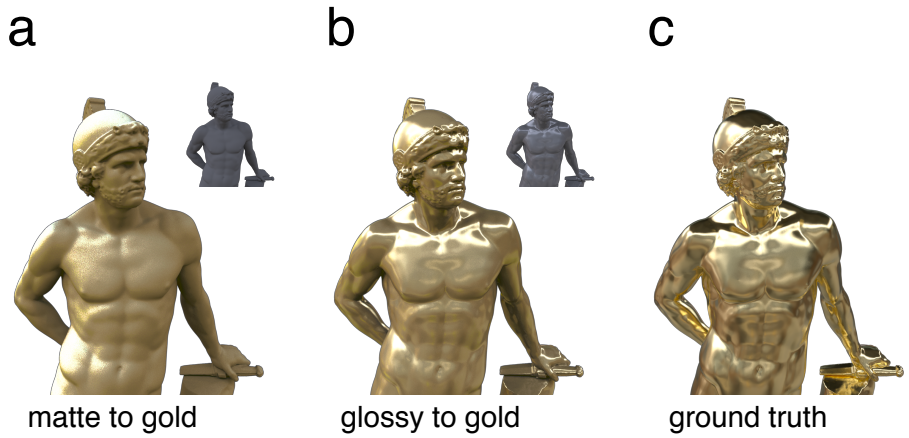


Figure 16.4: Results of the gold material transfer algorithm. a) Using a base image with a matte material. b) Using a base image with a glossy plastic material. c) The 'ground truth'—the same object rendered in real gold, which also served as the source of three-dimensional histogram to be transferred.

16.5 CONCLUSION

In summary, three-dimensional colour histogram warping may be a powerful technique to transfer not only colour palettes and overall aesthetics, but—when combined with object segmentation—may enable the transfer of material properties in some limiting cases. The success of this approach is limited by the properties of the base image, and whether the pre-existing chromatic variations are spatially concordant with shape-from-shading or material properties. Histogram warping may also prove to be a useful tool in the creation of novel stimuli for psychophysical investigations into the visual perception of materials, as it provides the experimenter the ability to generate images that are 'metameric' with respect to global image statistics, but differ in local image features.

Part VI

HIGH-SPEED SCHLIEREN ACOUSTICS

17

INTRODUCTION AND LITERATURE REVIEW

And the blind man said to the deaf man, 'Do you see what I hear?'.

—Wayne Gerard Trotman (Desalle and Wynne, 2018)

PART ABSTRACT

The research presented in this Part introduces a novel sound recovery system, that is both passive and indirect, while allowing simultaneous sensing of the sound field. This is achieved through a combination of schlieren imaging—a long-standing optical technique which renders visible the refractive inhomogeneities in transparent media—high-speed photography, and digital signal and image processing. The motivation for this work came from reading Robert Grosseteste's *De generatione sonorum* (discussed at length in Part 3), which explicitly discusses the visual sensing of sound in its opening paragraph (translation by the Ordered Universe research group):

> When an object capable of making sound [*sonativum*] is struck violently, parts of this *sonativum* move away from their natural position. The nature of the *sonativum* inclines these parts back to their natural position, and by the strength of this inclination, the parts that have gone out of the position natural to them, in returning fully to their natural position, move further to a position not natural to them; and the natural inclination,

> turned from this new direction, re-inclines [the parts] to their natural position and thus is created a subtle vibration in the smallest parts of the object. And this vibration is evident to touch and sight.

Grosseteste's description here is of the visible vibrations of a sounding body, perhaps a bell or other musical instrument. The next passage of the text describes how these vibrations of the body set up corresponding vibrations in the surrounding air, through which acoustic signals are transmitted to the ear:

> And when these parts of the sonativum vibrate, they move the air with which they are in contact to move according to their movement, and thus [the vibration] reaches the air of the same nature built up in the ear, and a passion of the body takes place, which is not hidden from the soul, and the sensation of hearing takes place.

Reading this text today prompted discussions of whether it could be possible to sense these vibrations in air optically, that is, whether such vibrations could be made 'evident to sight', and in such a manner that acoustic signals could be faithfully recovered. Chapter 18 describes not only the optical detection and visualisation of sound waves in air, but also the recovery of auditory information, using a *schlieren microphone*.

18

VISUALISATION OF ACOUSTIC WAVES IN AIR AND SUBSEQUENT AUDIO RECOVERY WITH A HIGH-SPEED SCHLIEREN IMAGING SYSTEM: EXPERIMENTAL AND COMPUTATIONAL DEVELOPMENT OF A SCHLIEREN MICROPHONE

Published in *Optics and Lasers in Engineering Science*.

AUTHORS

Joshua S. Harvey, Hannah E. Smithson, Clive R. Siviour.

ABSTRACT

We present a high-speed single-mirror double-pass coincident schlieren system and corresponding algorithms for the visualisation of acoustic waves and recovery of their associated audio signals. Schlieren systems are extensively used to visualise strong shockwaves, such as those from supersonic motion or explosions. Recently, they have also been used to visualise lower amplitude non-linear acoustic phenomena, such as the weak shockwaves arising from impact events including hand claps, belt snaps, and towel cracks. Time-invariant sounds produced by loudspeakers have also been imaged, in one case leading to frequency analysis, although these have been

limited to high-frequency signals at very high sound pressure levels. The research presented here shifts the focus from sound-field visualisation towards audio signal recovery. A comprehensive exploration of several parameters for imaging sound sources, including frequency, wave form, and amplitude, is presented. In addition, we address for the first time the recovery of phase information, which would be essential for speech intelligibility, and the more general case of non-contact sound field reconstruction. Through image and signal processing, it is shown that audio signals can be recovered from high-speed schlieren video whose acoustic waves appear to be below the limit of visibility, and were previously deemed unrecoverable by virtue of their frequency and sound pressure level. This includes sounds at frequencies and loudnesses relevant for human hearing, producing the first 'schlieren microphone'.

Published in *Optics and Lasers in Engineering Science*

18.1 INTRODUCTION

In this paper, a novel application of schlieren imaging is presented: a 'schlieren microphone', whereby the optical technique for visualising refractive inhomogeneities is exploited to indirectly recover audio signals from sound waves. An experimental setup specifically designed for the imaging of acoustic waves relevant to human hearing is combined with software built to recover audio signals from high-speed schlieren footage. This research goes beyond previous attempts to 'visualise' weak shockwaves or acoustic waves: there is an emphasis on signal fidelity in processing, and an evaluation of the limitations of this approach.

Previous attempts to visualise sound with schlieren techniques have mostly been limited to violent or extreme acoustic phenomena such as shock waves, as they give sharp variations of refractive indices in air. These have often relied on significant image processing that would corrupt information contained in the schlieren images. Despite the vast amounts of data present, schlieren images of sound waves have not routinely been analysed with digital signal processing, with the images instead being inspected and remarked upon qualitatively.

In the current research, the focus is shifted away from visualisation and towards signal processing and recovery. Using high-speed photography and an inexpensive schlieren apparatus, it is shown that audio signals can be recovered from schlieren high-speed video, at frequencies and sound pressure levels previously considered too optically weak for schlieren applications. The limitations of this setup are explored through an extensive array of tests, validated by independent microphone measurements, as well as simulated schlieren video.

18.2 BACKGROUND

18.2.1 *Theoretical principles of schlieren imaging*

Schlieren imaging is a classical-optics technique for making visible the refractive inhomogeneities in transparent media. Used extensively since the nineteenth century, the technique has found numerous applications in diverse academic and industrial fields such as optics, fluid dynamics (Freymuth, 1993), aerodynamics, ballistics (Biss et al., 2009; Bixler, 1968), ultrasonic acoustics (Möller et al., 2013), epidemiology (Tang et al., 2011), phonetics (Rowell et al., 2016), and marine biology (Westphal, 1965). The history, optical principles, and much of the modern research surrounding schlieren imaging is explored in Settles' comprehensive book on the topic (Settles, 2001), and Settles and Hargather also provide a review of recent research (Settles and Hargather, 2017).

The basic principle of schlieren imaging is that light is precisely focused to a spot by an optical element such as a lens or mirror, whereupon a razor or some other thin device blocks out a portion of the focal spot, after which the remaining light continues into the lens of a camera (see Figure 18.1). Rays undergoing refraction are deflected while travelling between the source and the focal spot; they consequently are either blocked by the razor instead of forming the image (if deflected in one direction) or avoid the razor instead of being blocked (if deflected in the other direction). The

Published in *Optics and Lasers in Engineering Science*

image therefore contains areas of increased or decreased pixel intensity, as a function of the variations in refractive index, n, of the transparent medium being imaged.

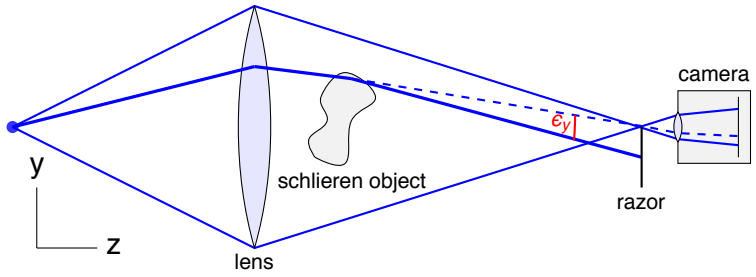


Figure 18.1: The basic principle of schlieren imaging, shown with a simple one-lens, (horizontally-oriented) knife-edge system. The y axis component of angular ray deflection is indicated by ϵ_y.

Using a coordinate system where z is the direction of undisturbed light propagation, and xy are planes orthogonal to this axis, it can be shown that inhomogeneities bend light rays in proportion to their gradients of refractive index, n, in the xy plane. The resulting ray curvatures in the x and y direction are given by:

$$\frac{\partial^2 x}{\partial z^2} = \frac{1}{\mathrm{n}} \frac{\partial \mathrm{n}}{\partial x}, \quad \frac{\partial^2 y}{\partial z^2} = \frac{1}{\mathrm{n}} \frac{\partial \mathrm{n}}{\partial y} \tag{18.1}$$

After integrating once, the x and y plane components of angular ray deflection (ϵ_x and ϵ_y) are found:

$$\epsilon_x = \frac{1}{\mathrm{n}} \int \frac{\partial \mathrm{n}}{\partial x} \partial z, \quad \epsilon_y = \frac{1}{\mathrm{n}} \int \frac{\partial \mathrm{n}}{\partial y} \partial z \tag{18.2}$$

Which for a two-dimensional schlieren system of extent ℓ along the z axis is:

$$\epsilon_x = \frac{\ell}{\mathrm{n}_0} \frac{\partial \mathrm{n}}{\partial x}, \quad \epsilon_y = \frac{\ell}{\mathrm{n}_0} \frac{\partial \mathrm{n}}{\partial y} \tag{18.3}$$

where n_0 is the refractive index of the surrounding medium. Therefore, the luminance level in a schlieren image relates to the first spatial derivative of the refractive index of the transparent medium. A schlieren video consisting of t frames, containing m by n pixels, can therefore be considered as $m \times n$ signals of length t,

Published in Optics and Lasers in Engineering Science

with the intensity of each signal varying proportionally with respect to the relative changes in air pressure along the axis perpendicular to the orientation of the razor.

18.2.2 *Previous attempts at schlieren imaging of acoustic waves*

Schlieren imaging is a contrast technique, and although quantitative approaches have been established whereby a system may be calibrated to absolute values of refractive index, the contrast in a schlieren image is an arbitrary function of the particular optical setup and its sensitivity. As such, a certain amount of image post-processing is reasonable, with linear normalisation of pixel intensity values to a certain range being the most straightforward and justifiable. Unsharp masking has often been used for visualisations of *schlieres* (streaks seen in a schlieren image) from shockwaves (Hargather et al., 2010) (Weinstein et al., 1997). While this can produce dramatic images, the processing involved can simultaneously lose information and introduce artifacts into the data recovered from the image. If one is producing schlieren images not just for the purpose of illustration or visualisation, but for quantitative analysis of refractive index variation, other processing approaches are required to improve the signal strength of the data.

One potential avenue of processing is the use of digital filters, which can be designed to reduce noise in the image without diminishing the signal, applied to either the temporal or spatial frequencies of intensity variation in the video. Chitanont *et al.* have implemented this kind of approach for the schlieren visualisation of sound fields (Chitanont et al., 2015a) (Chitanont et al., 2015b) (Chitanont et al., 2017). The blunt approach of designing a bandpass filter around the expected sound field frequency component can produce attractive images, but is not advisable for two reasons. Firstly, the frequencies present in the sound field must be known beforehand, and secondly, it is difficult to assess to what degree the signal obtained was actually a consequence of the sound field, and how much was simply filtered out of the noise in the image (which will invariably contain some of the sought-after frequencies). A

Published in *Optics and Lasers in Engineering Science*

more sophisticated approach uses filter banks. Without relying on exact knowledge of the sound field's frequency content, this approach has been found to successfully improve the signal-to-noise ratio (SNR) in both simulated (Chitanont et al., 2015b) and real (Chitanont et al., 2017) schlieren images of sound fields. The analysis method employed by Chitanont *et al.* allows for frequency component analysis of the raw and filtered video, but does not permit the phase-alignment of signals required for audio recovery.

18.2.3 *Other imaging modalities for acoustic wave visualisation and audio recovery*

Other imaging modalities have also been used to visualise acoustic waves and sound fields. In the National Physical Laboratory review of acousto-optical sound field visualisation (Piper et al., 2014), schlieren imaging is mostly overlooked in favor of coherent light techniques such as laser doppler anemometry (LDA) (Cullen et al., 1999; Koukoulas et al., 2008; Macgillivray et al., 2002; Thompson and Atchley, 2005), laser doppler velocimetry (LDV) (Gazengel and Poggi, 2005; Malkin et al., 2015), and photon correlation spectroscopy (PCS) (Sharpe and Greated, 1987, 1989). LDA and LDV have proved especially promising, with both techniques having been applied to the calibration of pressure microphones. However, these methods require lasers and other high-quality optical elements such as beam-splitters, so there are incentives for developing alternative modalities with less-stringent practical limitations. Existing methods may also require seeding particles to be introduced into the sound-carrying medium. Additionally, methods such as LDV may require repeated measurements of acoustic signals, prohibiting the mapping of transient sound fields.

While previous schlieren visualisations of sound fields may be useful for some acoustic applications, the various approaches tend not to prioritise signal fidelity and recovery. Once this is considered, it becomes clear that a schlieren video effectively comprises many hundreds or thousands of signal detectors, with every pixel theoretically sampling from the first spatial derivative of the sound field

Published in *Optics and Lasers in Engineering Science*

perpendicular to the razor's alignment. While each pixel's time series has a vanishingly small SNR, owing to the weak changes in refractive index caused by acoustic waves in the normal range of human hearing, if these numerous, related, time series could be combined, the limitation of each pixel's SNR could be overcome, producing a single one-dimensional time series signal from each video. In theory, integration of this signal would then recover the original audio signal in the acoustic wave, such as would be detected by a conventional microphone. This approach bears similarities to the 'visual microphone' developed to recover sound from high-speed video of everyday household objects (Davis et al., 2014).

In this paper, some simple processing and filtering is applied (such as to remove frequencies outside the range of human hearing), but a higher SNR is primarily achieved by the combining (averaging) of numerous data series. The resultant 'schlieren microphone' is then comprehensively evaluated through experimental tests, and the findings are validated through simulation work. Two different avenues of data recombination are presented, the first using a geometric model of the sound field and its acoustic waves derived from a point source assumption, and the second using a data-driven autocorrelation approach that more faithfully captures the actual propagation of acoustic waves.

18.3 EXPERIMENTAL SETUP

A representation of the single-mirror double-pass coincident schlieren apparatus used to collect all the data featured here is shown in Figure 18.2. The system has been optimised to be both affordable and highly sensitive, using a double-pass setup first described in 1933 by Taylor and Waldram (Taylor and Waldram, 1933), for which sensitivity is double that of a single-pass schlieren method, as light is deflected by refractive inhomogeneities for both the incoming illuminator and outgoing analyser beams.

Published in *Optics and Lasers in Engineering Science*

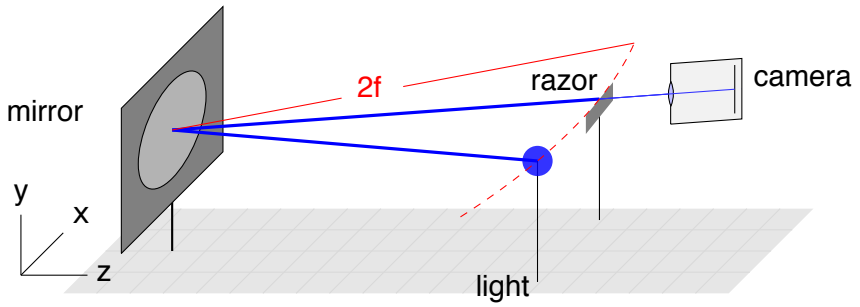


Figure 18.2: The single-mirror double-coincident setup for schlieren imaging. The angle between the illuminator and analyser beams is exaggerated for pictorial clarity.

As shown in Figure 18.2, just four components are required: a spherical mirror, in this case a 150 mm diameter focusing mirror with a focal length of 1.5 m; a light source; a razor; and a camera for collecting the image. While previous acousto-optic research exploring sound field visualisation has mostly required coherent light in the form of lasers (Piper et al., 2014), the schlieren system presented here uses 'white' non-coherent light emitted from a XENON NOVA® 300. The light source, by use of a liquid light guide, is placed at the mirror's radius of curvature, i.e. twice the focal length distance, f, such that $r = 2f$. Diverging light from the source illuminates the mirror, returning off-axis to a focus spot also at the mirror's radius of curvature, whereupon it is partially blocked by a razor. The degree of off-axis divergence between the illuminator and analyser beams is minimised to under 10 mm (and exaggerated in the figure) to reduce distortion in the image. The unblocked light continues past the razor into the lens of a camera, where it is focused to produce an image. For the imaging of acoustic waves, high-speed photography is necessary; here a Photron FASTCAM SA5 was used, fitted with a Nikkor 200 mm lens.

