Heron, S (2014) *From local constraints to global binocular motion perception.* PhD thesis.

http://theses.gla.ac.uk/5218/

FROM LOCAL CONSTRAINTS TO GLOBAL BINOCULAR MOTION PERCEPTION

Suzanne Heron

School of Psychology
University of Glasgow

Submitted for the Degree of Doctor of Philosophy to the Higher Degrees Committee of the
College of Science and Engineering, University of Glasgow

May, 2014

Abstract

Humans and many other predators have two eyes that are set a short distance apart so that an extensive region of the world is seen simultaneously by both eyes from slightly different points of view. Although the images of the world are essentially two-dimensional, we vividly see the world as three-dimensional. This is true for static as well as dynamic images.

We discuss local constraints for the perception of three-dimensional binocular motion in a geometric-probabilistic framework. It is shown that Bayesian models of binocular 3D motion can explain perceptual bias under uncertainty and predict perceived velocity under ambiguity. The models exploit biologically plausible constraints of local motion and disparity processing in a binocular viewing geometry.

Results from psychophysical experiments and an fMRI study support the idea that local constraints of motion and disparity processing are combined late in the visual processing hierarchy to establish perceived 3D motion direction. The methods and results reported here are likely to stimulate computational, psychophysical, and neuroscientific research because they address the fundamental issue of how 3D motion is represented in the human visual system.

Doubt is not a pleasant condition, but certainty is absurd.

Francois Marie Voltaire (1694-1778)

Declaration

I declare that this thesis is my own work, carried out under the normal terms of supervision and collaboration. Some of the work contained in this work has been previously published.

[1] Lages, M., & Heron, S. (2008). Motion and disparity processing informs Bayesian 3D motion estimation. *Proceedings of the National Academy of Sciences of the USA*, *105(51)*, e117.

[2] Lages, M., & Heron, S. (2009). Testing generalized models of binocular 3D motion perception [Abstract]. *Journal of Vision*, *9(8)*, 636a.

[3] Heron, S., & Lages, M. (2009). Measuring azimuth and elevation of binocular 3D motion direction [Abstract]. *Journal of Vision*, *9(8)*, 637a.

[4] Lages, M., &, Heron, S. (2010). On the inverse problem of local binocular 3D motion perception. *PLoS Computational Biology*, *6(11)*, e1000999.

[5] Heron, S., & Lages, M. (2012). Screening and sampling in binocular vision studies. *Vision Research*, *62*, 228-234.

[6] Lages, M., Heron, S., & Wang, H. (2013). Local constraints for the perception of binocular 3D motion (Chapter 5, pp. 90-120). In: *Developing and Applying Biologically-Inspired Vision Systems: Interdisciplinary Concepts* (M. Pomplun & J. Suzuki, Eds.) IGI Global: New York, NY.

[7] Wang, H., Heron, S., Moreland, J., & Lages, M. (in press). A Bayesian approach to the aperture problem of 3D motion perception. *Proceedings of IC3D 2012*, Liege BE.

Acknowledgements

I would like to express my heartfelt gratitude to my first supervisor Dr Martin Lages; without whose support, guidance and expertise, the writing of this thesis would not have been possible. Martin showed unfaltering patience and understanding throughout difficult times and encouraged me not to give up. I can only hope he understands what an integral role he played throughout my postgraduate studies.

I would also like to thank the others who contributed to the work in this thesis, in particular Dr Hongfang Wang for her contribution to the mathematical modelling work and my second supervisor Dr Lars Muckli, who waited patiently as I got to grips with Brainvoyager and was integral in the collecting and analysis of the brain imaging results. To Francis Crabbe, the research radiographer in the CCNi, thank you for helping to run the MRI experiment, for listening to my woes and for being full of good chat during the data gathering.

A general thanks towards all of the staff in the School of Psychology, Institute of Neuroscience and CCNi and to the teaching staff in the undergraduate psychology labs, who unknowingly provided relief from the rigours of academic study.

On a personal note I would like to thank all of my colleagues, and fellow graduate students, who have been such a valuable support network in the department. My officemates Dr Yui Cui, Lukasz Piwek and Emanuelle De Luca deserve a special mention for putting up with me for four years and for providing solace, chocolate and coffee when the going got tough. Thank you to Dr Rebecca Watson, Dr C.F. Harvey and Judith Stevenson for the unofficial therapy sessions and friendship.

Thank you also, to Dr David Simmons, who has been an unofficial mentor and friend throughout my studies and with whom I had many stimulating conversations about autism, philosophy and life in general.

A very special thank you to all of my family and friends, whose emotional support throughout my studies, and indeed life, has been immeasurable. In particular, my parents and grandparents and sister for giving such solid advice, financial assistance and for always letting me know I was loved unconditionally. A special mention to my late grandfather Patrick Heron, who I know would wish he could have been here to see the finished product. I should not forget to mention my close friend Sharan Tagore, who has seen me at my worst and continues to stand by me (be the change you wish to see in the world).

Finally, I would like to express my gratitude for the opportunity and financial assistance provided by the Engineering Physical Sciences Research Council (EPSRC) studentship. I would not have been able to undertake my postgraduate studies otherwise.

Table of Contents

Index of Figures

viii

Index of Tables

CHAPTER 1. LOCAL MOTION PERCEPTION

1.1 Introduction

Like many other predators in the animal kingdom humans have two eyes that are set a short distance apart so that an extensive region of the world is seen simultaneously by both eyes from slightly different points of view. Vision in this region of binocular overlap has a special quality that has intrigued artists, philosophers, and scientists.
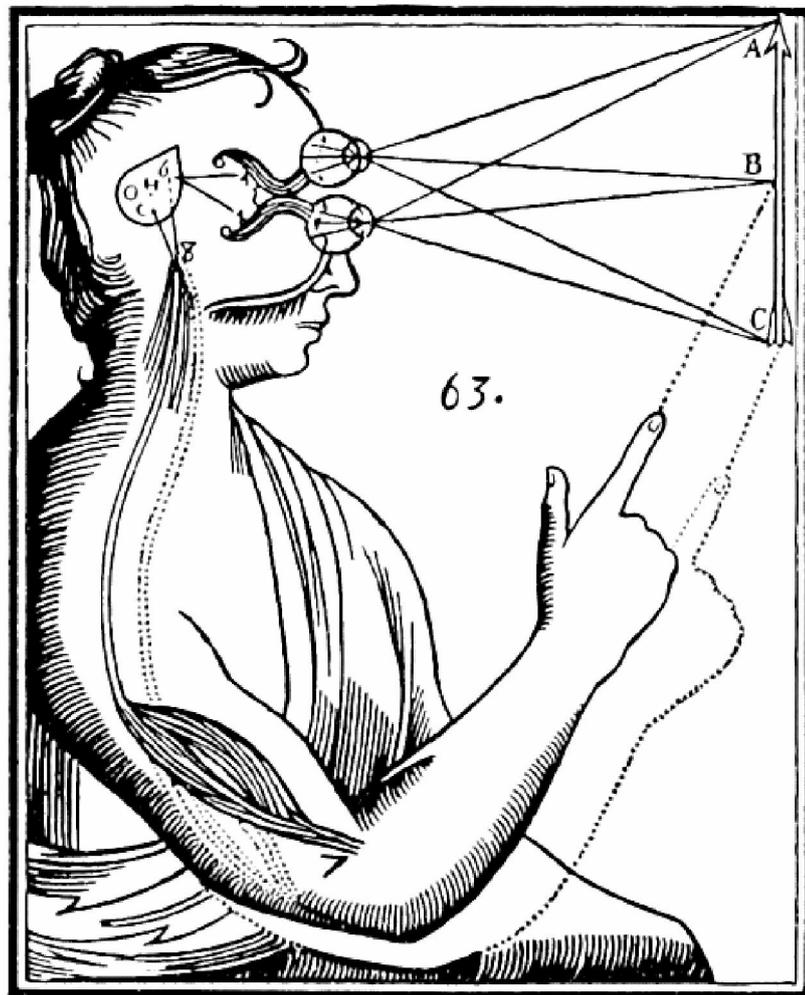
Figure 1.1 An early illustration of the binocular perceptual system after *René Descartes* (woodcut in *Traité de l'Homme*,1664 [*De Homine,* 1633/1662]).

Extromission theory, the notion that rays emanate from the eyes to inform about the external world, was proposed by a school of philosophers, known as 'extromissionists' in ancient Greece (Empedocles, 500 BCE; Plato, 400 BCE; Euclid, 300 BCE; Lucretius, 55 BCE; Ptolemy, 200 BCE). The idea has long been dismissed in favor of intromission theory, the concept that rays of light enter the eye. Similarly, René Descartes's concept of the mind as a spirit that communicates with the brain via the eyes has been refuted (see Fig. 1.1 for the original illustration). Contrary to what René Descartes (1641) believed, all the physiological evidence suggests that the mind is not situated outside the body in an ethereal metaphysical realm, but resides inside the head manifested as physical matter. Solving the inverse problem of visual perception however, highlights the need to infer a distal, physical world from proximal sensory information (Berkeley, 1709/1975). In this sense our mind ventures outside the body to create a metaphysical world – our perception of the external world.

The perceptual inference of the three-dimensional (3D) external world from two-dimensional (2D) retinal input is a fundamental problem (Berkeley, 1709/1975; von Helmholtz, 1910/1962) that the visual system has to solve through neural computation (Poggio,Torre, & Koch, 1985; Pizlo, 2001). This is true for static scenes as well as for dynamic events. For dynamic events the inverse problem implies that the visual system estimates motion in 3D space from local encoding and spatio-temporal processing.

Under natural viewing conditions the human visual system seems to effortlessly establish a 3D motion percept from local inputs to the left and right eye. The instantaneous integration of binocular input is essential for object recognition, navigation, action planning and execution. It appears obvious that many depth cues help to establish 3D motion perception under natural viewing conditions but local motion and disparity input features prominently in the early processing stages of the visual system (Howard & Rogers, 2002).

Velocity in 3D space is described by motion direction and speed. Motion direction can be measured in terms of azimuth and elevation angle, and motion direction together with speed is conveniently expressed as a vector in a 3D Cartesian coordinate system. Estimating local motion vectors is highly desirable for a visual system because local estimates in a dense vector field provide the basis for the perception of 3D object motion - that is direction and speed of a moving object. This information is essential for segmenting objects from the background, for interpreting objects as well as for planning and executing actions in a dynamic environment.

If a single moving point, corner, or other unique feature serves as binocular input then intersection of constraint lines or triangulation in a binocular viewing geometry provides a straightforward and unique geometrical solution to the inverse problem. If, however, the moving stimulus has spatial extent, such as an oriented line or contour inside a circular aperture or receptive field then local motion direction of corresponding receptive fields in the left and right eye remains ambiguous, and additional constraints are needed to solve the inverse problem in 3D.

The inverse optics and the aperture problem are well-known problems in computational vision, especially in the context of stereo processing (Poggio, Torre, & Koch, 1985; Mayhew & Longuet-Higgins, 1982), structure from motion (Koenderink & van Doorn, 1991), and optic flow (Hildreth, 1984). Gradient constraint and related methods (e.g., Johnston et al., 1999) belong to the most widely used techniques of optic-flow computation based on image intensities. They can be divided into local area-based (Lucas & Kanade, 1981) and into more global optic flow methods (Horn & Schunck, 1981). Both techniques usually employ brightness constancy and smoothness constraints in the image to estimate velocity in an over-determined equation system. It is important to note that optical flow only provides a constraint in the direction of the image gradient, the normal component of the optical flow. As a consequence some form of regularization or smoothing is needed. Various algorithms have been developed implementing error minimization and regularization for 3D stereo-motion detection (e.g., Bruhn, Weickert & Schnörr, (2005); Spies, Jähne & Barron, 2002; Min & Sohn,

2006; Scharr & Küsters, 2002). These algorithms effectively extend processing principles of 2D optical flow to 3D scene flow (Vedula, et al., 2005; Carceroni & Kutulakos, 2002).

However, computational studies on 3D motion are usually concerned with fast and efficient encoding. Here we are less concerned with the efficiency or robustness of a particular algorithm and implementation. Instead we want to understand local and binocular constraints in order to explain characteristics of human 3D motion perception such as perceptual bias under uncertainty and motion estimation under ambiguity. Ambiguity of 2D motion direction is an important aspect of biologically plausible processing and has been extensively researched in the context of the 2D aperture problem (Wallach, 1935; Adelson & Movshon, 1982; Sung, Wojtach, & Purves, 2009) but there is a surprising lack of studies on the 3D aperture problem (Morgan & Castet, 1997) and perceived 3D motion.

The entire perceptual process may be understood as a form of statistical inference (Knill, Kersten & Yuille, 1996) and motion perception has been modeled as an inferential process for 2D object motion (Weiss, Simoncelli & Adelson, 2002) and 3D surfaces (Ji & Fermüller, 2006). Models of binocular 3D motion perception on the other hand are typically deterministic and predict only azimuth or change in depth (Regan & Gray, 2009). In Chapter 3 we discuss probabilistic models of 3D motion perception that are based on velocity constraints and can explain perceptual bias under uncertainty as well as motion estimation under ambiguity.

For the sake of simplicity we exclude the discussion of eye, head and body movements of the observer and consider only passively observed, local motion. Smooth motion pursuit of the eyes and self-motion of the observer during object motion are beyond the scope of this thesis and have been considered elsewhere (Harris, 2006; Rushton & Warren, 2005; Miles, 1998).

1.2 BINOCULAR 3D MOTION

Any biologically plausible model of binocular 3D motion perception has to rely on binocular sampling of local spatio-temporal information (Beverley & Regan, 1973; 1974; 1975). There are at least three known cell types in primary visual cortex V1 that may be involved in local encoding of 3D motion: simple and complex motion detecting cells (Hubel & Wiesel, 1962; 1968; DeAngelis, Ohzawa, & Freeman, 1993; Maunsell & van Essen, 1983), binocular disparity detecting cells (Barlow et al, 1967; Hubel & Wiesel, 1970; Nikara et al, 1968; Pettigrew et al, 1986; Poggio & Fischer, 1977; Ferster, 1981; Le Vay & Voigt, 1988; Ohzawa, DeAngelis & Freeman, 1990), and joint motion and disparity detecting cells (Anzai, Ohzawa & Freeman, 2001; Bradley, Qian & Andersen, 1995; DeAngelis & Newsome, 1999).

It is therefore not surprising that three approaches to binocular 3D motion perception emerged in the literature: (i) interocular velocity difference (IOVD) is based on monocular motion detectors, (ii) changing disparity over time (CDOT) monitors output of binocular disparity detectors, and (iii) joint encoding of motion and disparity (JEMD) relies on binocular motion detectors also tuned to disparity.

These three approaches have generated an impressive body of results but psychophysical experiments have been inconclusive and the nature of 3D motion processing remains an unresolved issue (Regan & Gray, 2009; Harris, Nefs, & Grafton, 2008). Despite the wealth of empirical studies on 2D motion (x-y motion) and motion in depth (x-z motion) there is a lack of research on true 3D motion perception (x-y-z motion).

In psychophysical studies vision researchers have tried to isolate motion and disparity input by creating specific motion stimuli. These stimuli are rendered in stereoscopic view and typically consist of many random dots in so-called random dot kinematograms (RDKs) that give rise to the perception of a moving surface, defined by motion, disparity or both. However, psychophysical evidence based on detection and discrimination thresholds using these

stimuli has been inconclusive, supporting interocular velocity difference (Brooks, 2002; Fernandez & Farrell, 2005; Portfors-Yeomans & Regan, 1996; Shioiri, Saisho, & Yaguchi, 2000; Rokers, et al., 2008), changing disparity (Cumming & Parker, 1994; Tyler, 1971) or both (Brooks & Stone, 2004; Lages, Graf, & Mamassian, 2003; Rokers et al., 2009) as possible inputs to 3D motion perception.

Another limitation of random-dot stimuli is that random dots moving in depth may invoke intermediate and higher processing stages similar to structure from motion and global object motion. A surface defined by dots or other features can invoke mid-level surface and high-level object processing and therefore may not reflect characteristics of local motion encoding. Although the involvement of higher-level processing has always been an issue in psychophysical studies it is of particular concern when researchers relate behavioral measures of surface and object motion to characteristics of early motion processing as in binocular 3D motion perception.

In addition, detection and discrimination thresholds for RDKs often do not reveal biased 3D motion perception. Accuracy rather than precision of observers' perceptual performance needs to be measured to establish characteristics of motion and disparity processing in psychophysical studies (Harris & Dean, 2003; Welchman, Tuck & Harris, 2004; Rushton & Duke, 2007).

Lines and edges of various orientations are elementary for image processing because they signify either a change in the reflectance of the surface, a change in the amount of light falling on it, or a change in surface orientation relative to the light source. For these and other reasons, lines and edges are universally regarded as important image-based features or primitives (Marr, 1982). The departure from random-dot kinematograms (RDKs), typically used in stereo research and binocular motion in depth (Julesz, 1971), is significant because a line in a circular aperture effectively mimics the receptive field of a local motion detector. Local motion and disparity of a line, where endpoints are occluded behind a circular aperture, is highly ambiguous in terms of 3D motion direction and speed but it would be interesting to know how the visual system resolves this ambiguity and which constraints are employed to achieve estimates of local motion and global scene flow.

7

1.3 THE APERTURE PROBLEM

To represent local motion, the visual system matches corresponding image features on the retina over space and time. Due to their limited receptive field size, motion sensitive cells in the primary visual cortex (V1) sample only a relatively small range of the visual field. This poses a problem as the incoming motion signal remains ambiguous as long as there are no other features such as line terminators, junctions, and texture elements available. This phenomenon is known as the 'aperture problem' and has been extensively studied over the years (Wallach, 1935; Marr & Ullman, 1981; Marr, 1982). When observers view a moving grating or straight contour through a circular aperture, the motion direction is perceived as being orthogonal to the orientation of the line, edge, or contour. When neighbouring endpoints of the contour are occluded its motion direction is consistent with a 'family' of motions that can be described by a single constraint line in velocity space (Adelson & Movshon, 1982).

The aperture problem and the resulting 2D motion percepts and illusions have been modelled by Bayesian inference with a prior that favours a direction of motion with the least physical displacement of the stimulus (Weiss et al., 2002). This 'slow motion prior' is thought to constrain the percept under conditions of high ambiguity. A stereo analogue to the motion aperture problem has also been described. The occlusion of line end-points in a static binocular display results in ambiguity, leading to non-veridical stereo matching (van Ee & Schor, 2000; van Dam & van Ee, 2004; Read 2002).

Similar to local motion inputs, local stereo inputs are also subject to the 'stereo aperture problem' (Morgan & Castett, 1997). For stereo matching to occur, the visual system must combine retinal inputs by matching local feature information across space (Wheatstone, 1838). The information of local form is limited by the small receptive field cells of V1 neurons, so that matching between corresponding points in the left and right eye image can occur over a range of directions in two-dimensional space (Morgan & Castet, 1997; Farrell, 1998). To

recover depth, the visual system must arrive at an optimal percept from the available sensory information.

Van Ee & Schor (2001) measured stereo-matching of oblique line stimuli using an online depth probe method. When the end-points of the lines were clearly visible (short lines) observers made consistently veridical matches in response to depth defined by horizontal disparity (end-point matching) (Prazdny, 1983; Faugeras, 1993). As the length of the lines increased, matches became increasingly more consistent with 'nearest neighbour matching', orthogonal to the lines' orientation (Arditi et al; 1981; Arditi, 1982). Subsequently, the direction of stereo matching was shown to differ when the type of occluding border was defined as a single vertical line versus a grid (surface). When the occluder was perceived as a well-defined surface, a horizontal matching strategy was used. In the line occluder condition, response varied between observers; two observers used a horizontal match; two appeared to use line intersections (points where the line appears to intersect the aperture and a fifth observer matched in a direction with a perpendicular (nearest-neighbour) strategy (van Dam & van Ee, 2004). Response also varied with the aperture orientation.

When matching primitives, such as line endpoints, are weak or absent, the visual system appears to use a 'default strategy' to compute depth, in much the same way as it deals with motion ambiguity (Farrel, 1998). When computing local motion trajectories, the visual system faces two sources of ambiguity: the motion correspondence problem and the stereo correspondence problem. An important theoretical debate in the field of stereo-motion perception has centred around the role of local velocities (motion inputs) and disparities (depth inputs) in driving the early stages of motion-in-depth computation.

In the case of local binocular 3D motion perception we expect ambiguity for both motion and stereo due to local sampling. Figure 1.2 illustrates the 2D motion aperture problem in the left and right eye and the resulting 3D aperture problem where the motion signals have ambiguous disparity information.
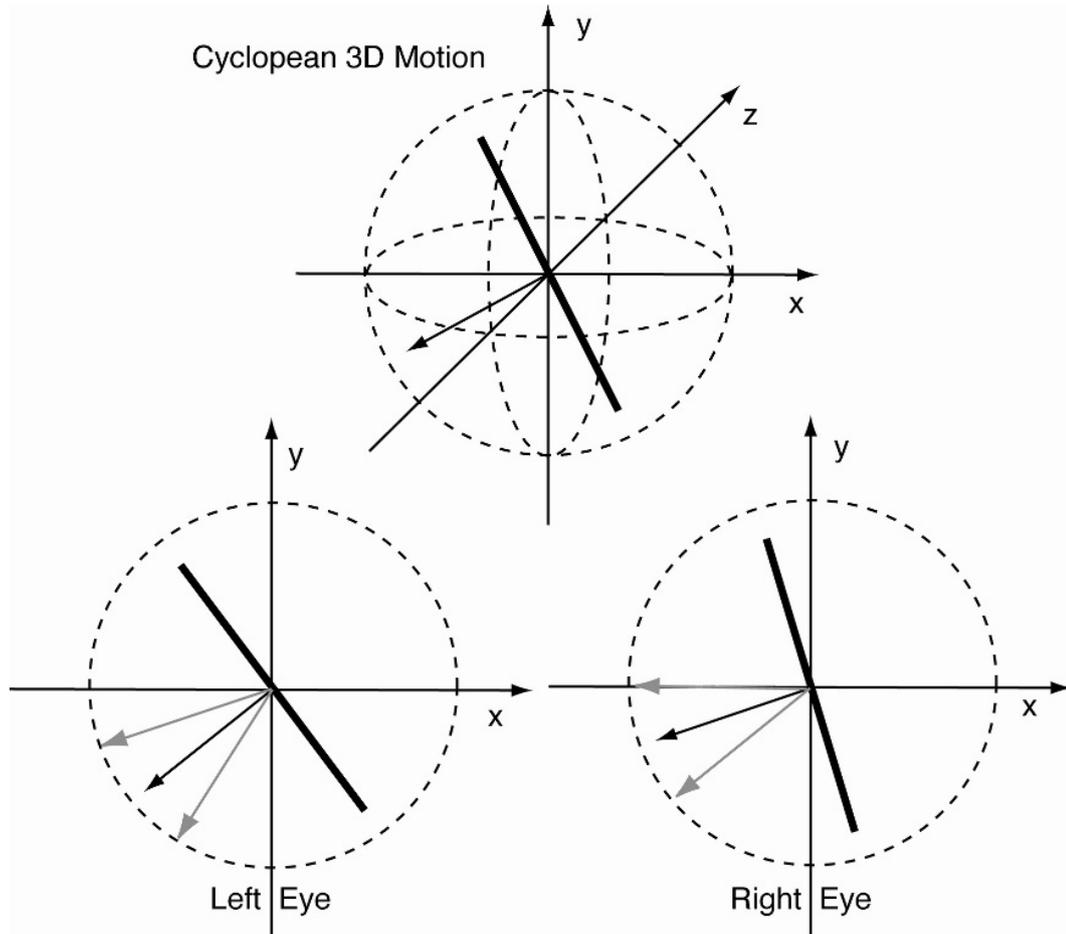
Figure 1.2 The basic 2D motion aperture problem for moving oriented line segments in the left and right eye. When viewed through an aperture, the visual signal is consistent with a range of motion directions and yet the visual system consistently selects the direction orthogonal to the lines' orientation. When binocular disparity is introduced by presenting differently oriented lines to the left and right eye, the 2D aperture problem is different for the left and right eye. The visual system has to resolve the ambiguous stereo-motion information to arrive at a (cyclopean) 3D motion estimate as illustrated above.

The binocular viewing geometry imposes obvious constraints for stimulus trajectory and velocity. For a moving dot for example the intersection of constraint lines in *x-z*

space determines trajectory angle and speed of the target moving in depth as illustrated in Fig. 1.2.
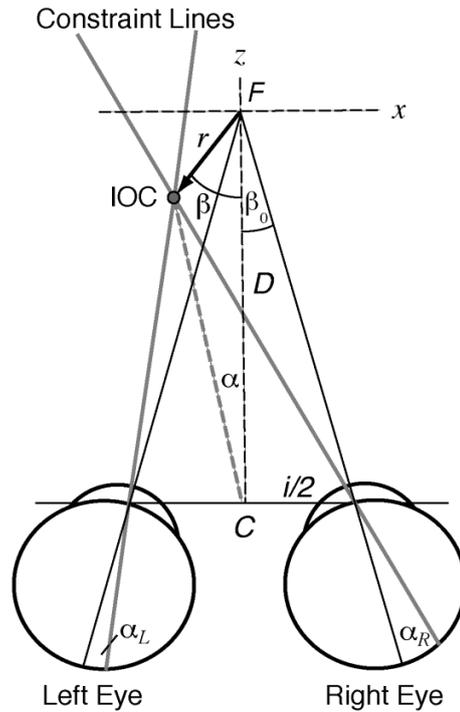
Figure 1.3 Binocular viewing geometry in top view. If the two eyes are verged on a fixation point at viewing distance D with angle β then projections of a moving target (arrow) with angle α_L in the left eye and α_R in the right eye constrain motion of the target in x-z space. The intersection of constraints (IOC) determines stimulus trajectory β and radius r.

So far models and experiments on 3D motion perception have only considered horizontal 3D motion trajectories of dots or unambiguous features that are confined to the x-z plane. In the next three chapters we investigate velocity estimates in the context of the 3D aperture problem.

The 3D aperture problem arises when a line or edge moves in a circular aperture while endpoints of the moving stimulus remain occluded. Such a motion stimulus closely resembles local motion encoding in receptive fields of V1 (Hubel & Wiesel, 1968) but disambiguating motion direction and speed may reflect characteristics of motion and disparity integration in area V5/MT and possibly beyond (DeAngelis & Newsome, 2004). Similar to the 2D aperture problem (Adelson & Movshon, 1982; Wallach, 1935) the 3D aperture problem requires that the visual system resolves motion correspondence but at the same time it needs to establish stereo correspondence between binocular receptive fields.