Several parameters were varied during data collection, to explore the potential and limits of acoustic wave visualisation and signal recovery; these are summarised in Table 18.1. For all conditions, sounds were produced by a Scanspeak D2604/833000 Tweeter. Positioned directly in front of the mirror, the speaker output signals produced by either a function generator or an online application capable of producing various waveforms, and was driven by an audio amplifier system to a range of volumes. Hence, the sound pressure level (SPL) was varied between conditions to experimentally

Published in Optics and Lasers in Engineering Science

PARAMETER	VALUES
Waveform	Sine or square wave
SPL / dB	80, 85, 90, 95, 100, 105, 110, 115, 120
Frequency / kHz	0.11, 0.22, 0.44, 1, 2, 5, 7, 10, 15, 20, 25
Framerate / fps	40, 50, 60 $\times$1000
Exposure time / µs	4, 7, 10, 16, 20

Table 18.1: Imaging and acoustic wave parameters varied for data collection.

determine limits of visibility, being read by a Tenma sound-level meter positioned 50 mm away from the diaphragm of the speaker. For the higher frequency sound waves tested this distance is sufficient for measurements to be in the far-field, where SPL varies according to the inverse square law, and hence characterises the whole sound field. While this assumption cannot be upheld for lower frequency sound waves (<10 kHz), the measurement still provides an indication of SPL in the region of air being imaged. All SPLs are given in dB according to the standard convention that 0 dB = auditory threshold at 1 kHz. Waveforms were generated as either sine waves or square waves, for a range of frequencies between 110 Hz and 25 kHz. However, as the acoustic wave outputs are limited by the frequency response of the speaker (which sharply drops off at around 40 kHz), it should be noted that at higher frequencies fewer harmonics of square waves are feasible. Therefore while a 20 kHz square wave may have been generated and sent to the speaker, one would expect a sound field containing only a 20 kHz sine wave. Additionally, it was not possible to image acoustic waves between 4 and 5 kHz, as these appeared to set up corresponding resonant vibrational modes in the mirror itself, distorting the image.

All of the footage collected by varying the parameters of both the speaker output and imaging setup is published open-source, freely available for download (Harvey et al., 2017a).

Published in *Optics and Lasers in Engineering Science*

18.3.1 *Defining the limit of visibility*

The 'limit of visibility' is a term that has been employed without much explanation or definition in the discussion of schlieren images. It is broadly interpreted as the minimum variations in contrast of an image required to perceive the differentials in refractive index present. This is a function of both the refractive index variations in the transparent medium and the sensitivity of the schlieren apparatus, as well as imaging parameters pertaining to the camera.

For acoustic waves in the range of human hearing, a highly sensitive schlieren apparatus is required to give clear contrast variations in the individual images collected. However, when reviewing unprocessed video footage it becomes clear that there is a visible, yet subtle, 'shimmering' present for many acoustic waves of weaker SPLs, for which variations are not apparent in static images. This illustrates a perceptual mechanism of the human visual system, which can extract motion information when seemingly random intensity values have a correlation in their movement. This can be accomplished even when there is a strong background illuminant dominating the image, with the visual system adapting to very low-frequency variations in background intensity, analogous to a high-pass filter. Here, quantitative data will be recovered from these sequences.

18.4 IMAGE PROCESSING AND DATA ANALYSIS

18.4.1 *Visibility of acoustic waves and image denoising*

Although the motion of acoustic waves propagating through the image is perceptible when viewing the raw schlieren images as a video, for individual frames it is not possible to clearly discern the contrast variations. There are several post-processing routes to improving visibility, each with different advantages and disadvantages.

Published in *Optics and Lasers in Engineering Science*

Unsharp masking is the simplest way to improve the visibility of intensity variations. However, due to the sensitivity of the schlieren system, there are many other sources of intensity variations such as thermal currents and the grinding imperfections of the mirror. The results of unsharp masking with the inbuilt MATLAB function `imsharpen` can be shown in Figure 18.3 b.

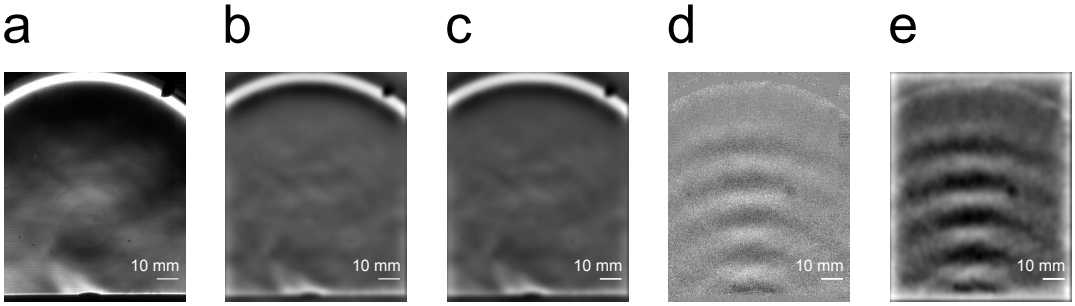


Figure 18.3: Different post-processing for a 20 kHz sine wave, 115 dB. a) Raw footage. b) After unsharp masking. c) After spatial filtering with a 2D Gaussian kernel. d) After a high-pass temporal filter is applied. e) With temporal and spatial filtering. Note that the best visibility is obtained using the latter post-processing method.

Spatial filtering with 2D Gaussian kernels also did little to improve the visibility of the acoustic waves in the image. High-pass and low-pass Gaussian filters were designed based on the expected spatial frequency of the sound waves in the image, as shown in Figure 18.4. Although some improvements can be seen, especially when the footage is viewed as a video, many of the artifacts in the image caused by thermal currents have a similar spatial frequency to the variations in intensity caused by acoustic waves.

In contrast, temporal high-pass filtering (Figure 18.5 d) results in a marked improvement of visibility, as thermal currents move slowly compared to propagating sounds. The resultant image still contains high-frequency noise from the camera, which can be removed by subsequent spatial filtering (Figure 18.3 e). High-pass temporal filters, either in the form of an 'ideal' filter (zeroing the lowest bins of the signal in the frequency domain) or a Butterworth filter, readily give visible results for many signals.

The SPL of acoustic waves is a major limiting factor for their visibility. Frequency is another factor; waves of higher frequency are more visible as refractive index

Published in Optics and Lasers in Engineering Science

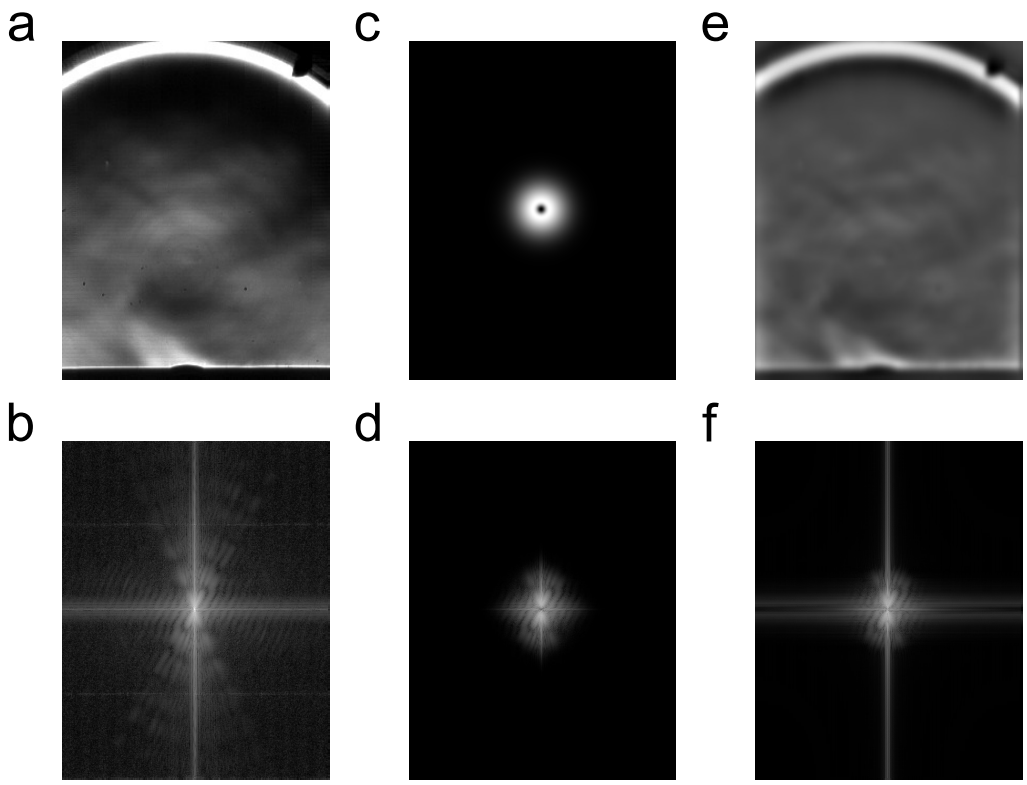


Figure 18.4: Filtering a frame of schlieren video with 2D Gaussian filters; 20 kHz sine wave, 50,000 fps. a) an unprocessed frame and b) its FFT; c) the Gaussian filter response in 2D Fourier space and d) the product of the raw image FFT and the filter in Fourier space; e) the frame after convolution with the filter and f) the FFT of the resultant image.

Published in *Optics and Lasers in Engineering Science*

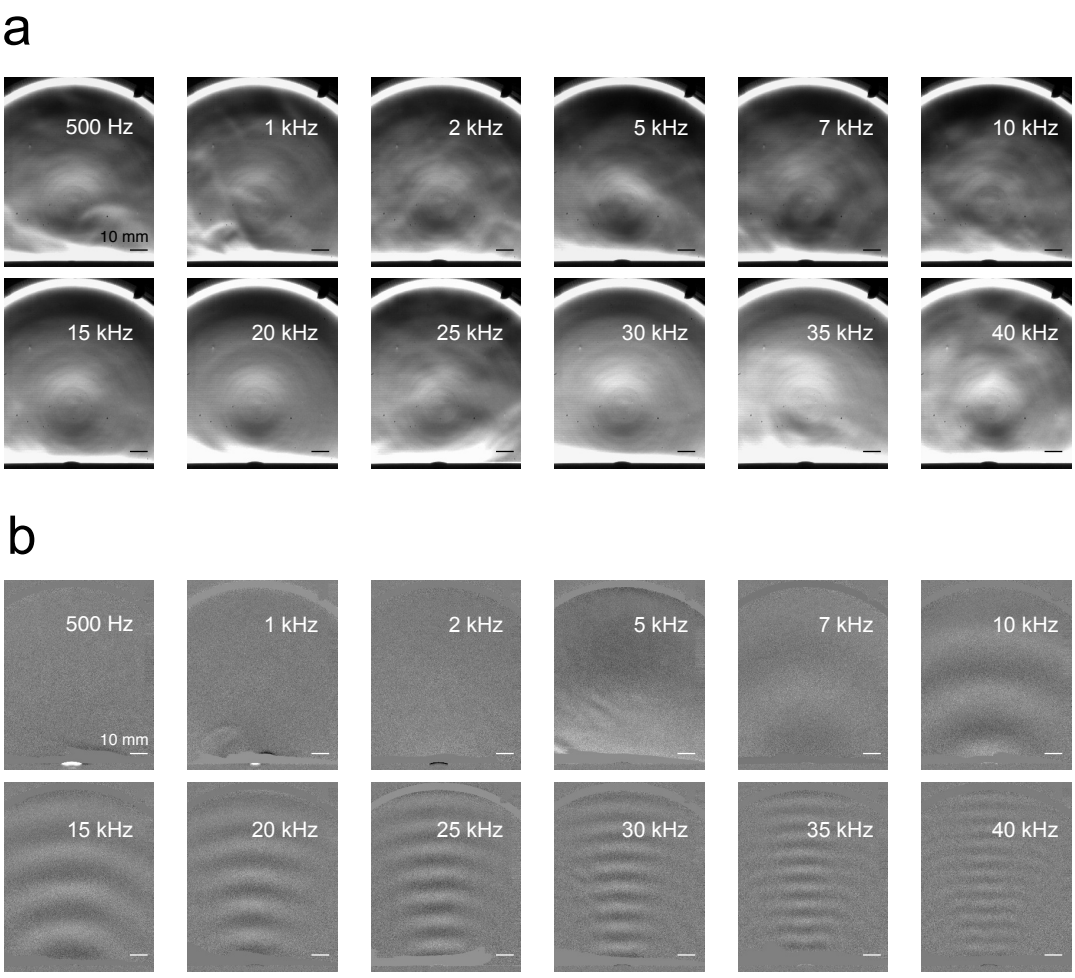


Figure 18.5: a) Raw footage. The bright arc at the top of the images is the boundary of the mirror, the grinding imperfections of which can be spotted as circular details in the centre of the images. b) Footage after a high-pass temporal filter (cut off = 100 Hz) has been applied. Top left to bottom right: 500 Hz (115 dB), 1 kHz (120 dB), 2 kHz (120 dB), 5 kHz (120 dB), 7 kHz (120 dB), 10 kHz (120 dB), 15 kHz (120 dB), 20 kHz (115 dB), 25 kHz (100 dB), 30 kHz (105 dB), 35 kHz (90 dB), 40 kHz (75 dB).

gradients are greater. Even with the spatial and temporal filtering discussed above, only acoustic waves above certain frequencies and registering certain SPLs can be made visible, as shown in Figure 18.6, as higher-pitched and louder sounds have sufficiently strong pressure, and hence refractive index, gradients. A more sensitive system, perhaps with a longer optical axis, or a camera with reduced sensor noise, would extend this range of visibility.

Published in Optics and Lasers in Engineering Science

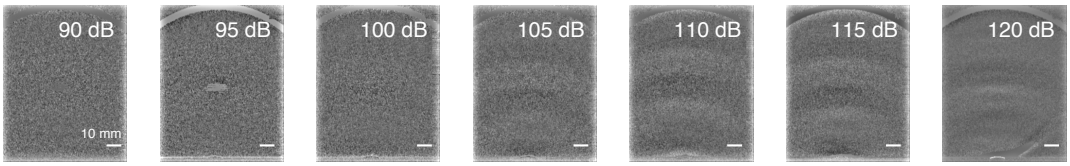


Figure 18.6: Visibility of acoustic waves is dependent on a high enough SPL for a given schlieren sensitivity. Left to right: a 10 kHz square wave measured at SPLs of 90, 95, 100, 105, 110, 115 and 120 dB.

18.4.2 *Data recombination through modelling of the sound field*

The signal of each pixel in the sound field image can be defined by a shift in time (or phase, for periodic signals) and change in amplitude relative to the source signal. While some limited frequency component analysis has previously been performed on high-speed schlieren video of acoustic waves (Chitanont et al., 2017), this did not address the phase realignment of signals from individual pixels required to recover an audio signal. Magnitude spectra (e.g. spectrograms) are often presented without phase data for discussions of sound, and particularly phonetics and the perception of speech. However, it has been found that phase can contribute significantly to speech intelligibility (Paliwal and Alsteris, 2003), especially for weak or noisy signals (Guangji Shi et al., 2006).

If the sound field is modelled as spherically emanating from a point source, the time shift for a pixel's time series corresponds to its radial distance from the sound source. The amplitude can then be modelled by the spherical decay of the acoustic wave in three-dimensional space. For each pixel, I_{x_i,y_i}, at a radial distance r_i from a sound source modelled at x_0, y_0, the time shift τ_i of the signal and its relative amplitude A_i are given by:

$$r_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \tag{18.4}$$

$$\tau_i = \frac{r_i}{c \cdot \psi}, \quad A_i = \frac{1}{r_i^2} \tag{18.5}$$

where ψ is the scale of the schlieren video in pixels per meter and c is the wave speed in meters per second.

Published in Optics and Lasers in Engineering Science

This point-source assumption can therefore be used to produce a phase map for each schlieren image, allowing the time series from pixels to be combined once shifted back in phase. An example phase map is given in Figure 18.10. The data can then be shifted in a pixelwise manner so that their time series are all aligned. Taking the mean for each realigned frame will then give a single time series signal, with a sampling frequency equal to the frame rate of the footage.

However, compared to the spatial resolution of the image, the temporal resolution of the schlieren video with respect to acoustic waves is poor. This is illustrated in Figure 18.7. To retain the high resolution of the data within each frame, the data can be reorganised by radial distance rather than pixel index, ordered by increasing distance from the sound source, and resampled to give a uniform distribution of data. The intensity values in images across frames can then be combined according to the time shift between frames (i.e. the frame rate, λ). The time shift for any intensity value τ is therefore a function of both its radial position r_i and its frame k_n:

$$\tau(r_i, k_n) = c \cdot r_i + \lambda \cdot k_n \tag{18.6}$$

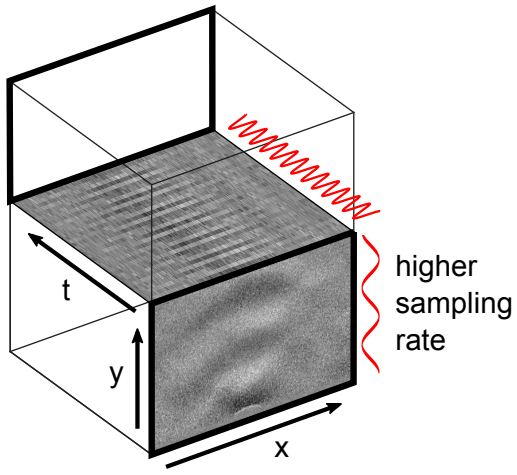


Figure 18.7: A visualisation of footage taken of a 20 kHz sine wave. The first and last frames of a 35 frame sequence are shown, connected by a single row of pixels varying through time. Approximate variations in pixel intensity through the centre of the image are shown in red. It can be seen that the video has a far higher sampling rate spatially than temporally.

Published in Optics and Lasers in Engineering Science

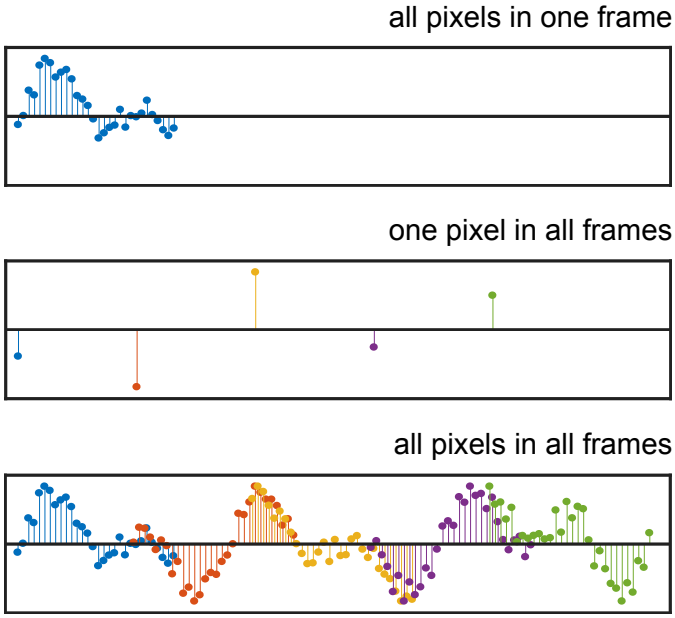


Figure 18.8: An illustration of the effective sampling rates for intra-frame spatial resolution and inter-frame temporal resolution. The intensity values for pixels in a single frame give dense data within a limited time period. Depending on the frame rate, these can overlap with data across other frames, and be combined by calculating the time shift between frames.

With every data point from the footage now identified by a time shift in signal from the sound source, a one dimensional time series is obtained with a very high sampling rate, as illustrated in Figure 18.8. Smoothing, by cubic spline interpolation or a moving average, can then be used to obtain a signal at a desired sampling rate with a high SNR.

18.5 DATA RECOMBINATION THROUGH AUTOCORRELATION OF SIGNALS

While modeling the sound field as circular waves propagating radially from a point source gives good results, it is clear that it does not perfectly account for the actual propagation of acoustic waves. As can be seen in Figure 18.5 b, the point source approximation is increasingly inaccurate as frequency of the acoustic waves increase, because waves spread out less and retain more of the flat geometry of the extended sound source (i.e. the speaker).

Published in Optics and Lasers in Engineering Science

For schlieren videos with sufficiently strong contrast changes throughout the sound field, a truer mapping of wave propagation can be obtained through the cross-correlation of pixel time series. First, the time series data for each pixel, $I_{x,y}$, are upsampled by an integer rate L with an interpolation filter:

$$U_{x,y}[j+nL] = \sum x[n-k] \cdot h[j+kL] \tag{18.7}$$

where h corresponds to the impulse response of the low-pass filter.

This upsampling will allow each time series to be phase shifted at a higher temporal resolution than that of the frame rate. Then, the data for every pixel in the image are cross-correlated to a dataset, $M_{\mu,y}$, obtained by averaging a small number of high-contrast pixels whose upsampled data are already aligned (i.e. neighboring pixels taken from a single row in the center of the wavefront). For each pixel the signal lag, γ, giving rise to the maximal cross-correlation is then converted as a function of upsampling rate and frame rate to give the phase shift (as the signal is periodic):

$$M_{\mu,y} = \sum_{\mu=-n}^{n} \frac{1}{2n+1} U_{\mu,y} \tag{18.8}$$

$$\phi_{x,y} = \arg\max_{\gamma} \Big((U_{x,y} \star M_{\mu,y})(\gamma) \Big) \frac{1}{L \cdot \lambda} \tag{18.9}$$

where $\phi_{x,y}$ is the phase shift for pixel $I_{x,y}$ in the footage. An example phase-shift map is shown in Figure 18.10 of the Results section.

This phase-shift map is converted to a map of each pixel's time shift by adding the maximal lag multiplied by the wave crest index, w, of the pixel, ultimately producing a more accurate time shift map of the video:

$$\tau_{x,y} = \phi_{x,y} + \max(\gamma) \cdot w_{x,y} \tag{18.10}$$

Published in Optics and Lasers in Engineering Science

18.6 SIMULATED SCHLIEREN VIDEO OF ACOUSTIC WAVES

Generating simulated schlieren video from specified signals and noise can validate the analysis of experiment-derived footage. A sound field can be modelled as sinusoidal waves propagating from a point source. The amplitude of pressure, P, throughout the sound field at time t and a radial distance r from the source is given by:

$$\begin{aligned} P(r,t) =& \frac{P_0}{r^2} \exp\left\{i(\kappa r - \omega t)\right\} \cdot \exp\left\{\phi(\omega)\right\} \\ \propto& \frac{1}{r^2} \exp\left\{i(\kappa r - \omega t)\right\} \end{aligned} \tag{18.11}$$

$$\text{where} \quad \omega = 2\pi f$$

For a harmonic series of sine waves, with the h-th harmonic corresponding to an angular frequency ω_h, the pressure function for the total series approximated to σ harmonic components is given simply by the sum of each harmonic:

$$P(r,t) = \sum_{h=1}^{\sigma} P(\omega_h, r, t) \tag{18.12}$$

For a schlieren image, S, the first derivative of pressure in one direction is required, obtained by differentiating with respect to x. For a given waveform:

$$S = \frac{\partial P}{\partial x} = \frac{\partial P}{\partial r} \cdot \frac{\partial r}{\partial x} \tag{18.13}$$

$$\frac{\partial P(\omega_h, r, t)}{\partial r} = H_h \left(\frac{-2}{r^3} + \frac{i\kappa}{r^2} \right) \exp\left\{i(\kappa r - \omega t)\right\} \tag{18.14}$$

$$\frac{\partial P(\omega_h, r, t)}{\partial r} = H_h \left(\frac{\kappa \cos(\kappa r - t\omega_h)}{r^2} - \frac{2\sin(\kappa r - t\omega_h)}{r^3} \right) \tag{18.15}$$

$$\frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}} \tag{18.16}$$

where H_h is the Fourier coefficient for the h-th harmonic component. Terms can be expressed either as complex numbers (eq. 18.14) or the sums of sines and cosines

Published in Optics and Lasers in Engineering Science

(eq. 18.15, showing only real components). Combining eq. 18.15 and eq. 18.16, the component from each harmonic in a sound field is therefore:

$$S(\omega_h) = H_h\left(\frac{\kappa\cos(\kappa r - t\omega_h)}{r^2} - \frac{2\sin(\kappa r - t\omega_h)}{r^3}\right).\left(\frac{x}{\sqrt{x^2+y^2}}\right) \tag{18.17}$$

With the total *schliere* contrast for a wave approximated to σ harmonic components given by:

$$S(r,t) = \sum_{h=1}^{\sigma} S(\omega_h, r, t) \tag{18.18}$$

Once calculated, simulated schlieren images may then be combined with zero-mean Gaussian noise to simulate the heat noise of the camera's image sensor. Results from the simulated schlieren analysis confirmed that the software was performing as desired, and example images can be found in Figure 18.9.

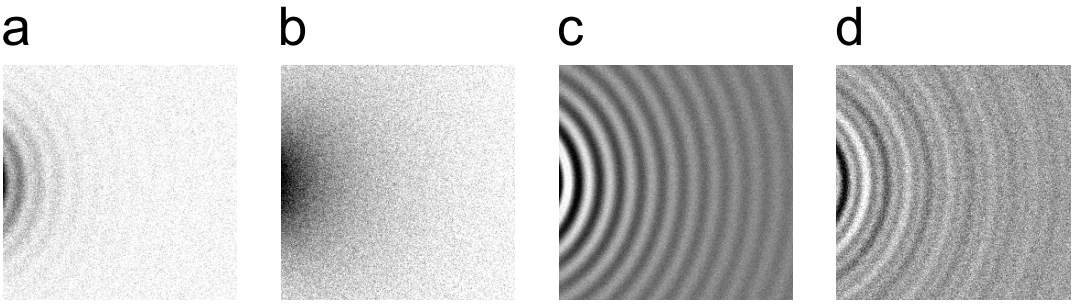


Figure 18.9: Simulated schlieren images, produced with different waveforms and SNRs. a) 440 Hz square wave, 25 Fourier components, SNR = 0. b) 1 kHz sine wave, SNR = 4. c) 10 kHz 'square' wave (only 2 Fourier components), SNR = −2. d) 20 kHz sine wave, SNR = 10.

18.7 RESULTS

Figure 18.10 shows the overall work flow of the schlieren microphone; from high-speed schlieren footage of a sound field to the recovered audio signal. Both the point-source assumption model and the auto-correlation of signals are shown as possible routes to relate and combine each pixel's time series data into a single signal.

Published in Optics and Lasers in Engineering Science

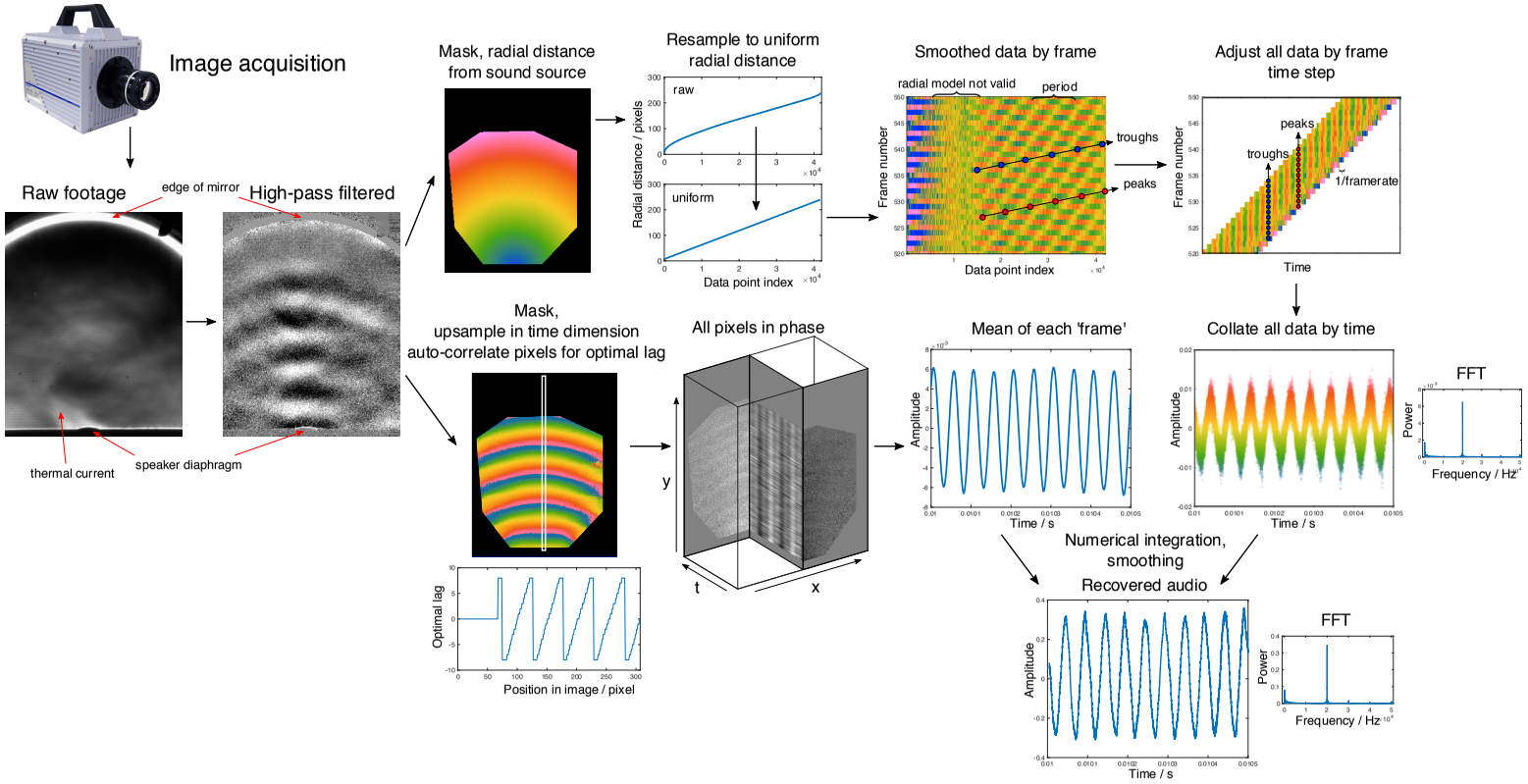


Figure 18.10: The general work flow for extracting audio signals from schlieren footage. Two methods are shown for relating the time series across multiple pixels: the radial propagation model, and the auto-correlation of signals. Note that in the expanded video view for footage where pixels have been aligned by auto-correlation, the temporal dimension now has a higher resolution due to prior upsampling (compare with Figure 18.7).

Three example results are provided in Figure 18.11, showing the recovered audio for both high-frequency and low-frequency waveforms, and both sine and square waveforms. For the data presented here, signal processing and analysis was performed on a Macbook Pro (Early 2015) with 16 GB RAM, running Matlab 2017a. For a single schlieren video file of 1000 frames, the total time required from data import to audio export was approximately ten minutes.

18.7.1 *Visibility of acoustic waves is dependent on frequency, waveform, and amplitude*

In accordance with previous literature and mathematical models, the limit of visibility was found to be dependent on both the amplitude and frequency of the acoustic wave, with greater amplitudes and higher frequencies giving rise to increasingly clear contrast differentials.

As discussed above, there is no strict definition of visibility, and while contrast differentials may be visible when footage is inspected as a video, individual schlieren photographs may not appear to feature any signal. Figure 18.12 shows the approximate limits of visibility as a function of two parameters—amplitude and frequency.

However, more objective evaluations of the recoverability of the audio signal can be found in metrics such as root mean square (RMS) and SNR. Figure 18.13 illustrates how the RMS for each pixel throughout the video can be used to visualise where the sound field is most visible. Notably, lower frequency acoustic waves, down to 110 Hz, may still yield audio signals of high SNR even at lower SPLs, beyond the limits of visibility, as shown in Figure 18.14. By comparing with Figure 18.12, it is clear that the visibility of acoustic waves in raw and/or filtered footage is not a good predictor of the recoverability of an audio signal and its SNR across different frequencies. This makes intuitive sense when the limits of the human visual system are taken into consideration. A schlieren video of a low-frequency wave may have just as strong contrast differentials as a high-frequency wave, but as these variations in contrast are distributed across many frames, they may go unnoticed. This is due to the contrast

Published in *Optics and Lasers in Engineering Science*

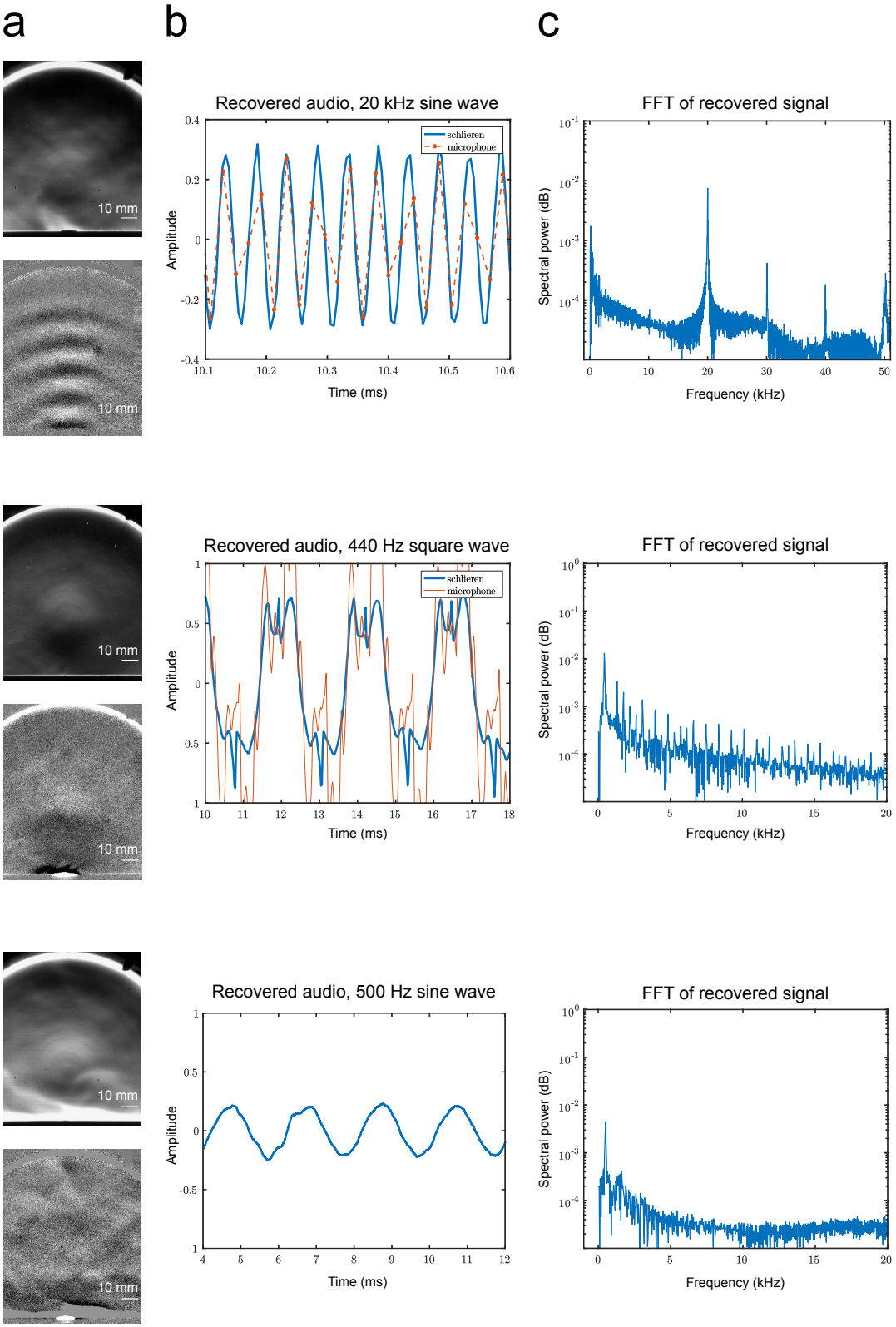


Figure 18.11: Continued on the following page.

Published in Optics and Lasers in Engineering Science

Figure 18.11: (From the previous page.) Examples of different frequencies and waveforms. a) A frame from the raw schlieren footage, with the same frame after filtering below. b) The audio signal recovered by the schlieren microphone (blue) and a conventional microphone (red). d) The spectral make-up of the recovered audio. Top panel: 20 kHz wave recorded at 115 dB. It is clear that the schlieren microphone has a far higher sampling rate than a conventional microphone, which is visibly undersampled. Middle panel: 440 Hz square wave at 120 dB. Many Fourier components, corresponding to the odd harmonics of the square wave, are recovered. Lower panel: 500 Hz sine wave at 115 dB. Audio signals were recovered by use of the point-source assumption model.