When an oriented line stimulus moves in depth at a given azimuth angle then local motion detectors tuned to different speeds may respond optimally to motion normal or perpendicular to the orientation of the line. If the intensity gradient or normal from the left and right eye serves as a default strategy, similar to the 2D aperture problem (Adelson & Movshon, 1982; Sung, Wojtach & Purves, 2009), then the resulting vectors in each eye may have different lengths. Inverse perspective projection of the retinal motion vectors reveals that monocular velocity constraint lines are usually skew so that an intersection of line constraints (IOC) does not exist. Since adaptive convergence of skew constraint lines is computationally expensive, it seems plausible that the visual system uses a different strategy to solve the aperture problem in 3D. The inverse problem will be discussed in detail in Chapter 2.
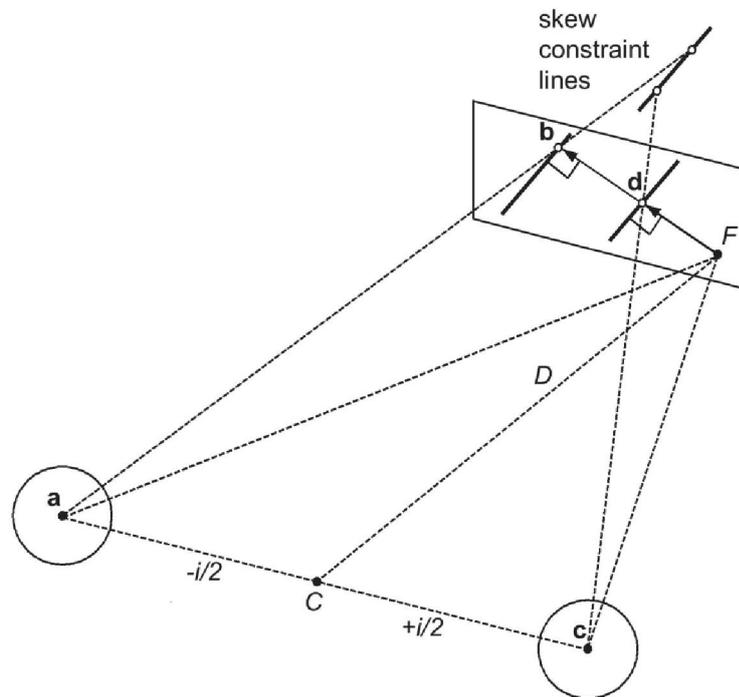
Figure 1.4 Illustration of the inverse problem for local binocular 3D motion perception. Note that left and right eye velocity constraints of a line derived from vector normals in 2D, depicted here on a common fronto-parallel screen rather than the left and right retina, do not necessarily intersect in 3D space. If the constraint lines are skew the inverse problem remains ill-posed.

In Chapter 3 we extend the geometric considerations of Chapter 2 on line stimuli moving in 3D space. Lines and contours have spatial extent and orientation reflecting properties of local encoding in receptive fields (Hubel & Wiesel, 1962; 1968; 1970). We suggest a generalized Bayesian model that provides velocity estimates for arbitrary azimuth and elevation angles. This model requires knowledge about eye positions in a binocular viewing geometry together with 2D intensity gradients to establish velocity constraint planes for each eye. The velocity constraints are combined with a 3D motion prior to estimate local 3D velocity. In the absence of 1D features such as points, corners, and T-junctions and without noise in the likelihoods, this approach approximates the shortest distance in 3D. This Bayesian approach is flexible

because additional constraints or cues from moving features can be integrated to further disambiguate motion direction of objects under uncertainty or ambiguity (Weiss et al., 2002).

These generalized motion models capture perceptual bias in binocular 3D motion perception and provide testable predictions in the context of the 3D aperture problem. In Chapter 4 we test specific predictions of line motion direction in psychophysical experiments. Chapter 5 we investigate some implications of late motion and disparity integration using neuro-imaging methods (fMRI). In Chapter 6 we provide a literature survey on stereo deficiencies and suggest that there are inter-individual differences in stereo and stereo-motion perception. In the final Chapter 7 we discuss future research directions and draw conclusions.

CHAPTER 2. INVERSE PROBLEM OF BINOCULAR 3D MOTION PERCEPTION

Abstract

It is shown that existing processing schemes of 3D motion perception such as interocular velocity difference, changing disparity over time, as well as joint encoding of motion and disparity do not offer a general solution to the inverse optics problem of local binocular 3D motion. Instead we suggest that local velocity constraints in combination with binocular disparity and other depth cues provide a more flexible framework for the solution of the inverse problem. In the context of the aperture problem we derive predictions from two plausible default strategies: (1) the vector normal prefers slow motion in 3D whereas (2) the cyclopean average is based on slow motion in 2D. Predicting perceived motion directions for ambiguous line motion provides an opportunity to distinguish between these strategies of 3D motion processing. Our theoretical results suggest that velocity constraints and disparity from feature tracking are needed to solve the inverse problem of 3D motion perception. It seems plausible that motion and disparity input is processed in parallel and integrated late in the visual processing hierarchy.

2.1 INTRODUCTION

The representation of the three-dimensional (3D) external world from two-dimensional (2D) retinal input is a fundamental problem that the visual system has to solve (Berkely, 1709/1965; von Helmholtz, 1910/1962; Poggio, Torre & Koch, 1985; Pizlo, 2001). This is true for static scenes in 3D as well as for dynamic events in 3D space. For the latter the inverse problem extends to the inference of dynamic events in a 3D world from 2D motion signals projected into the left and right eye. In the following we exclude observer movements and only consider passively observed motion.

Velocity in 3D space is described by motion direction and speed. Motion direction can be measured in terms of azimuth and elevation angle, and motion direction together with speed is conveniently expressed as a 3D motion vector in a Cartesian coordinate system. Estimating such a vector locally is highly desirable for a visual system because the representation of local estimates in a dense vector field provides the basis for the perception of 3D object motion- that is direction and speed of moving objects. This information is essential for interpreting events as well as planning and executing actions in a dynamic environment.

If a single moving point, corner or other unique feature serves as binocular input then intersection of constraint lines or triangulation together with a starting point provides a straightforward and unique geometrical solution to the inverse problem in a binocular viewing geometry. If, however, the moving stimulus has spatial extent, such as an edge, contour, or line inside a circular aperture (Morgan & Castet, 1997) then local motion direction in corresponding receptive fields of the left and right eye remains ambiguous and additional constraints are needed to solve the aperture and inverse problem in 3D.

2.2 FROM 2D TO 3D APERTURE PROBLEM

We investigate geometric constraints for velocity estimation in the context of the aperture problem. The 2D aperture problem arises when a line or edge moves in a circular aperture while endpoints of the moving stimulus remain occluded. As pointed out in Chapter 1 such a motion stimulus closely resembles local motion encoding in receptive fields of V1 (Hubel & Wiesel, 1968) but disambiguating motion direction and speed may involve motion and disparity integration in area hMT+/V5 and possibly beyond (DeAngelis & Newsome, 2004).

Lines and edges of various orientations are elementary for image processing (Marr, 1982). Local motion and disparity of a line, where endpoints are occluded behind a circular aperture, is highly ambiguous in terms of 3D motion direction and speed but it would be interesting to know how the visual system resolves this ambiguity and which constraints are employed to achieve estimates of local motion and global scene flow.

Consider, for example, a local feature with spatial extent such as an oriented line inside a circular aperture so that the endpoints of the line are occluded. Stereo correspondence between oriented lines or edges remains ambiguous (Morgan & Castet, 1997; van Ee & Schor, 2000). If a binocular observer maintains fixation at a close or moderate viewing distance then the oriented line stimulus projects differently onto the left and right retina (see Fig. 2.1 for an illustration with projections onto a single fronto-parallel screen). When an oriented line stimulus moves in depth at a given azimuth angle then local motion detectors tuned to different speeds may respond optimally to motion normal or perpendicular to the orientation of the line. If the intensity gradient or normal in 3D from the left and right eye serves as a default strategy, similar to the 2D aperture problem (Adelson & Movshon, 1982; Sung, Wojtach & Purves, 2009), then the resulting vectors in each eye may have approximately the same direction but different lengths. Inverse perspective projection of the retinal motion vectors through the nodal points of the left

and right eye reveals that monocular velocity constraint lines may be skew so that an intersection of line constraints (IOC) often does not exist.

Another violation occurs when the line is slanted in depth and projects with different orientations into the left and right eye. The resulting misalignment on the *y*-axis between motion vectors in the left and right eye is reminiscent of vertical disparity and the induced effect (Ogle, 1940; Banks & Backus, 1998). However, an initially small vertical disparity between motion gradients increases with motion in depth. The stereo system can extract depth from input with vertical disparity (Hinkle & Connor, 2002) and possibly orientation disparity (Greenwald & Knill, 2009) but it seems unlikely that the 3D motion system is based on combinations of motion detectors tuned to different orientations and speeds in the left and right eye. Since adaptive convergence of skew constraint lines is computationally expensive, it seems plausible that the visual system uses a different strategy to solve the aperture problem in 3D.
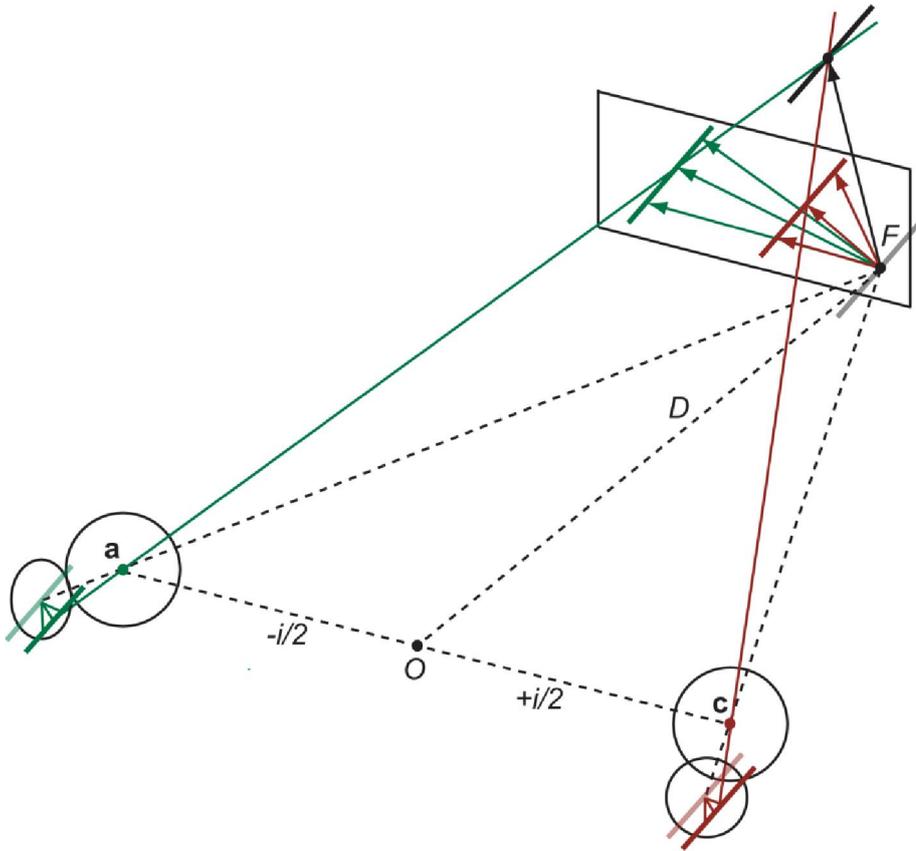
Figure 2.1 Illustration of the aperture problem of 3D motion with projections of an oriented line or contour moving in depth. The left and right eye with nodal points **a** and **c,** separated by interocular distance *i*, are verged on a fixation point *F* at viewing distance *D*. If an oriented stimulus (diagonal line) moves from the fixation point to a new position in depth along a known trajectory (black arrow) then perspective projection of the line stimulus onto local areas on the retinae or a fronto-parallel screen creates 2D aperture problems for the left and right eye (green and brown arrows).

The inverse optics and the aperture problem are well-known problems in computational vision, especially in the context of stereo (Poggio, Torre & Koch, 1985; Mayhew & Longuet-Higgins, 1982), structure from motion (Koenderink & van Doorn, 1991), and optic flow (Hildreth, 1984). Gradient constraint methods belong to the most widely used

techniques of optic-flow computation from image sequences. They can be divided into local area-based (Lucas & Kanade, 1981) and into more global optic flow methods (Horn & Schunck, 1981). Both techniques employ brightness constancy and smoothness constraints in the image to estimate velocity in an over-determined equation system. It is important to note that optical flow only provides a constraint in the direction of the image gradient, the normal component of the optical flow. As a consequence some form of regularization or smoothing is needed that can be computationally expensive. Similar techniques in terms of error minimization and regularization have been offered for 3D stereo-motion detection (Spies, Jahne & Barron, 2002; Min & Sohn, 2006; Scharr & Küsters, 2002). Essentially these algorithms extend processing principles of 2D optic flow to 3D scene flow but face similar problems.

Computational studies on 3D motion algorithms are usually concerned with fast and efficient encoding when tested against ground truth. Here we are less concerned with the efficiency or robustness of a particular implementation. Instead we want to understand and predict behavioral characteristics of human 3D motion perception. 2D motion perception has been extensively researched in the context of the 2D aperture problem (Wallach, 1935; Adelson & Movshon, 1982; Sung, Wojtach & Purves, 2009) but there is a surprising lack of studies on the aperture problem and 3D motion perception.

Three approaches to binocular 3D motion perception have emerged in the literature: Interocular velocity difference (IOVD), changing disparity over time (CDOT), and joint encoding of motion and disparity (JEMD).

(i) The *motion-first* model postulates monocular motion processing followed by stereo processing (Lu & Sperling, 1995; Regan & Beverley, 1973; Regan, et al., 1979). In this model monocular motion is independently detected in the left and right eye before interocular velocity difference (IOVD) establishes motion in depth.

(ii) The *stereo-first* model assumes disparity encoding followed by binocular motion processing (Cumming & Parker, 1994; Peng & Shi, 2010). This model first extracts binocular disparities and then computes change of disparity over time (CDOT). Note that tracking of spatial position is also required to recover a 3D motion trajectory.

(iii) Finally, the *stereo-motion* model suggests joint encoding of motion and disparity (JEMD) or binocular disparity and interocular delay (Carney, Paradiso, & Freeman, 1989; Morgan & Fahle, 2000; Qian, 1994; Qian & Andersen, 1997). In neurophysiological studies it was shown that a number of binocular complex cells in cats (Anzai, Ohzawa, & Freeman, 2001) and cells in V1 and MT of monkey (Pack, Born, & Livingstone, 2003) are tuned to interocular spatial-temporal shifts but the significance of these findings has been questioned (Read & Cumming, 2005a,b). Pulfrich like stimuli, in which the sensation of depth is produced through interocular delay, are often used as evidence in favour of joint encoding of motion and disparity. It is suggested that Pulfrich-like phenomena could only be encoded by a small number of direction selective disparity cells. This is often cited as evidence for joint encoding theories. However, Read & Cumming (2005a,b) show mathematically that the depth component of such displays can be encoded by pure disparity cells and the motion component by pure motion cells. In particular, they show that Pulfrich stimuli contain spatial disparities that can be used to derive depth, separately to the temporal integration process which underlies motion. They also state that physiological and not psychophysical studies should be used to investigate joint encoding, since there are no stimuli that completely cancel out motion or disparity information.

These three approaches have generated an extensive body of research but psychophysical results have been inconclusive and the nature of 3D motion processing remains an unresolved issue (Harris, Nefs & Grafton, 2008; Regan & Grey, 2009). Despite a wealth of empirical studies on motion in depth there is a lack of studies on true 3D motion stimuli. Previous psychophysical and neurophysiological studies typically employ stimulus dots with unambiguous motion direction or fronto-parallel random-dot surfaces moving in

depth. The aperture problem and local motion encoding however, which features so prominently in 2D motion perception (Wallach, 1935; Adelson & Movshon, 1982; Sung, Wojtach & Purves, 2009) has been neglected in the study of 3D motion perception.

The aim of this chapter is to evaluate existing models of 3D motion perception and to gain a better understanding of the underlying principles of binocular 3D motion perception. Following Lages and Heron (2010) we first show that existing models of 3D motion perception are insufficient to solve the inverse problem of binocular 3D motion. Second, we establish velocity constraints in a binocular viewing geometry and demonstrate that additional information is necessary to disambiguate local velocity constraints and to derive a velocity estimate. Third, we compare two default strategies of perceived 3D motion when local motion direction is ambiguous. It is shown that critical stimulus conditions exist that can help to determine whether 3D motion perception favors slow 3D motion or averaged cyclopean motion.

2.3 ANALYTIC GEOMETRY

In the following we give a general and intuitive overview of the mathematical concepts that are needed to build the models in Chapter 3 and that have been derived elsewhere (Lages & Heron, 2010; Appendix A2). Throughout we assume a fixed binocular viewing geometry with the cyclopean origin $O = (0,0,0)$ centered between the nodal points of the left and right eye and the eyes verged on a fixation point *F* straight ahead at viewing distance *D* (see Fig. 2.1). More complicated geometries arise if we take into account version, cyclovergence, and cyclotorsion of the eyes (Read, Phillipson & Glernnerster, 2009; Schreiber et al. 2008). For the sake of simplicity we ignore the non-linear aspects of visual space (Lüneburg, 1947) and represent perceived 3D motion as a linear vector in a three-dimensional Euclidean space where the fixation point is also the starting point of the motion stimulus.

Intersection of Constraint Lines

In the following we consider the simple case of projections onto a fronto-parallel screen in front of the eyes (rather than but equivalent to coplanar planes on the back of the eyes) at a fixed viewing distance D (see Fig. 2.2). In this simplified case epipolar lines (in epipolar geometry, this is defined as the intersection of the epipolar plane with the image plane) are horizontal with equivalent z-values $z_L = z_R = z_C$ on the screen.

It is obvious from the geometry that an intersection between the left and right eye constraint line exists only if they also have equivalent values on the y-axis of the screen

$$y_L - y_R = 0 \tag{2.1}$$

For an intersection to exist the left and right eye motion vector must have equivalent horizontal y co-ordinates or zero vertical disparity on the screen.

If the y co-ordinates do not correspond the constraint lines are skew and no intersection exists (see Fig. 2.2). This occurs, for example, when an oblique line moves on a horizontal trajectory to the left and in depth so that the projections into the left and right eye (red and green) have different horizontal velocity on the screen. The 2D gradient orthogonal to the moving line points in the same direction but has different lengths and as a consequence no intersection can be established.

If the eyes remain verged on a fixation point F in a binocular viewing geometry then the constraint line in the left and right eye can be defined by pairs of points $(\mathbf{a}, \mathbf{b})$ and $(\mathbf{c}, \mathbf{d})$, respectively. The nodal point in the left eye $\mathbf{a} = (-i/2, 0, 0)$ and a projection point $\mathbf{b} = (x_L, y_L, z_L)$ of the motion vector on the left retina define a constraint line for the left eye. Similarly, points $\mathbf{c} = (+i/2, 0, 0)$ and $\mathbf{d} = (x_R, y_R, z_R)$ determine a constraint line in the right eye. The definition of vectors and operations are derived in Appendix A2 (Lages & Heron, 2010).

24

Intersection of Constraint Lines

If the eyes remain verged on a fixation point in a binocular viewing geometry then the constraint line in the left and right eye can be defined by pairs of points $(\mathbf{a},\mathbf{b})$ and $(\mathbf{c},\mathbf{d})$, respectively (Fig 2.1). The nodal point in the left eye $\mathbf{a} = (-i/2,0,0)$ and a projection point $\mathbf{b} = (x_L, y_L, z_L)$ of the motion vector on the fronto-parallel screen (simulating the left retina) define a constraint line for the left eye. Similarly, points $\mathbf{c} = (+i/2,0,0)$ and $\mathbf{d} = (x_R, y_R, z_R)$ determine a constraint line for the right eye. If an intersection exists it can be determined by triangulation and the corresponding vector operations (see Appendix A2).

Intersection of Constraint Planes

Monocular line motion gives rise to a constraint plane defined by three points: the nodal point of an eye and two points defining the position of the line projected on a screen at a given time. These are illustrated as shaded green and brown triangles in Fig. 2.3 for the left and right eye, respectively. If the planes are not parallel the two constraint planes intersect in 3D. This is illustrated by the oriented black line (IOC) in Fig. 2.3. The intersection coincides with the position of the moving line at a given time point.

In order to find the intersection of the left and right eye constraint plane we use the plane normal in the left and right eye. The computation of the two constraint planes and their intersection is detailed in Appendix A2.

Vector Normal (VN)

The shortest distance in 3-D (velocity) space between a starting point $\mathbf{p}_0 = (0,0,D)$ of the stimulus line and the constraint line $\mathbf{p}$ is the line or vector normal through point $\mathbf{p}_0$. In order to determine the intersection point of the vector normal with the constraint line we

pick two arbitrary points $\mathbf{p}_1$ and $\mathbf{p}_2$ on intersection constraint line $\mathbf{p}$ by choosing a scalar u (e.g., 0.5).

Cyclopean Average (CA)

We can define a cyclopean constraint line in terms of the cyclopean origin $O = (0,0,0)$ and projection point $\mathbf{p}_C = (x_C, y_C, z_C)$ on the fronto-parallel screen where $x_C = (x_L + x_R)/2$ and $y_C = (y_L + y_R)/2$ are the averages of the 2D vector normal co-ordinates for the left and right eye projections.

If we measure disparity δ at the same retinal coordinates as the horizontal offset between the left and right eye anchored at position $\mathbf{p}_C$ then we can define new points $\mathbf{b}$ with $x'_L = x_C - \delta/2$ and $\mathbf{d}$ with $x'_R = x_C + \delta/2$. (Alternatively, we may establish an epipolar or more sophisticated disparity constraint.) The resulting two points together with the corresponding nodal points $\mathbf{a}$ and $\mathbf{c}$ define two constraint lines, one for the left and the other for the right eye. By inserting the new co-ordinates we can then find the intersection of constraint lines. The intersection and start point determine the perceived trajectory.

2.4 APPLICATION OF THE GEOMETRIC RESULTS

In the following we summarize shortcomings for each of the three main approaches to binocular 3D motion perception in terms of stereo and motion correspondence, 3D motion direction, and speed. We also provide a counterexample to illustrate the limitations of each approach.

Interocular velocity difference (IOVD)

This influential processing model assumes that monocular spatio-temporal differentiation or motion detection (Adelson & Bergen, 1985) is followed by a difference computation between velocities in the left and right eye (Beverley & Regan, 1973; 1975; Regan &

Beverley, 1973). The difference or ratio between monocular motion vectors in each eye, usually in a viewing geometry where interocular separation *i* and viewing distance *D* is known, provides an estimate of motion direction in terms of azimuth angle only.

We argue that the standard IOVD model (Welchman, Lam & Bulthoff, 2008; Brooks, 2002; Shioiri, Saisho, Yaguchi, 2000; Fernandez & Farell, 2005; Rokers, Cormack & Huk, 2008) is incomplete and ill-posed if we consider local motion encoding and the aperture problem. In the following the limitations of the IOVD model are illustrated.

Stereo Correspondence. The first limitation is easily overlooked: IOVD assumes stereo correspondence between motion in the left and right eye when estimating 3D motion trajectory. The model does not specify which motion vector in the left eye should correspond to which motion vector in the right eye before computing a velocity difference. If there is only a single motion vector in the left and right eye then establishing a stereo correspondence appears trivial since there are only two positions in the left and right eye that signal dynamic information. Nevertheless, stereo correspondence is a necessary pre-requisite of IOVD processing which quickly becomes challenging if we consider multiple stimuli that excite not only one but many local motion detectors in the left and right eye. It is concluded that without explicit stereo correspondence between local motion detectors the IOVD model is incomplete.

3D Motion Direction. The second problem concerns 3D motion trajectories with arbitrary azimuth and elevation angles. Consider a local contour with spatial extent such as an oriented line inside a circular aperture so that the endpoints of the line are occluded. This is known as the aperture problem in stereopsis (Morgan & Castet, 1997; van Ee & Schor, 2000). If an observer maintains fixation at close or moderate viewing distance then the oriented line stimulus projects differently onto the left and right retina (see Fig. 2.2 for an illustration with projections onto a single fronto-parallel plane).

27
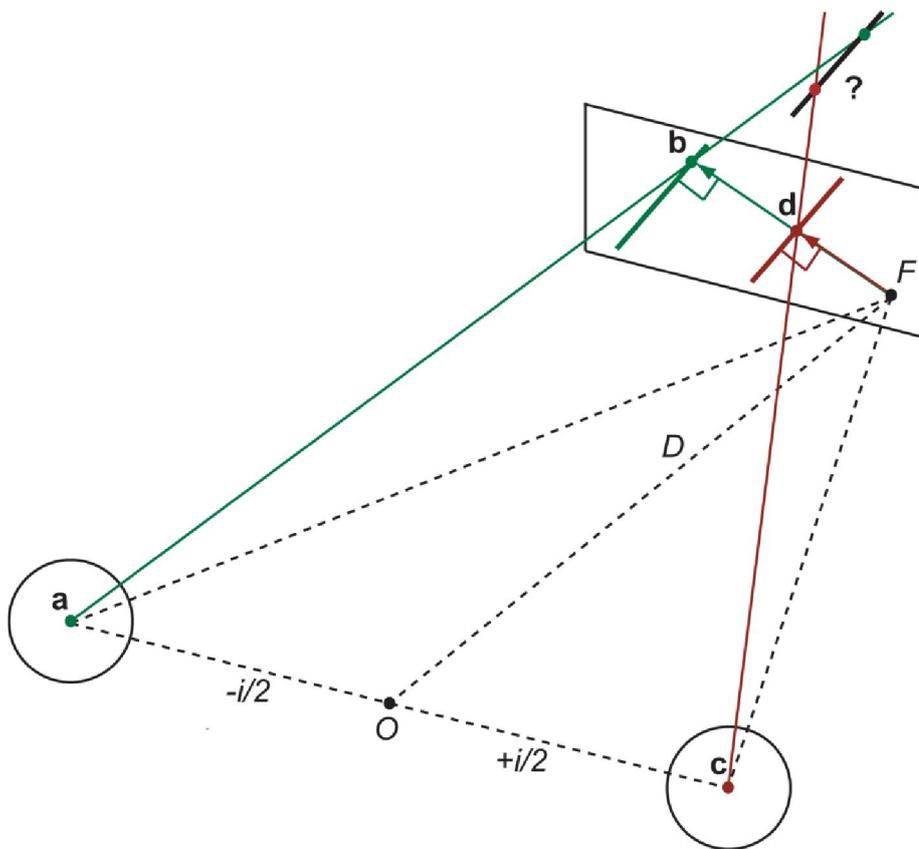
Figure 2.2 Inverse projection of constraint lines preferring slow 2D motion in the left and right eye. Constraint lines (shown in green and red) through projection point **b** and **d** do not intersect at a single point on the 3D motion constraint line for the oriented line stimulus (shown here as a black line). This line represents a range of plausible motion directions in 3D and, as shown the red and green constraint lines do not converge on a single point along this line. So there is no unique intersection of constraints (IOC) solution in 3D and therefore 3D motion cannot be determined (see text for details).

When the oriented line moves horizontally in depth at a given azimuth angle then local motion detectors tuned to different speeds respond optimally to motion normal (perpendicular) to the orientation of the line. If the normal in the left and right eye serves as a default strategy for the aperture problem in 2D (Wallach, 1935; Sung, Wojtach &

Purves, 2009) then these vectors may have different lengths (as well as orientations if the line or edge is oriented in depth). Inverse perspective projection of the retinal motion vectors reveals that the velocity constraint lines are skew and an intersection of line constraints (IOC) does not exist. In fact, an intersection only exists if the following constraint for the motion vector in the left and right eye holds (see Appendix A2):

$$\frac{y_L}{y_R} - \frac{z_L}{z_R} = 0 \, . \tag{2.2}$$

(If the image planes are fronto-parallel so that $z_L = z_R$ then the condition is simply $y_L - y_R = 0$). However, this constraint is easily violated as illustrated in Fig. 2.2 and Counterexample 1 below.

Speed. It is worth pointing out that IOVD offers no true estimate of 3D speed. This is surprising because the model is based on spatial-temporal or speed-tuned motion detectors. The problem arises because computing motion trajectory without a constraint in depth does not solve the inverse problem. As a consequence speed is typically approximated by motion in depth along the line of sight (Brooks, 2002).