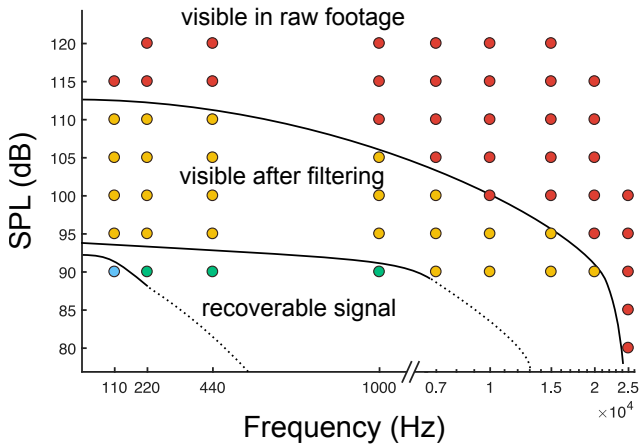


Figure 18.12: Summary of results, showing which samples are visible in raw and processed footage, or only recoverable after spatio-temporal processing. Where example footage is available, markers have been color-coded according to data quality (visible in raw footage = red; visible after filtering = yellow; recoverable signal after processing = green; no recoverable signal = blue). These findings are particular to the schlieren system used for imaging, and the parameter space of visible and recoverable waveforms may be extended with a schlieren system of greater sensitivity or reduced noise.

sensitivity function of the human visual system, which has a bandpass filter for spatial frequency (at low temporal frequency): visibility of image features drops both above and below peak sensitivity at 3 to 4 cycles per degree of visual angle. This is illustrated by a Campbell-Robson contrast sensitivity chart (Campbell and Robson, 1968), shown in Figure 18.15, using the contrast sensitivity function described by Mannos and Sakrison (Mannos and Sakrison, 1974):

$$A(f_r) \approx 2.6[0.0192 + 0.114 f_r] \exp\left\{ -(0.114 f_r)^{1.11} \right\} \tag{18.19}$$

Published in Optics and Lasers in Engineering Science

where $A(f_r)$ is the contrast sensitivity function for an image feature of spatial frequency f_r, defined in cycles per degree of visual field.

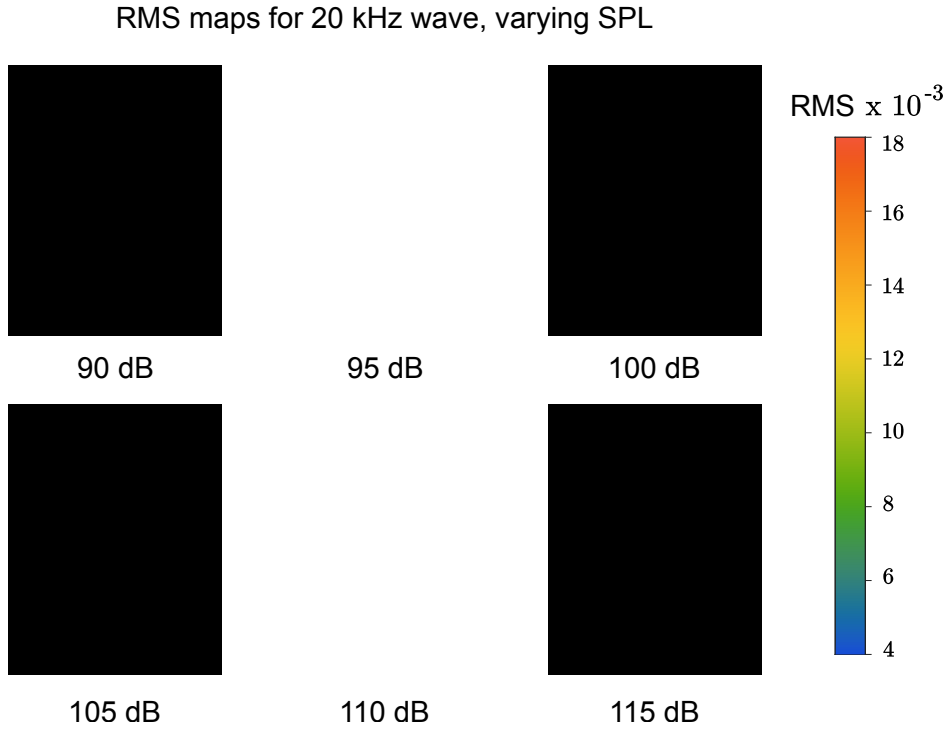


Figure 18.13: Variations of pixel RMS values according to acoustic wave amplitude.

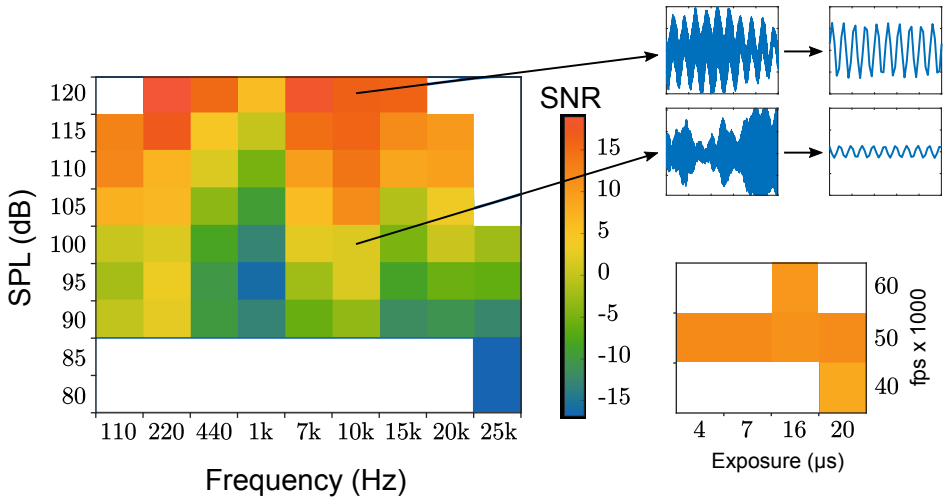


Figure 18.14: Plots of SNR to evaluate the quality of recovered audio signal. The audio signals recovered from footage of a 10 kHz waveform, measured at two SPLs are shown to illustrate the different SNRs. Data are also shown for the effect of varying imaging parameters—both exposure time and frame rate—although this had a minimal effect on the SNR of audio recovered from a 10 kHz square wave measured at an SPL of 105 dB.

Published in Optics and Lasers in Engineering Science

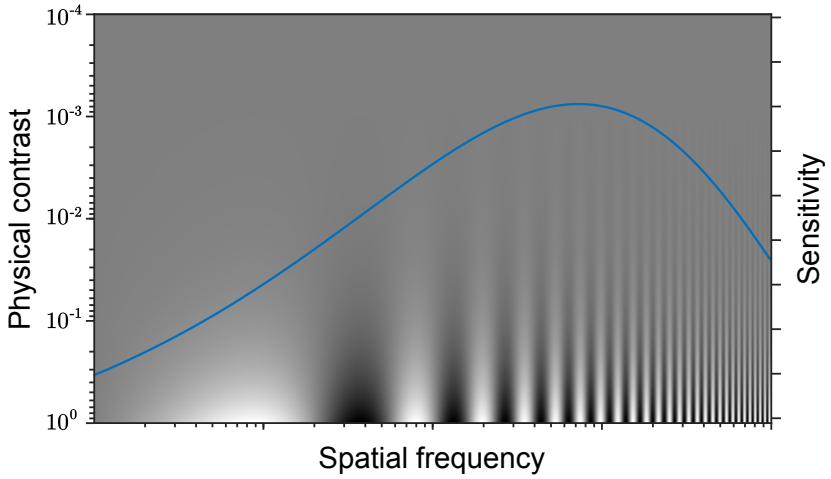


Figure 18.15: The human visual system is limited by a contrast sensitivity function, illustrated here. For different spatial frequencies of image contrast variation, illustrated here by a sine wave of logarithmically-increasing frequency, the sensitivity of the human visual system (indicated by the right axis, and shown in blue) varies for equivalent changes in physical intensity contrasts (indicated by the left axis). The overall effect is an envelope of visibility of the sine wave, with a peak sensitivity at between 3 and 4 cycles per degree of visual field. Sensitivity according to the contrast sensitivity function is overlaid in blue. (This plot will be subject to image size and viewing distance.)

18.7.2 *Impact of imaging parameters on recovered data*

With high-speed imaging, there is a trade-off between temporal and spatial resolution, as higher frame rates are only achievable with reduced image resolution. There is an additional compromise between frame rate and maximal exposure, with higher imaging speeds suffering from reduced image contrast. After preliminary tests, a frame rate of 50,000 fps was settled on for collecting schlieren footage, as this allowed a good spatial resolution and the image was not limited by a lack of light when a maximal exposure time was used (20 µs). It should be noted that there is not a direct relationship between the frame rate of the footage and the sampling rate of the recovered audio, and 50,000 fps does not limit audio recovery to a Nyquist frequency of 25 kHz. Rather, the effective sampling rate will be dependent on the spatial sampling rate of the image, itself a function of the image magnification and wave speed, as well as the smoothing and downsampling carried out in the processing.

Published in *Optics and Lasers in Engineering Science*

To experimentally evaluate the impact of imaging parameters on signal recoverability, exposure time, frame rate, and spatial resolution were varied. Variations in exposure (4-20 µs) and frame rate (40,000-60,000 fps) were found to have little effect on recovering a 10 kHz wave; with SNRs remaining between 8 and 12. Footage was also collected with a different setup that prioritised spatial resolution over temporal resolution (frame rate), with an image of 424 x 512 pixels recorded at 31,000 fps, instead of the standard resolution of 256 x 480 pixels recorded at 50,000 fps. While signals were still recoverable, this resulted in a dramatic worsening of SNR to -10 and below. This suggests that the system could be further optimised by changing the imaging parameters in the other direction, increasing temporal resolution at the cost of spatial resolution.

18.8 DISCUSSION

The origins of sound visualisation go back as far as the phonautograph of Édouard-Léon Scott de Martinville, patented in 1857. While the phonautograph produced visual outputs for the inspection of acoustic information, it was only after Thomas Edison's invention of the phonograph that the focus of acoustic engineering shifted from visualising sound to recovering audio information for playback and analysis. With the work presented here, a similar shift is made for a different imaging modality of acoustic phenomena.

For a broad spectrum of acoustic waves, we show that schlieren imaging performed with an inexpensive apparatus is capable of recovering significant amounts of original audio information. Our approach has several benefits and potential applications. As a general exercise in how to gather quantitative data from schlieren images, we present a detailed consideration of the image manipulation, processing, and analysis that retains and recovers the original information within the image. The schlieren microphone is capable of detecting a sound field while external to it, and therefore does not disrupt the propagation of acoustic waves. This may have particular relevance

Published in *Optics and Lasers in Engineering Science*

to acoustic refraction or other studies of acoustic wave propagation, as well as in assessing the qualities of acoustic output for audio devices such as speakers. The work described here was performed an inexpensive mirror, and without the use of optical benches. With higher quality optical components and a more tunable apparatus, a schlieren microphone of greater sensitivity would have many more potential applications. Even with components of higher specification, a high-speed schlieren imaging system may be more affordable than coherent light-based techniques such as LDA and LDV, as well as presenting fewer safety risks.

As a general caveat to this work, we wish to state clearly that this is not a microphone in the traditional sense, and would be unsuitable for recording the audio signals most often received by conventional transducer microphones. While the SPLs of sound waves recovered in this work are significantly below the expected threshold of schlieren imaging 'visibility', they are still above the typical conversational speaking volume of 60 dB. Rather, the schlieren microphone will have particular applications where a non-contact method of sound field reconstruction is required.

More abstractly, the schlieren microphone provides another visual modality for the consideration of sound waves, as previously established by techniques such as LDV. This may prove especially valuable for pedagogical considerations of acoustics, optics, and digital signal processing. Both sound and vision are phenomena for which we have natural, though different, intuitions. Sound particularly lends itself to considerations of harmonics and Fourier analysis, as mathematical principles such as filters and superposition can be related to aural experiences of music (resonant acoustic spaces and combining individual notes, respectively). Visual experience and imagination is fundamental for many people's understanding of the world, and particularly useful in comparing numerical data. By uniting both perceptual modalities, the schlieren microphone may serve as a useful conceptual tool for illustrating mathematical and physical principles.

Published in *Optics and Lasers in Engineering Science*

18.9 CONCLUSION

We have found that schlieren imaging is a promising imaging modality for the study of acoustic phenomena occupying the domain of human hearing. Previous work has largely been confined to discussions of visibility, which is subject to the limitations of the human visual system, particularly for lower-frequency waveforms. By focusing on a computational and image-processing approach, we have extended the scope of acoustic phenomena for which schlieren imaging may have potential applications.

FUNDING

JSH is funded by the Andrew W. Mellon Foundation and Clarendon Foundation through the Oxford interdisciplinary research centre, the Oxford Research Centre in the Humanities (TORCH). The equipment used was developed with funding from the Oxford University John Fell Fund.

ACKNOWLEDGMENTS

The authors wish to thank the researchers of the interdisciplinary Ordered Universe Project (supported by Arts and Humanities Research Council grant AH/N001222/1) for many stimulating discussions on sound, prompted by studying the writings of thirteenth century scientist and bishop, Robert Grosseteste (c. 1170-1253). The initial inspiration for this work came from studying Grosseteste's pioneering text on sound and phonetics, *De generatione sonorum* [*On the Generation of Sounds*] (*DGS*). We also thank Rebekah White and Sarah Regan for comments on the manuscript.

Published in *Optics and Lasers in Engineering Science*

18.J APPENDIX

It is mentioned in section 18.2.1 that schlieren imaging probes the variations in refractive index along the optical path of the system. The technique is therefore a line-integration of these variations through the imaging field. For spherically-propagating sound waves emitted by a speaker acting as an effective point source this may initially raise concerns, as the refractive index variations will be highly heterogeneous throughout the sound field along the optical path. Indeed, this is a significant issue for laser Doppler methods that probe absolute refractive index (the scalar field n), which confounds the task of audio recovery and sound field reconstruction. A discussion of this may be found in (Malkin et al., 2015).

Here, schlieren imaging has an inherent advantage. By probing the first derivative of refractive index (the vector field $\nabla \mathrm{n}$), and only in the direction perpendicular to the razor in axes orthogonal to the optical path, the sensitivity of a schlieren system to these heterogeneous refractive inhomogeneities is reduced. We may consider a local change in refractive index encountered by the optical path through the schlieren system, $\nabla \mathrm{n}$, as a three-dimensional vector. The schlieren contribution, s, is then the scalar product of this vector and the property of the filter, f:

$$s = f_y \cdot \nabla \mathrm{n} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \delta \mathrm{n}/\delta x \\ \delta \mathrm{n}/\delta y \\ \delta \mathrm{n}/\delta z \end{bmatrix} \qquad \text{for a } y\text{-perpendicular filter.} \tag{18.20}$$

The sensitivity of a schlieren system to $\delta \mathrm{n}$ is therefore scaled according to the decomposition of its vector with respect to the axis perpendicular to the razor, giving $\delta \mathrm{n}(\cos\theta)$ where θ is the angle to the axis. For a double-pass schlieren system such as is used in this research, this scaling occurs twice, giving $\delta \mathrm{n}(\cos\theta)^2$.

This $(\cos\theta)^2$ scaling function as modelled for spherical sound waves emitted from a point source is shown in Figure 18.16, with a cross-section of the yz plane at $x = 0$.

Published in Optics and Lasers in Engineering Science

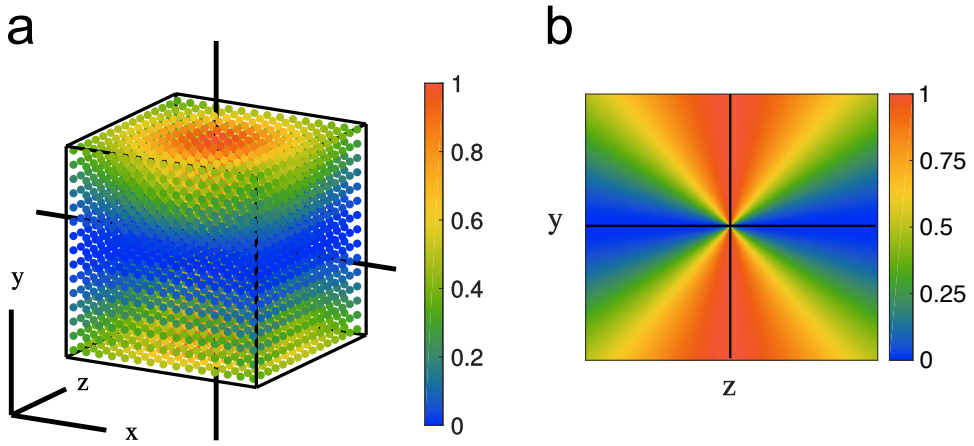


Figure 18.16: a) The $(\cos\theta)^2$ scaling function for gradients of refractive index along the optical path of the schlieren system. b) A cross-section of the $(\cos\theta)^2$ scaling function through the yz plane at $x = 0$.

Published in Optics and Lasers in Engineering Science

Part VII

CONCLUSION

19

MEETING IN THE MIDDLE

You were not there for The Beginning.

You will not be there for The End...

Your knowledge of what is going on can only be superficial and relative...

—William S. Burroughs, Naked Lunch

19.1 SUMMARY OF CONTRIBUTIONS

This thesis has attempted to engage with medieval science in a fresh and radically interdisciplinary way, bringing this engagement to bear on questions of sensory perception. Here, I summarise the key contributions that resulted.

19.1.1 *History of science*

A significant proportion of this research project has been embedded within the international Ordered Universe project, and its numerous symposia and conferences. Central to these, and the methodology underpinning this thesis project, has been the large-scale collaborative readings, which worked through the scientific treatises of Robert Grosseteste. The first volume of *The Scientific Works of Robert Grosseteste* was recently published (Gasper et al., 2019). This volume comprises a new critical edition of two Latin treatises: the *De artibus liberalibus* [*On the Liberal Arts*] (*DAL*) and

the *De generatione sonorum* [*On the Generation of Sounds*] (*DGS*), translations of the text, which were discussed at length by the full interdisciplinary team of historians, philosophers, scientists and engineers, as well as commentary and scientific analyses.

The contributions of this thesis to history of science are most keenly seen in part ii and part iii, which deal with Grosseteste's theories on the rainbow and on speech. In the former, I have demonstrated through physical experiment and optics simulations that his pioneering theory of rainbow formation expressed in the *De iride* [*On the Rainbow*] (*DI*)—the first in medieval Europe to utilise refraction, as opposed to reflection only—was most probably inspired by direct empirical observation. This is an exciting finding, as here we see a medieval scholar departing from tradition and the accepted authority of the day, Aristotle, on the basis of sensory data. Intriguingly, Grosseteste does not allude to this observation, as he does in other discussions on optics. Perhaps, at this early time in the history of science, and long before the 'seeing is believing' cultural norms of the Enlightenment, empirical observation was still subject to scepticism. This finding cements the *DI* as a stepping stone embedded within the evolution of not just theories of the rainbow, but western scientific method in general.

In part iii, the *DGS* is explored from a mathematical and perceptual perspective. Through a close reading of the text, and especially by situating it in the intellectual-historical context of epicyclic astronomy, the classes and combinations of movements described by Grosseteste are reformulated as linear transformations of the two-dimensional plane. The overarching paradigm of speech production and perception that Grosseteste claims relies on these movements as 'mental images', as well as their realisation in the vocal tract and in the writing of letters, is found to be commensurable with recent theories of multisensory integration and crossmodal correspondence. Specifically, Grosseteste's application of Aristotelian mimesis to the relationship between phonemes and graphemes, by which speech sounds may be represented by external figures that share the geometric properties of the vocal apparatus during vocalisation, is contrasted with modern explorations into glyph iconicity and non-arbitrary representation.

19.1.2 *Visual perception of medieval sculpture*

Part iv, presents medieval polychrome sculpture from the perspective of its optical, material, and visual qualities. Through physical spectroradiometric measurements, and physically-based hyperspectral rendering, the properties of an 'imitation gold' material are explored, and contrasted to real gold. The results show that the material features distinct visual properties, suggesting that its use in polychrome sculpture, instead of real gold, may have been an aesthetic, and not simply economic, decision.

19.1.3 *Material perception and metallicity*

Part v directly responds to the research into medieval polychrome sculpture with experiments and modelling concerning the visual percept of metallicity. Previous studies have related metallic appearance to physical or rendering parameters; this work provides an extension by directly relating material appearance to visual features, specifically low-level elements which are represented early in the visual pathway. To go about this, 'synthetic natural images' are used as stimuli, generated through physically-based rendering. A transparent coating atop the metal object, with configurable geometry, permits predictable control over low-level image features, while ensuring each stimulus appears convincingly real. Observer results from a conjoint measurement task, in tandem with steerable pyramid-based analysis and computational model, supports the notion that metallicity is not simply an extreme case of surface smoothness and glossiness, i.e. the presence of sharply-defined highlights. Instead, it is generally correlated with a high global contrast over the surface of objects, which may be estimated as a function of the coarse luminance gradients represented in higher levels of the steerable pyramid.

19.1.4 *Schlieren imaging of acoustics*

Part vi details a high-speed schlieren imaging system that is used to recover sound from acoustic waves in air—a 'schlieren microphone'. This constitutes a new category of microphone, with its own advantages as well as limitations. It functions indirectly, without coming into contact with or disrupting acoustic waves, and samples throughout the full sound field. With current camera technology, it is limited to acoustic waves with a sound pressure level of 90 dB or higher, but with improvements this may be reduced to give a wider gamut of applications.

19.2 FUTURE WORK

The medieval scientific world is expansive and for the most part unexplored—it seems inevitable that there lie in wait many texts, theories, and empirical descriptions that would merit the attention of scholars of the sciences and the humanities alike, within an interdisciplinary framework.

From the work presented in this thesis, there are four specific avenues of research which could be or are being further explored. First, ongoing investigations into the multisensory mechanisms at play in speech and reading, which Grosseteste keenly alludes to in his theory of speech set out in the *DGS*. There is evidence that auditory stimuli can affect shape perception—with different vowel sounds affecting the judged proportions of ovals (Sweeny et al., 2012), but it remains to be seen whether there is an effect of geometric shape perception, or indeed motion perception, on vowel listening, and what the role might be of mental imagery (Berger and Ehrsson, 2013).

Second, the 'imitation gold' material used in medieval polychrome sculpture, remains a subject of both physical and psychophysical study. In collaboration with the Norwegian University of Science and Technology and the University of Oslo, the optical properties of this material—reflectance chromaticity, gloss, and goniochronism, and how they may alter during the life cycle of an object, are being

studied experimentally with imaging techniques and accelerated ageing. Meanwhile, the effect of material (e.g. polychrome 'imitation gold' vs real gold) on looking behaviours is being tested with eye tracking, to validate the DeepGaze II predictions presented in chapter 13.

Third, the area of material perception in visual neuroscience, and particularly the judgement of metallicity, remains to be explored in greater depth. Experiments are planned that use the three-dimensional colour histogram warping described in chapter 16 to generate stimuli, to probe the relationship between global image statistics, local image features, and metallicity judgement.

Finally, the schlieren microphone developed in part vi presents numerous research possibilities. As an instrument to visualise sound fields, it could be used to investigate acoustic phenomena such as reflection and refraction, with respect to material properties. The recovery of human speech from schlieren imaging, and other more complex auditory signals, remains a tempting prospect. Beyond acoustics, the schlieren system could be applied to many other invisible phenomena, and a current collaboration with the Oxford Flight Group into the behaviour of phasmids (stick insects) is already yielding surprising results.

19.3 BEGINNINGS, MIDDLES, AND ENDS

Understandably, science is preoccupied with achievements of today, and how they build on recent findings. But engaging with any historical science, and especially the medieval period, forces us to reckon with our own 'middle place' in history. Science as a system is distributed socially, but also historically, and the act of doing science is as much socially constructed as it is historically developed, and future-determining. How cognisant are we of modern science's inheritance of Aristotle and other academic ancestors, even as we seem at times to wrestle with the very same questions? Too often, 'future directions' and the like are mere afterthoughts, yet surely they are of paramount consequence. Whether medieval scholars such as Grosseteste were

aware of their own 'middle place' we shall never know, but undoubtedly they felt no chagrin engaging with theories that originated many centuries before their own, and indeed saw the benefits this could provide for their own understanding of the natural world. It is my hope that this body of work, with its feet planted in both the medieval context and our own, can inspire that same vision today.

To the reader who has journeyed through this thesis to the end, tracing a meandering path through intellectual history, atmospheric optics, psychophysics, phonetics, computer graphics, and high-speed imaging, I wish to thank you for your generosity and good faith. Hopefully, having made it this far, I have piqued your interest in the mental and material world of the thirteenth century, and the relative merits—and challenges—of radically interdisciplinary research.

BIBLIOGRAPHY

Adams, W. J., Kucukoglu, G., Landy, M. S. and Mantiuk, R. K. (2018). "Naturally glossy: Gloss perception, illumination statistics, and tone mapping". *Journal of Vision* 18.13: 4–4. DOI: 10.1167/18.13.4.

Adelson, E. H., Anderson, C. H., Bergen, J. R., Burt, P. J. and Ogden, J. M. (1984). "Pyramid methods in image processing": 9.

Aguilar, G., Wichmann, F. A. and Maertens, M. (2017). "Comparing sensitivity estimates from MLDS and forced-choice methods in a slant-from-texture experiment". *Journal of Vision* 17.1: 37–37. DOI: 10.1167/17.1.37.

Allott, R. (2000). "The articulatory basis of the alphabet". In: *Becoming Loquens: More Studies in Language Origins*. Ed. by B. H. Bichakjian, T. Chernigovskaya, A. Kendon and A. Moller. 1st ed. Vol. 1. Bochum Publications in Evolutionary Cultural Semiotics. Frankfurt am Main: Peter Lang: 18.

Anderson, B. L. and Kim, J. (2009). "Image statistics do not explain the perception of gloss and lightness". *Journal of Vision* 9.11: 10–10. DOI: 10.1167/9.11.10.

Arai, T., Usuki, N. and Murahara, Y. (2001). "Prototype of a vocal-tract model for vowel production designed for education in speech science." In: *INTERSPEECH*: 2791–2794.

Arend, L. and Reeves, A. (1986). "Simultaneous color constancy". *Journal of the Optical Society of America A* 3.10: 1743. DOI: 10.1364/JOSAA.3.001743.

Aristotle (1984a). *Complete Works of Aristotle, Volume 1: The Revised Oxford Translation: Revised Oxford Translation v. 1*. Ed. by J. Barnes. ed. Barnes, J. Princeton, NJ: Princeton University Press.

Aristotle (1984b). "De Sensu et Sensato". In: *Complete Works of Aristotle: The Revised Oxford Translation*. Ed. by J. Barnes. Two-volume set edition. Vol. 1. Princeton, N.J: Princeton University Press.

Aristotle (1984c). "Meteorologica III [Meteorology: Book III]". In: *Complete Works of Aristotle: The Revised Oxford Translation*. Ed. by J. Barnes. Trans. by E. W. Webster. Two-volume set edition. Vol. 1. Princeton, N.J: Princeton University Press.

Aristotle (1987). *De Anima*. Trans. by H. Lawson-Tancred. Reissue edition. Harmondsworth, Middlesex, England; New York, N.Y., U.S.A: Penguin Classics.

Atal, B. S., Chang, J. J., Mathews, M. V. and Tukey, J. W. (1978). "Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer-sorting technique". *The Journal of the Acoustical Society of America* 63.5: 1535–1555. DOI: 10.1121/1.381848.

Bakalla, M. H. (1982). *Ibn Jinni: An Early Arab Muslim Phonetician : An Interpretive Study of His Life and Contribution to Linguistics*. European Language Publications.

Barney, S. A., Lewis, W. J., Beach, J. A. and Berghof, O. (2006). "Book VIII: The Church and sects (De ecclesia et sectis)". In: *The Etymologies of Isidore of Seville*. New York, USA: Cambridge University Press.

Baur, L. (1912). *Die philosophischen Werke des Robert Grosseteste, Bischofs von Lincoln*. Münster i. W. : Aschendorff.

Bell, A. M. (1867). *Visible speech: the science of universal alphabetics; or Self-interpreting physiological letters, for the writing of all languages in one alphabet*. London, Simpkin, Marshall & Co.

Bell, S., Upchurch, P., Snavely, N. and Bala, K. (2015). "Material recognition in the wild with the Materials in Context Database". In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston, MA, USA: IEEE: 3479–3487. DOI: 10.1109/CVPR.2015.7298970.

Benton, J. R. (2009). *Materials, Methods, and Masterpieces of Medieval Art*. ABC-CLIO.

Berger, C. C. and Ehrsson, H. H. (2013). "Mental Imagery Changes Multisensory Perception". *Current Biology* 23.14: 1367–1372. DOI: 10.1016/j.cub.2013.06.012.

Berzock, K. B. (2019). *Caravans of Gold, Fragments in Time: Art, Culture, and Exchange Across Medieval Saharan Africa*. Princeton University Press.

Bex, P. J. and Makous, W. (2002). "Spatial frequency, phase, and the contrast of natural images". *Journal of the Optical Society of America A* 19.6: 1096. DOI: 10.1364/JOSAA.19.001096.

Beyerer, J., Puente León, F. and Frese, C. (2016). "Image Pyramids, the Wavelet Transfm and Multiresolution Analysis". In: *Machine Vision: Automated Visual Inspection: Theory, Practice and Applications*. Ed. by J. Beyerer, F. Puente León and C. Frese. Berlin, Heidelberg: Springer: 721–750. DOI: 10.1007/978-3-662-47794-6_15.

Biss, M. M., Settles, G. S., Hargather, M. J., Dodson, L. J. and Miller, J. D. (2009). "High-speed digital shadowgraphy of shock waves from explosions and gunshots". In: *Shock Waves*. Springer, Berlin, Heidelberg: 91–96. DOI: 10.1007/978-3-540-85168-4_13.

Bixler, O. C. (1968). "Color Schlieren as a Quantitative Tool for Acoustic Measurements in Ballistics". *The Journal of the Acoustical Society of America* 44.1: 375–375. DOI: 10.1121/1.1970556.

El-Bizri, N. (2016). "Grosseteste's Meteorological Optics: Explications of the Phenomenon of the Rianbow after Ibn al-Haytham". In: *Robert Grosseteste and the pursuit of Religious and Scientific Learning in the Middle Ages*. Ed. by J. P. Cunningham and M. Hocknull. Vol. 18. London: Springer International Publishing: 21–39.

Blake, W. (1790). "Proverbs of Hell". In: *The Marriage of Heaven and Hell*. London, UK.

Blender Online Community (2016). *Blender - a 3D modelling and rendering package*. Blender Institute, Amsterdam: Blender Foundation.

Bobin, O., Schvoerer, M., Ney, C., Rammah, M., Pannequin, B., Platamone, E. C., Daoulatli, A. and Gayraud, R. P. (2003). "The role of copper and silver in the colouration of metallic luster decorations (Tunisia, 9th century; Mesopotamia, 10th century; Sicily, 16th century): A first approach". *Color Research & Application* 28.5: 352–359. DOI: 10.1002/col.10183.

Bonato, F. and Cataliotti, J. (2000). "The effects of figure/ground, perceived area, and target saliency on the luminosity threshold". *Perception & Psychophysics* 62.2: 341–349.

Bonato, F. and Gilchrist, A. L. (1999). "Perceived area and the luminosity threshold". *Perception & Psychophysics* 61.5: 786–797.

Borgia, I., Brunetti, B., Giulivi, A., Sgamellotti, A., Shokouhi, F., Oliaiy, P., Rahighi, J., Lamehi-Rachti, M., Mellini, M. and Viti, C. (2004). "Characterisation of decorations on Iranian (10th–13th century) lustreware". *Applied Physics A* 79.2: 257–261. DOI: 10.1007/s00339-004-2519-z.

Bower, R. G., McLeish, T. C. B., Tanner, B. K., Smithson, H. E., Panti, C., Lewis, N. and Gasper, G. E. M. (2014). "A Medieval Multiverse: Mathematical Modelling of the 13th Century Universe of Robert Grosseteste". *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 470.2167: 20140025–20140025. DOI: 10.1098/rspa.2014.0025.

Boyer, C. B. (1959). *The Rainbow: from Myth to Mathematics*. New York: Thomas Yoseloff.

Boyer, C. B. (1954). "Robert Grosseteste on the Rainbow". *Osiris* 11: 247–258. DOI: 10.1086/368581.

Boyer, C. B. (1987). *The Rainbow: From Myth to Mathematics*. Reissue edition. Princeton, N.J: Princeton University Press.

Brainard, D. H. (1997). "The Psychophysics Toolbox". *Spatial Vision* 10.4: 433–436. DOI: 10.1163/156856897X00357.

Brett, G. S. (2014). *A History of Psychology: Ancient and Patristic*. Routledge.

Brettel, H., Viénot, F. and Mollon, J. D. (1997). "Computerized simulation of color appearance for dichromats". *JOSA A* 14.10: 2647–2655.

Brown, A. M., Lindsey, D. T. and Guckes, K. M. (2011). "Color names, color categories, and color-cued visual search: Sometimes, color perception is not categorical". *Journal of Vision* 11.12: 2–2. DOI: 10.1167/11.12.2.

Buccolieri, G., Buccolieri, A., Bracci, S., Carnevale, F., Falletti, F., Palam, G., Cesareo, R. and Castellano, A. (2009). "Gold leafs in 14th century Florentine painting". *ArcheoSciences. Revue d'archéométrie* 33: 409–415. DOI: 10.4000/archeosciences.2532.

Burnett, C. (1996). "The Introduction of Aristotle's Natural Philosophy into Great Britain : A Preliminary Survey of the Manuscript Evidence". In: *Aristotle in Britain during the Middle Ages*. Vol. 5. Rencontres de Philosophie Médiévale 5. Brepols Publishers: 21–50. DOI: 10.1484/M.RPM-EB.4.000076.

Burnett, C., Lindberg, D. C. and Shank, M. H. (2013). "Translation and Transmission of Greek and Islamic Science to Latin Christendom". In: *The Cambridge History of Science*. Vol. 2. The Cambridge History of Science. Cambridge University Press: 341–364. DOI: 10.1017/CHO9780511974007.016.

Caiger-Smith, A. (1985). *Lustre Pottery: Technique, Tradition and Innovation in Islam and the Western World*. First Edition edition. London; Boston: Faber & Faber.

Campbell, F. W. and Robson, J. G. (1968). "Application of fourier analysis to the visibility of gratings". *The Journal of Physiology* 197.3: 551–566.

Cant, J. S., Arnott, S. R. and Goodale, M. A. (2009). "fMR-adaptation reveals separate processing regions for the perception of form and texture in the human ventral stream". *Experimental Brain Research* 192.3: 391–405. DOI: 10.1007/s00221-008-1573-8.

Catford, J. C. (2001). *A practical introduction to phonetics*. 2nd ed. Oxford textbooks in linguistics. Oxford; New York: Oxford University Press.

Cavina-Pratesi, C., Kentridge, R., Heywood, C. and Milner, A. (2010). "Separate Channels for Processing Form, Texture, and Color: Evidence from fMRI Adaptation and Visual Object Agnosia". *Cerebral Cortex* 20.10: 2319–2332. DOI: 10.1093/cercor/bhp298.

Chadwick, A. C., Cox, G., Smithson, H. E. and Kentridge, R. W. (2018). "Beyond scattering and absorption: Perceptual unmixing of translucent liquids". *Journal of Vision* 18.11: 18–18. DOI: 10.1167/18.11.18.

Chang, H. (2017). "Who cares about the history of science?" *Notes and Records of the Royal Society of London* 71.1: 91–107. DOI: 10.1098/rsnr.2016.0042.

Charrier, C., Knoblauch, K., Maloney, L. T., Bovik, A. C. and Moorthy, A. K. (2012). "Optimizing multiscale SSIM for compression via MLDS". *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society* 21.12: 4682–4694. DOI: 10.1109/TIP.2012.2210723.

Charrier, C., Maloney, L. T., Cherifi, H. and Knoblauch, K. (2007). "Maximum likelihood difference scaling of image quality in compression-degraded images". *Journal of the Optical Society of America A* 24.11: 3418. DOI: 10.1364/JOSAA.24.003418.

Chitanont, N., Yaginuma, K., Yatabe, K. and Oikawa, Y. (2015a). "Visualization of sound field by means of Schlieren method with spatio-temporal filtering". In: *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*: 509–513. DOI: 10.1109/ICASSP.2015.7178021.

Chitanont, N., Yatabe, K., Ishikawa, K. and Oikawa, Y. (2017). "Spatio-temporal filter bank for visualizing audible sound field by Schlieren method". *Applied Acoustics* 115: 109–120. DOI: 10.1016/j.apacoust.2016.08.028.