Counterexample 1. If an edge or line tilted from horizontal by $0 < \theta < 90°$ moves in depth at a fixed azimuth angle so that horizontal translations of the projected images into the left and right eye are unequal $h_L \neq h_R$, it follows from basic trigonometry that the local motion vectors normal to the oriented line have *y*-co-ordinates $y_L = h_L (\sin \theta)^2$ and $y_R = h_R (\sin \theta)^2$, thus $y_L \neq y_R$ (see Fig. 2.2 and Appendix A2).

Another violation occurs when the line is slanted in depth and projects with different orientations into the left and right eye. The resulting misalignment on the *y*-axis between motion vectors in the left and right eye is reminiscent of vertical disparity and the induced effect (Ogle, 1940; Banks & Backus, 1998) with vertical disparity increasing over time. The stereo system can reconstruct depth from input with orientation disparity and even

vertical disparity (Hinkle & Connor, 2002) but it seems unlikely that the binocular motion system can establish similar stereo correspondences.

It is concluded that the IOVD model is incomplete and easily leads to ill-posed inverse problems. These limitations are difficult to resolve within a motion processing system and point to contributions from disparity or depth processing.

Changing disparity over time (CDOT)

This alternative processing scheme uses disparity input and monitors changing disparity over time (CDOT). Disparity between the left and right image is detected (Ohzawa, De Angelis & Freeman, 1990) and changes over time give rise to motion-in-depth perception (Cumming & Parker, 1994; Beverley & Regan, 1974; Julesz, 1971; Peng & Shi, 2010). We argue that this approach also has limitations when the inverse problem of local 3D motion is considered.

Motion correspondence. Assuming CDOT can always establish a suitable stereo correspondence between features including lines (Morgan & Castet, 1997; Ogle, 1940) then the model still needs to resolve the motion correspondence problem. It needs to integrate disparity not only over time but also over 3D position to establish a 3D motion trajectory. Although this may be possible for a global feature tracking system it is unclear how CDOT arrives at estimates of local 3D motion.

3D Motion Direction. Detecting local disparity change alone is insufficient to determine an arbitrary 3D trajectory. CDOT has difficulties to recover arbitrary 3D motion direction because only motion-in-depth along the line of sight is well defined. 3D motion direction in terms of arbitrary azimuth and elevation requires a later global mechanism that has to solve the inverse problem by tracking not only disparity over time but also position in 3D space over time.

Speed. As a consequence the rate of change of disparity provides a speed estimate for motion-in-depth along the line of sight but not for arbitrary 3D motion trajectories.

Counterexample 2. In the context of local surface motion consider a horizontally slanted surface moving to the left or right behind a circular aperture. Without corners or other unique features CDOT can only detect local motion in depth along the line of sight. Similarly in the context of local line motion, the inverse problem remains ill posed for a local edge or line moving on a slanted surface because additional motion constraints are needed to determine a 3D motion direction.

In summary, CDOT does not provide a general solution to the inverse problem of local 3D motion because it lacks information on motion direction. Even though CDOT is capable of extracting stereo correspondences over time, additional motion constraints are needed to represent arbitrary motion trajectories in 3D space.

Joint encoding of motion and disparity (JEMD)

This approach postulates that early binocular cells are both motion and disparity selective and physiological evidence for the existence of such cells was found in cat striate cortex (Anzai, Ohzawa & Freeman, 2001) and monkey V1 (Pack, Born & Livingstone, 2003; see however Read & Cumming, 2005). Model cells in this hybrid approach extract motion and disparity energy from local stimulation. A read-out from population activity and population decoding is needed to explain global 3D motion phenomena such as transparent motion and Pulfrich-like effects (Qian, 1997; Qian & Andersen, 1997). Although JEMD is physiologically plausible it shares two problems with IOVD.

3D Motion Direction. Similar to cells tuned to binocular motion, model cells of JEMD prefer corresponding velocities in the left and right eye. Therefore a binocular model cell can only establish a 2D fronto-parallel velocity constraint at a given depth. Model cell

activity remains ambiguous because it can be the result of local disparity or motion input (Lages, Dolia & Graf, 2007). A later processing stage, possibly at the level of human V5/MT (DeAngelis & Newsome, 2004) needs to read out population cell activities across positions and depth planes and has to approximate global 3D motion. Similar to CDOT, the model defers the inverse problem to a later global processing stage.

Speed. Again, similar to IOVD and CDOT, JEMD provides no local 3D speed estimate. It also has to rely on sampling across depth planes in a population of cells in order to approximate speed.

Counterexample 3. Consider local 3D motion with unequal velocities in the left and right eye but the same average velocity, e.g. diagonal trajectories to the front and back through the same point in depth. JEMD has no mechanism to discriminate between these local 3D trajectories when monitoring binocular cell activity across depth planes in a given temporal window.

In the following we introduce general velocity constraints for 3D motion and suggest two default strategies of 3D motion perception that are based on different processing principles.

Velocity Constraints and Two Default Strategies
Which constraints does the visual system use to solve the inverse as well as aperture problem for local 3D line motion where endpoints are invisible or occluded? This is a critical question because it is linked to local motion encoding and the possible contribution from depth processing.

The 3D motion system may establish constraint planes rather than constraint lines to capture all possible motion directions of a contour or edge, including motion in the direction of the edge's orientation. Geometrically the intersection of two constraint

planes in a given binocular viewing geometry defines a constraint line oriented in 3D velocity space (see Fig. 2.3).
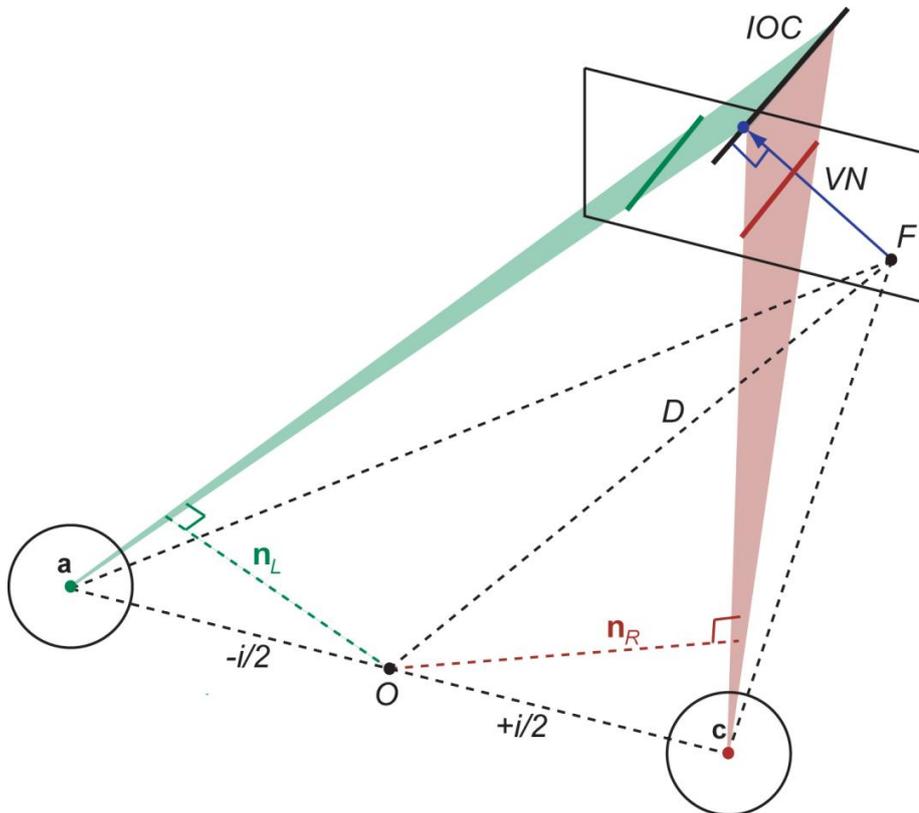
Figure 2.3 Illustration of vector normal (VN) as a default strategy for local 3D motion perception (see text for details). Rather than constraint lines, here we define constraint planes for the monocular velocities. Monocular inputs therefore remain ambiguous in that motion direction has a range of possibilities (represented by the red and green lines). When two monocular constraint planes intersect, they define an intersection of constraint planes, a line, rather than a single point in 3D space. This is represented by the black line which also indicates the end-points for the motion of an oblique line in 3D space. The intersection of constraint planes (IOC) together with the assumption of slow motion determines the shortest vector in 3D space (blue arrow) that fulfills the velocity constraints.

We suggest that in analogy to 2D motion perception (Adelson & Movshon, 1982; Weiss, Simoncelli & Adelson, 2002) tracking of features in depth coupled with binocular velocity constraints from motion processing provides a flexible strategy to disambiguate 3D motion direction and to solve the inverse problem of 3D motion perception.

But which principles or constraints are used? Does the binocular motion system prefer slow 3D motion or averaged 2D motion? Does it solve stereo correspondence before establishing binocular velocity constraints or does it average 2D velocity constraints from the left and right eye before it solves stereo correspondence? We derive predictions for two alternative strategies to address these questions.

Vector Normal (VN). Velocity constraints in the left and right eye provide velocity constraint planes in 3D velocity space. In Fig. 2.3 they are illustrated as translucent green and brown triangles in a binocular viewing geometry. The intersection of constraint planes defines a velocity constraint line in 3D that also describes the true end-position of the moving line or contour (black line). The vector or line normal from the oriented constraint line to the starting point gives a default 3D motion estimate (blue arrow). It is the shortest distance in 3D velocity space and denotes the slowest motion vector that fulfills both constraints. Note that this strategy requires that the 3D motion system has established some stereo correspondence so that the intersection of constraints as well as the vector normal can be found in 3D velocity space.

The VN strategy is a generalization of the vector normal and IOC in 2D (Adelson & Movshon, 1982) and it is related to area-based regression and gradient constraint models (Lucas & Kanade, 1981) where the local brightness constancy constraint ensures a default solution that is normal to the orientation of image intensity.

Cyclopean Average (CA). If the motion system computes normal 2D motion vectors independently in the left and right eye then the cyclopean average provides an alternative velocity constraint (Harris & Drga, 2005; Harris & Rushton, 2003). In this case the monocular constraint planes, shown as translucent green and brown triangles in figure

34

2.4, are averaged to define a single constraint line. There is therefore a unique point of intersection or IOC. Averaging of monocular constraints increases robustness of the motion signal at the expense of binocular disparity information. Thus, a cyclopean average constrains velocity but gives no default estimate of velocity. However, if we attach (dynamic) disparity to the cyclopean average then the resulting CA strategy provides a default estimate of 3D velocity (see Fig. 2.4).

Figure 2.4 Illustration of Cyclopean Average (CA) as a default strategy for local 3D motion perception (see text for details). Monocular velocities are defined as constraint planes (shown here as translucent green and brown triangles), but unlike the VN model, a unique intersection point is derived at a 2D level by averaging early motion inputs to give a single constraint line extending from the mid-point O, through the vector average IOC and into

3D space. Combining the cyclopean velocity constraint (averaged left and right velocities) with horizontal disparity determines a vector in 3D space (red arrow). The end-point of this vector lies along the intersection of constraint planes (shown as a black line).

The CA strategy is a generalized version of the vector average strategy for 2D motion (Wilson, Ferrera & Yo, 1992) and can be linked to computational models of 3D motion that use global gradient and smoothness constraints (Horn & Schunck, 1981). These global models amount to computing the average flow vector in the neighborhood of each point and refining the scene flow vector by the residual of the average flow vectors in the neighborhood. Interestingly, tracking the two intersection points or T junctions of a moving line with a circular aperture in the left and right eye and averaging the resulting vectors gives predictions that are equivalent to the CA strategy.

Predictions for VN and CA Strategy. We use the Vector Normal (VN) and Cyclopean Average (CA) as default strategies to predict 3D velocity of an oriented line or contour moving in depth inside a circular aperture.

The 3D plot in Fig. 2.5 shows predictions of the VN strategy (blue) and the CA strategy (red) for a diagonal line stimulus moving on two trajectories in depth at a viewing distance D=57 cm and interocular distance of i=6.5 cm. The line stimulus has a trajectory to the front and left with azimuth +57.2 deg and elevation 0 deg, and a trajectory to the back and left with azimuth -57.2 deg and elevation 0 deg. Azimuth and elevation of 0 deg denotes a horizontal and fronto-parallel trajectory to the left. The starting point of each trajectory is the origin of the vector fields in the 3D plot. An open circle denotes the endpoint of a predicted motion vector. For each default strategy and stimulus trajectory a field of 120 vectors are shown with orientation disparity of the line stimulus ranging from -6° to +6° in steps of 0.1°. Orientation disparity changes perceived slant of the diagonal

36

line so that at -6° the bottom-half of the line is slanted away from the observer and the top-half is slanted towards the observer.

If the diagonal line is fronto-parallel and has zero orientation disparity both strategies make equivalent predictions (intersection of red and blue vector fields in Fig. 2.5). If, however, the stimulus line has orientation disparity and is slanted in depth then predictions clearly discriminate between the two strategies. The VN strategy always finds the shortest vector between starting point and moving line so that velocity predictions approximate a semi-circle for changing orientation disparity. Please note that for the VN predictions the sign of orientation disparity reverses for the stimulus trajectory to the front and back. The CA strategy on the other hand computes an average vector and as a consequence the endpoints of the predictions approximate a velocity constraint line through the cyclopean origin.
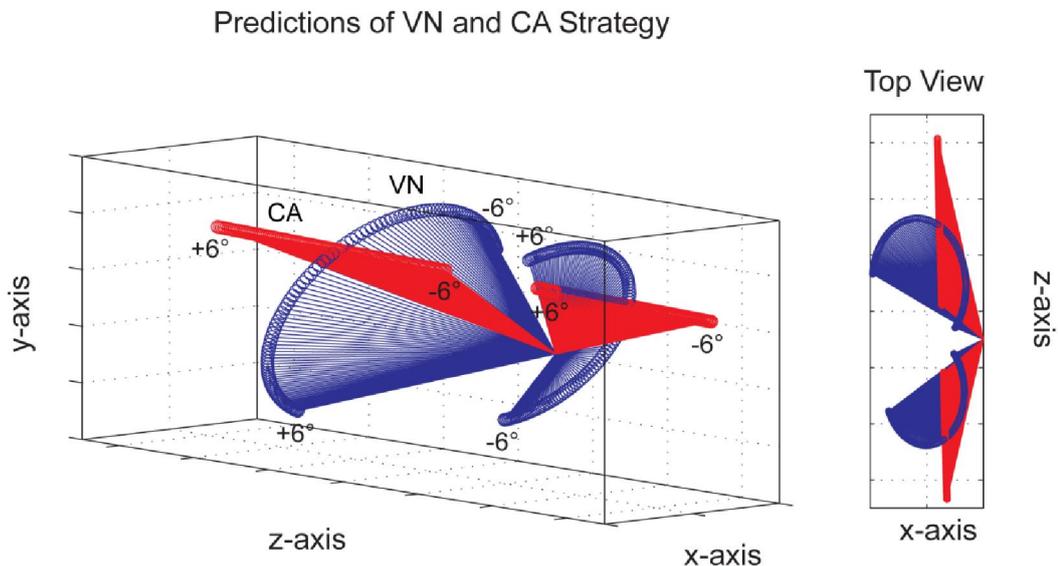
Figure 2.5 Velocity predictions of Vector Normal (VN, blue) and Cyclopean Average (CA, red) as default strategies of perception of local 3D line motion. Predictions are shown for an oblique stimulus line moving on a fixed trajectory to the front left and to the back left.

Predicted velocities show characteristic differences when the moving stimulus line or contour is slanted in depth (orientation disparities range between -6° to +6°).

In Chapter 3 we introduce probabilistic models of 3D motion perception and in Chapter 4 we report on a series of psychophysical experiments. We used a psychophysical matching task to measured perceived 3D motion direction of an oriented line moving behind a circular aperture.

2.5 DISCUSSION

If motion-only or disparity-only input determines 3D motion perception as suggested by IOVD and CDOT then processing of other input needs to be silenced or disengaged. This requires top-down interference rather than bottom-up processing by the visual system. Instead, we suggest that the visual system takes advantage of motion and disparity input (Bradshaw & Cumming, 1997; Lages & Heron, 2008) as well as additional cues. We favour parallel over serial processing and late integration over early joint encoding (Lages, Dolia, & Graf, 2007). The inverse problem of local 3D motion remains ill-posed for joint early encoding and JEMD needs a population read-out at some later stage to approximate 3D motion.

Combining disparity or depth information with velocity constraints at a later stage solves the inverse problem of local 3D motion and provides a flexible scheme that can exploit intermediate depth processing such as relative and orientation disparity in V2 and V4 (Hinkle & Connor, 2002; Thomas, Cumming, & Parker, 2002). Velocity constraints may be processed in the dorsal stream and binocular disparity together with other depth cues in the ventral stream (Neri, 2005; Ponce et al., 2008). Recent physiological evidence suggests that the ventral stream supports the processing of relative disparity, whilst the dorsal stream is involved in processing absolute disparity (See Neri, 2005). It seems anatomically and neurophysiologically plausible that integration of motion and disparity occurs late in

38

subregions of human MT+/V5 (DeAngelis & Newsone, 2004; Orban, 2008; Majaj, Carandini, & Movshon, 2007; Rokers et al., 2009) if not in areas beyond hMT+/V5 (Likova & Tyler, 2007).

What enables the visual system to instantaneously perceive 3D motion and to infer direction and speed of a moving object? It seems likely that the visual system exploits many cues in concert to make this difficult inference as reliable and veridical as possible. The diverse set of effective local and global cues demonstrated in psychophysical studies (Bradshaw & Cumming, 1997; van Ee & Anderson, 2001) points at late integration within the visual processing hierarchy.

More specifically, we suggest that binocular 3D motion perception is based on parallel streams of motion and disparity processing. Thereby motion processing captures coarse spatio-temporal constraints in the scene whereas disparity processing provides a fine and frequently updated depth map that helps to disambiguate motion direction and to maintain a detailed spatial representation of the scene. Late integration of local motion and disparity constraints in combination with other cues solves the inverse problem of local 3D motion and allows the visual system to remain flexible when binding and segmenting local inputs from different processing stages into a global 3D motion percept. Parallel processing and late integration may explain why, compared to 2D motion perception, 3D motion perception shows reduced spatio-temporal tuning characteristics (Lages, Mamassian, & Graf, 2003; Tyler, 1971) and why motion perception can retain relatively fine spatial detail at slow speeds. The combination of local motion constraints with a more global dynamic depth map from higher-order features may even explain the perception of different types of non-linear motion, such as non-rigid and 2nd order motion (Ledgeway & Smith, 1994; Lu & Sperling, 2001).

The notion of parallel pathways feeding functionally different aspects of motion perception into a later stage is not new and has been advanced in the context of 2D speed

perception (Braddick, 1974; 1980), 2D pattern motion (Adelson & Movshon, 1982; Weiss et al., 2002; Wilson, Ferrera, Yo, 1992), eye movements (Rashbass & Westheimer, 1961; Masson & Castet, 2002), and the processing of higher order motion (Ledgeway & Smith, 1994; Lu & Sperling, 2001). Surprisingly however, it has not sufficiently been addressed in the context of binocular 3D motion perception (Lu & Sperling, 2001; Regan, Beverley, Cynader, & Lennie, 1979).

Considering the ill-posed inverse problem of existing approaches and the underdetermined characteristics of local binocular motion constraints, parallel processing and late integration of motion and disparity as well as other cues appears promising. Solving the inverse problem for local 3D motion adds a functional significant aspect to the notion of parallel streams of dynamic disparity and motion processing.
It will require considerable efforts to unravel the entire process but geometric-probabilistic models can achieve motion and disparity integration under uncertainty and ambiguity and will be described in Chapter 3.

CHAPTER 3. PROBABILISTIC 3D MOTION MODELS

Abstract

We first discuss existing Bayesian models of 3D motion perception before we provide a probabilistic extension for the geometric considerations on 3D line motion in Chapter 2. Lines and contours have spatial extent and orientation reflecting properties of local encoding in receptive fields (Hubel & Wiesel, 1962; 1968; 1970). We suggest a generalized Bayesian model that provides velocity estimates for arbitrary azimuth and elevation angles. Again, this model requires knowledge about eye positions in a binocular viewing geometry together with 2D intensity gradients to establish velocity constraint planes for each eye. The velocity constraints are then combined with a 3D motion prior to estimate local 3D velocity. In the absence of 1D features such as points, corners, and T-junctions and with little noise in the likelihoods, this approach approximates the VN strategy and provides a dense array of local velocity estimates. A dense array of local estimates is a desirable feature of any 3D motion model because it models scene flow. The present geometric-statistical approach is flexible because additional constraints or cues from moving features can be integrated to further disambiguate motion direction of objects under uncertainty or ambiguity (Weiss et al., 2002).

3.1 INTRODUCTION

Perception may be understood as a form of statistical inference (Knill, Kersten & Yuille, 1996). The Bayesian framework provides an optimal way of combining the uncertain information extracted from images on the retina with prior assumptions about the nature of objects in the world. This approach has been successful when modeling human visual performance in a range of tasks (Mamassian, Landy & Maloney, 2002).

In 2D motion perception for example, perceived direction and speed of a stimulus can be inferred from noisy velocity constraints that are combined with a prior assumption for slow speed (Weiss et al., 2002). Similarly, it has been suggested that horizontal disparity processing may be exposed to a zero disparity prior under uncertainty (Prince& Eagle, 2000; Read, 2002).

Large and persistent perceptual bias has been found for dot stimuli with unambiguous motion direction (Harris & Drga, 2005; Lages, 2006; Welchman, Lam & Bülthoff, 2008) suggesting processing strategies that are different from IOVD, CDOT and JEMD (Lages, 2006; Welchman, Lam & Bülthoff, 2008; Ji & Fermüller, 2006). It seems promising to investigate local motion stimuli with ambiguous motion direction such as a line or contour moving inside a circular aperture (Heron & Lages, 2009) because they relate to local encoding (Hubel & Wiesel, 1962; 1968; DeAngelis, Ohzawa & Freeman, 1993; Maunsell & van Essen, 1983; Hubel & Wiesel, 1970; Anzai, Ohzawa & Freeman, 2001; Bradley, Qian & Andersen, 1995; DeAngelis & Newsome, 1999) and may reveal principles of 3D motion processing (Lages & Heron, 2008; 2009; 2010).

The purpose of this chapter is to introduce Bayesian models of binocular 3D motion perception because they can provide local estimates of 3D velocity even under uncertainty and ambiguity. We will focus on local constraints of motion perception using

basic features such as single dots, lines or edges moving in 3D space because these features are regarded as primitives of local encoding in the early stages of visual processing.

3.2 BINOCULAR MOTION PERCEPTION UNDER UNCERTAINTY

The binocular viewing geometry imposes obvious constraints for stimulus trajectory and velocity. For a dot moving on a horizontal plane for example the intersection of constraint lines in x-z space determines trajectory angle and speed of the target moving in depth as illustrated in Fig. 3.1.

Figure 3.1. Binocular viewing geometry in top view. If the two eyes are verged on a fixation point at viewing distance D with angle β then projections of a moving target (arrow) with angle α_L in the left eye and α_R in the right eye constrain motion of the target

in *x-z* space. The intersection of constraints (IOC) determines stimulus trajectory β and radius *r*. A Gaussian motion prior is indicated by the gray circle around the fixation point that is also the start point of motion.

If we assume that the eyes remain accommodated and verged at angle β_0 on a fixation point *F* straight ahead then motion information is projected onto the retina of the left and right eye as illustrated in Fig. 3.1. The projection angles onto the retinae depend on the azimuth β and speed r of the motion stimulus, as well as viewing distance *D* and interpupillary distance *i*. The average of the left and right projection angle approximates the visual angle α in cyclopean view through point *C* and the difference defines binocular horizontal disparity δ. If the projection angles are interpreted as angular velocities then their difference describes interocular velocity difference or changing disparity over time.

Motion and disparity constraint lines intersect at the same point in space, regardless of whether they are based on the computation of angular velocities or binocular disparities. Although binocular motion and changing disparity input share the same geometry and are therefore mathematically equivalent, different neural encoding and processing of these inputs are subject to noise, resulting in characteristic perceptual biases. If noise or uncertainty is introduced together with a motion or disparity prior then the intersection of constraints lines generates different predictions. As an illustration of the Bayesian approach we will show in detail how to derive characteristics and how to model data as reported in a previously published paper (Lages, 2006). However, at the end we add new results on Bayesian model selection.

3.2.1 Bayesian Models of 3D Motion Perception

Some promising Bayesian models have been developed in vision (Knill & Richards, 1996). Bayesian inference is based upon a simple formula known as Bayes' rule. Assume that we

have an agent who is attempting to infer a scene *S* in the world for generating a perceived image *I*. Let *S* be a possible scene (hypothesis) and *p(S)* indicate the probability that the agent would have ascribed to *S* being the true scene, before seeing *I*. This is known as the prior probability. How should the agent go about changing his belief in the light of the evidence provided by *I*? To answer this question we need a procedure for computing the posterior probability *p(S|I)*. Following Bayes's Rule the likelihoods *p(I|S)* and prior *p(S)* of a scene S and image I are combined to produce a posterior $p(S|I) = \frac{p(I|S)p(S)}{p(I)}$.

The denominator *p(I)* is obtained by summing over possible scenes, a procedure known as marginalization $p(I) = \sum_i p(I|S_i)p(S_i)$. The posterior probability is proportional to the product of the prior and likelihood. The sum in the denominator simply ensures that the resulting probabilities are normalized to 1. A non-normalized posterior is obtained by dropping the denominator

$$p(S|I) \propto p(I|S)p(S)$$

(3.1)

Various quantities given in the images, such as motion and disparity, can be used to infer aspects of a scene. Weiss, Simoncelli, and Adelson (2002), for example, combined motion constraints of local motion detectors with a Gaussian prior for slow motion to predict perceived motion direction and velocity of luminance-defined objects in 2D space. With this elegant approach they could explain a range of 2D motion illusions.

Most objects in natural scenes are stationary. If we assume that objects tend to move slowly on an arbitrary trajectory in *x-z* space then a bivariate Gaussian probability distribution centered on the starting point of a stimulus provides a plausible prior for 3D motion perception in *x-z* space. Symmetric perspective projections of this world prior into the left and right eye give rise to marginal Gaussian distributions defining motion priors centered on zero velocity. Similarly, the difference of the marginal distributions in the left and right eye defines a prior for disparity (change) centered on zero disparity. Thus, the

same 3D motion prior in the world results in a Gaussian velocity and disparity prior on the retinae.

There are several potential sources for uncertainty and noise in binocular motion processing. For example, local moving targets in a sparse 3D environment offer limited motion and disparity input and other depth cues thereby introducing different degrees of uncertainty in the observer. Mini-saccades during fixation, or early noise in the encoding system are possible sources of uncertainty (Hogervorst & Eagle, 1998). In the following we extend the motion-first (IOVD) and stereo-first (CDOT) processing models to probabilistic models by adding Gaussian noise to the input and postulating a plausible prior for each processing scheme.

Bayesian Motion-first Model (BIOVD)

First assume that noise is present in the activation of monocular motion detectors optimally tuned to velocities in the left and right eye. The representation of angular velocity in each eye is therefore not exact but subject to noise (Ascher & Grzywacz, 2000). The corresponding likelihood distributions for angular velocity in the left and right eye are conveniently expressed as Gaussian distributions with equal variance centered on the true angular velocity of the stimulus in each eye. Each likelihood distribution is then combined with the motion prior. Motion priors favoring slow motion have been suggested in the context of 2D motion (Ascher & Yuille, 2000; Ullman & Yuille, 1989; Weiss et al., 2002).