Chitanont, N., Yatabe, K. and Oikawa, Y. (2015b). "Audible sound field visualization by using Schlieren technique". In: *12th West. Pac. Acoust. Conf.(WESPAC)*: 5–9.

Christie, J. S. (1969). "An instrument for the geometric attributes of metallic appearance". *Applied Optics* 8.9: 1777–1785. DOI: 10.1364/AO.8.001777.

CIE (2014). *Chromaticity Difference Specification for Light Sources*. Technical report CIE TN 001-2014. Commission Internatinoale de l'Eclairage.

Clegg, B. (2004). *The First Scientist: A Life of Roger Bacon*. 1st Carroll & Graf Ed edition. Carroll & Graf Publishers Inc.

Cohen, R. W., Gorog, I. and Carlson, C. R. (1975). *Image Descriptors for Displays*. Tech. rep. PRRL-75-CR-2. RCA LABS PRINCETON NJ.

Cooke, P. (2014). "Transversal or linear? Knowledge externalities and the complexity of knowledge interactions". In: *Routledge Handbook of the Economics of Knowledge*. Ed. by C. Antonelli and A. N. Link. London: Routledge.

Corning Museum of Glass (2002). *http://www.CMoG.org/*. New York: Accessed: 2017-02.

Crombie, A. C. (1953). *Robert Grosseteste and the Origins of Experimental Science 1100-1700*. Oxford: Oxford University Press.

Crombie, A. C. (2002). *Robert Grosseteste and the Origins of Experimental Science 1100-1700*. 1st Edition edition. Oxford: British Academy.

Cullen, J. S., Greated, C. A. and Campbell, D. M. (1999). "LDA measurement of sound: amplitude modulation of laser Doppler signals". *Measurement Science and Technology* 10.9: 812.

Daimon, M. and Masumura, A. (2007). "Measurement of the refractive index of distilled water from the near-infrared region to the ultraviolet region". *Applied optics* 46.18: 3811–3820.

D'Ausilio, A., Bufalari, I., Salmas, P. and Fadiga, L. (2012). "The role of the motor system in discriminating normal and degraded speech sounds". *Cortex* 48.7: 882–887. DOI: 10.1016/j.cortex.2011.05.017.

Davis, A., Rubinstein, M., Wadhwa, N., Mysore, G., Durand, F. and Freeman, W. T. (2014). "The Visual Microphone: Passive Recovery of Sound from Video". *ACM Transactions on Graphics (Proc. SIGGRAPH)* 33.4: 79:1–79:10.

Davy, C. (1772). *Conjectural observations on the origin and progress of alphabetic writing*. London: T. Wright.

De Leemans, P. (2010). "Aristotle Transmitted: Reflections on the Transmission of Aristotelian Scientific Thought in the Middle Ages". *International Journal of the Classical Tradition* 17.3: 325–353. DOI: 10.1007/s12138-010-0200-9.

Desalle, R. and Wynne, P. J. (2018). *Our Senses: An Immersive Experience*. New Haven: Yale University Press.

Dinkova-Bruun, G., Gasper, G. E. M., Huxtable, M., McLeish, T. C. B., Panti, C. and Smithson, H. (2013). *The Dimensions of Colour: Robert Grosseteste's De colore*. Toronto: PIMS.

Drösler, J. (2013). "The invariances of Weber's and other laws as determinants of psychophysical structures". In: *Geometric Representations of Perceptual Phenomena: Papers in Honor of Tarow indow on His 70th Birthday*. Ed. by R. D. Luce, M. D'Zmura, D. D. Hoffman, G. J. Iverson and A. K. Romney. Google-Books-ID: rC6aGllNr3sC. Psychology Press.

Dudley, H. and Tarnoczy, T. H. (1950). "The Speaking Machine of Wolfgang von Kempelen". *The Journal of the Acoustical Society of America* 22.2: 151–166. DOI: 10.1121/1.1906583.

Eastwood, B. S. (1966). "Robert Grosseteste's theory of the rainbow: A chapter in the history of non-experimental science". *Archives Internationales d'Histoire des Sciences* 19: 313–332.

Eastwood, B. S. (1967). "Grosseteste's "Quantitative" Law of Refraction: A Chapter in the History of Non-Experimental Science". *Journal of the History of Ideas* 28.3: 403–414.

Ennis, R., Schiller, F., Toscani, M. and Gegenfurtner, K. R. (2018). "Hyperspectral database of fruits and vegetables". *Journal of the Optical Society of America. A, Optics, Image Science, and Vision* 35.4: B256–B266. DOI: 10.1364/JOSAA.35.00B256.

Farrant, T. and Harvey, J. S. (2017). "Understanding and feeling: theological and psychological reflections through the lens of Aspectus and Affectus". *Bulletin de Philosophie Médiévale* 59: 262–272.

Ferguson, A., Myers, C. S., Bartlett, R. J., Banister, H., Bartlett, F. C., Brown, W., Campbell, N. R., Drever, J., Guild, J., Houstoun, R. A., Irwin, J. C., Kaye, G. W. C., Philpott, S. J. F., Richardson, L. F., Shaxby, J. H., Smith, T., Thouless, R. H. and Tucker, W. S. (1938). "Quantitative Estimates of Sensory Events: Interim Report of the Committee appointed to consider and report upon the possibility of Quantitative Estimates of Sensory Events". In: *Report of the Annual Meeting, 1938*. Vol. 108. Cambridge, UK: Burlington house, London: 277–334.

Ferguson, A., Myers, C. S., Bartlett, R. J., Banister, H., Bartlett, F. C., Brown, W., Campbell, N. R., Craik, K., Drever, J., Guild, J. et al. (1940). "Final report of the committee appointed to consider and report upon the possibility of quantitative estimates of sensory events". *Report of the British Association for the Advancement of Science* 2: 331–349.

Field, D. J. (1987). "Relations between the statistics of natural images and the response properties of cortical cells". *JOSA A* 4.12: 2379–2394. DOI: 10.1364/JOSAA.4.002379.

Fisher, L. (2015). "Perceptual thresholds: Music inspired Newton's rainbow". *Nature* 520: 436. DOI: 10.1038/520436a.

Fleming, R. W., Holtmann-Rice, D. and Bülthoff, H. H. (2011). "Estimation of 3D shape from image orientations". *Proceedings of the National Academy of Sciences of the United States of America* 108.51: 20438–20443. DOI: 10.1073/pnas.1114619109.

Fleming, R. W., Wiebel, C. and Gegenfurtner, K. (2013). "Perceptual qualities and material classes". *Journal of Vision* 13.8: 9–9. DOI: 10.1167/13.8.9.

Foster, D. H. (2011). "Color constancy". *Vision Research* 51.7: 674–700. DOI: 10.1016/j.visres.2010.09.006.

Foucault, M. (1971). *The order of things: an archaeology of the human sciences*. OCLC: 112096. New York: Pantheon Books.

Franz, M. O. and Schölkopf, B. (2005). "Implicit Wiener Series for Higher-Order Image Analysis". In: *Advances in Neural Information Processing Systems 17*. Ed. by L. K. Saul, Y. Weiss and L. Bottou. MIT Press: 465–472.

Freeman, W. T. and Adelson, E. H. (1991). "The Design and Use of Steerable Filters". *IEEE Trans. Pattern Anal. Mach. Intell.* 13.9: 891–906. DOI: 10.1109/34.93808.

Freymuth, P. (1993). "Flow visualization in fluid mechanics". *Review of Scientific Instruments* 64.1: 1–18. DOI: 10.1063/1.1144433.

Galantucci, B., Fowler, C. A. and Turvey, M. T. (2006). "The motor theory of speech perception reviewed". *Psychonomic bulletin & review* 13.3: 361–377.

Gasper, G. E. M., McLeish, T. C. B., Panti, C. and Smithson, H. E., eds. (forthcoming). *Colour and the Refraction of Rays: Robert Grosseteste's On Colour and On the Rainbow*. Oxford: Oxford University Press. Forthcoming.

Gasper, G. E. M., Panti, C., McLeish, T. C. B. and Smithson, H. E., eds. (in press 2019). *Knowing and Speaking: Robert Grosseteste's De artibus liberalibus 'On the Liberal Arts' and De generatione sonorum 'On the Generation of Sounds'*. Vol. 1 of 6. The Scientific Works of Robert Grosseteste. Oxford: Oxford University Press.

Gasper, G. E. M., Panti, C., McLeish, T. and Smithson, H. E. (2019). *The Scientific Works of Robert Grosseteste, Volume 1: Knowing and Speaking: Robert Grosseteste's De artibus liberalibus 'On the Liberal Arts' and De generatione sonorum 'On the Generation of Sounds'*. Oxford, New York: Oxford University Press.

Gatys, L. A., Ecker, A. S. and Bethge, M. (2016). "Image Style Transfer Using Convolutional Neural Networks". In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA: IEEE: 2414–2423. DOI: 10.1109/CVPR.2016.265.

Gazengel, B. and Poggi, S. (2005). "Measurement of acoustic particle velocities in enclosed sound field: Assessment of two Laser Doppler Velocimetry measuring systems". *Applied Acoustics* 66.1: 15–44. DOI: 10.1016/j.apacoust.2004.07.001.

Gerardin, P., Devinck, F., Dojat, M. and Knoblauch, K. (2014). "Contributions of contour frequency, amplitude, and luminance to the watercolor effect estimated by conjoint measurement". *Journal of Vision* 14.4: 9–9. DOI: 10.1167/14.4.9.

Giesel, M. and Zaidi, Q. (2013). "Frequency-based heuristics for material perception". *Journal of Vision* 13.14. DOI: 10.1167/13.14.7.

Green, D. M. and Swets, J. A. (1966). *Signal detection theory and psychophysics*. Signal detection theory and psychophysics. Oxford, England: John Wiley.

Greenler (1980). *Rainbows, Halos and Glories*. 1 edition. Cambridge; New York: Cambridge University Press.

Grosholz, E. R. (1991). *Cartesian Method and the Problem of Reduction*. Oxford University Press.

Guangji Shi, Shanechi, M. and Aarabi, P. (2006). "On the importance of phase in human speech recognition". *IEEE Transactions on Audio, Speech and Language Processing* 14.5: 1867–1874. DOI: 10.1109/TSA.2005.858512.

Hansmann-Roth, S. and Mamassian, P. (2017). "A Glossy Simultaneous Contrast: Conjoint Measurements of Gloss and Lightness". *i-Perception* 8.1. DOI: 10.1177/2041669516687770.

Harden, D. B. (1978). "Anglo-Saxon and Later Medieval Glass in Britain: Some Recent Developments". *Medieval Archaeology* 22.1: 1–24. DOI: 10.1080/00766097.1978.11735405.

Hargather, M. J., Settles, G. S. and Madalis, M. J. (2010). "Schlieren imaging of loud sounds and weak shock waves in air near the limit of visibility". *Shock Waves* 20.1: 9–17. DOI: 10.1007/s00193-009-0226-6.

Harvey, J. S., Smithson, H. E., Siviour, C. R., Gasper, G. E. M., Sønnesyn, S. O., McLeish, T. C. B. and Howard, D. M. (2019a). "A thirteenth-century theory of speech". *The Journal of the Acoustical Society of America* 146.2: 937–947. DOI: 10.1121/1.5119126.

Harvey, J. S. and Farrant, T. (2017). "Aspectus and Affectus: Robert Grosseteste, Understanding and Feeling. Georgetown University,Washington, D.C. (USA), 30 March-1 April 2017". *Revista Española de Filosofía Medieval* 24: 279–281. DOI: 10.21071/refime.v24i.10463.

Harvey, J. S., Kollandsrud, K. and Smithson, H. E. (2018a). "Not All That Glitters: Probing Material Perception With a Physical Imitation Gold Stimulus". In: *Perception*. Vol. 47. SAGE Journals: 572. DOI: 10.1177/0301006618756416.

Harvey, J. S., Morimoto, T. and Spitschan, M. (2019b). "The Neon Fruit Illusion: A Fresh Recipe for Colour Science Demonstrations". *Perception* 48.3: 242–247. DOI: 10.1177/0301006618824484.

Harvey, J. S., Siviour, C. R. and Smithson, H. E. (2018b). "From Spectra to Perceptual Color: Visualization Tools for the Dimensional Reduction Achieved by the Human Color Sense". In: *EG 2018 - Posters*. Ed. by E. Jain and J. Kosinka. The Eurographics Association. DOI: 10.2312/egp.20181025.

Harvey, J. S., Smithson, H. E. and Siviour, C. R. (2017a). "Data for: Visualization of acoustic waves in air and subsequent audio recovery with a high-speed schlieren imaging system (Harvey et al.)" 1. DOI: 10.17632/mjrx5467mv.1.

Harvey, J. S., Smithson, H. E. and Siviour, C. R. (2018c). "Raw high-speed schlieren footage of acoustic waves in air for subsequent computational analysis and audio recovery". *Data in Brief* 19: 1647–1649. DOI: 10.1016/j.dib.2018.06.069.

Harvey, J. S., Smithson, H. E. and Siviour, C. R. (2018d). "Visualization of acoustic waves in air and subsequent audio recovery with a high-speed schlieren imaging system: Experimental and computational development of a schlieren microphone". *Optics and Lasers in Engineering* 107: 182–193. DOI: 10.1016/j.optlaseng.2018.03.015.

Harvey, J. S., Smithson, H. E., Siviour, C. R., Gasper, G. E. M., Sønnesyn, S. O., Tanner, B. K. and McLeish, T. C. B. (2017b). "Bow-shaped caustics from conical prisms: a 13th-century account of rainbow formation from Robert Grosseteste's De iride". *Applied Optics* 56.19: G197–G204. DOI: 10.1364/AO.56.00G197.

Harvey, J. S., White, R. C., Smithson, H. E., McLeish, T. C. B., Howard, D. and Coleman, J. (2019c). "Instrumental motions: Shaping and perceiving speech sounds". In: *Knowing and Speaking: Robert Grosseteste's De artibus liberalibus 'On the Liberal Arts' and De generatione sonorum 'On the Generation of Sounds'. Edition, Translation, and Interdisciplinary Analysis*. Vol. 1. The Scientific Works of Robert Grosseteste. Oxford: Oxford University Press: 336–366.

Harvey, L. O. and Gervais, M. J. (1978). "Visual texture perception and Fourier analysis". *Perception & Psychophysics* 24.6: 534–542. DOI: 10.3758/BF03198780.

Hassan, Z. M. and Heselwood, B., eds. (2011). *Instrumental Studies in Arabic Phonetics*. John Benjamins.

Heasly, B. S., Cottaris, N. P., Lichtman, D. P., Xiao, B. and Brainard, D. H. (2014). "RenderToolbox3: MATLAB tools that facilitate physically based stimulus rendering for vision research". *Journal of Vision* 14.2: 6–6. DOI: 10.1167/14.2.6.

Helmont, F. M. van (1667). *Alphabeti veri naturalis hebraici brevissima delineatio*. Sulzbach: Abraham Lichtenthaler.

Hering, E. (1920). *Outlines of a Theory of the Light Sense*. Harvard University Press.

Heselwood, B. (2013). *Phonetic Transcription in Theory and Practice*. Edinburgh University Press.

Hickok, G. and Poeppel, D. (2007). "The cortical organization of speech processing". *Nature Reviews Neuroscience* 8.5: 393–402. DOI: 10.1038/nrn2113.

Hiramatsu, C., Goda, N. and Komatsu, H. (2011). "Transformation from image-based to perceptual representation of materials along the human ventral visual pathway". *NeuroImage* 57.2: 482–494. DOI: 10.1016/j.neuroimage.2011.04.056.

Ho, Y.-X., Landy, M. S. and Maloney, L. T. (2008). "Conjoint Measurement of Gloss and Surface Texture". *Psychological Science* 19.2: 196–204. DOI: 10.1111/j.1467-9280.2008.02067.x.

Holakooei, P. (2016). "A Mediaeval Persian Treatise on Coloured and Enamelled Glass: Bayan Al-Sana'at". *Iran* 54.2: 95–106. DOI: 10.1080/05786967.2016.11879215.

Hope, E. R. (1957). "Letter Shapes in Korean Önmun and Mongol hPhagspa Alphabets". *Oriens* 10.1: 150–159. DOI: 10.2307/1578766.

Howard, T. (2012). *Why Isn't the Sky Blue?* Radiolab (podcast). WNYC Studios.

Huxley, A. (1961). "Visionary experience". In: *14th International Congress on Applied Psychology*. Copenhagen.

Jakob, W. (2010). *Mitsuba renderer*.

John Wilkins (1668). *An Essay towards a Real Character and a Philosophical Language*. London: Samuel Gellibrand.

Junge, M. and Reisenzein, R. (2013). "Indirect scaling methods for testing quantitative emotion theories". *Cognition and Emotion* 27.7: 1247–1275. DOI: 10.1080/02699931.2013.782267.

Kastan, D. and Farthing, S. (2018). *On Color*. New Haven, CT: Yale University Press.

Katz, D. (2014). *The World Of Colour*. 6th edition. Abingdon; New York: Routledge: 27–28.

Kelemen, C., Szirmay-Kalos, L., Antal, G. and Csonka, F. (2002). "A simple and robust mutation strategy for the metropolis light transport algorithm". *Computer Graphics Forum (Eurographics)* 21.3: 531–540.

Kendall, C. B. (2016). *Isidore of Seville, On the Nature of Things*. Translation edition. Liverpool: Liverpool University Press.

Kendall, C. and Wallis, F. (2010). *Bede: On the Nature of Things and on Times*. Liverpool University Press - Translated Texts for Historians. Liverpool University Press.

Keynes, J. M. (1947). "Newton, The Man". In: *Proceedings of the Royal Society Newton Tercentenary Celebrations*. Cambridge University Press.

Kider, J. T., Fletcher, R. L., Yu, N., Holod, R., Chalmers, A. and Badler, N. I. (2009). "Recreating Early Islamic Glass Lamp Lighting". In: *Proceedings of the 10th International Conference on Virtual Reality, Archaeology and Cultural Heritage*. VAST'09. St. Julians, Malta: Eurographics Association: 33–40. DOI: 10.2312/VAST/VAST09/033-040.

Kim, J., Tan, K. and Chowdhury, N. S. (2016). "Image Statistics and the Fine Lines of Material Perception". *i-Perception* 7.4: 1–11. DOI: 10.1177/2041669516658047.

Kim-Renaud, Y.-K. (1997). *The Korean Alphabet: Its History and Structure*. University of Hawaii Press.