In this framework perceived angular velocity of motion-first processing may be described as a product of likelihood and prior for the left and right eye

$$p(v_L \mid \beta) \propto p(\beta \mid v_L)p(v_L); \quad p(v_R \mid \beta) \propto p(\beta \mid v_R)p(v_R) \tag{3.2}$$

using the same prior $p(v_L) = p(v_R) = p(v)$.

The likelihood for the left eye is modeled as a Gaussian distribution of angular velocities centered on the true angular velocity with $d\alpha_L(t)/dt$ abbreviated as $\dot{\alpha}_L$. The standard deviation σ_v of the likelihood distribution is left as a free parameter.

$$p(\beta \mid v_L; \sigma_v) = \frac{1}{\sqrt{2\pi}\sigma_v} \exp\left(\frac{-\left(v_L - \dot{\alpha}_L\right)^2}{2\sigma_v^2}\right)$$

(3.3)

The likelihood for the right eye is modeled accordingly. The prior for slow motion is a Gaussian distribution centered on zero velocity with unknown but fixed standard deviation σ.

$$p(v; \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-\left(v\right)^2}{2\sigma^2}\right)$$

(3.4)

The product of the Gaussian likelihood distribution with a normal (conjugate) prior $\mathcal{N}(0, \sigma)$ defines a posterior distribution, that is the probability of each possible angular velocity taking into account both prior and likelihood of the trajectory is also normally distributed. Through differentiation the *maximum a posteriori* (MAP) estimates of angular velocity are found for the left eye and right eye, respectively. These are the points of the posterior distributions with maximal probability density.

$$\hat{\alpha}_L = \frac{1/\sigma_v^2 \, \dot{\alpha}_L}{1/\sigma_v^2 + 1/\sigma^2} = \frac{\dot{\alpha}_L}{1 + \left(\sigma_v/\sigma\right)^2}; \quad \hat{\alpha}_R = \frac{1/\sigma_v^2 \, \dot{\alpha}_R}{1/\sigma_v^2 + 1/\sigma^2} = \frac{\dot{\alpha}_R}{1 + \left(\sigma_v/\sigma\right)^2}$$

(3.5)

The noise ratio between likelihood and prior standard deviations σ_v/σ. If σ is fixed to 1 then σ_v is the only free parameter in this model (Hürlimann, Kiper & Carandini, 2002).

Bayesian Disparity-first Model (BCDOT)

Alternatively, internal noise may be introduced by the activation of binocular disparity detectors tuned to different disparities. The likelihood distribution for disparity (change) is

also conveniently expressed as a Gaussian distribution centered on the true disparity (change) of the stimulus. The disparity likelihood is then combined with the disparity prior favoring zero disparity. A similar disparity prior has been suggested in the context of sustained and transient stereo images (Read, 2002a,b).

The Bayesian stereo-first model describes perceived binocular disparity (change) as the product of likelihood and prior.

$$p(d \mid \beta) \propto p(\beta \mid d)p(d) \tag{3.6}$$

The likelihood for binocular disparity (change) is modeled as a Gaussian distribution centered on the true disparity δ (or disparity change) measured at the endpoint of stimulus motion. The standard deviation σ_d of the distribution is left as a free parameter.

$$p(\beta \mid d; \sigma_d) = \frac{1}{\sqrt{2\pi}\sigma_d} \exp\left(\frac{-(d-\delta)^2}{2\sigma_d^2}\right) \tag{3.7}$$

The preference or prior for small disparity (change) is modeled as a Gaussian distribution centered on zero disparity.

$$p(d; \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(d)^2}{2\sigma^2}\right) \tag{3.8}$$

The MAP estimate for disparity is then given by

$$\hat{\delta} = \frac{1/\sigma_d^2 \, \delta}{1/\sigma_d^2 + 1/\sigma^2} = \frac{\delta}{1 + (\sigma_d / \sigma)^2} \tag{3.9}$$

Changing disparity information needs to be coupled with spatial position to recover 3-D motion. The cyclopean azimuth $\dot{\alpha}$ can be approximated by $(\dot{\alpha}_L + \dot{\alpha}_R)/2$ and disparity constraints are then given relative to angle $\dot{\alpha}$.

$$\hat{\alpha}_L = \dot{\alpha} - \hat{\delta}/2, \quad \hat{\alpha}_R = \dot{\alpha} + \hat{\delta}/2 \tag{3.10}$$

These estimates determine perceived constraints of the stimulus and their intersection gives a biased estimate of perceived motion direction and speed.

Bayesian stereo-motion model (BJEMD)

In the present framework velocity and disparity input can be combined in a Bayesian model with different noise ratios for motion and disparity processing. Uncertainty in velocity and disparity processing are combined and both uncertainty parameters are estimated together.

If we estimate cyclopean azimuth $\hat{\alpha}$ by $(\hat{\alpha}_L + \hat{\alpha}_R)/2$ and insert the velocity estimates from Eq (5) into Eq (10) then velocity and disparity input can be combined in a single Bayesian model

$$\hat{\alpha}_L = \hat{\alpha} - \hat{\delta}/2, \quad \hat{\alpha}_R = \hat{\alpha} + \hat{\delta}/2 \tag{3.11}$$

with estimates $(\hat{\alpha}_L, \hat{\alpha}_R, \hat{\delta})$ based on noise parameters for velocity σ_v and disparity σ_d processing.

Following Bayes' rule, likelihoods and priors are combined to establish a posterior distribution for each model and trajectory. Applying a simple decision rule, such as the *maximum a posteriori* rule, provides *a posteriori* estimates of angular velocity and disparity. The estimates describe biased constraint lines and their intersection determines an azimuth angle and radial distance in *x-z* co-ordinates.

3.2.2 Perception of Horizontal Velocities

In this Bayesian framework uncertainty is modeled by the ratio of standard deviations between likelihood and prior. If uncertainty is negligible model predictions of azimuth

angle are veridical but with increasing uncertainty model predictions approximate a shrinking circle for the motion-first and a compressed ellipse for the stereo-first Bayesian model. In Fig. 3.2 predictions over the full range of 360° of stimulus azimuths are plotted for the two Bayesian models.
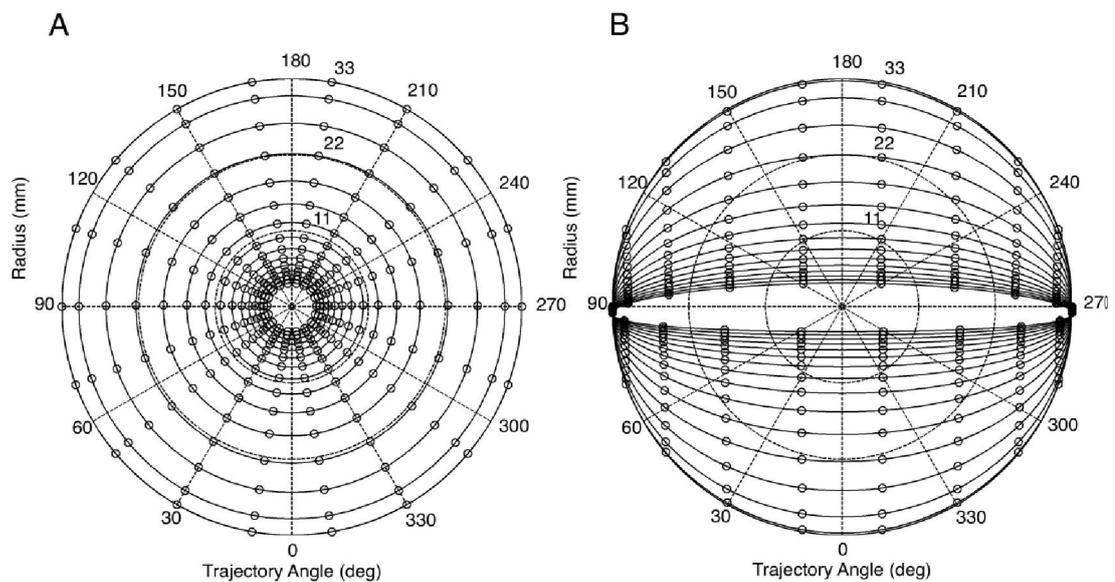
Figure 3.2. Simulation results for (A) motion-first Bayesian IOVD and (B) stereo-first Bayesian CDOT and Bayesian JEMD. Plots show model predictions of trajectory angle and velocity in polar co-ordinates for azimuth angles of 10 to 350° in steps of 20° at a viewing distance of 114 cm. Uncertainty is modeled by the ratio of likelihood and prior ranging from 0.1 to 3.0 in steps of 0.2.

The unbiased prediction of azimuth angles, its circular shape, is the result of multiplying left and right angular velocity by the same factor. The increasingly flat elliptical shape is the consequence of a stronger bias for larger disparities near the *z*-axis or azimuths of 0°

51

and 180° and weaker bias for smaller disparities near the *x*-axis or azimuths of 90° and 270°.

Under natural viewing conditions there are many monocular cues to 3D motion but in a sparse environment only uncertain binocular motion and disparity cues may be available. In a psychophysical experiment Lages (2006) investigated perceived bias of motion trajectories of small target dots and used the Bayesian models of interocular velocity difference (BIOVD) and disparity change (BCDOT) as well as joined encoding (BJEMD) to explain the results.
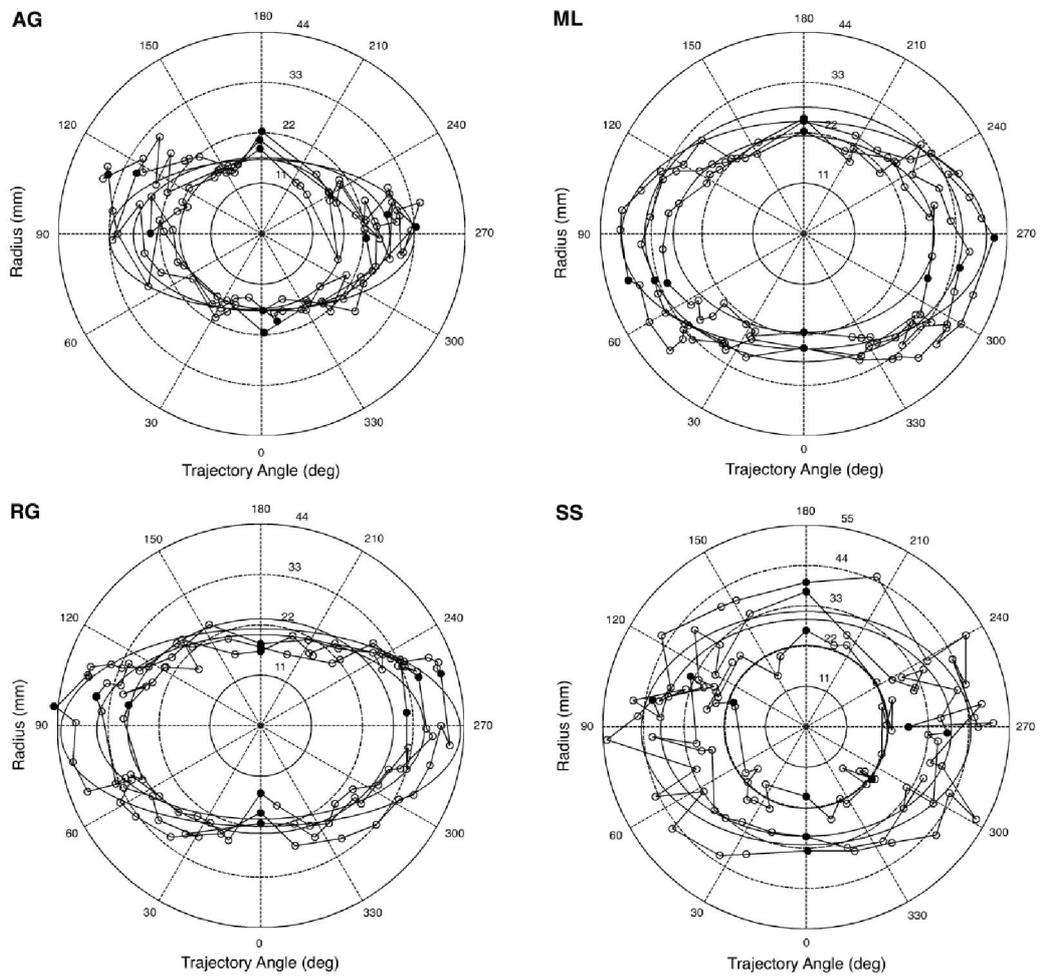
Figure 3.3 Illustration of empirical results for four observers (adapted from Lages 2006, Fig. 8). Polar plots for perceived azimuth and radial distance (speed) for stimulus velocities of 0.02 m/s (16.6 mm), 0.03 m/s (25.0 mm) and 0.04 m/s (33.3 mm) and best model fits of the stereo-first Bayesian model (BCDOT). Filled data points correspond to cardinal stimulus trajectories (0, 90, 180, 270 deg). With increasing stimulus speed estimates of radius and uncertainty increase and model fits and data assume a more compressed elliptical shape.

Stimuli were presented to the left and right eye using a split-screen Wheatstone configuration. The observer viewed three anti-aliased dots presented above and below a

53

fixation cross, surrounded by a rectangular fusion lock, at a viewing distance of 114 cm. Each dot subtended less than 4.4 arcmin at 27.7, 38.8 and 50 arcmin above and below fixation. In randomly intermixed trials the dots moved 16.6, 25, or 33.3 mm on the horizontal *x-z* plane for 833 ms (0.02, 0.03 and 0.04 m/s) on 36 different trajectories.

On each trial the observer verged on the fixation cross before they initiated motion of target dots by key-press. Azimuth angle of the target ranged between 0° and 350° in steps of 10°. Each observer attended a total of eight separate blocks (2 tasks x 4 repetitions) each comprising 108 trials (3 velocities x 36 trajectories). In each block of trials observers judged either motion azimuth or radial distance. Adjustments to 36 trajectories and 3 velocities were repeated four times in randomly intermixed trials.

Using maximum likelihood (ML) fits individual adjustments of azimuth and radial distance were averaged across trials with the same stimulus velocity. Bayesian models with one free parameter (noise ratio between likelihood and prior determines perceived speed) for the motion-first model and two free parameters (perceived speed and noise ratio between likelihood and prior) for the stereo first model were fitted to data of each observer and three stimulus speeds and results are summarized in Table 3.1. The stereo-first Bayesian model gives slightly better fits and parameter values assume more plausible values than the motion-first model.

Obs.	Radius	Motion-first Model			Stereo-first Model		
	R	$\sigma_v{:}\sigma$	r_v	$\chi^2(34)$	$\sigma_d{:}\sigma$	r_d	$\chi^2(33)$
A.G.	16.6	0.0	17.6	2.40	0.32	17.7	2.29
	25.0	0.0	20.6	6.63	0.81	27.7	2.12
	33.3	0.0	25.2	17.2	1.01	32.9	5.79
M.L.	16.6	0.0	26.3	2.89	0.56	28.2	1.92
	25.0	0.0	30.8	4.07	0.49	34.3	3.09
	33.3	0.0	34.4	9.24	0.77	39.5	3.00
R.G.	16.6	1.83	113	1.40	0.48	28.8	1.32
	25.0	2.84	274	3.40	0.87	35.4	0.58
	33.3	2.22	213	8.97	1.00	43.2	1.45
S.S.	16.6	0.0	22.2	2.31	0.0	22.2	2.31
	25.0	0.0	36.1	1.45	0.58	39.3	1.14
	33.3	0.82	64.5	1.74	0.64	44.5	1.35

Table 3.1. Individual results from four observers and three stimulus speeds (radius in mm) from Lages (2006, Exp. 2). Parameter estimates and goodness-of-fit for motion-first and stereo-first Bayesian model.

Likelihood fits and parameters for the stereo-motion Bayesian model (BJEMD) are not reported because results were almost identical to the stereo-first Bayesian model with no significant improvement for any of the individual data sets.

As illustrated in Fig. 3.3 ML fits of the stereo-first Bayesian model for each stimulus velocity suggest that an increase in stimulus velocity systematically raised uncertainty in the noise ratio. This trend appears in all observers except for Observer M.L. in the 16.6 mm condition.

In a standard log-likelihood test (see -2log(Λ) in Table 3.2), the χ^2-distribution was badly approximated and therefore did not indicate any significant differences between BCDOT and BIOVD model fits. Bayes Information Content (BIC), however, approximates the distribution-free Bayes Factor (BF) and suggests weak to positive evidence in favor of the stereo-first model (Raftery, 1995). The BIC for model H_i is defined as

$$BIC(H_i) = -2\log(L_i) + k_i \log(n) \qquad (3.12)$$

where n is the number of observations, k_i is the number of free parameters of model H_i and L_i is the maximum likelihood for model H_i (Raftery, 1995, 1999). The BIC approximation to the prior predictive probability is obtained by $P(D|H_i) = exp(-BIC(H_i)/2)$. In the case of two models, H and H, the Bayes factor is defined as the ratio of the prior predictive probabilities; hence, the BIC approximation of the Bayes factor is given by

$$BF_{01} \approx \frac{P_{BIC}(D|H_0)}{P_{BIC}(D|H_1)} = exp(\Delta BIC_{10}/2) \qquad (3.13)$$

where $\Delta BIC_{10} = BIC(H_1) - BIC(H_0)$. Here, the probability associated with BF$_{BIC}$ indicates that given the empirical data the stereo-first model is on average 3 times more probable than the motion-first model.

Obs.	Radius	Model Selection		
	r	$-2\log(\Lambda)$	BF_{BIC}	Pr_{BIC}
A.G.	16.6	0.04	2.22	0.69
	25.0	0.99	3.57	0.78
	33.3	0.95	3.49	0.78
M.L.	16.6	0.36	2.60	0.72
	25.0	0.24	2.45	0.71
	33.3	0.98	3.55	0.78
R.G.	16.6	0.05	2.23	0.69
	25.0	1.54	4.69	0.82
	33.3	1.58	4.80	0.83
S.S.	16.6	0.0	2.18	0.69
	25.0	0.21	2.42	0.71
	33.3	0.22	2.43	0.71

Table 3.2 Model selection for Bayesian motion-first and disparity-first model based on log-likelihood ratio test $-2\log(\Lambda)$ and BIC approximation of Bayes Factor BF_{BIC} and corresponding probability Pr_{BIC}. Model selection results for four observers and three stimulus speeds are shown (radius in mm; Lages, 2006, Exp. 2).

Rendering 3D motion in a stereoscopic set-up is difficult and can introduce various artifacts and cue conflicts. In this experiment constant size and blur of the target stimuli moving in depth may have influenced perceived depth (Watt, Akeley, Ernst, & Banks, 2005). On the other hand, cue conflicts due to looming and accommodative cues are probably too small to account for the substantial and systematic bias found for small

blurred targets that move maximally ±3.3 cm in depth at a viewing distance of 114 cm. Using LEDs moving in depth Harris and Dean (2003), as well as Welchman et al. (2004) reported systematic overestimation of perceived azimuths near the fronto-parallel plane, confirming that perceptual bias also exists for real-world stimuli at various trajectory angles.

Bayesian model fits for perceived azimuth angle and speed (radial distance) and Bayesian model selection promotes the idea that bias in 3D motion perception is introduced by disparity processing. This confirms previous findings in psychophysical studies that used different stimuli and methods (e.g., Cumming & Parker, 1994; Lages, Mamassian, & Graf, 2003). It is possible, however, that interocular velocity difference or optical flow contributes to 3D motion perception, especially when stimuli define surfaces that move on a trajectory near the observer's line of sight (Brooks & Stone, 2004).

In the stereo-first Bayesian model disparity estimates are derived from the endpoint of stimulus motion rather than integrated over time. As a consequence the stereo-first Bayesian model may be interpreted as (i) temporal integration of biased disparities or (ii) biased temporal integration of disparity. The latter interpretation appears more plausible since uncertainty estimates increased systematically with stimulus velocity as reported in Table 3.1 and Figure 3.3.

If 3D motion perception is based on velocity-tuned processing the relatively small change of stimulus velocity in our experiment should have very little effect on uncertainty. Disparity-tuned processing on the other hand may increase uncertainty levels for faster stimuli due to the temporal limits of disparity integration (Read & Cumming, 2005b; Tyler, 1971) in a transient stereo-system (Edwards & Schor, 1999).

One of the main goals of visual processing is to segregate and identify objects in space and time. With increasing proximity or size of a moving object local motion detectors signal a wider range of velocities. As a consequence a system that processes motion input first needs to establish correspondence between rather different monocular motions before it can build a percept of 3D object motion. Computationally it appears more parsimonious to solve the stereo correspondence problem before deriving a 3D motion percept. This argument also applies to joint encoding of motion and disparity (JEMD) because early encoding of true 3D motion would require a large number of detectors specifically tuned to all combinations of spatial frequency, orientation, and interocular spatio-temporal offsets to capture all possible local 3D motions.

It is concluded that under the present experimental conditions perceptual bias in 3D velocity is most likely the result of limited temporal integration when processing disparity change. This points to stereo-first or stereo-motion processing and rules out a motion-first mechanism that relies on interocular velocity difference only.

3.2.3 Bayesian Model Selection

In a standard log-likelihood test (see -2log(Λ) in Table 3.2), the test statistic is badly approximated by the χ^2-distribution and therefore has difficulties to detect significant differences between BCDOT and BIOVD model fits. We therefore apply Bayesian hypothesis testing or model selection (Raftery, 1995; 1999). After observing the data, the posterior odds in favour of model H_0 versus the alternative model H_1 are given by the Bayes Factor (BF or likelihood odds) times the prior odds:

$$\frac{\Pr(H_0|D)}{\Pr(H_1|D)} = \frac{Pr(D|H_0)}{\Pr(D|H_1)} \times \frac{\Pr(H_0)}{\Pr(H_1)} \tag{3.12}$$

Usually the prior odds for both hypotheses or models are assumed to be the same. Bayes Information Content (BIC) approximates the distribution-free Bayes Factor in Equation 3.12

$$BF_{01} = \frac{\Pr(D|H_0)}{\Pr(D|H_1)} \qquad\qquad (3.13)$$

The BIC for model H_i is defined as

$$BIC(H_i) = -2\log(L_i) + k_i\log(n) \qquad\qquad (3.14)$$

where n is the number of observations, k_i is the number of free parameters of model H_i and L_i is the maximum likelihood for model H_i (Raftery, 1995, 1999). The BIC approximation to the prior predictive probability can be obtained by $P(D|H_i) = exp(-BIC(H_i)/2)$. In the case of two models, H_0 and H_1, the Bayes factor is defined as the ratio of the prior predictive probabilities; hence, the BIC approximation of the Bayes factor is given by

$$BF_{01} \approx \frac{P_{BIC}(D|H_0)}{P_{BIC}(D|H_1)} = exp(\Delta BIC_{10}/2) \qquad\qquad (3.15)$$

where $\Delta BIC_{10} = BIC(H_1) - BIC(H_0)$. If both models are equally likely ($P(H_0)=P(H_1)=0.5$) then the posterior probability $P(H_1|D)$ can be approximated by $Pr_{BIC} = BF_{01}/(1+ BF_{01})$.

Here, the BF_{BIC} and the posterior probabilities Pr_{BIC} in Table 3.1 indicate weak to positive evidence in favor of the stereo-first model (Raftery, 1995). Given the empirical data the stereo-first model is on average about 3 times more probable than the motion-first model (with an average posterior probability of 0.76). On average this constitutes positive evidence in favour of the stereo-first model (Raftery, 1995).

Rendering 3D motion in a stereoscopic set-up is difficult and can introduce various artifacts and cue conflicts. In this experiment constant size and blur of the target stimuli moving in depth may have influenced perceived depth (Watt, Akeley, Ernst, & Banks, 2005). On the other hand, cue conflicts due to looming and accommodative cues are probably too small to account for the substantial and systematic bias found for small blurred targets that move maximally ±3.3 cm in depth at a viewing distance of 114 cm. Using LEDs moving in depth Harris and Dean (2003), as well as Welchman et al. (2004)

also reported systematic overestimation of perceived azimuths near the fronto-parallel plane, confirming that perceptual bias also exists for real-world stimuli at various trajectory angles.

Bayesian model fits for perceived azimuth angle and speed (radial distance) and Bayesian model selection promotes the idea that bias in 3D motion perception is introduced by disparity processing. This confirms previous findings in psychophysical studies that used different stimuli and methods (e.g., Cumming & Parker, 1994; Lages, Mamassian, & Graf, 2003). It is possible, however, that interocular velocity difference or optical flow contributes to 3D motion perception, especially when stimuli define surfaces that move on a trajectory near the observer's line of sight (Brooks & Stone, 2004).

In the stereo-first Bayesian model disparity estimates are derived from the endpoint of stimulus motion rather than integrated over time. As a consequence the stereo-first Bayesian model may be interpreted as (i) temporal integration of biased disparities or (ii) biased temporal integration of disparity. The latter interpretation appears more plausible since uncertainty estimates increased systematically with stimulus velocity as reported in Table 3.1 and Figure 3.3.

If 3D motion perception is based on velocity-tuned processing the relatively small change of stimulus velocity in our experiment should have very little effect on uncertainty. Disparity-tuned processing on the other hand may increase uncertainty levels for faster stimuli due to the temporal limits of disparity integration (Read & Cumming, 2005b; Tyler, 1971) in a transient stereo-system (Edwards & Schor, 1999).

One of the main goals of visual processing is to segregate and identify objects in space and time. With increasing proximity or size of a moving object local motion detectors signal a

wider range of velocities. As a consequence a system that processes motion input first needs to establish correspondence between rather different monocular motions before it can build a percept of 3D object motion. Computationally it appears more parsimonious to solve the stereo correspondence problem before deriving a 3D motion percept. This argument also applies to joint encoding of motion and disparity (JEMD) because early encoding of true 3D motion would require a large number of detectors specifically tuned to all combinations of spatial frequency, orientation, and interocular spatio-temporal offsets to capture all possible local 3D motions.

It is concluded that under the present experimental conditions perceptual bias in 3D velocity is most likely the result of limited temporal integration when processing disparity change. This points to stereo-first or stereo-motion processing and rules out a motion-first mechanism that relies on interocular velocity difference only.

3.3 GENERALIZED BAYESIAN APPROACH

In an extended geometric-statistical approach we explore binocular 3D motion perception under ambiguity (see Fig 3.4; Wang, et al., 2012; Lages, Heron & Wang, 2013). The parameters in this Bayesian model describe how strongly perceived motion direction may be influenced by noise from motion processing and noise from (orientation) disparity processing. We focused on a probabilistic extension of the VN model because the CA strategy (see chapter 2 for in depth description of models) did not capture essential characteristics of the empirical data (see chapter 4 for comparisons between models from chapter 2/3 and empirical data).

Following Weiss and Fleet (2001), we incorporate intensity constraints into a Bayesian model of 3D motion perception. However, the present approach extends Bayesian models of *x-y* (Weiss et al., 2002) and *x-z* motion (Lages, 2006; Welchman, et al., 2008) to true 3D

motion in *x-y-z* thereby providing default estimates for the 3D aperture problem (Lages & Heron, 2010; Wang et al., 2012; Lages, Heron, & Wang, 2013).