Kirchner, E. (2013). "Color theory and color order in medieval Islam: A review". *Color Research & Application* 40.1: 5–16. DOI: 10.1002/col.21861.

Kirchner, E. and Bagheri, M. (2013). "Color Theory in Medieval Islamic Lapidaries: Nıshābūrı, Tūsı and Kāshānı". *Centaurus* 55.1: 1–19. DOI: 10.1111/1600-0498.12000.

Kirchner, E., Kieboom, G.-J. van den, Njo, L., Supèr, R. and Gottenbos, R. (2007). "Observation of visual texture of metallic and pearlescent materials". *Color Research & Application* 32.4: 256–266. DOI: 10.1002/col.20328.

Kirp, D. L. (2004). *Shakespeare, Einstein, and the Bottom Line: The Marketing of Higher Education*. New Ed edition. Cambridge, Mass.: Harvard University Press.

Koenderink, J. J. (2010). *Color for the Sciences*. Cambridge, Mass: MIT Press.

Koenderink, J. J. (2013). "Color spaces". In: *Color Imaging XVIII: Displaying, Processing, Hardcopy, and Applications*. Vol. 8652. California, USA: International Society for Optics and Photonics. DOI: 10.1117/12.2020725.

Koenderink, J., Valsecchi, M., Doorn, A. v., Wagemans, J. and Gegenfurtner, K. (2017). "Eidolons: Novel stimuli for vision research". *Journal of Vision* 17.2: 7–7. DOI: 10.1167/17.2.7.

Kollandsrud, K. (2018). "Evoking the Divine: The Visual Vocabulary of Sacred Polychrome Wooden Sculpture in Norway between 1100 and 1350". PhD thesis. University of Oslo.

Komatsu, H., Nishio, A., Okazawa, G. and Goda, N. (2013). "'Yellow' or 'Gold'?: Neural Processing of Gloss Information". In: *Computational Color Imaging*. Ed. by S. Tominaga, R. Schettini and A. Trémeau. Lecture Notes in Computer Science. Springer Berlin Heidelberg: 1–12.

Koukoulas, T., Theobald, P., Schlicke, T. and Barham, R. (2008). "Towards a future primary method for microphone calibration: Optical measurement of acoustic velocity in low seeding conditions". *Optics and Lasers in Engineering* 46.11: 791–796. DOI: 10.1016/j.optlaseng.2008.06.005.

Krantz, D. H. and Tversky, A. (1971). "Conjoint-measurement analysis of composition rules in psychology". *Psychological Review* 78.2: 151–169. DOI: 10.1037/h0030637.

Kruskal, J. B. (1964). "Nonmetric multidimensional scaling: A numerical method". *Psychometrika* 29.2: 115–129. DOI: 10.1007/BF02289694.

Kuehn, D. (2013). "Keynes, Newton and the Royal Society: the events of 1942 and 1943". *Notes and Records of the Royal Society of London* 67.1: 25–36. DOI: 10.1098/rsnr.2012.0053.

Kuehni, R. G. and Schwarz, A. (2008). *Color Ordered: A Survey of Color Order Systems from Antiquity to the Present*. Oxford, New York: Oxford University Press.

Kuhl, P. K., Kiritani, S., Deguchi, T., Hayashi, A., Stevens, E. B., Dugger, C. D. and Iverson, P. (1997). "Effects of language experience on speech perception: American and Japanese infants' perception of /ra/ and /la/". *The Journal of the Acoustical Society of America* 102.5: 3135–3136. DOI: 10.1121/1.420646.

Kuhn, T. S. (1970). *The Structure of Scientific Revolutions*. 2nd ed. International encyclopedia of unified science. Foundations of the unity of science, v. 2, no. 2. Chicago: University of Chicago Press.

Kümmerer, M., Wallis, T. S. A. and Bethge, M. (2016). "DeepGaze II: Reading fixations from deep features trained on object recognition". *arXiv:1610.01563 [cs, q-bio, stat]*.

Kümmerer, M., Wallis, T. S. A. and Bethge, M. (2017). "Saliency Benchmarking: Separating Models, Maps and Metrics". *arXiv:1704.08615 [cs, stat]*.

Laric, O. (2016). *http://threedscans.com*.

Lautié, A., Sagon, G. and Froment, F. (2005). "Case Study: Fifteen Years of Artefact Investigations by Raman Microscopy at the CNRS Laboratoire de Dynamique, Interactions et Réactivité". In: *Raman Spectroscopy in Archaeology and Art History*. Ed. by H. G. M. Edwards, J. M. Chalmers and C. Coupry. Cambridge, UK: Royal Society of Chemistry.

Le Blon, J. C. (1720). *Coloritto, or, The harmony of colouring in painting*. London, UK.

Lee, R. L. and Fraser, A. B. (2001). *Rainbows in Art, Myth, and Science*. University Park, PA: Pennsylvania State University Press.

Lenz, R. (2002). "Spaces of Spectral Distributions and Their Natural Geometry". In: *CGIV*.

Leroi, A. M. (2015). *The Lagoon: How Aristotle Invented Science*. Reprint edition. New York: Penguin Books.

Lindberg, D. C. (1966). "Roger Bacon's Theory of the Rainbow: Progress or Regress?" *Isis* 57: 235–48.

Lindberg, D. C. (1992). *The Beginnings of Western Science: the European Scientific Tradition in Philosophical, Religious and Institutional Context, c.600 B.C. to A.D. 1450*. Chicago: University of Chicago Press.

Lindberg, D. (1976). *Theories of Vision - from al-Kindi to Kepler*. Chicago: University of Chicago Press.

Lindberg, D. C. (1967). "Alhazen's Theory of Vision and Its Reception in the West". *Isis* 58.3: 321–341.

Liu, C., Sharan, L., Adelson, E. H. and Rosenholtz, R. (2010). "Exploring features in a Bayesian framework for material recognition". In: *2010 IEEE Computer Society Conference*

on Computer Vision and Pattern Recognition. San Francisco, CA, USA: IEEE: 239–246. DOI: 10.1109/CVPR.2010.5540207.

Luce, R. D. and Edwards, W. (1958). "The derivation of subjective scales from just noticeable differences." *Psychological Review* 65.4: 222–237. DOI: 10.1037/h0039821.

Luce, R. D. and Tukey, J. W. (1964). "Simultaneous conjoint measurement: A new type of fundamental measurement". *Journal of Mathematical Psychology* 1.1: 1–27. DOI: 10.1016/0022-2496(64)90015-X.

Macgillivray, T., Campbell, D., Greated, C. and Barham, R. (2002). "The Development of a Microphone Calibration Technique Using Laser Doppler Anemometry". *Acta Acustica united with Acustica* 88: 135–141.

Malkin, R., Todd, T. and Robert, D. (2015). "Quantitative imaging of acoustic reflection and interference". *Journal of Physics: Conference Series* 581: 012007. DOI: 10.1088/1742-6596/581/1/012007.

Maloney, L. T. and Yang, J. N. (2003). "Maximum likelihood difference scaling". *Journal of Vision* 3.8: 5–5. DOI: 10.1167/3.8.5.

Mannos, J. and Sakrison, D. (1974). "The effects of a visual fidelity criterion of the encoding of images". *IEEE Transactions on Information Theory* 20.4: 525–536. DOI: 10.1109/TIT.1974.1055250.

Matsumoto, T., Fukuda, K. and Uchikawa, K. (2016). "Appearance of Gold, Silver and Copper Colors of Glossy Object Surface". *International Journal of Affective Engineering* 15.3: 239–247. DOI: 10.5057/ijae.IJAE-D-16-00003.

McCamy, C. S. (1996). "Observation and measurement of the appearance of metallic materials. Part I. Macro appearance". *Color Research & Application* 21.4: 292–304. DOI: 10.1002/(SICI)1520-6378(199608)21:4<292::AID-COL4>3.0.CO;2-L.

McCamy, C. S. (1998). "Observation and measurement of the appearance of metallic materials. Part II. Micro appearance". *Color Research & Application* 23.6: 362–373. DOI: 10.1002/(SICI)1520-6378(199812)23:6<362::AID-COL4>3.0.CO;2-5.

McClelland, J. L. and Elman, J. L. (1986). "The TRACE model of speech perception". *Cognitive Psychology* 18.1: 1–86. DOI: 10.1016/0010-0285(86)90015-0.

Meese, T. S., Baker, D. H. and Summers, R. J. (2017). "Perception of global image contrast involves transparent spatial filtering and the integration and suppression of local contrasts (not RMS contrast)". *Royal Society Open Science* 4.9: 170285. DOI: 10.1098/rsos.170285.

Meese, T. S. and Summers, R. J. (2007). "Area summation in human vision at and above detection threshold". *Proceedings of the Royal Society B: Biological Sciences* 274.1627: 2891–2900. DOI: 10.1098/rspb.2007.0957.

Meisami, J. S. and Starkey, P. (1998). *Encyclopedia of Arabic Literature*. Taylor & Francis.

Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D. and Iacoboni, M. (2007). "The Essential Role of Premotor Cortex in Speech Perception". *Current Biology* 17.19: 1692–1696. DOI: 10.1016/j.cub.2007.08.064.

Michell, J. (1997). "Quantitative science and the definition of measurement in psychology". *British Journal of Psychology* 88.3: 355–383. DOI: 10.1111/j.2044-8295.1997.tb02641.x.

Michell, J. (1999). *Measurement in Psychology: A Critical History of a Methodological Concept*. Cambridge University Press.

Molera, J., Bayés, C., Roura, P., Crespo, D. and Pradell, T. (2007). "Key Parameters in the Production of Medieval Luster Colors and Shines". *Journal of the American Ceramic Society* 90.7: 2245–2254. DOI: 10.1111/j.1551-2916.2007.01563.x.

Möller, D., Degen, N. and Dual, J. (2013). "Schlieren visualization of ultrasonic standing waves in mm-sized chambers for ultrasonic particle manipulation". *Journal of Nanobiotechnology* 11: 21. DOI: 10.1186/1477-3155-11-21.

Monge, G. (1781). *Mémoire sur la théorie des déblais et des remblais*. Paris: De l'Imprimerie Royale.

Motoyoshi, I., Nishida, S., Sharan, L. and Adelson, E. H. (2007). "Image statistics and the perception of surface qualities". *Nature* 447.7141: 206–209. DOI: 10.1038/nature05724.

Möttönen, R. and Watkins, K. E. (2009). "Motor Representations of Articulators Contribute to Categorical Perception of Speech Sounds". *Journal of Neuroscience* 29.31: 9819–9825. DOI: 10.1523/JNEUROSCI.6018-08.2009.

Munro, J. H. (1983). "Medieval Monetary Problems: Bimetallism and Bullionism". *The Journal of Economic History* 43.1: 294–298.

Nadolny, J. M. (2000). "The techniques and use of gilded relief decoration by northern European painters, c.1200-1500." OCLC: 272511484. PhD thesis. Courtauld Institute of Art, University of London.

Nascimento, S. M., Amano, K. and Foster, D. H. (2016). "Spatial distributions of local illumination color in natural scenes". *Vision Research* 120: 39–44. DOI: 10.1016/j.visres.2015.07.005.

Neri, E., Jackson, M., O'Hea, M., Gregory, T., Blet-Lemarquand, M. and Schibille, N. (2017). "Analyses of glass tesserae from Kilise Tepe: New insights into an early Byzantine

production technology". *Journal of Archaeological Science: Reports* 11: 600–612. DOI: 10.1016/j.jasrep.2016.12.036.

Neri, E. and Verità, M. (2013). "Glass and metal analyses of gold leaf tesserae from 1st to 9th century mosaics. A contribution to technological and chronological knowledge". *Journal of Archaeological Science* 40.12: 4596–4606. DOI: 10.1016/j.jas.2013.07.017.

Neri, E., Verità, M., Biron, I. and Guerra, M. F. (2016). "Glass and gold: Analyses of 4th–12th centuries Levantine mosaic tesserae. A contribution to technological and chronological knowledge". *Journal of Archaeological Science* 70: 158–171. DOI: 10.1016/j.jas.2016.05.003.

Neruda, P. (1991). *The Book of Questions*. Trans. by W. O'Daly. Port Townsend: Copper Canyon Press.

Newman, S. D., Klatzky, R. L., Lederman, S. J. and Just, M. A. (2005). "Imagining material versus geometric properties of objects: an fMRI study". *Cognitive Brain Research* 23.2: 235–246. DOI: 10.1016/j.cogbrainres.2004.10.020.

Nishio, A., Goda, N. and Komatsu, H. (2012). "Neural Selectivity and Representation of Gloss in the Monkey Inferior Temporal Cortex". *Journal of Neuroscience* 32.31: 10780–10793. DOI: 10.1523/JNEUROSCI.1095-12.2012.

Nolan, E. and Hirsch, S. A. (1902). *The Greek grammar of Roger Bacon and a fragment of his Hebrew grammar;* Cambridge University Press.

Obein, G., Knoblauch, K. and Viéot, F. (2004). "Difference scaling of gloss: Nonlinearity, binocularity, and constancy". *Journal of Vision* 4.9: 4–4. DOI: 10.1167/4.9.4.

O'Grady, P. F. (2002). *Thales of Miletus: The Beginnings of Western Science and Philosophy*. 1 edition. Aldershot, Hants, England; Burlington, VT: Routledge.

Ohno, Y. and Blattner, P. (2014). *Chromaticity Difference Specification for Light Sources*. Technical Note CIE TN 001:2014. Kuala Lumpur, Malaysia: Commission Internatinoale de l'Eclairage.

Okazawa, G., Koida, K. and Komatsu, H. (2011). "Categorical properties of the color term "GOLD"". *Journal of Vision* 11.8: 4–4. DOI: 10.1167/11.8.4.

Oppenheim, A. V. and Lim, J. S. (1981). "The importance of phase in signals". *Proceedings of the IEEE* 69.5: 529–541. DOI: 10.1109/PROC.1981.12022.

Owens, J. (2013). *The Oxford Handbook of Arabic Linguistics*. Oxford University Press.

Paliwal, K. K. and Alsteris, L. (2003). "Usefulness of phase spectrum in human speech perception". In: *Eighth European Conference on Speech Communication and Technology*.

Pereira-Pardo, L., Tamburini, D. and Dyer, J. (2019). "Shedding light on the colours of medieval alabaster sculptures: Scientific analysis and digital reconstruction of their original polychromy". *Color Research & Application* 44.2: 221–233. DOI: 10.1002/col.22323.

Pérez-Arantegui, J., Molera, J., Larrea, A., Pradell, T., Vendrell-Saz, M., Borgia, I., Brunetti, B. G., Cariati, F., Fermo, P., Mellini, M., Sgamellotti, A. and Viti, C. (2001). "Luster Pottery from the Thirteenth Century to the Sixteenth Century: A Nanostructured Thin Metallic Film". *Journal of the American Ceramic Society* 84.2: 442–46. DOI: 10.1111/j.1151-2916.2001.tb00674.x.

Petrov, Y. (2014, Retrieved January 2017). *Optometrika*. MATLAB Central File Exchange.

Pharr, M., Jakob, W. and Humphreys, G. (2016). *Physically Based Rendering*. 3rd ed. Elsevier. DOI: 10.1016/C2013-0-15557-2.

Piotrowski, L. N. and Campbell, F. W. (1982). "A Demonstration of the Visual Importance and Flexibility of Spatial-Frequency Amplitude and Phase". *Perception* 11.3: 337–346. DOI: 10.1068/p110337.

Piper, B., Koukoulas, T. and Butterworth, I. (2014). "Illuminating sound: a review of how sound can be measured using light". In: *Institute of Acoustics 40th Anniversary Conference*. Vol. 36. DOI: 10.13140/2.1.2483.0406.

Pitie, F. and Kokaram, A. (2007). "The linear Monge-Kantorovitch linear colour mapping for example-based colour transfer". In: *4th European Conference on Visual Media Production*: 1–9. DOI: 10.1049/cp:20070055.

Plahter, U. (2006). "The Crucifix from Hemse - Analyses of the painting technique". In: *Medieval Painting in Northern Europe: Techniques, Analysis, Art History*. Ed. by J. Naldony. London: Archetype Publications Ltd.

Portilla, J. and Simoncelli, E. P. (2000). "A Parametric Texture Model Based on Joint Statistics of Complex Wavelet Coefficients". *International Journal of Computer Vision* 40.1: 49–70. DOI: 10.1023/A:1026553619983.

Pradell, T., Molera, J., Roque, J., Vendrell-Saz, M., Smith, A. D., Pantos, E. and Crespo, D. (2005). "Ionic-Exchange Mechanism in the Formation of Medieval Luster Decorations". *Journal of the American Ceramic Society* 88.5: 1281–1289. DOI: 10.1111/j.1551-2916.2005.00223.x.

Pradell, T. (2016). "Lustre and Nanostructures—Ancient Technologies Revisited". In: *Nanoscience and Cultural Heritage*. Ed. by P. Dillmann, L. Bellot-Gurlet and I. Nenner. Paris: Atlantis Press: 3–39. DOI: 10.2991/978-94-6239-198-7_1.

Pulvermuller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O. and Shtyrov, Y. (2006). "Motor cortex maps articulatory features of speech sounds". *Proceedings of the National Academy of Sciences* 103.20: 7865–7870. DOI: 10.1073/pnas.0509989103.

Pyykko, P. and Desclaux, J. P. (1979). "Relativity and the periodic system of elements". *Accounts of Chemical Research* 12.8: 276–281.

Qin, C. and Carreira-Perpiñán, M. Á. (2009). "The geometry of the articulatory region that produces a speech sound". In: *2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*: 1742–1746. DOI: 10.1109/ACSSC.2009.5469741.

Radonjić, A., Cottaris, N. P. and Brainard, D. H. (2015). "Color constancy supports cross-illumination color selection". *Journal of Vision* 15.6: 13–13. DOI: 10.1167/15.6.13.

Rashed, R. (1990). "A Pioneer in Anaclastics: Ibn Sahl on Burning Mirrors and Lenses". *Isis* 81.3: 464–491. DOI: 10.1086/355456.

Reinhard, E., Ashikhmin, M., Gooch, B. and Shirley, P. (2001). "Color Transfer between Images". *IEEE Computer Graphics and Applications*: 8.

Reinhard, E., Stark, M., Shirley, P. and Ferwerda, J. (2002). "Photographic tone reproduction for digital images". *ACM Transactions on Graphics* 21.3. DOI: 10.1145/566654.566575.