Local 2D motion constraints from the left and right eye in a binocular viewing geometry are sufficient to establish two velocity constraint planes or hyperplanes in 3D space: the moving line projects on the left and right retina. The projected lines together with each nodal point establish velocity constraint planes or hyperplanes for the left and right eye (see Fig. 3.4). The left and right constraint plane are determined by a 3D gradient through the cyclopean point *C* or the fixation point *F*. Note that stereo correspondence in a 3D binocular viewing geometry needs to be established before velocity constraint planes can be defined. It seems plausible that vergence, accommodation and disparity cues as well as other depth cues help to establish stereo correspondence between images in the left and right eye.
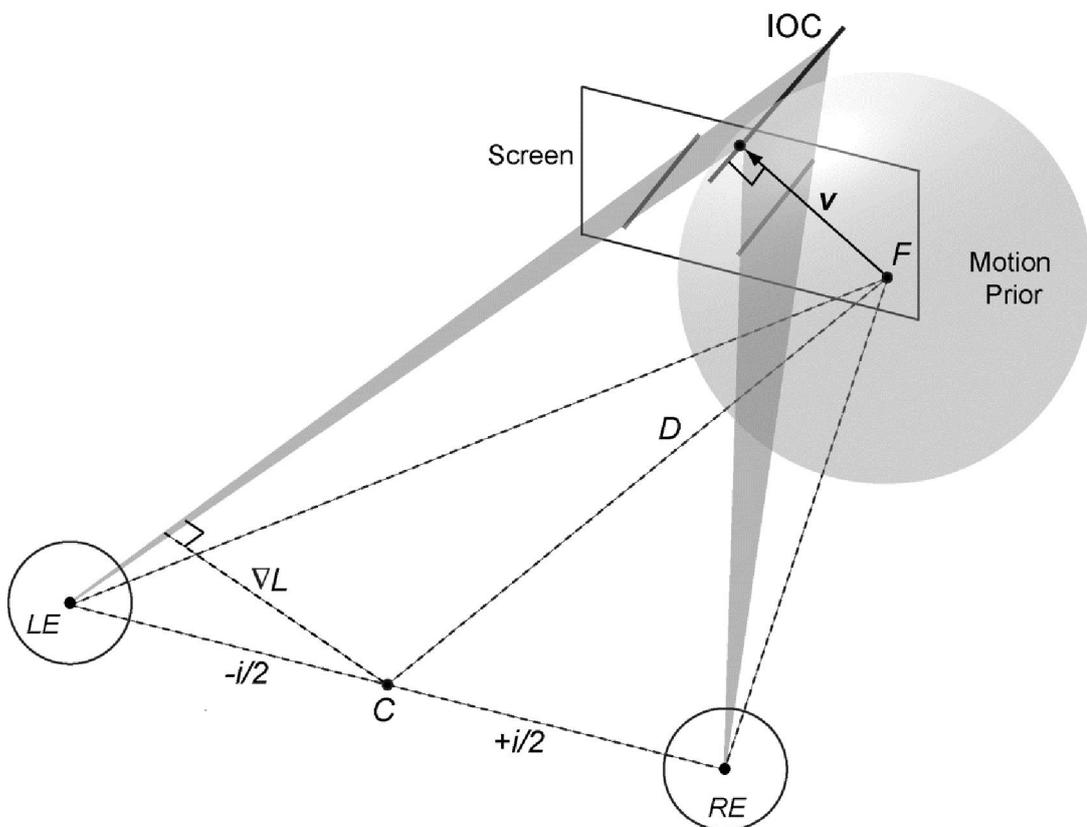
Figure 3.4. Illustration of binocular Bayesian model with left and right eye velocity constraint planes (shaded triangles) and intersection of constraints (IOC) for a line moving in 3D. A prior for slow motion in 3D is indicated by the sphere centered on point F which is also the fixation point. If the constraints contain little noise compared to the prior then the resulting 3D velocity estimate $\mathbf{v}$ approximates the vector normal of the IOC through F (arrow).

Similar to 2D motion, intensity gradients in 3D may be approximated by first-order Taylor series expansion using brightness constraints in the left and right eye. For corresponding constraints in the left and right eye we assume that

$$\mathbf{v}^T \nabla L + L_t = 0, \quad \mathbf{v}^T \nabla R + R_t = 0 \tag{3.12}$$

where $\nabla L = (L_x, L_y, L_z)$ and $\nabla R = (R_x, R_y, R_z)$ denote the spatial gradients corresponding to the constraint planes derived from local line or contour projections in the left and right eye, respectively (see Appendix A3). As before, we assume a binocular viewing geometry where the position of the projected fixation point *F* on the retina and the nodal points of the eyes (and therefore the cyclopean point) are known and only velocity of a local line or edge needs to be estimated. Since the visual system continuously monitors accommodation and vergence of both eyes it is reasonable to assume that local velocity constraints are available in form of intensity gradients.

3.3.1 Generalized Bayesian Binocular 3D Motion Model

The present geometric-statistical approach builds on existing Bayesian models of 2D motion perception (Weiss et al., 2002). It extends binocular Bayesian models introduced above (Lages, 2006) to derive a 3D velocity estimate under ambiguity in a binocular

viewing geometry (Wang & Lages, 2012; Lages, Heron & Wang, 2013). However, the binocular Bayesian model presented here should not be understood as a processing model but instead as a framework to quantify noise on stereo and velocity constraints for the left and right eye.

For the sake of simplicity we assume a fixed viewing geometry in a Cartesian coordinate system where positions on the image plane, nodal points of the eyes and fixation point are known and nodal points and fixation remain constant. Since the visual system continuously monitors accommodation and vergence of the eyes in a binocular viewing geometry, it is reasonable to assume that the visual system may not only compute local 2D intensity gradients in the image but something akin to 3D intensity gradients, constraining velocity for each eye.

The left and right eye velocity constraints can be expressed through the nodal points and two-dimensional motion gradients in the image plane (see Appendix A2). As a matter of convenience the fronto-parallel image plane is located at the fixation point which is also the starting point of motion. We used two points on the image plane to define a line at known distance D for each eye (rather than gradients at a focal length on the retina) to derive each constraint plane.

Bayesian Vector Normal (BVN) Model

The velocity constraint planes may be noisy due to microsaccades, missing information (e.g., occlusion) and neural encoding. In the following we make the simplifying assumption that spatial derivatives of the constraint planes are precise but temporal derivatives have additive noise.

If the constraint planes through the left and right eye are not coincident or parallel their intersection constrains 3D velocities but does not provide a unique solution or local

65

velocity estimate. As a consequence additional constraints are needed to disambiguate local velocity. It seems plausible that disparity information from feature tracking, together with other depth cues, helps to disambiguate 3D motion perception (Lages & Heron, 2008; 2010). However, if disparity features and other depth information is unavailable a weak prior for 3D motion resolves ambiguity and provides a local default estimate.

In their influential paper on 2D motion illusions Weiss et al. (2002) suggested a 2D Gaussian motion prior for slow motion perception in x-y space. Similarly, Lages (2006) introduced a bivariate Gaussian motion prior on the x-z axes to explain bias in perceived azimuth and speed of a target moving on a horizontal plane. If we assume that most features and objects in a scene are stationary or tend to move slowly on an arbitrary trajectory in 3D space then a symmetric 3D Gaussian provides a plausible world prior for binocular 3D motion perception of 3D velocity $\mathbf{v}$.

Here we propose the 3D Gaussian as conjugate motion prior of Gaussian likelihoods

$$p(\mathbf{v}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-\mathbf{v}^T\mathbf{v}}{2\sigma^2}\right) \tag{3.20}$$

This world prior simply reflects a preference for slow motion in every direction. This is a plausible assumption as most features in natural scenes remain static and moving objects tend to move slowly.

Similar to Equation 3.1, the posterior distribution is the result of combining likelihood constraints and prior using Bayes' Rule where $L(\mathbf{x},t)$ and $R(\mathbf{x},t)$ describe intensities in world co-ordinates associated with the left and right eye. The denominator is dropped because it is independent of $\mathbf{v}$ and only scales the posterior by a constant factor.

$$p(\mathbf{v}\,|\,L(\mathbf{x},t), R(\mathbf{x},t)) \propto p(L(\mathbf{x},t)\,|\,\mathbf{v})p(R(\mathbf{x},t)\,|\,\mathbf{v})p(\mathbf{v}) \tag{3.21}$$

The posterior distribution gives a random normal variable as an estimate. In order to find the most probable velocity or MAP estimate, we take the negative logarithm of the posterior, differentiate it with respect to 3D velocity $\mathbf{v}$ and set the derivative equal to zero.

The logarithm of the posterior is quadratic in $\mathbf{v}$ so that the solution can be written in closed form using standard linear algebra (see Appendix A3).

Bayesian Disparity and Vector Normal (BDVN) Model

In an extension of the BVN model we introduce orientation disparity computation in a separate stage. As a consequence the corresponding likelihoods are adjusted accordingly. As for the BVN Model, the adjusted likelihoods are then combined with the 3D motion prior to derive a MAP estimate of perceived 3D velocity.

Line orientation θ_L for the left eye and θ_R for the right eye measured from the horizontal on the image plane. If the likelihood for orientation disparity $d = (\theta_L - \theta_R)$ is a Gaussian centered on true orientation disparity δ with standard deviation σ_d

$$p(\delta \,|\, d) = \frac{1}{\sqrt{2\pi}\sigma_d} \exp\left(\frac{-(d-\delta)^2}{2\sigma_d^2} \right) \tag{3.24}$$

The orientation-adjusted lines on the image plane together with the nodal points also define intersecting 3D velocity constraint planes. The estimation of 3D velocity is then achieved as for the BVN Model by combining the adjusted likelihoods with the 3D motion prior.

With negligible noise for orientation disparity σ_d and noise for motion σ_v the BDVN model provides vector normal (VN) estimates as a default solution. With increasing orientation disparity noise however, velocity estimates show characteristics similar to the observed data in our experiments (see chapter 4 for description of experiments and comparison

67

between models and empirical data).

3.2.2 Simulation of 3D Motion under Uncertainty

The previously discussed bias in binocular 3D motion perception describes a perceptual bias in azimuth and speed. More specifically, as noise or uncertainty in the likelihoods increases perceived azimuth is biased towards the fronto-parallel fixation plane and perceived speed is reduced (see Fig. 3.4). We have implemented the Generalized Bayesian model in MatLab (Mathworks, Natick MA) to simulate velocity estimation under uncertainty. We compared the Bayesian MAP estimates with the vector normal (VN), which is equivalent to the shortest distance between starting point F and the line defined by the IOC planes in 3D space.

In the first simulation the line stimulus was vertical and moved on horizontal trajectories with azimuth ranging between 10° to 360° in steps of 10°. Viewing distance D was set to 55 cm and interocular distance i to 6.5 cm. In the degenerate case of a moving vertical line the elevation angle of motion direction remains at 0°. In this case predictions are equivalent to a single dot moving in the horizontal *x-z* plane as discussed in Section 3.2 (compare Fig. 3.3 and 3.5).
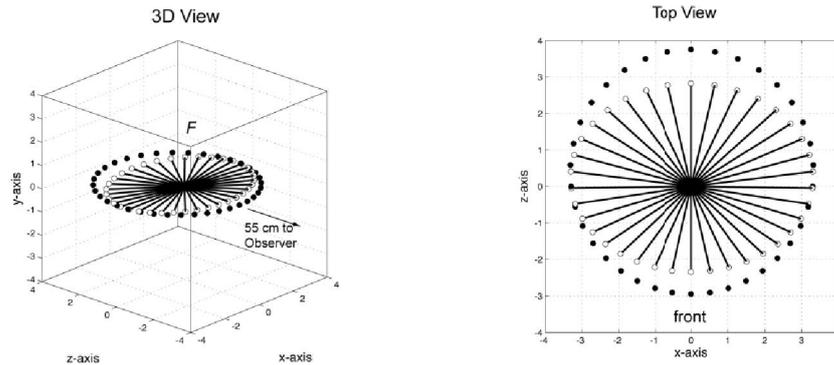
Figure 3.5. Simulation results for the generalized 3D Bayesian model: 3D view in left plot and top view in right plot. Predicted velocity of a vertical line moving on horizontal azimuths ranging from 0° to 360° in steps of 10° are shown. Bayesian estimates (MAP) of predicted trajectories are shown as thick lines with open circles attached for a noise ratio between likelihood and prior of 1:32. Endpoints of geometric vector normal (VN) predictions are indicated by filled circles.

The MAP estimates approximate VN predictions if likelihoods have very little noise compared to the prior (noise ratio $\sigma_v : \sigma$ <1:100). For a noise ratio of 1:32 MAP estimates are compressed in depth leading to trajectories that are biased towards the fronto-parallel fixation plane. If trajectories point away from the fixation plane then speed is increasingly underestimated. These model characteristics match empirical results on perceived azimuth and speed of stimulus dots under uncertainty (see Fig. 3.5 and Table 3.1).

3.2.3 Simulation of 3D Motion under Ambiguity

We also computed Bayesian MAP estimates for the 3D aperture problem at different noise ratios and compared the velocity estimates with vector normal (VN) predictions as a

plausible default. The moving line stimuli were oblique or vertical (45 and 90 deg). In addition the line stimuli were slanted in depth due to systematic manipulation of orientation disparity between the projections into the left and right eye. Orientation disparity in the stimulus display ranged from -6° to +6°. As in the first simulation, viewing distance *D* was set to 55 cm and interocular distance *i* to 6.5 cm. At 0° elevation the line stimulus moved on a fixed azimuth of +52.5° (37.5°) to the front and -52.5° (142.5°) to the back. Note that an azimuth of 0° denotes motion frontoparallel to the observer (directly towards the observer).

We then varied variability or noise of the likelihood σ_v while keeping the prior σ constant; the resulting ratio $\sigma_v : \sigma$ assumed values of 1:10, 1:32, and 1:100.
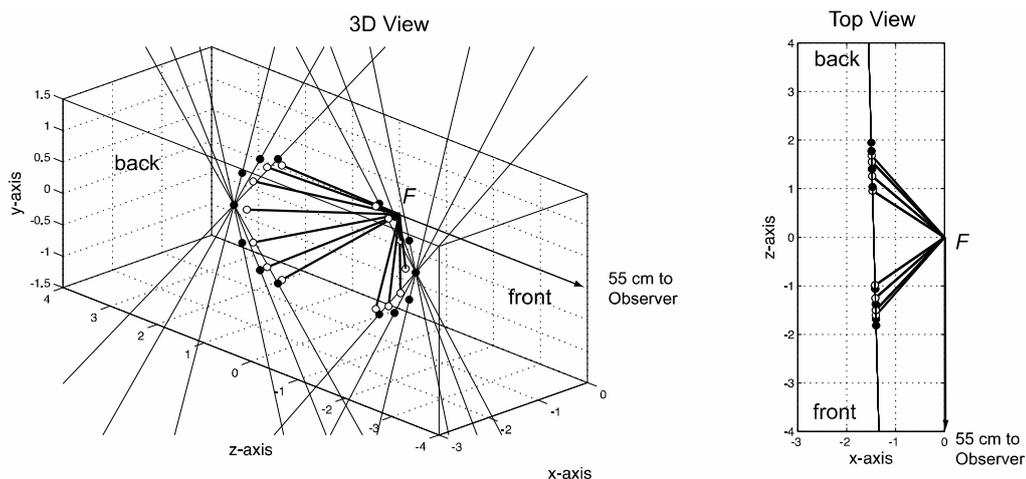
Figure 3.6. Simulation results for generalized 3D Bayesian model: 3D view in left plot and top view in right plot. Predicted trajectories to front (+iovd) and back (-iovd) as a function of orientation disparity (slant in depth varied between ±6° in steps of 2°) of a moving oblique stimulus line (thin line). Noise ratio between likelihood and prior is set to 1:100 and Bayesian estimates (MAP) are shown as thick lines, originating from fixation point *F* with open circles attached. Endpoints of geometric vector normal (VN) predictions are indicated by filled circles.

As in the first simulation MAP estimates approximate the VN solution if the noise ratio is less than 1:100. MAP estimates are only slightly biased away from the vector normal if the likelihoods have little noise compared to the prior. As noise in the likelihoods increases MAP estimates shorten and move away from the IOC line towards the surface normal of the right or left eye constraint plane, whichever is closer to the starting point *F* (see Fig. 3.6 and 3.7). These results may approximate local velocity estimates in the absence of other disambiguating cues.
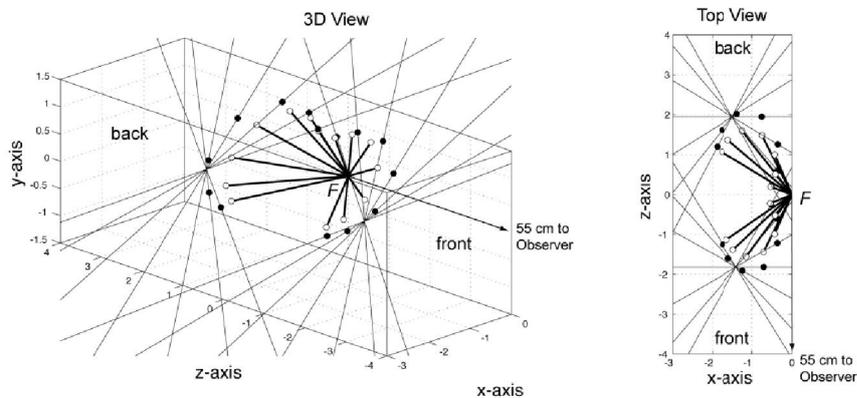
Figure 3.7. Simulation results for generalized 3D Bayesian model: 3D view in left plot and top view in right plot. Predicted trajectories to front (+iovd) and back (-iovd) as a function of orientation disparity (slant in depth varied between ±6° in steps of 2°) of a moving vertical stimulus line (thin line) slanted in depth. Noise ratio between likelihood and prior is set to 1:32 and Bayesian estimates (MAP) are shown as thick lines, originating from fixation point *F* with open circles attached. Endpoints of geometric vector normal (VN) predictions are indicated by filled circles.

3.4 DISCUSSION

In this chapter we have described two Bayesian models of local trajectory computation, based on the Vector Normal model (VN) described in chapter 2. The parameters in these models describe how strongly perceived motion trajectory may be influenced by noise

from motion processing (BVN) and noise from motion processing in conjunction with noise from (orientation) disparity processing (BDVN). These models predict that under uncertainty and ambiguity (3D aperture problem, see chapter 1 and 2 for details and chapter 4 for details on empirical work) perceived motion trajectory will be more influenced by the motion prior (perceived speed is reduced, BVN model) and both motion and disparity priors (perceived speed is reduced, perceived azimuth is underestimated, BDVN model).

We describe the results of two model simulations. The first simulation compared predictions of the generalized Bayesian model with the VN model (chapter 2), for a vertical line moving on horizontal azimuths ranging from 0° to 360° in steps of 10° when a) the noise ratio between likelihood and prior was almost zero (likelihoods have low noise compared to prior) b) the noise ratio between likelihood and prior was 1:32. For a small noise ratio, Bayesian predictions are similar to those of the VN model; when there is more noise in the likelihood i.e. under increased uncertainty, the Bayesian predictions are compressed in depth.

The second simulation compared predictions of the generalized Bayesian model with the VN model for the case of a line of a given orientation (45, 90) and orientation disparity (-6, -4, -2, 0, 2, 4, 6) oscillating in depth behind a circular aperture (based on the experimental set-up in chapter 4) with interocular velocity difference (IOVD) either to the front or back of fixation. In this case the noise of the prior was again kept constant, whilst the noise of the likelihood varied. There were 3 noise ratio settings, 1:10, 1:32, and 1:100. Again when noise ratio between likelihood and prior was negligible Bayesian model predictions approximated VN predictions. As the likelihood noise increases, Bayesian model predictions shorten and move away from the IOC line towards the surface normal of the right or left eye constraint plane.

The second simulation is based on the empirical work carried out in chapter 4, where we consider the case of an oblique or vertical line oscillating in depth behind a circular aperture. It shows that Bayesian models presented here enjoy more flexibility than the computational models discussed in chapter 2. In chapter 4, the Bayesian models will be compared with empirical results from four observers in order to investigate their utility in describing empirical data. From the simulation results described here, it appears that Bayesian models will be better suited when modeling perception under ambiguity.

One of the strengths of the Bayesian approach is that it can incorporate multiple cues at the same or neighboring locations. For example, local velocity constraints for the left and right eye can be stabilized in an area-based approach (Lucas & Kanade, 1981) by using multiple overlapping apertures and likelihoods tuned to different spatial frequencies. In addition, the model can be extended to the perception of 3D object motion using a sufficiently dense array of local estimates to capture features such as endpoints, junctions, and corners. Constraints from depth processing can therefore disambiguate object motion direction similar to the Bayesian approach to 2D motion in Weiss et al. (2002).

In the present model intensity gradients, likelihood constraints, and motion prior are expressed in world co-ordinates. Transformations from image or retinal gradients into world co-ordinates may be achieved using the epipolar constraint (Hansard & Haroud, 2008) and the fundamental matrix (Faugeras, 1992; Hartley & Zisserman, 2004). [The fundamental matrix defines the transformations between corresponding projections onto verged retinal or image planes.]

It is tempting to assume that intensity gradients, velocity constraints, and motion prior are exclusively the result of motion encoding and processing. However, solving the aperture problem locally requires stereo correspondence suggesting a significant contribution from depth processing. Consider for example a moving line or edge that also

changes size and orientation over time. If tilt, slant and size of the moving line changes over time then velocity constraint planes and their intersection are no longer sufficient. They need to be updated frequently. Estimating motion trajectory and speed of such a non-linear motion stimulus is only possible if stereo correspondence and depth is resolved with sufficient temporal resolution. Sampling of IOC constraints over time is equivalent to transient (orientation) disparity processing. Here, the term 'transient' refers to changes in orientation disparity over time. A system that can encode rotational as well as translational line motion locally is also capable of capturing non-rigid object motion. Therefore, it seems plausible that the 3D motion system employs motion and disparity processing and late integration (DeAngelis & Newsome, 2004; Ponce et al., 2008) to overcome the inverse problem of 3D motion perception (Lages & Heron, 2010).

In summary, the present Bayesian model extends existing models of 3D motion perception. It captures some bias in 3D motion perception and provides testable predictions in the context of the 3D aperture problem. In addition, however, this model would require updates from disparity processing to capture non-linear 3D motion trajectories.

CHAPTER 4. PSYCHOPHYSICAL EXPERIMENTS

Abstract

Local 2D motion direction of a line inside a circular aperture has ambiguous velocity but is typically perceived perpendicular (normal) to its line orientation. This vector normal solution minimizes distance and speed of a traveling line or edge in 2D. Here we investigate for the first time whether this basic principle extends to binocular 3D motion perception.

In two experiments we varied orientation and orientation disparity of a line stimulus that moved at a given horizontal interocular velocity difference behind a circular aperture. Using a psychophysical matching task we measured perceived 3D motion direction. Although human observers resolved ambiguity in a systematic way perceived motion directions did not follow geometric predictions. A geometric-statistical model however that favors slow motion and small disparity under ambiguity gave reasonable fits to individual data sets. Parameter estimates of this Bayesian model suggest small temporal noise in motion processing but large noise or uncertainty in the processing of (orientation) disparity.

We discuss implications of our results for the integration of motion and disparity information in the human visual system. It is concluded that the visual system minimizes speed and distance but that disparity processing introduces strong perceptual bias. This suggests late rather than early integration of motion and disparity information.

4.1 INTRODUCTION

Under natural viewing conditions the human visual system seems to effortlessly establish a 3D motion percept from local inputs to the left and right eye. The instantaneous integration and segmentation of binocular input is essential for object recognition, action planning and execution. It seems obvious that the visual system exploits many cues when establishing 3D motion perception in a natural environment. In a typical experimental setting however, motion and disparity input features prominently because both inputs have been related to early processing stages within the visual system (Howard & Rogers, 2002).

Any biologically plausible solution to binocular 3D motion perception has to rely on sampling of local spatio-temporal information in the left and right eye (Beverley & Regan, 1973; 1974; 1975). There are at least three known cell types in primary visual cortex (V1) that are involved in local encoding of 3D motion: motion detecting cells (DeAngelis, Ohzawa, & Freeman, 1993; Maunsell & van Essen, 1983), binocular disparity detecting cells (Hubel & Wiesel, 1970; Ohzawa, DeAngelis & Freeman, 1990), and joint motion and disparity detecting cells (Anzai, Ohzawa & Freeman, 2001; Bradley, Qian & Andersen, 1995; Carney, Paradiso, & Freeman, 1989; DeAngelis & Newsome, 1999; 2004; Pack, Born & Livingston, 2003).

It is therefore not surprising that different approaches to binocular 3D motion perception have emerged in the literature: (1) interocular velocity difference (IOVD) is based on monocular motion detectors (Lu & Sperling, 1995; Regan, Beverley, & Cynader, 1979; Shioiri, Saisho, & Yaguchi, 2000; Fernandez & Farell, 2005; Czuba et al., 2010), (2) changing disparity over time (CDOT) monitors binocular disparity detectors (Cumming & Parker, 1994; Regan, 1993; Peng & Shi, 2010), and (3) joint encoding of motion and

disparity (JEMD) relies on binocular motion and disparity detectors (Qian, 1994; Qian & Andersen, 1997; Morgan & Fahle, 2000).

If only motion or disparity input determines 3D motion perception as suggested by IOVD and CDOT then processing of other input needs to be silenced. This would require suppression of either motion or disparity input. Instead, the visual system may take advantage of both inputs (Bradshaw & Cumming, 1997; van Ee & Anderson, 2001; Brooks, 2002; Lages, Mamassian & Graf, 2003; Lages & Heron, 2008) and possibly a range of other cues (e.g., size, texture, shading, blur). This could be achieved through early joint encoding of motion and disparity inputs (JMED) or through parallel processing and late integration. As we show later in the chapter, the inherent ambiguity in early motion and depth encoding is reflected in systematic bias in perceived motion trajectory (see section 4.3). This favours a parallel encoding account, as availability of both IOVD and CDOT cues to motion trajectory does not necessarily result in a veridical representation of motion direction. It is more likely that the two sources of information are integrated further along in the visual processing hierarchy.

The Aperture Problem

Motion direction of a line moving inside a circular aperture is typically perceived as perpendicular to its orientation. This seminal finding on 2D line motion (Stumpf, 1911; Korte, 1915) and subsequent studies on the 2D aperture problem of motion (Wallach, 1935; Adelson & Movshon, 1982) suggest minimal displacement over time as a basic principle of motion perception (Hildreth & Koch, 1987; Marr & Ullman, 1981)

Under binocular viewing conditions a line or edge that moves behind a circular aperture on a 3D trajectory in depth also has ambiguous motion direction because not only the *x* and *y* component but also the *z* component of motion is unknown.

A moving oriented line or edge may not only have interocular velocity difference but also orientation disparity between the left and right eye. As a consequence motion direction and speed of an oriented line stimulus can be described by an infinite number of vectors.