Rogers, M., Knoblauch, K. and Franklin, A. (2016). "Maximum likelihood conjoint measurement of lightness and chroma". *Journal of the Optical Society of America A* 33.3: A184. DOI: 10.1364/JOSAA.33.00A184.

Roser-Catach, I. (1997). "Roger Bacon and Grammar". In: *Roger Bacon and the Sciences: Commemorative Essays*. Ed. by J. Hackett. BRILL.

Rowe, G. and Frewer, L. J. (2005). "A Typology of Public Engagement Mechanisms". *Science, Technology, & Human Values* 30.2: 251–290. DOI: 10.1177/0162243904271724.

Rowell, J., Noguchi, M., Bernhardt, B. M., Herdman, A., Gick, B. and Schellenberg, M. (2016). "Schlieren study of external airflow during the production of nasal and oral vowels in French". *Canadian Acoustics* 44.3.

Rubner, Y. and Tomasi, C. (1998). "Texture metrics". In: *SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No.98CH36218)*. Vol. 5: 4601–4607 vol.5. DOI: 10.1109/ICSMC.1998.727577.

Rubner, Y., Tomasi, C. and Guibas, L. J. (1998). "A metric for distributions with applications to image databases". In: *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*. Bombay, India: Narosa Publishing House: 59–66. DOI: 10.1109/ICCV.1998.710701.

Rubner, Y., Tomasi, C. and Guibas, L. J. (2000). "The Earth Mover's Distance as a Metric for Image Retrieval". *International Journal of Computer Vision* 40.2: 99–121. DOI: 10.1023/A:1026543900054.

Ruderman, D. L. and Bialek, W. (1994). "Statistics of natural images: Scaling in the woods". *Physical Review Letters* 73.6: 814–817. DOI: 10.1103/PhysRevLett.73.814.

Ruderman, D. L., Cronin, T. W. and Chiao, C.-C. (1998). "Statistics of cone responses to natural images: implications for visual coding". *JOSA A* 15.8: 2036–2045. DOI: 10.1364/JOSAA.15.002036.

Rust, N. C. and Movshon, J. A. (2005). "In praise of artifice". *Nature Neuroscience* 8.12: 1647–1650. DOI: 10.1038/nn1606.

Sadeghi, I., Munoz, A., Laven, P., Jarosz, W., Seron, F., Gutierrez, D. and Jensen, H. W. (2012). "Physically-based simulation of rainbows". *ACM Transactions on Graphics* 31.1: 1–12. DOI: 10.1145/2077341.2077344.

Sakitt, B. (1972). "Counting every quantum". *The Journal of Physiology* 223.1: 131–150. DOI: 10.1113/jphysiol.1972.sp009838.

Sarton, G. (1916). "The History of Science". *The Monist* 26.3: 321–365.

Sato, M., Tremblay, P. and Gracco, V. L. (2009). "A mediating role of the premotor cortex in phoneme segmentation". *Brain and Language* 111.1: 1–7. DOI: 10.1016/j.bandl.2009.03.002.

Sawayama, M. and Nishida, S. (2018). "Material and shape perception based on two types of intensity gradient information". *PLOS Computational Biology* 14.4: e1006061. DOI: 10.1371/journal.pcbi.1006061.

Sayili, A. M. (1939). "The Aristotelian Explanation of the Rainbow". *Isis* 30.1: 65–83.

Schmidtke, D. S., Conrad, M. and Jacobs, A. M. (2014). "Phonological iconicity". *Frontiers in Psychology* 5. DOI: 10.3389/fpsyg.2014.00080.

Schnitzler, A. D. (1976a). "Theory of spatial-frequency filtering by the human visual system. I. Performance limited by quantum noise*". *JOSA* 66.6: 608–617. DOI: 10.1364/JOSA.66.000608.

Schnitzler, A. D. (1976b). "Theory of spatial-frequency filtering by the human visual system. II. Performance limited by video noise*". *JOSA* 66.6: 617–625. DOI: 10.1364/JOSA.66.000617.

Seetzen, H., Heidrich, W., Stuerzlinger, W., Ward, G., Whitehead, L., Trentacoste, M., Ghosh, A. and Vorozcovs, A. (2004). "High Dynamic Range Display Systems". In: *ACM SIGGRAPH*

2004 Papers. SIGGRAPH '04. event-place: Los Angeles, California. New York, NY, USA: ACM: 760–768. DOI: 10.1145/1186562.1015797.

Selmke, M. and Selmke, S. (2017). "Artificial circumzenithal and circumhorizontal arcs". *American Journal of Physics* 85.8: 575–581. DOI: 10.1119/1.4984802.

Settles, G. S. (2001). *Schlieren and Shadowgraph Techniques: Visualizing Phenomena in Transparent Media*. Springer Science & Business Media.

Settles, G. S. and Hargather, M. J. (2017). "A review of recent developments in schlieren and shadowgraph techniques". *Measurement Science and Technology* 28.4: 042001. DOI: 10.1088/1361-6501/aa5748.

Sharan, L., Li, Y., Motoyoshi, I., Nishida, S. and Adelson, E. H. (2008). "Image statistics for surface reflectance perception". *JOSA A* 25.4: 846–865. DOI: 10.1364/JOSAA.25.000846.

Sharan, L., Rosenholtz, R. and Adelson, E. H. (2014). "Accuracy and speed of material categorization in real-world images". *Journal of Vision* 14.9: 12–12. DOI: 10.1167/14.9.12.

Sharpe, J. P. and Greated, C. A. (1987). "The measurement of periodic acoustic fields using photon correlation spectroscopy". *Journal of Physics D: Applied Physics* 20.4: 418.

Sharpe, J. P. and Greated, C. A. (1989). "A stochastic model for photon correlation measurements in sound fields". *Journal of Physics D: Applied Physics* 22.10: 1429. DOI: 10.1088/0022-3727/22/10/003.

Sharpe, L. T., Stockman, A., Jägle, H. and Nathans, J. (1999). "Opsin genes, cone photopigments, color vision, and color blindness". In: *Color Vision*. Ed. by K. R. Gegenfurtner and L. T. Sharpe. Cambridge, UK: Cambridge University Press: 50.

Sibson, R. (1978). "Studies in the Robustness of Multidimensional Scaling: Procrustes Statistics". *Journal of the Royal Statistical Society. Series B (Methodological)* 40.2: 234–238.

Simoncelli, E. P. and Olshausen, B. A. (2001). "Natural Image Statistics and Neural Representation". *Annual Review of Neuroscience* 24.1: 1193–1216. DOI: 10.1146/annurev.neuro.24.1.1193.

Simoncelli, E. and Freeman, W. (1995). "The steerable pyramid: a flexible architecture for multi-scale derivative computation". In: *Proceedings., International Conference on Image Processing*. Vol. 3. Washington, DC, USA: IEEE Comput. Soc. Press: 444–447. DOI: 10.1109/ICIP.1995.537667.

Sinnott, J. M., Brown, C. H., Malik, W. T. and Kressley, R. A. (1997). "A multidimensional scaling analysis of vowel discrimination in humans and monkeys". *Perception & Psychophysics* 59.8: 1214–1224.

Smith, A. M. (2015). *From Sight to Light: The Passage from Ancient to Modern Optics*. Chicago: University of Chicago Press.

Smithson, H. E. (2016). "All the colours of the Rainbow: Robert Grosseteste's Three-Dimensional Colour Space". In: *Robert Grosseteste and the pursuit of Religious and Scientific Learning in the Middle Ages*. Ed. by J. P. Cunningham and M. Hocknull. Vol. 18. London: Springer International Publishing: 59–84.

Smithson, H. E., Anderson, P. S., Dinkova-Bruun, G., Fosbury, R. A. E., Gasper, G. E. M., Laven, P., McLeish, T. C. B., Panti, C. and Tanner, B. (2014a). "A color coordinate system from a 13th century account of rainbows". *Journal of the Optical Society of America. A, Optics, image science, and vision* 31.4: A341–A349.

Smithson, H. E., Dinkova-Bruun, G., Gasper, G. E. M., Huxtable, M., McLeish, T. C. B. and Panti, C. (2012). "A three-dimensional color space from the 13th century". *Journal of the Optical Society of America A* 29.2: A346. DOI: 10.1364/JOSAA.29.00A346.

Smithson, H. E., Gasper, G. E. M. and McLeish, T. C. B. (2014b). "All the colours of the rainbow". *Nature Physics* 10.8: 540–542. DOI: 10.1038/nphys3052.

Snobelen, S. D. (2016). ""A Time and Times and the Dividing of Time": Isaac Newton, the Apocalypse, and 2060 A.D." *Canadian Journal of History*. DOI: 10.3138/cjh.38.3.537.

Sønnesyn, S. O. and Gasper, G. E. M. (2019). "Aristotle, Priscian, and Isidore". In: *Knowing and Speaking: Robert Grosseteste's De artibus liberalibus 'On the Liberal Arts' and De generatione sonorum 'On the Generation of Sounds'*. Ed. by G. E. M. Gasper, C. Panti, T. C. B. McLeish and H. E. Smithson. Vol. 1. The Scientific Works of Robert Grosseteste. Oxford: Oxford University Press.

Speigle, J. M. and Brainard, D. H. (1996). "Luminosity thresholds: effects of test chromaticity and ambient illumination". *JOSA A* 13.3: 436–451. DOI: 10.1364/JOSAA.13.000436.

Steffens, B. (2006). *Ibn Al-Haytham: First Scientist*. Greensboro, N.C: Morgan Reynolds Publishing.

Stephens, G. J., Mora, T., Tkačik, G. and Bialek, W. (2013). "Statistical Thermodynamics of Natural Images". *Physical Review Letters* 110.1: 018701. DOI: 10.1103/PhysRevLett.110.018701.

Stevens, K. N. (2002). "Toward a model for lexical access based on acoustic landmarks and distinctive features". *The Journal of the Acoustical Society of America* 111.4: 1872–1891. DOI: 10.1121/1.1458026.

Stevens, S. S. (1936). "A scale for the measurement of a psychological magnitude: loudness". *Psychological Review* 43.5: 405–416. DOI: 10.1037/h0058773.

Stevens, S. S. (1957). "On the psychophysical law." *Psychological Review* 64.3: 153–181. DOI: 10.1037/h0046162.

Sweeny, T. D., Guzman-Martinez, E., Ortega, L., Grabowecky, M. and Suzuki, S. (2012). "Sounds exaggerate visual shape". *Cognition* 124.2: 194–200. DOI: 10.1016/j.cognition.2012.04.009.

Tamura, H., Higashi, H. and Nakauchi, S. (2018). "Dynamic Visual Cues for Differentiating Mirror and Glass". *Scientific Reports* 8.1: 8403. DOI: 10.1038/s41598-018-26720-x.

Tanaka, M. and Horiuchi, T. (2017). "Perception of gold materials by projecting a solid color on black materials". *Color Research & Application* 42.4: 522–530. DOI: 10.1002/col.22107.

Tang, J. W., Nicolle, A. D. G., Pantelic, J., Jiang, M., Sekhr, C., Cheong, D. K. W. and Tham, K. W. (2011). "Qualitative Real-Time Schlieren and Shadowgraph Imaging of Human Exhaled Airflows: An Aid to Aerosol Infection Control". *PLOS ONE* 6.6: e21392. DOI: 10.1371/journal.pone.0021392.

Tape, W. and Moilanen, J. (2006). *Atmospheric Halos and the Search for Angle x*. American Geophysical Union.

Taylor, H. G. and Waldram, J. M. (1933). "Improvements in the Schlieren method". *Journal of Scientific Instruments* 10.12: 378. DOI: 10.1088/0950-7671/10/12/304.

Thompson, M. W. and Atchley, A. A. (2005). "Simultaneous measurement of acoustic and streaming velocities in a standing wave using laser Doppler anemometry". *The Journal of the Acoustical Society of America* 117.4 Pt 1: 1828–1838.

Thomson, M. G. A. and Foster, D. H. (1997). "Role of second- and third-order statistics in the discriminability of natural images". *Journal of the Optical Society of America A* 14.9: 2081. DOI: 10.1364/JOSAA.14.002081.

Tian, X. and Poeppel, D. (2012). "Mental imagery of speech: linking motor and perceptual systems through internal simulation and estimation". *Frontiers in Human Neuroscience* 6. DOI: 10.3389/fnhum.2012.00314.

Todd, J. T. and Norman, J. F. (2018). "The visual perception of metal". *Journal of Vision* 18.3: 9–9. DOI: 10.1167/18.3.9.

Todorovic, D. (2018). *Visual illusions and contextual effects: definitions, varieties, explanations*. European Conference on Visual Perception; Trieste, Italy.

Tolhurst, D. J., Tadmor, Y. and Chao, T. (1992). "Amplitude spectra of natural images". *Ophthalmic & Physiological Optics: The Journal of the British College of Ophthalmic Opticians (Optometrists)* 12.2: 229–232.

Toscani, M., Guarnera, D., Claudio Guarnera, G., Hardeberg, J. Y. and Gegenfurtner, K. R. (2018). *A role for metallicity in the perception of surface reflectance*. Trieste, Italy.

Turoman, N. and Styles, S. J. (2017). "Glyph guessing for 'oo' and 'ee': spatial frequency information in sound symbolic matching for ancient and unfamiliar scripts". *Royal Society Open Science* 4.9: 170882. DOI: 10.1098/rsos.170882.

Tversky, A. (1967). "A general theory of polynomial conjoint measurement". *Journal of Mathematical Psychology* 4.1: 1–20. DOI: 10.1016/0022-2496(67)90039-9.

Uchikawa, K., Koida, K., Meguro, T., Yamauchi, Y. and Kuriki, I. (2001). "Brightness, not luminance, determines transition from the surface-color to the aperture-color mode for colored lights". *JOSA A* 18.4: 737–746. DOI: 10.1364/JOSAA.18.000737.

Urbach, P. (1987). *Francis Bacon's Philosophy of Science: An Account and a Reappraisal*. 2nd Revised edition edition. La Salle, Ill: Open Court Publishing Co ,U.S.

Vangorp, P., Barla, P. and Fleming, R. W. (2017). "The perception of hazy gloss". *Journal of Vision* 17.5: 19. DOI: 10.1167/17.5.19.

Vangorp, P., Laurijssen, J. and Dutré, P. (2007). "The influence of shape on the perception of material reflectance". *ACM Transactions on Graphics* 26.99: 77. DOI: 10.1145/1239451.1239528.

Walter, B., Marschner, S. R., Li, H. and Torrance, K. E. (2007). "Microfacet Models for Refraction Through Rough Surfaces". In: *Proceedings of the 18th Eurographics Conference on Rendering Techniques*. EGSR'07. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association: 195–206. DOI: 10.2312/EGWR/EGSR07/195-206.

Wanat, R., Petit, J. and Mantiuk, R. (2012). "Physical and Perceptual Limitations of a Projector-based High Dynamic Range Display". In: *TPCG*. DOI: 10.2312/LocalChapterEvents/TPCG/TPCG12/009-016.

Watkins, K., Strafella, A. and Paus, T. (2003). "Seeing and hearing speech excites the motor system involved in speech production". *Neuropsychologia* 41.8: 989–994. DOI: 10.1016/S0028-3932(02)00316-0.

Weidhorn, M. (2005). *The Person of the Millennium: The Unique Impact of Galileo on World History*. New York; Lincoln, NE: iUniverse.

Weinstein, L. M., Stacy, K., Vieira, G. J., Haering Jr, E. A. and Bowers, A. H. (1997). "Visualization and image processing of aircraft shock wave structures". In: *1st Pacific Symposium on Flow Visualization and Image Processing, Feb. 23-26, 1997, Honolulu, Hawaii*.

Westheimer, G. (2001). "The Fourier Theory of Vision". *Perception* 30.5: 531–541. DOI: 10.1068/p3193.

Westphal, J. A. (1965). "Schlieren Technique for Studying Water Flow in Marine Animals". *Science* 149.3691: 1515–1516. DOI: 10.1126/science.149.3691.1515-a.

Whitrow, G. J. (1973). *Einstein, the Man and His Achievement*. Courier Corporation.

Wichmann, F. A., Braun, D. I. and Gegenfurtner, K. R. (2006). "Phase noise and the classification of natural images". *Vision Research* 46.8: 1520–1529. DOI: 10.1016/j.visres.2005.11.008.

Wiebel, C. B., Aguilar, G. and Maertens, M. (2017). "Maximum likelihood difference scales represent perceptual magnitudes and predict appearance matches". *Journal of Vision* 17.4: 1–1. DOI: 10.1167/17.4.1.

Wiebel, C. B., Toscani, M. and Gegenfurtner, K. R. (2015). "Statistical correlates of perceived gloss in natural images". *Vision Research*. Perception of Material Properties (Part II) 115: 175–187. DOI: 10.1016/j.visres.2015.04.010.

Wilks, S. S. (1938). "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". *The Annals of Mathematical Statistics* 9.1: 60–62. DOI: 10.1214/aoms/1177732360.

Woolgar, C. M. (2006). *The Senses in Late Medieval England*. New Haven, CT, and London, UK: Yale University Press.

Wootton, D. (2015). *The Invention of Science: A New History of the Scientific Revolution*. First Edition edition. London: Allen Lane.

Yamauchi, Y. and Uchikawa, K. (2000). "Upper-limit luminance for the surface-color mode appearance". *JOSA A* 17.11: 1933–1941. DOI: 10.1364/JOSAA.17.001933.

Yamauchi, Y. and Uchikawa, K. (2005). "Depth information affects judgment of the surface-color mode appearance". *Journal of Vision* 5.6: 3–3. DOI: 10.1167/5.6.3.

Yang, J., Kanazawa, S. and Yamaguchi, M. K. (2013). "Can Infants Tell the Difference between Gold and Yellow?" *PLOS ONE* 8.6: e67535. DOI: 10.1371/journal.pone.0067535.

Young Thomas (1802). "II. The Bakerian Lecture. On the theory of light and colours". *Philosophical Transactions of the Royal Society of London* 92: 12–48. DOI: 10.1098/rstl.1802.0004.

Zele, A. J., Feigl, B., Adhikari, P., Maynard, M. L. and Cao, D. (2018). "Melanopsin photoreception contributes to human visual detection, temporal and colour processing". *Scientific Reports* 8.1: 1–10. DOI: 10.1038/s41598-018-22197-w.

Zilsel, E. (1941). "The Origins of William Gilbert's Scientific Method". *Journal of the History of Ideas* 2.1: 1–32. DOI: 10.2307/2707279.