Similar to the 2D aperture problem, the visual system needs to establish motion correspondence. In addition, it has to solve the stereo correspondence between the stimulus in the left and right eye to arrive at a 3D velocity estimate. Therefore, it appears straightforward that both motion and disparity processing are involved in the computation of local 3D velocity.
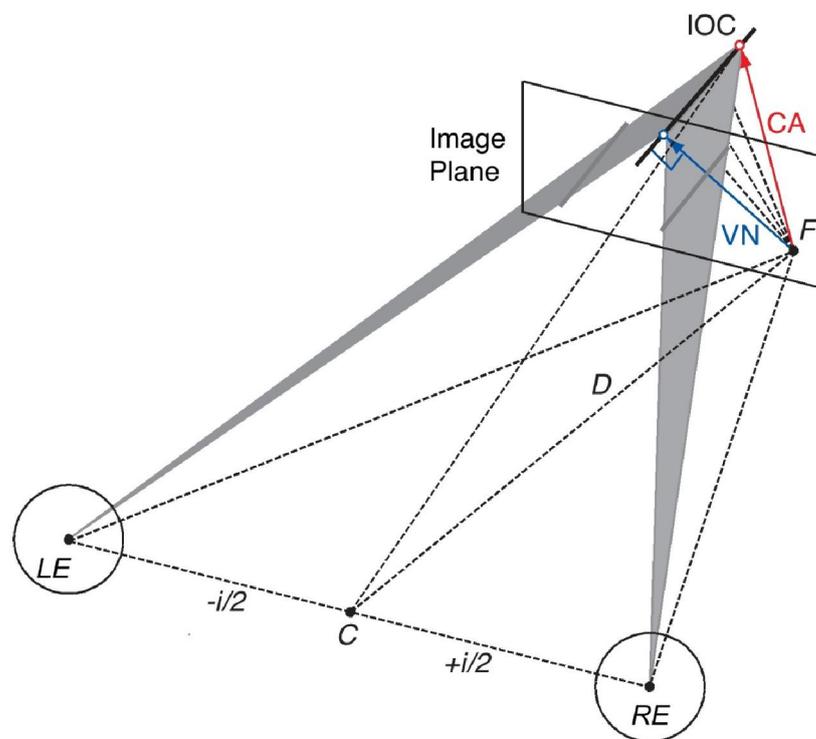
Figure 4.1 Illustration of binocular viewing geometry with left eye (LE) and right eye (RE) constraint planes (shaded triangles) and intersection of constraints (IOC) for a line moving in 3D starting at fixation point F. For better illustration the eyes are set apart by interocular distance *i* and the image plane is placed on the fixation point *F* at viewing distance *D*. Ambiguity in the aperture problem of 3D motion is illustrated by dashed arrows, vector normal (blue arrow), and cyclopean average (red arrow) describing

plausible default strategies for an IOC line slanted and tilted in 3D.

Many psychophysical studies have addressed the 2D aperture problem for motion (Wallach 1935; Adelson & Movshon, 1982; Welch, 1989; Yo & Wilson, 1992) and the 3D aperture problem for stereo (Morgan & Castet, 1997; Banks & Backus, 1998; Farell, 1998; van Ee & Schor, 2001; van Dam & van Ee, 2004).

Studies on structure from motion typically use random-dot stimuli, plaids or textured surfaces to measure surface orientation (Koendering, van Doorn & Kappers (1995); Ji & Fermuller, 2006), perceived shape (Hogervorst & Eagle, 2000) or volume of 3D objects (van Ee & Anderson, 2001). There is also a range of studies that have investigated the integration of stereo and motion cues (Richards, 1985; Johnston, Cumming & Landy, 1994; Todd & Norman, 2003; Domini, Caudek & Tassinari, 2006; Maloney & Landy, 1989; Scarfe & Hibbard, 2011).

Surprisingly however, there are no psychophysical studies that have systematically investigated perceived 3D motion direction of lines or edges in a circular aperture. The few existing studies used complex plaid stimuli and were restricted to the special case of frontoparallel motion in depth (Adelson & Movshon, 1984; Ito, 2003; Sakai, Ogiya & Hirai, 2011) or looming motion along the line of sight (Rokers et al., 2011).

In the following we try to fill this gap by studying perceived 3D motion direction of a single line moving in 3D behind a circular aperture. Since there are no explicit endpoints or texture elements and no other depth cues, local line motion direction remains highly ambiguous. From a computational point of view the aperture problem of binocular 3D motion is ill-posed because stereo and motion correspondence is underdetermined (Faugeras, 1992; Waxman & Duncan, 1986; Mayhew & Longuet-Higgins, 1982). If however

the visual motion system combines local motion constraints and disparity input to derive a local velocity estimate then perceived line motion direction may reveal principles of binocular 3D motion processing (Lages & Heron, 2010; see also, Chapter 2).

Geometric models of 2D motion perception for line and plaid patterns are based on intersection of constraints (IOC) or vector averaging (VA). Both IOC and VA models have been used to predict 2D motion direction of different plaid patterns (Adelson & Movshon, 1982; Yo & Wilson, 1992). In chapter 2, we suggest two plausible extensions of the geometric IOC and VA model in order to predict binocular 3D motion perception (see Fig. 4.1)

(1) First we establish left and right eye constraint planes as defined by the projections of the moving line stimulus onto the image plane and the nodal point of the left and right eye. The intersection of the two constraint planes defines an oriented constraint line but motion direction remains ambiguous. The shortest distance between the starting point and the constraint line gives the vector normal as a default estimate of 3D motion direction. Since this binocular IOC model uses the vector normal in 3D we simply refer to it as the VN model.

(2) Alternatively, the left and right eye vector normal may be extracted in the 2D image plane. If the 2D vector normal of the left and right eye are averaged a single motion constraint line through the cyclopean eye is obtained. If this vector average is then combined with binocular disparity between the left and right eye then this strategy provides a default estimate of local 3D motion direction. Since this binocular model uses the vector average and cyclopean point we refer to it as the cyclopean average or CA model.

If the line stimulus has fronto-parallel orientation with zero orientation disparity both default strategies make the same prediction. However, if we systematically vary orientation disparity between left and right eye projections on the image plane then the VN and CA model give different predictions (see Fig. 2.5).

Large and persistent perceptual bias have been reported for dot stimuli with unambiguous motion direction (Harris & Dean, 2003; Harris & Drga, 2005) suggesting processing strategies that do not follow simple geometric predictions (Lages, 2006, Ji & Fermuller, 2006; Welchman, Lam & Bulthoff, 2008). In addition to this, much larger bias has been reported for 2D motion direction of line stimuli, as a function of line tilt and line length, than random dot stimuli (Loffler & Orbach, 2001). Deviations from veridical were systematic and substantial, suggesting that bias was not fixed/ inherent in the visual system, but reflected the activity of a flexible processing mechanism (Loffler & Orbach, 2001).

Thus we try to model perception of local motion stimuli with ambiguous 3D motion direction, such as a line or contour moving inside a circular aperture, using the probabilistic approach outlined in Chapter 3.

4.2 MATERIALS AND METHODS

4.2.1 Ethics

Informed written consent was obtained from all observers before participation. Experiments and experimental procedures were approved by the Faculty Ethics Committee at Glasgow University in compliance with national legislation and the Code of Ethical Principles for Medical Research Involving Human Subjects of the World Medical Association (Declaration of Helsinki).

4.2.2 Participants

Observers were a convenience sample of two naive students (T.N. and S.W.) from Glasgow University and two authors (S.H. and M.L.). A fifth observer completed only a few sessions and the data were excluded. All observers had normal or corrected-to-normal visual acuity and were screened for stereo deficiencies (Random dot E test, Heron & Lages, 2012). Before testing each observer attended several training blocks. In the training blocks they received auditory feedback when their setting of the probe was opposite in depth to the actual 3D motion direction of the stimulus.

4.2.3 Apparatus

The stimulus and task was programmed in MatLab (MathWorks, Natick, MA) using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) and run on a Macintosh G4 computer with two 21 in Sony GDM-F500R cathode-ray tube flat screen monitors in a Wheatstone configuration. The monitors were calibrated for luminance using a Minolta photometer (Cambridge Research Systems). Stimuli were presented stereoscopically through haploscopic mirrors at a viewing distance of 55 cm at a frame rate of 120 Hz. Stimuli were shown at 50% Michelson contrast. Observers were comfortably seated in front of the mirrors with their head supported by a chin- and head-rest. The experimental room remained dark with lights switched off during testing.

4.2.4 Stimulus

As illustrated in Fig. 4.2 observers fixated a hairline cross at the centre of a black circular aperture surrounded by a uniform mid-gray screen (mean luminance on screen was 34 cd/m^2). The fixation cross was flanked by vertical nonius lines in order to maintain vergence. An oriented line stimulus of the same mid-gray as the surround moved back and forth on a horizontal 3D trajectory inside a circular aperture with diameter 4.83 deg. The line had oblique (45°, 135°) or vertical (90°) orientation, was blurred (Gaussian with SD=4 pixels) and blended perfectly with the gray surround revealing no explicit line endpoints. The midpoint of the line was positioned 2.48 cm behind the fixation cross at a distance of 57.48 cm. A depth probe consisting of a string of red dots could be called up

by the observer (via button press). Observers could indicate the perceived direction of motion of the line stimulus by adjusting the orientation and slant of the probe. The probe could also be hidden in order to provide an unobstructed view (again via button press). On any given trial, the stimulus remained on-screen until the observer had completed the matching task. To ensure maximum accuracy, the observer could view the stimulus without the depth probe as many times as required and controlled the length of trial by pressing a key when satisfied with the position of the probe.
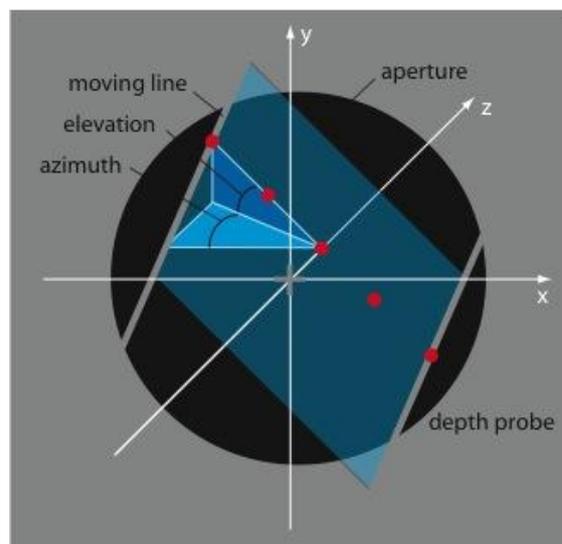
Figure 4.2 Illustration of stimulus display for motion direction-matching task. In each trial an oriented line (mid-grey) moved back and forth behind a circular aperture on an invisible plane (translucent blue) slanted and tilted in 3D space. The observers' task was to fixate the cross at the center and to adjust orientation and horizontal disparity of a string of red dots until they matched the perceived direction of the moving line. No endpoints of the line stimulus were visible. Translucent plane and axes did not appear in the stimulus display and are shown to illustrate azimuth and elevation angle of stimulus and probe in

3D.

The stimulus line oscillated horizontally from left to right and back with a horizontal velocity of 3.0±0.23 deg/s on the left and right screen of the stereoscope. The horizontal interocular velocity difference (iovd) was ±0.46 deg/s. In spherical co-ordinates this motion corresponds to an azimuth angle of 52.5° from fronto-parallel at 0° elevation from horizontal. Note that this represents the ground truth for perceived motion direction in all conditions and experiments.

4.2.5 Procedure

A matching task was used to measure perceived 3D motion direction in open-loop trials. After observing repeated oscillations of a single line behind the aperture, the observer pressed the spacebar to reveal a string of five red dots equally spaced by 18 pixels and centered on the midpoint, inside the aperture. The observer adjusted orientation (between 0 and 180°) and horizontal disparity until the probe matched the perceived motion direction of the line stimulus. Adjustments spanned +36.5 arcmin crossed to -30.9 arcmin uncrossed for the two outer dots of the depth probe. The two inner dots always assumed half the disparity of the neighboring outer dots and the midpoint remained centered at -9.1 arcmin behind fixation. After adjustments by the observer the line appeared to move through the string of red dots with minimal displacement and shearing in depth. Each observer made online adjustments of orientation and horizontal disparity of the probe by pressing corresponding keys on a keyboard. When observers adjusted orientation the string of dots rotated around the midpoint and when they adjusted horizontal disparity the string of probe dots changed slant about the vertical axis with the midpoint anchored at the same depth. The observer could toggle between appearance and disappearance of the probe dots by pressing the spacebar. Once the observer was confident that the adjusted probe matched perceived motion direction or surface

orientation of the line stimulus they pressed a separate key to confirm orientation and disparity setting and to continue with the next trial.

The slant about the vertical axis of the depth probe was not perceived veridically and disparity bias in the probe was assessed in a separate task. In this calibration task observers had to adjust the probe to the perceived motion direction before they indicated perceived horizontal azimuth at 0° elevation on a protractor aligned with the fronto-parallel image plane. On average observers underestimated azimuth of the probe by -20°, or by a factor of 0.6. This corresponds to previously reported underestimation of surface slant about the vertical axis (Cagnello & Rogers, 1993; Mitchison & McKee, 1990; Ryan & Gillam, 1994).

Disparity settings of the probe in Experiment 1 and 2 were calibrated accordingly before the data were transformed into perceived motion directions (azimuth and elevation in spherical co-ordinates) using a ray-tracing method.

4.2.6 Design

In four experiments, observers adjusted an online depth probe to indicate either the perceived direction of motion of an oscillating line stimulus or the surface slant of a static display of three parallel lines of known orientation disparity:

Experiment 1(A): 45°/135° motion condition. Orientation disparity ranged between -6 and +6° of orientation disparity. IOVD was (3.0±0.23 deg/s). In randomly inter-mixed trials the observer adjusted the orientation and disparity of the online depth probe to indicate the perceived direction of motion-in-depth.

Experiment 1(B): 45°/135° static condition. Orientation disparity ranged between -6 and +6° of orientation disparity. In randomly inter-mixed trials the observer adjusted the

orientation and disparity of the online depth probe to indicate the perceived surface slant of the display. This allowed us to investigate the involvement of disparity in biasing trajectory judgments in experiment 1(A).

Experiment 2(A): 90° motion condition. Orientation disparity ranged between -6 and +6° of orientation disparity. IOVD was (3.0±0.23 deg/s). In randomly inter-mixed trials the observer adjusted the orientation and disparity of the online depth probe to indicate the perceived direction of motion-in-depth.

Experiment 2(B): 90° static condition. Orientation disparity ranged between -6 and +6° of orientation disparity. In randomly inter-mixed trials the observer adjusted the orientation and disparity of the online depth probe to indicate the perceived surface slant of the display. This allowed us to investigate the involvement of disparity in biasing trajectory judgments in experiment 2(A).

Matching of motion direction of an oblique line in Exp. 1A and a vertical line in experiment 2A or surface orientation of oblique lines in Exp. 1B and vertical lines in experiment 2B was repeated four to five times in blocks of 28 trials with randomized iovd (3.0±0.23 deg/s), line orientation (Exp. 1A and 1B: 45° and 135°) and orientation disparity (6°) across trials.
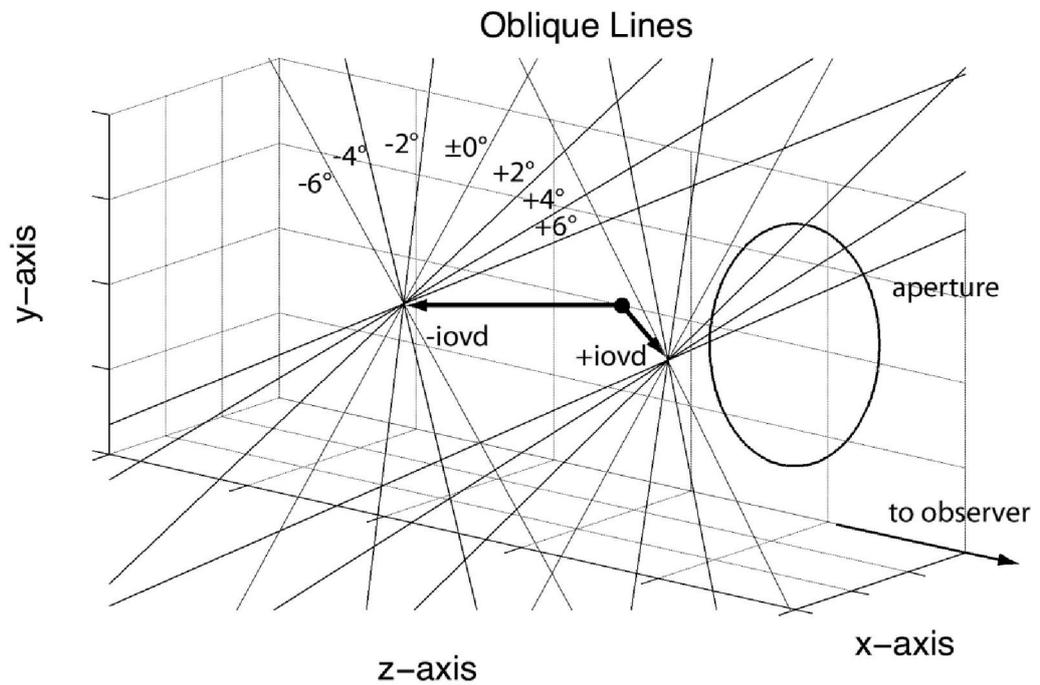
Figure 4.3 Illustration of horizontal trajectories (black arrows) for oblique line stimulus(-iovd/back and +iovd/front) with orientation disparity ranging from -6° to +6° in steps of 2°.

Conditions in which the stimulus was mirrored along the vertical axis (e.g., 45° line with +iovd and 135° line with -iovd) gave very similar settings and these trials were collapsed for each observer, iovd, and orientation disparity. Results are always shown for collapsed data and displayed for stimuli left of fixation (front/+iovd, back/-iovd; see Fig. 4.3).

4.2.7 Model Fits and Model Selection

The orientation and disparity settings made by observers were transformed using a ray-tracing algorithm into perceived motion directions (azimuth and elevation in spherical co-ordinates). We were, therefore, able to generate a visual representation of the motion vectors by plotting the co-ordinates of the end-point of the lines' motion in 3D space and showing the motion path from fixation to this point. The azimuth and elevation settings over the range of motion trajectories were plotted for individual observers. For experiments 1(B) and 2(B), these corresponded to the perceived surface slant.

We fit the generalized Bayesian models from Chapter 3 to the data from each of the experiments, in order to quantify how much perceived motion direction is influenced by temporal noise from motion processing and spatial noise from orientation disparity processing. We focused on probabilistic extensions of the VN model (BVN, BDVN) since the CA predictions did not correspond to the observed data.

Maximum-likelihood fits were obtained by the fminsearch() routine in MatLab (MathWorks, Natick MA) which minimized the difference between observed and predicted azimuth and elevation angles. By fitting noise parameters for velocity (σ_v) and orientation disparity (σ_d) processing in the Bayesian Disparity and Vector Normal (BDVN) model while keeping the motion and disparity prior constant (σ= 1:0) we can quantify the ratios between noise in the likelihood and prior (Hurlimann, Kiper, & Carandini, 2002) for each observer, condition (2 iovds; 7 orientation disparities), and experiment (vertical and oblique; moving and static lines). Similarly, we fitted only the noise parameter for velocity processing (σ_v) in the Bayesian Vector Normal (BVN) model.

We compared the two-parameter BDVN model with the one-parameter BVN model approximating the Bayes Factor (BF) by the Bayesian Information Criterion (BIC) for nested models (Raftery, 1995, 1999). The BF is a method for comparing two models,

based on the ratio of the marginal likelihoods of the two models rather than the maximum likelihoods. Importantly this comparison does not assume normality or large samples. The BIC is applied to control for the possibility of increasing the likelihood for the BDVN model, which can occur simply by the addition of a second parameter. The potential for over-fitting is overcome by imposing a penalty depending on the number of parameters in the model. This ensures that the BDVN model is not falsely selected. See section 3 in chapter 3 for more details

4.3 PSYCHOPHYSICAL RESULTS

In two psychophysical experiments with moving stimuli and two controls with static stimuli we studied perceived 3D motion direction under ambiguity. In an (open-loop) matching task without time limit observers indicated perceived motion direction (azimuth and elevation in spherical co-ordinates) of an oblique (Exp1A) and vertical (Exp.2A) line stimulus while orientation disparity between the left and right eye image varied across trials. In two control experiments (Exp.1B and 2B) observers adjusted perceived surface orientation for three static lines whose orientation disparity varied across trials.

4.3.1 Experiment 1A. Motion Direction of Oblique Line

The stimulus was a single moving line with an oblique orientation of 45 and 135 from horizontal inside a circular aperture. Monocular velocities were set to 3:0 0:23 deg/s with an interocular velocity difference (iovd) of 0.46 deg/s on the horizontal axis. Orientation disparity between lines on the image plane was systematically varied between 6 in steps of 2 giving rise to the perception of a line tilted about the horizontal x-axis. Line orientation, sign of iovd, and orientation disparity was randomly inter-mixed in blocks of trials. In each trial an observer adjusted a probe of dots in a stereoscopic display to indicate perceived motion direction (see Methods and Materials). Adjustments for each

observer and trial were transformed into corresponding azimuth (slant about the vertical axis) and elevation (tilt about the horizontal axis) using a ray-tracing algorithm.

Predictions of VN and CA for Moving Oblique Line
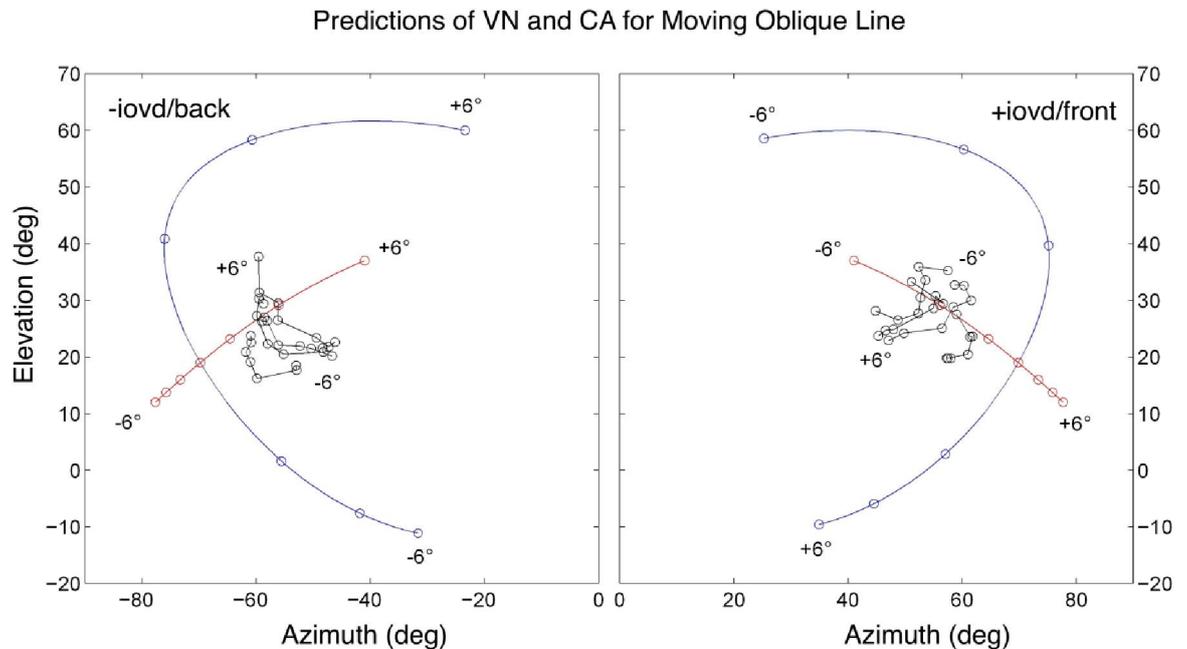
Figure 4.4 Overview of geometric predictions and results from 3D motion direction adjustments for moving oblique lines (45 and 135°) with orientation disparity ranging between ±6° in steps of 2°. The data of four observers (black circles) are shown for motion trajectory -iovd/back and +iovd/front in separate plots. Superimposed are the predictions from the geometric VN (blue circles) and CA (red circles) strategy.

Conditions in which stimulus orientation and motion direction were mirrored about the vertical axis (e.g., line orientation 45 with +iovd and 135 with -iovd) gave very similar settings and these trials were collapsed for each observer.

First, we compared individual data settings to the horizontal stimulus direction and found systematic departures from ground truth in all conditions. We then compared the data to predictions of the geometric VN and CA model (see Fig. 4.4) and found little agreement with the CA strategy as data points (black circles) are not aligned with the CA predictions

(red circles). Instead, the data points show a curvature characteristic for the VN strategy (blue circles) but with reduced azimuth and elevation angles.

We applied the Bayesian Vector Normal (BVN) model with one and the Bayesian Disparity Vector Normal (BDVN) model with two free parameters (see Chapter 3). In this processing approach the first parameter captures noise due to motion processing whereas the second parameter reflects noise from (orientation) disparity processing. The BVN model assumes noisy velocity inputs are responsible for biasing perceptual judgments and therefore the response pattern will favour the basic VN strategy of selecting the shortest motion vector in 3D (see chapter 2 for a full description of the VN model). The BDVN model assumes noisy monocular motion inputs coupled with noisy disparity processing. If observers use a strategy similar to the BDVN model, then perceptual judgments should deviate markedly from predictions of the VN model and should reflect a tendency towards zero disparity (or tendency to underestimate orientation disparity in the display).See also the model simulations in chapter 3 (section 3.2.2 and 3.2.3)

Figure 4.5 plots the azimuth (deg) against the elevation (deg) over the six orientation disparities (-6, -4, -2, 0, 2, 4, 6) for each of the four observers at the two IOVD settings (-iovd/back; +iovd/front) (Black circles). The predictions of the BDVN model over the range of orientation disparities and IOVD settings are plotted alongside (Blue circles). It can be seen from the plots that the BDVN model reasonably predicts the empirical data.

We fitted the BVN and BDVN model to individual data and the results are summarized in Table 4.1. The parameter estimates suggest a strong influence of a disparity prior centered on zero orientation disparity ($\sigma_d : \sigma > 1.0$) and a weak influence of a motion prior centered on zero velocity ($\sigma_v : \sigma < 0.1$) in all four observers. Note that a BVN (BDVN) with small noise σ_v (and small noise σ_d) approximates the geometric VN strategy.

We also measured monocularly perceived 2D motion direction of the left and right eye

stimulus in separate conditions and blocks. As expected the results suggest that each observer perceived 2D line motion approximately perpendicular to line orientation (results not shown).
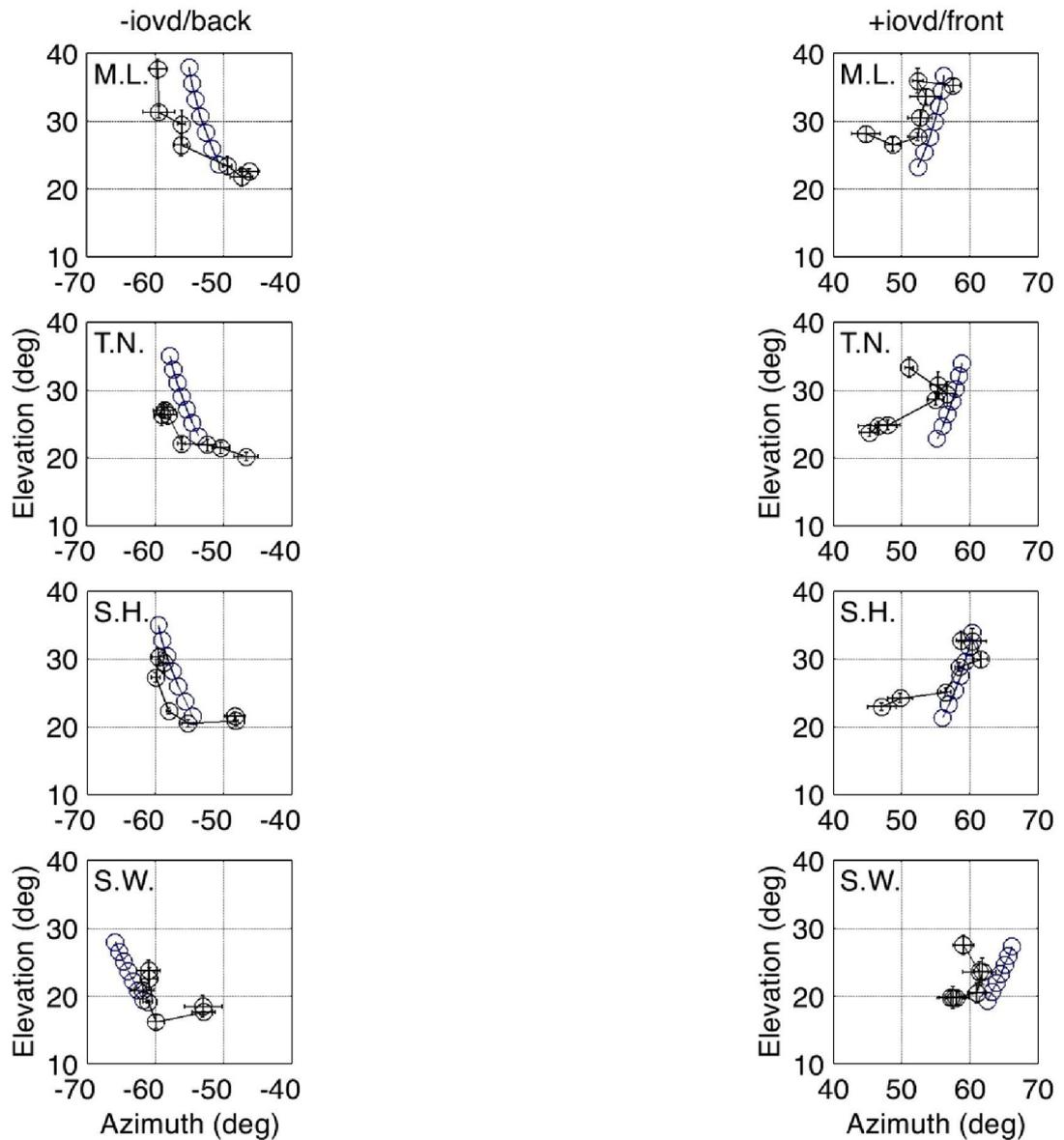
Figure 4.5 Results from Experiment 1A: 3D motion direction adjustments for an oblique line (45 and 135) with orientation disparity ranging between ±6° in steps of 2°. Data points

averaged across four or five repeated trials are shown as black circles with horizontal and vertical error bars (1 SEM) for azimuth and elevation, respectively. The data of four observers are plotted in rows and for the two motion trajectories (-iovd/back and +iovd/front) in columns. The BDVN model with two parameters (blue circles) gives reasonable results (see Table 4.1).

4.3.2 Experiment 1B. Surface Orientation of Static Oblique Lines

In this control experiment we investigated the effect of orientation disparity on perceived surface orientation of static oblique lines. Three copies of the line stimulus in Exp.1A were displayed side by side suggesting an oriented surface slanted and tilted in 3D space. The three lines were sampled from the line motion experiment. They corresponded to the start and the two inflection points of the line motion

As before, we varied orientation disparity of the lines. In open-loop trials each observer was instructed to adjust azimuth (slant about the vertical axis) and elevation (tilt or direction of slant about the horizontal axis) of the probe so that the string of five dots matched the perceived orientation of the surface.

Oblique Static Lines
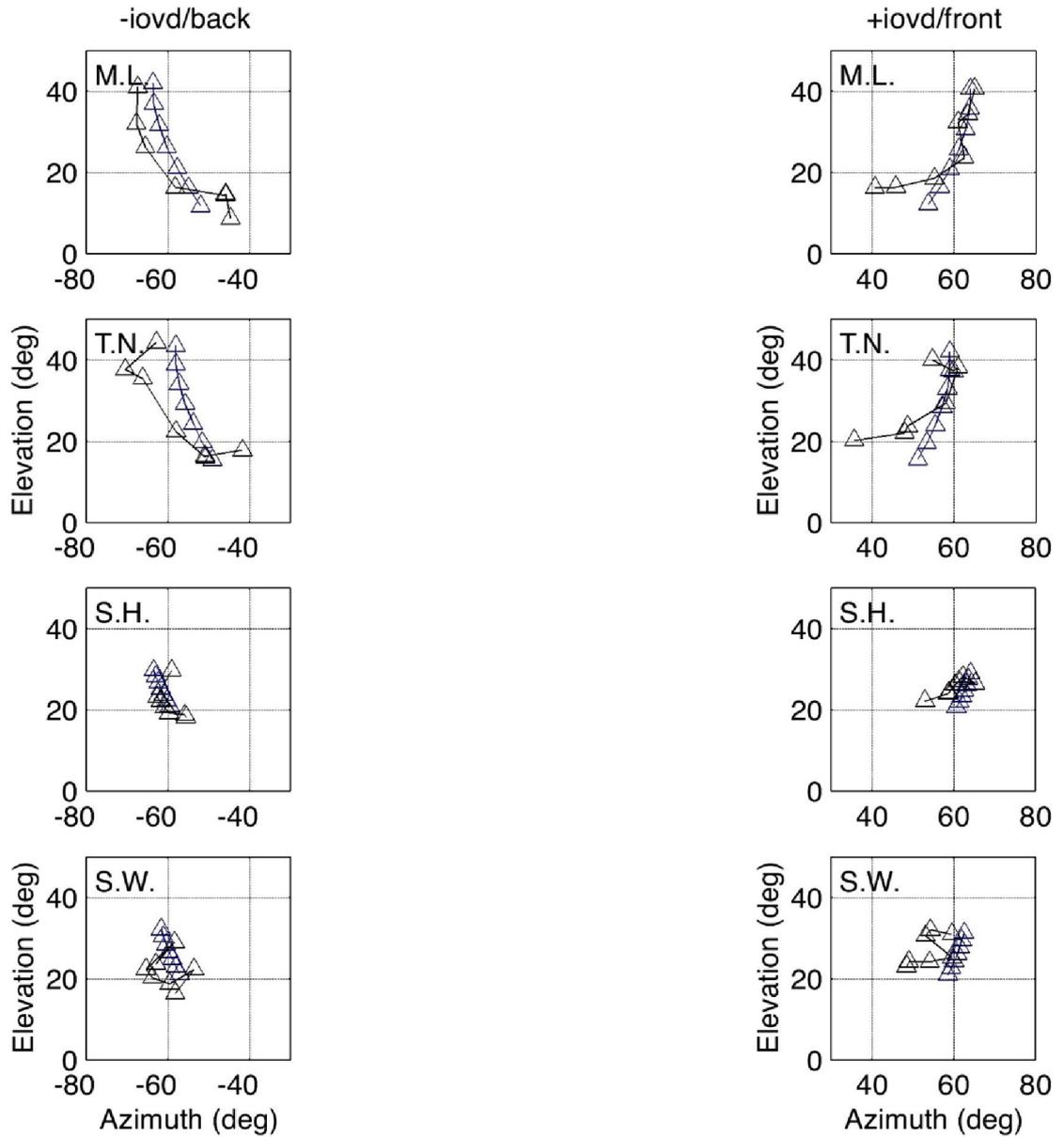
Figure 4.6 Results from Experiment 1B. 3D surface orientation adjustments for three oblique lines (45 and 135) with orientation disparity ranging between ±6° in steps of 2°. Data points averaged across four or five repeated trials are shown as black circles with horizontal and vertical error bars (1 SEM) for azimuth and elevation, respectively. The data of four observers are plotted in rows and for two motion trajectories (-iovd/back and

+iovd/front) in columns. The BDVN model with two parameters (blue circles) gives reasonable results (see Table 4.1).

Perceived orientation of the static surface closely followed perceived direction of the moving line as is evident when comparing the adjustments in Fig. 4.5 and 4.6, respectively. However, compared to the motion task in Exp. 1A, uncertainty in orientation disparity σ_d was reduced for Observer T.N. and M.L. Estimates of σ_v on the other hand remained almost constant across observers and experiments (see Table 4.1). In the context of surface orientation parameter σ_v of the Bayesian model may be interpreted as noise in perceived position and as preference for shorter distances from the centre of the oriented surface.

Table 4.1. Bayesian estimates (for σ = 1:0) and model selection for Exp. 1A motion direction of oblique line, and Exp. 1B surface orientation of static oblique lines

Obs.	BDVN			BVN		Mod Sel
Exp. 1A	$\sigma_v : \sigma$	$\sigma_d : \sigma$	$\chi^2(11)$	$\sigma_v : \sigma$	$\chi^2(12)$	BF
M.L.	0.056	2.59	8.26	0.088	831.5	376.6
T.N.	0.050	2.95	15.4	0.086	766.1	186.5
S.H.	0.047	2.76	8.71	0.085	850.9	365.5
S.W.	0.031	3.66	12.5	0.083	923.7	275.6
Exp. 1B	$\sigma_v : \sigma$	$\sigma_d : \sigma$	$\chi^2(11)$	$\sigma_v : \sigma$	$\chi^2(12)$	BF
M.L.	0.041	1.67	25.0	0.080	854.6	127.7
T.N.	0.051	1.71	20.3	0.083	779.9	143.5
S.H.	0.037	3.54	7.51	0.083	944.7	470.7
S.W.	0.042	3.13	16.2	0.084	901.3	208.2

For both experiments we compared the one-parameter Bayesian Motion (BVN) with the two-parameter Bayesian Disparity and Motion (BDVN) model. Similar to a conventional

likelihood-ratio test we used Bayesian Information Content (BIC) that approximates the Bayes Factor (BF) as a statistical measure of model evidence.

Similar to a conventional likelihood-ratio test we used Bayesian Information Content (BIC) that approximates the Bayes Factor (BF) as a statistical measure of model evidence. Here we made the prediction that the BDVN model would provide a better fit to the data than the BVN model, even when controlling for the number of parameters (BIC penalty for the two-parameter model). A Bayes Factor >10 is considered to be strong evidence in favour of the BDVN model. Thus in both experiments and in all observers the BDVN model was strongly favoured. Assuming that the two model variants are equally plausible the posterior probability accumulated over observers was less than .99 in each condition of the two experiments (see chapter 3 for a full description of the model selection process). This constitutes very strong evidence in favor of the BDVN model (Raftery, 1995, 1999). Thus, a Bayesian VN strategy in 3D space with a strong bias for zero orientation disparity provides a far better account of perceived motion direction as well as perceived surface orientation than the BVN model without bias in disparity processing.

4.3.3 Experiment 2A. Motion Direction of Vertical Line

The following experiment was identical to Exp. 1A except that the stimulus was a single moving line with vertical orientation of 90° inside the circular aperture. Please note that in this condition the predictions for the CA strategy remain more or less constant whereas the VN strategy describes a symmetric arc (see Fig. 4.8). As in Exp. 1A monocular velocities were set to 3.0±0.23 deg/s with an interocular velocity difference (iovd) of 0.46 deg/s along the horizontal axis. Orientation disparity between the line stimuli on the image plane was systematically varied, ranging between ±6° in steps of 2°. Observers adjusted the probe in the stereoscopic display in order to indicate perceived motion direction (azimuth and elevation).

As before, we compared the adjustments of each observer to the ground truth and predictions of the geometric VN and CA model and found little overall agreement (see

97

Fig.4.6). Compared to VN predictions (blue circles) the individual data points (black circles) show reduced azimuth and elevation. In contrast CA predictions (red circles) are almost constant due to the symmetric disparity offsets in the image plane.

Predictions of VN and CA for Moving Vertical Line
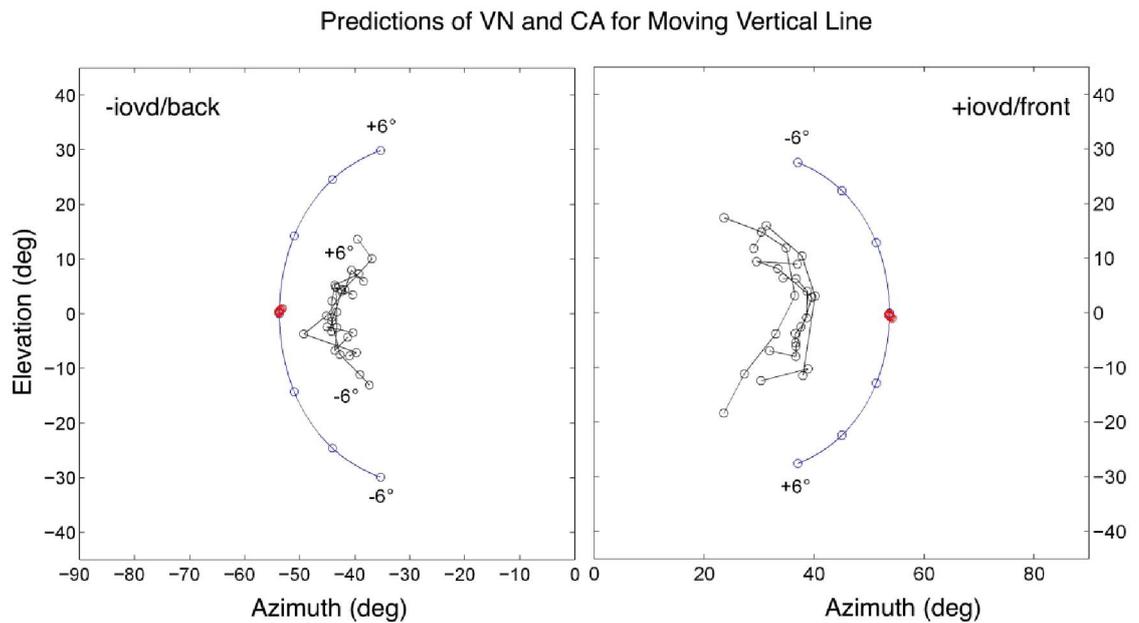
Figure 4.7 Overview of predictions and results from 3D motion direction adjustments for moving vertical line (90°) with orientation disparity ranging between ±6° in steps of 2°. The data of four observers (black circles) are shown for motion trajectory -iovd/back and +iovd/front in separate plots. Superimposed are the predictions from the geometric VN (blue circles) and CA (red circles) model. Empirical data adheres more closely to the pattern predicted by the VN model, but observer judgements are extremely conservative in terms of azimuth and elevation. Worthy of note is the consistency in response between the observers, which suggests that the settings are not arbitrary and are the result of some underlying computational strategy.

In Figure 4.8 the azimuth (deg) and elevation (deg) adjustments for four observers over the range of orientation disparity settings for IOVD to the back (-iovd) and to the front of fixation (+iovd), are plotted as black circles.

Alongside, the predictions of the BDVN model over the range of conditions are plotted as blue circles. The BDVN model predicts the empirical data reasonably well in comparison to the simple computational models (CA, VA) previously described. This was also the case for the oblique line condition and so the utility of the BDVN model for predicting responses to the range of 3D trajectories, appears to extend to lines of different orientations.

We fitted the Bayesian VN model with one or two free parameters. Again, the first parameter describes temporal noise whereas the second parameter indicates noise for orientation disparity between the left and right image. Parameter estimates and model selection are summarized in Table 2. As in Exp. 1A, the estimates indicate a strong influence of the small orientation disparity prior but relatively weak influence of the slow motion prior. Note that relatively small temporal noise $\sigma_v < 0.1$ approximates the VN strategy in 3D.

Vertical Moving Line
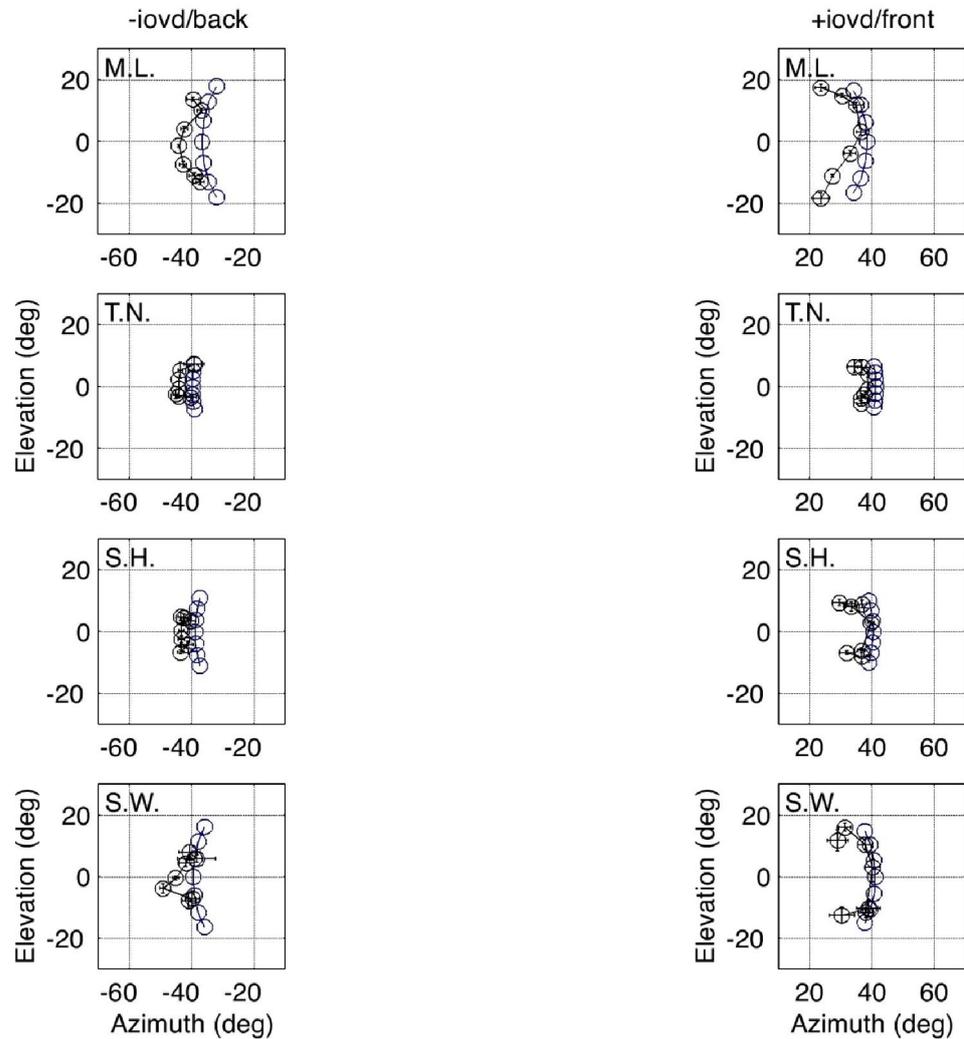
Figure 4.8 Results from Experiment 2A: 3D motion direction adjustments for a vertical line (90°) with orientation disparity ranging between ±6° in steps of 2°. Data points averaged across four or five repeated trials are shown as black circles with horizontal and vertical error bars (±1 SEM) for azimuth and elevation, respectively. The data of four observers are plotted in rows and for two motion trajectories (-iovd/back and +iovd/front) in columns. The BDVN model with two parameters (blue circles) gives reasonable results (see Table 4.2).

100

4.3.4 Experiment 2B. Surface Orientation of Vertical Lines

As in Exp.1B copies of the line stimulus from three time points were displayed side by side suggesting an oriented surface slanted about the vertical axis and tilted about the horizontal axis. In open-loop trials each observer was instructed to adjust azimuth (slant about the vertical axis) and elevation (tilt or direction of slant) of the probe so that the string of dots matched the perceived orientation of the surface. Similar to a single moving line these adjustments may be the result of biased orientation disparity as well as a preference for the shortest distance in 3D space.

Indeed, perceived orientation of the surface show similar characteristics as perceived direction of the moving vertical line. The corresponding results of Exp.2A and 2B suggest that the vector normal strategy for line motion may also apply to perceived orientation of a static surface, minimizing distance in 3D space. In addition, perceived azimuth (slant about the vertical axis) and elevation (tilt or direction of slant) of the static surface is similarly biased by orientation disparity as perceived motion direction.

Following the results of the earlier experiments (1(A/B), 2A), it was anticipated that the azimuth and elevation settings (deg) of observers in response to the range of orientation disparity settings (-6, -4, -2, 0, 2, 4, 6°) would closely follow the pattern predicted by the two parameter BDVN model. In figure 4.9, the azimuth and elevation settings for each of the experimental conditions are plotted against the predictions of the BDVN model. The pattern of responses of our observers seems reasonably close to the predictions of the model.

Similar to our findings for moving and static oblique stimuli, the response pattern for static vertical stimuli does not vary significantly from its moving counterpart. Together with the proven utility of the BDVN model, which assumes a strong influence of the

disparity prior, it seems that trajectory judgments are strongly influenced by a disparity prior.
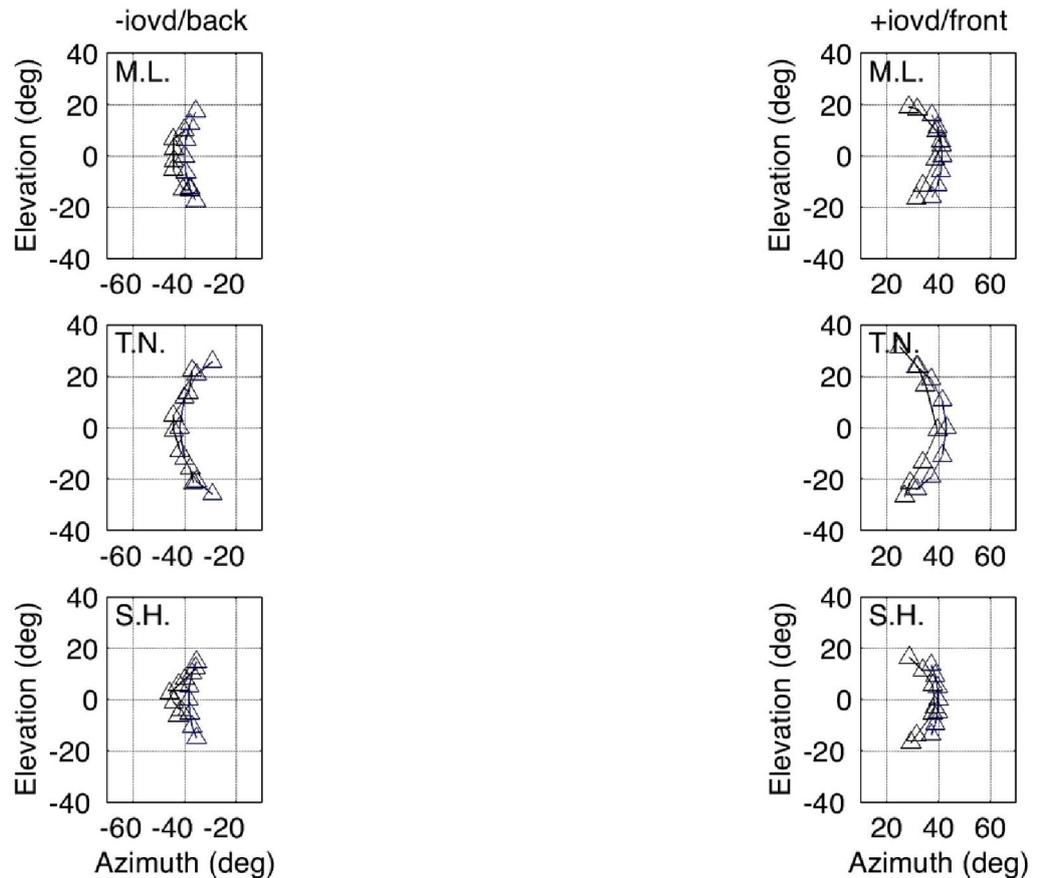
Figure 4.9 Results from Experiment 2B: 3D surface orientation adjustment for three vertical lines (90°) with orientation disparity ranging between ±6° in steps of 2°. Data points averaged across up to five repeated trials are shown as black triangles with horizontal and vertical error bars (±1 SEM) for azimuth and elevation, respectively. The data for three observers are plotted in rows and for two motion trajectories (-iovd/back and +iovd/front) in columns. The BDVN model with two parameters (blue triangles) gives reasonable results (see Table 4.2).

Table 4.2. Bayesian estimates for $\sigma = 1{:}0$ and model selection for Exp. 2A motion direction of vertical line, and Exp. 2B surface orientation of static vertical lines.

Obs.	BDVN			BVN		Mod Sel
Exp. 2A	$\sigma_v : \sigma$	$\sigma_d : \sigma$	$\chi^2(11)$	$\sigma_v : \sigma$	$\chi^2(12)$	BF
M.L.	0.061	1.98	10.8	0.062	358.2	123.6
T.N.	0.070	0.79	27.8	0.070	599.9	80.71
S.H.	0.064	1.44	30.5	0.064	418.6	51.32
S.W.	0.062	1.00	40.1	0.062	521.4	49.62
Exp. 2B	$\sigma_v : \sigma$	$\sigma_d : \sigma$	$\chi^2 11)$	$\sigma_v : \sigma$	$\chi^2(12)$	BF
M.L.	0.061	0.92	28.8	0.055	910.1	115.6
T.N.	0.056	0.19	29.5	0.060	575.3	74.64
S.H.	0.065	1.08	26.8	0.066	708.3	98.93

For both experiments we compared the one-parameter Bayesian Motion (BVN) with the two-parameter Bayesian Disparity and Motion (BDVN) model approximating the Bayes Factor (BF) by the Bayesian Information Content (BIC). Here we made the prediction that the BDVN model would provide a better fit to the data than the BVN model, even when controlling for the number of parameters (BIC penalty for the two-parameter model). A Bayes Factor >10 is considered to be strong evidence in favour of the BDVN model. Thus in both experiments and in all observers the BDVN model was strongly favoured. Assuming that the two model variants are equally plausible the posterior probability accumulated over observers was less than .99 in each condition of the two experiments. Again, this constitutes very strong evidence in favor of the BDM model (Raftery, 1995, 1999). Thus, a Bayesian VN strategy in 3D space with a strong bias for zero orientation

disparity of vertical lines slanted about the horizontal axis provides a better account of perceived motion direction as well as perceived surface orientation than the BVN model without disparity bias.

4.4 DISCUSSION

What enables the visual system to instantaneously perceive 3D motion and to infer direction and speed of a moving object? It seems likely that the visual system exploits many cues in concert (multiplexing) to make this difficult inference as reliable and veridical as possible. The diverse set of effective local and global cues as documented in psychophysical studies (Bradshaw & Cumming, 1997; van Ee & Anderson, 2001; Lorenceau, Shifrar, Wells & Castet, 1993) already points at late rather than early integration within the visual processing hierarchy.

In four experiments we measured observer perceptions of lines oscillating in depth behind a circular aperture (Experiment 1A, 2A) and lines arranged to define a surface slanted in depth (Experiment 1B, 2B). Observers made adjustments to show the perceived direction in depth of the displays and from this we derived the perceived azimuth (deg) and elevation (deg). The adjustments made by observers across a range of orientation disparities (-6, -4, -2, 0, 2, 4, 6) at two IOVD settings (front/back of fixation) were compared to the predictions of two computational models.

The first the Vector Normal model (VN) assumed that monocular velocities, defined as constraint planes converge on an intersection of constraints line and the solution is given by the slowest motion path in 3D space. The Cyclopean Average model (CA), assumes early motion averaging of monocular motion inputs, the vector average is then combined with disparity information in a subsequent stage to give the cyclopean average.

The empirical data from four observers was compared to the predictions of these two models, in all cases we showed that the VN model gave a more reasonable fit to the data than the CA model. However, our observers showed a strong perceptual bias towards underestimating (orientation) disparity in the display. To understand the results further we compared the data to the predictions of two Bayesian models, both based on the VN model, the BVN and BDVN models are described fully in chapter 3. In the BVN model, we estimated only one free parameter, reflecting a preference for slow motion. This was simply a probabilistic version if the VN model.

In terms of the Bayesian disparity vector normal model (BDVN) we estimated two free parameters. The first parameter expresses a weak preference for slow motion whereas the second parameter reflects a strong disparity bias under ambiguity. For some conditions and observers the ratio between disparity noise and prior exceeds 1.0. This indicates a very strong perceptual bias towards the fronto-parallel plane at fixation. The BDVN model gave the most reasonable fit to the data, across all experiments and observers. This suggests that under conditions of ambiguity there is a bias towards zero disparity, or a strong tendency to underestimate disparity. Given this information, we cannot discount the role of disparity information in computing motion-in depth.

Underestimation of surface slant is a well-known phenomenon in depth perception (Cagnello & Rogers, 1993; Mitchison & McKee, 1990; Ryan & Gillam, 1994). Our finding that perceptual bias for motion direction is comparable to surface orientation suggests that (orientation) disparity processing significantly contributes to binocular 3D motion perception. Disparity in intermediate processing stages in V2 (Bakin Nakayama, Gilbert, 2000; Thomas, Cumming & Parker, 2002) and V4 (Hinkle & Connor, 2002) may influence motion processing within the visual hierarchy (Maunsell & van Essen, 1983b). The separate pathways that connect V1 with MT/V5 directly or via V2 and/or V4 (Maunsell & van Essen, 1983b) are good candidates for parallel processing and late integration of disparity and motion information.

The BDVN model therefore suggests that perceived 3D motion direction in our experiments is the result of two processing stages or streams (Lages & Heron, 2008; 2010): (1) a dynamic disparity system that tracks features such as corners, endpoints and junctions and (2) a 3D motion system that disambiguates velocity constraints according to the vector normal.

The combination of motion and dynamic feature tracking provides a biologically plausible explanation of the behavioral characteristics and estimated parameters.

The notion of parallel pathways feeding functionally different aspects of motion into a later stage has been advanced in the context of 2D speed perception (Braddick, 1974; 1980), 2D pattern motion (Adelson & Movshon, 1982; Weiss et al., 2002; Wilson, Ferrera, & Yo, 1992), eye movements (Rashbass & Westheimer, 1961; Masson & Castet, 2002), and the processing of higher-order motion (Ledgeway & Smith, 1994; Lu & Sperling, 2001). Surprisingly however, the functional aspects of motion and (dynamic) disparity processing were not sufficiently addressed in the context of binocular 3D motion perception (Lu & Sperling, 2001; Regan, Beverley, Cynader, & Lennie, 1979).

Combining disparity or depth information with velocity constraints at a later stage provides a flexible scheme that can exploit intermediate depth processing such as relative disparity in V2 (Bakin Nakayama, Gilbert, 2000; Thomas, Cumming & Parker, 2002) and orientation disparity in V4 (Hinkle & Connor, 2002). Furthermore, it seems possible that velocity constraints may be processed in the dorsal stream and (dynamic) binocular disparity and other depth cues in the ventral stream (Ponce, Lomber & Born (2008). Although as discussed in chapter 5 it is thought that dorsal and ventral streams differ only in their perceptual goals but are almost identical in terms of signal content. This is supported by the fact that the dorsal and ventral stream show extensive crosstalk which can be traced back to early visual processing regions (Lewis & Van Essen, 2000; Nassi & Callaway, 2009). Therefore parallel processing of IOVD and CDOT information need not

occur in strictly segregated pathways and may be more likely to be integrated gradually over several levels of processing, before a veridical representation of motion-in-depth is achieved in a later region.

It is also neuroanatomically and neurophysiologically plausible that integration of motion and disparity occurs late in subregions of human MT/V5 (Movshon, Adelson, Gizzi, & Newsome, 1985; DeAngelis & Newsome, 2004; Majaj, Carandini & Movshon, 2007; Orban 2008; Rokers, Cormack & Huk, 2009) if not in areas beyond MT (Likova & Tyler, 2007).

It is tempting to assume that gradients, velocity constraints, and motion prior as postulated in the binocular Bayesian model are the immediate result of motion encoding and processing. However, stereo correspondence between oriented lines in a binocular viewing geometry requires a contribution from (orientation) disparity processing. Therefore, the intersection of velocity constraints may also be understood as frequently tracking binocular disparity input rather than smooth constraint planes intersecting in 3D velocity space.

The comparable adjustments for moving and static line stimuli also suggest that (dynamic) disparity processing of extrinsic terminators and related features disambiguate not only 3D surface orientation but also 3D motion direction. In particular, the pattern of response bias for moving and static stimuli follows a similar pattern, suggesting that a default strategy for matching the static stereo component of the stimulus (position of the three lines displaced in depth) may also be applied when the line is displaced over time i.e. tracking change in disparity over time of unique stimulus features. Thus, the motion prior in our Bayesian model may not necessarily reflect a preference for slow velocity but a preference for shortest distance in 3D space. Note that the only features that are available in both the static and moving line stimulus are orientation disparity between lines and disparity between junctions with the aperture (extrinsic terminators or endpoints).

Due to the size of our experimental stimulus it seems unlikely that visual processing occurred locally within a single receptive field. The line stimulus and aperture suggest that not only multiple receptive fields contribute to perceived motion direction, but also higher-order features and their "read-out" at intermediate and higher processing levels (DeAngelis & Newsome, 2004; Majaj, Carandini, & Movshon, 2007). For example, the moving line stimuli in our monocular displays create a moving lower and upper junction (endpoint or terminator) with the aperture. It is plausible, based on what we know about the 2D aperture problem, that the visual system can use these line terminators (points where the line meets the aperture) to disambiguate motion direction. In this case, as shown previously in 2D motion, the shape of the aperture is likely to influence perceived motion direction and as the line moves away from the center and towards the aperture the two junctions describe curves that converge on corresponding points in the image plane. If a monocular motion system tracks both junctions and computes the average motion vector relative to the cyclopean eye then this strategy coincides with the geometric CA predictions.

Alternatively, if a binocular system uses the disparity of the two endpoints to establish tilt about the horizontal of the (moving) line then the shortest vector that connects the start point with the oriented line stimulus describes the geometric VN strategy. If the line is slanted about the horizontal axis and moves on a trajectory in depth one junction is always nearer in depth whereas the other is further in depth from the aperture. Since features closer to an occluder are less extrinsic they influence perceived motion direction more (Shimojo, Silverman, & Nakayama, 1989; Graf, Adams & Lages, 2004; Anderson, 1999; van der Smagt & Stoner, 2008). As a consequence perceived motion direction of the slanted line would be biased towards the nearer junction as a function of orientation disparity. Indeed tracking more, and less extrinsic junctions or endpoints in depth would also explain the limited range of adjustments in our data sets (see Fig. 4.5 and 4.8).

In order to confirm our Bayesian approach it would be important to validate the 3D motion prior and disparity prior through real-world measurements of scene flow and through psychophysical experimentation (Stocker & Simoncelli, 2006; Hibbard, 2007). The motion prior may not be Gaussian as assumed here. There is behavioral and physiological evidence that the motion prior in 2D velocity has heavier tails than a Gaussian (Stocker & Simoncelli, 2006) and may be better expressed as a lognormal (Nover et. al., 2005) or Laplace distribution (Lu et al., 2010).

In the present experiments we have used point estimates of motion direction to select between models. The Bayesian models make predictions about the posterior distribution of velocity estimates that can be tested not only in terms of motion direction (MAP) but also in terms of speed and other characteristics. However, speed is notoriously difficult to measure and higher-order statistics of posteriors would require a much larger number of observations per subject and condition in our experiments.

So far we have only considered translation in 3D space but a moving line or edge may also travel on a curved trajectory and may change orientation over time. It is immediately clear that velocity constraint planes in our model cannot encode curved and rotational line movement. Again, it seems likely that encoding of rotation of a line or edge is captured by feature tracking and disparity processing with a frequent update of relative depths, whereas translational motion at a fixed depth may be encoded almost instantaneously by local motion filters.

In conclusion our psychophysical results support the idea that the visual system employs the vector normal principle to solve the aperture problem of binocular 3D motion but that (orientation) disparity processing exerts a strong perceptual bias when disambiguating motion direction. A geometric-statistical approach provides a suitable framework to model perception of binocular 3D motion under ambiguity and offers opportunities to quantify the influence of different processing streams. However, more experimental and

computational studies are needed to confirm and understand the principles of binocular 3D motion perception.

CHAPTER 5: GLOBAL MOTION PERCEPTION

5.1 INTRODUCTION

In the previous chapters we have investigated the contribution of low level cues to stereo-motion perception in accuracy judgements of 3D trajectory. We concluded that CDOT and IOVD cues are likely to be processed in parallel pathways in the early stages of vision, with integration further along the visual hierarchy. In the following we are concerned with the integration of motion and depth cues into a percept of object motion at intermediate processing stages in the brain. Using fMRI we investigate the involvement of human MT in processing stereo-motion information when establishing a percept of global three-dimensional object motion.

5.1.1 Functional Specificity in the Visual Cortex: Anatomy and Physiology

Since the pioneering work by Hubel & Wiesel (1968), describing the functional properties of simple and complex cells in monkey, vision researchers have distinguished two processing streams; they can be discriminated at the level of retinal ganglion cells and are preserved throughout early retinotopic visual regions. Further anatomical separation of the streams was observed in the lateral geniculate nucleus (LGN), a six-layered structure that receives its primary input from the retina. The layers in the LGN can be grouped according to the structure and response properties of the cells of which they are composed. The outer four layers (3-6) of the LGN are known as the parvocellular layers and receive input from the midget ganglion (P-ganglion) cells of the retina; the inner two layers (1,2) are known as the magnocellular layers and

receive their primary input from the Parasol ganglion (M-ganglion) cells of the retina (Perry et al. 1984; Sincich & Horton, 2005).

Cells in the magnocellular or M-pathway and the contrasting parvocellular or P-pathway are identified by unique structural and functional properties that make them suitable to compute motion and form information, respectively (Livingstone & Hubel, 1988). Cells in the M-pathway are characterised by larger receptive field sizes, a fast and transient response to visual input, sensitivity to low contrast stimuli and little or no sensitivity to difference in wavelength. These cells are, therefore, low in visual acuity and are essentially colour blind, but are highly suited to detecting motion (Livingstone & Hubel, 1988).

Cells in the P-pathway exhibit red/green colour opponency with implications for colour vision. They have smaller receptive field sizes (and therefore greater spatial resolution) than magnocellular cells. For these reasons they have been linked to form perception, although, as pointed out in a recent review, this involved splitting parvocellular cells into two sub-types whereby form cells were stripped of their colour coding capabilities (Sincich & Horton, 2005). P-cells show a relatively slow, sustained response to visual input, making them unsuitable for the analysis of motion signals (Livingstone & Hubel, 1988).

A third type of LGN cell which is found in between the magno and parvo layers, receives a blue-on, yellow-off colour opponent signal from small and large bistratified ganglion cells on the retina (Kass et al., 1978; Dacey & Lee, 1994; Lawrence et al., 2005). This led to the proposal of a third pathway; the 'koniocellular' pathway; suggested to be involved in blue/yellow colour vision (Nassi & Callaway, 2009). The response properties of koniocellular cells are not well documented but projections to layer 1 and the cytochrome oxidase blobs in layer 2/3 of the primary visual cortex have been identified (Hendry & Reid, 2000).

In the traditional view, this functional organisation of the visual system is maintained in V1 and beyond. Magnocellular projections to layer 4Cα, and 6 and parvocellular projections to 4Cβ and 6 of V1 (Hendrickson et al., 1978; Livingstone & Hubel, 1988; Blasdell & Lund, 1983; Chaterjee & Callaway, 2003; Nassi & Callaway, 2009) feed directly into extrastriate regions responsible for pooling motion and form information respectively.

 This view is partially supported by research involving cytochrome oxidase (CO), staining a mitochondrial enzyme that is present to some degree in all living cells. CO staining in V1 reveals regions of high concentration of the enzyme, known as 'blobs'; interspersed with areas of low concentration known as 'interblobs'. These light and dark patches appear to be functionally distinct, with CO blobs showing high sensitivity to colour but little or no response to orientation and interblobs showing fine grain

114

orientation tuning but no selectivity for colour (Horton & Hubel, 1981; Horton, 1984; Hendry & Yoshioka, 1994; Sincich & Horton, 2005; Adams, Sincich & Horton, 2007). This evidence lends support to the notion of parallel processing streams in area V1 in so far as it reveals functionally distinct regions. Together with evidence of distinct projections from magno-, parvo- and konio- layers of the LGN, researchers have speculated that distinct processing streams are maintained in the intracortical projections through area V1.

Until recently, efforts to trace intracortical projections have involved methods such as Golgi staining or other dying techniques, or, the injection of tracer chemicals into isolated layers to identify potential synaptic connections between layers of neurons. These methods provide information about potential neuronal connections only - they offer no conclusive proof of how neurons in different layers communicate (Sincich & Horton, 2005). Recent developments in methodology include, the transmission of the rabies virus across a synapse, allowing the identification of intercellular connections via chemical labelling of infected cells (Ugolini, 1995; Sincich & Horton, 2005) and the release of caged glutamate via laser photo-stimulation, to study the inputs from various layers onto a single cell (Callaway & Katz, 1993; Sincich & Horton, 2005). These modern methods, have allowed researchers to trace the propagation of signals between the layers in V1.

The results of several studies strongly suggest that some V1 cells receive input from multiple LGN channels. In particular, the notion that the CO blobs and interblobs represent functionally distinct regions (in keeping with signal segregation in LGN) has been called into question. Input from layer 4Cβ (thought to carry a parvocellular signal) has been found in both the blobs and interblobs of layer 3 of the striate cortex (Lachica et al, 1992; Yabuta & Callaway, 1998b). Input from the mangocellular layer 4Cα, can be traced in all superficial layers of V1; particularly in layer 2/3, which may receive more input than layer 4B -the most commonly highlighted projection site for 4Cα in the literature (Callaway & Wiser, 1996; Sincich & Horton, 2005).

As mentioned previously, early anatomical evidence suggests that magnocellular input from 4Cα would be expected to converge mainly on orientation selective interblobs (Horton & Hubel, 1981; Horton, 1984; Hendry & Yoshioka, 1994; Sincich & Horton, 2005). However, physiological studies in awake, anesthetized monkeys have identified cells that are both orientation and colour tuned (Horwitz et al. 2004; Nassi & Calloway, 2009). In-fact, as little as 17-21% of V1 'colour' cells have been found to be unselective for orientation (Johnson et al., 2001; Friedman et al., 2003; Sincich & Horton, 2005).

From the research mentioned it seems clear that even the feed-forward segregation of form and motion in early visual cortex is not as clear-cut as previously thought. There is evidence that intermixing of signals may occur in addition to the preservation of cue specific processing (Nassi & Calloway, 2009).

If form and motion signals are found in early retinotopic regions, we can be relatively certain of cross-talk and feedback between processing streams in regions further along the retinotopic hierarchy and in extrastriate areas. This is important, as it calls into question a strict hierarchical model of visual processing.

5.1.2 Projections from V1 to V2 and beyond (basic feed-forward physiology)

We return to evidence for the classical, feed-forward view of early visual processing, Again, anatomical evidence from cytochrome oxidase (CO) staining in area V2, gives reason to assume that geniculate processing streams are preserved beyond V1. V2 contains functionally distinct regions of CO staining known as thick, pale and thin stripes (Tootell et al., 1983; Horton, 1984, DeYoe & Van Essen, 1985; Shipp & Zeki, 1985, 2002; Sincich & Horton, 2005; Nassi & Calloway, 2009).

A review of studies concerning the receptive field properties of the different V2 stripe types (Sincich & Horton, 2005), concluded that whilst early evidence suggested that colour sensitive cells were concentrated in thin stripes, orientation selective cells in pale stripes and disparity tuned cells in thick stripes (Hubel & Livingstone, 1987), subsequent studies differed widely in their methods, and definition of 'selective' and degrees of functional specialisation (Sincich & Horton, 2005). In general, evidence for orientation selectivity in pale and thick stripes is by far the most convincing finding across studies. This is bolstered by results from optical imaging showing in squirrel

money, macaque and owl monkey that orientation columns are confined to thick and pale stripes (Malach et al., 1994; Vanduffel et al., 2002; Xu et al., 2004; Sincich & Horton, 2005).

Subsequently, it was shown that the thick stripes in V2 project to area V5/MT, which contains a large concentration of direction and disparity tuned neurons (Born & Bradely, 2005); whilst thin and pale stripes project to V4, linked to colour and from processing (DeYoe & Van Essen, 1985; Ship & Zeki, 1985; Sincich & Horton, 2005).

This evidence led to the assumption of three functionally distinct processing channels, where layer 2/3 blob to V2 thin stripe projections carried colour information; 2/3 interblob to pale stripe projections carried form information and 4B to thick stripe projections continued the pure magnocellular pathway responsible for carrying stereo and motion information to area MT (Livingstone & Hubel, 1988: tripartite model). Modern physiological data has again shown that in reality, the stripes of V2 do not receive input exclusively from a single layer of V1 instead multiple layers of V1 project onto the layers of V2. These connections from layer 2/3 are 'bipartite' in that blobs connect to thin stripes, whilst both pale and thick stripes receive input from interblobs.

This so-called 'double labelling' of interblob projection neurons suggests that information about form and stereo/motion may reach V2 via a single processing stream (Sincich & Horton, 2002a; Sincich & Horton, 2005). The possibility of parallel

118

streams, carrying form and stereo/motion information, within the interblob to thick

stripe channel has been discussed, but there is currently no strong evidence for this

(Anderson & Martin, 2009; Nassi & Callaway, 20*09; Sincich & Horton, 2005).*

Based on known projections from thick stripes to motion area V5/MT, the bipartite

model offers a plausible way for form and stereo/motion information to be combined.

Theoretically speaking, the gradual pooling of information from distinct processing

streams should be possible, particularly in the later stages of visual processing. The

majority of MT neurons are tuned to binocular disparity (Maunsell & Van Essen, 1983a,

Born & Bradely, 2005). In addition motion and disparity selectivity has been shown to

be linked directly to stereopsis (Bradely et al., 1998; DeAngelis et al., 1998; Dodd et al.,

2001; Born & Bradely, 2005); where before it was assumed to play an ancillary role in

motion encoding.

5.1.3 Dorsal and Ventral Streams

The classical, parallel processing view of vision maintains that early segregation of

different forms of visual signal culminates in functionally distinct processing streams

that feed into separate higher order brain pathways. As mentioned previously, the

outputs of cells in the thin, pale and thick stripes of V2 feed into specialised

extrastriate areas. V5/MT receives primarily motion input from the thick stripes (Born

& Bradely, 2005) and area V4 receives form/colour input from pale and thin stripes

(DeYoe & Van Essen, 1985; Ship & Zekki, 1985; Sincich & Horton, 2005).

These functionally distinct areas are assumed to represent the origin of two

extrastriate, parallel processing streams. The dorsal stream, which also receives input

from area V5/MT, is associated with defining the spatial location of objects. It is often

casually referred to as the 'where' pathway and has been implicated in visuo-motor

control and navigation (Milner & Goodale, 2008; Nassi & Callaway, 2009). More

recently a further subdivision was suggested within the dorsal stream. The dorsal-

dorsal pathway runs through V6 and the superior parietal lobule (SPL) (Rizzolatti &

Matelli, 2003; Nassi & Callaway, 2009) and may be involved in 'online' motor control,

the control of goal directed movements whilst they are ongoing. The ventral-dorsal

pathway runs through MT and the inferior parietal lobule (IPL) and may be involved in

action recognition (of others) and spatial awareness (Rizzolatti & Matelli, 2003).

The ventral stream, or the 'what' pathway, which receives its primary input from visual

area V4, is thought to be involved in creating a fine-grain internal representation of

objects, allowing for identification, and encoding the spatial relationships between

objects in the world (Rizzolatti & Matelli, 2003). Although most researchers accept a

theoretical distinction between perception and action streams in visual processing, it is

widely acknowledged that extensive communication between dorsal and ventral

pathways occurs (Lewis & Van Essen, 2000; Nassi & Callaway, 2009). An extreme view

is that the signal content in both streams is likely to be almost identical, with only the

goals of perception distinguishing one stream from another (Nassi & Callaway, 2009).

Again, the study of these streams in terms of processing goals means that theories of

functional specialisation and theories of early signal integration need not be considered as mutually exclusive. In linking evidence for cross-talk between ventral and dorsal sub-systems with the apparent mixing of signals in early visual areas V1 and V2, Nassi & Callaway (2009) make the important point that intermixing of early visual signals may represent the gradual integration of signals in order to reach specific perceptual goals at later stages of processing. Anatomical, physiological and lesion studies in monkeys showed a double dissociation between the ventral and dorsal streams in terms of their functional role in perception (Milner & Goodale, 2008; Nassi & Callaway, 2009).

One purpose for having a separate visual stream for processing perceptual information relevant to motor output is that organisms require a quick, streamlined visual system to facilitate fast motor responses to a constantly changing visual environment. As we have seen, the dorsal stream is well equipped for this function, but shows deficits when the support of the ventral system is removed. Tasks involving internal object representations, such as grasping based on memory are a good example of this. Recognising objects, making relative judgements between objects and forming and storing useful memories for frequently encountered objects, requires perception that is independent of a particular context (Goodale et al., 2005). As a consequence, the ventral system would not be expected to operate in real-time, as the perceptual representations of this system are less transient and more object-centred.

5.1.4 Basic Physiological Challenge of Global 3D Motion Perception

The primary physiological challenge for the integration of motion and depth inputs

starts with the cortical segregation of form and motion pathways. To simplify this

problem for the purposes of explanation, it is necessary to re-visit the basic bottom-up

view of visual processing. Whilst acknowledging that the true organisation of the visual

cortex is far more complex than this framework allows, due to crosstalk between

processing streams and cortical feedback (as discussed previously in this chapter).

 It is convenient to begin with a simple model of the visual system in order to frame

the problem. To summarise, selectivity for spatial and temporal (form and motion)

inputs begins with activity of retinal ganglion cells and is preserved in two distinct

processing streams, which continue in the lateral geniculate nucleus (LGN), visual area

V1 and subsequent layers of the striate cortex. These streams are termed the

magnocellular and parvocellular pathways (M-pathway and P-pathway, respectively)

and are characterised by the response properties of cells within the individual

pathways.

The segregation of form and motion information, which begins in early visual

processing regions, continues beyond the striate cortex. The form or P-pathway

projects to the temporal lobe whilst the motion or M-pathway projects mainly to the

parietal lobe (Jeffrey et al., 2007). These parallel processing pathways are referred to

as the ventral (what) and dorsal (where) stream, respectively, due to their involvement

in specific perceptual goals.

Motion inputs are processed along the magnocellular pathway and largely converge on the middle temporal region MT/hMT+ (in humans). Studies in monkeys show that MT forms one of the main inputs to the dorsal/posterior parietal processing stream (Maunsell & Newsome, 1987). It outputs primarily to structures involved in optic flow (MST, VIP) and those controlling eye movements (LIP, FEF, SC, dorsolateral pons) (Born & Bradely, 2005). Direct inputs to MT come mostly from the striate cortex, in particular area V1, V2, V3 (Maunsell & Van Essen, 1983c; Shipp & Zeki, 1985, 1989b; Anderson & Martin, 2002). The main feed-forward connection originates in area V1. A single MT neuron often receives input via multiple synapses to VI cells (Anderson & Martin, 2002). MT and the analogous region in the human brain hMT+/V5, may provide an important link between motion and depth processing in the brain. The response properties of cells in this region suggest that it is a possible region of crossover and integration between visual processing streams, where dual processing of motion and stereo inputs and multiplexing with other cues is highly likely.

Located in both hemispheres, MT is retinotopically organised with each side containing a full map of the contralateral visual hemifield. The central 15% of the fovea occupies more than half of the total surface area of MT, suggesting a strong emphasis on inputs from this region (Van Essen et al., 1981). The functional organisation of MT is columnar. Neurons can be grouped into oblique Columns running through the surface, dependent on preferred direction (Dubner & Zeki, 1971; Albright et al. 1984). Columnar organisation for binocular disparity selectivity, is also present (DeAngelis &

Newsome, 1999) and there is evidence of the clustering of neurons with regard to speed preference (Liu & Newsome, 2003b).

5.1.5 The Role of MT in Depth Perception

The response properties of MT neurons, therefore, suggest specialisation for motion processing and also significant disparity preference. In-fact, visual responses in this region are determined by a number of factors: retinal position, direction of motion, speed of motion, binocular disparity and stimulus size (Orban, 1997; Britten, 2003). Most neurons display disparity selectivity (Maunsell & Van Essen, 1983a: using moving bars) with twice as many being tuned to near disparities as far disparities and few tuned to zero disparity (Bradley & Anderson, 1998).

The involvement of MT in depth perception has been investigated in monkeys, with MT firing rates corresponding to direction judgments of a bi-stable, cyclopean cylinder rotating in depth (Dodd et al. 2001; Britten et al. 1996). MT is also involved in depth discrimination tasks (Uka & DeAngelis, 2004). Some neurons in this area are selective for surface orientation defined by combined motion and disparity cues. In these neurons tuning is finer for gradients defined by a combination of motion and disparity cues than those defined by either cue alone (Nguyinkim & DeAngelis, 2004).

The precise role of MT in the integration of motion and disparity information is not clear. Although the ability of this region to process both types of input is not disputed, the role of cells that respond optimally to combined motion and disparity input is not necessarily one of integration. This area may just carry both sources of information for integration elsewhere.

Existing accounts of stereo-motion processing assume early integration of interocular velocity difference (IOVD) and changing disparity over time (CDOT) cues although experimental evidence for either of the two main processing routes (stereo-first, motion-first) does not clearly favour one account over the other (Harris, Nefs & Grafton, 2008; Lages, 2006; Lages, Mamassian & Graf, 2003; Lages & Heron, 2010; Regan & Gray, 2009).

This is consistent with the view that neither cue in isolation is sufficient to drive motion-in-depth perception as outlined in Chapter 2. Whilst both IOVD and CDOT displays can produce the sensation of motion-in-depth, isolating one cue is not clear evidence that the other cue is not important for the perception of 3D motion. Regan & Gray (2009) refer to the question of whether IOVD contributes to motion-in-depth perception as 'ill-posed', since it is not possible to dismiss the role of monocular motion inputs altogether (Regan & Gray, 2009).

This suggests that motion-in-depth perception, which includes detection, speed discrimination, and motion direction, may rely on intact processing of both monocular velocities and binocular disparities (Brooks & Stone, 2004; Regan & Gray, 2009).

There is growing evidence that motion and stereo inputs are processed in parallel, with integration occurring further along the visual pathway.

5.1.6 The Role of hMT+/V5 in Motion-in-Depth Perception

The response properties of MT neurons in the monkey brain, suggest that integration of motion and disparity information may occur at this stage. Much less is known about the properties of the analogous region in human hMT+/V5.

Since the perceptual systems of humans are not assumed to differ markedly from other species of primate (see however Orban et al., 2003) this area has recently been targeted in brain imaging research to investigate its' selectivity for motion and disparity changes (Smith & Wall, 2008) and its' involvement in processing cyclopean motion-in-depth (Likova & Tyler, 2007; Rokers et al. 2009; Ban et al., 2012).

Using an fMRI adaptation paradigm, Smith & Wall (2008) showed sequential displays of rotating dot patterns that moved in a) the same or opposite direction b) the same or different depth planes. They found adaptation effects in human MT and MST when stimuli moved in the same depth plane but not when they were in different depth planes. This effect was also observed in areas V1-V4. In MT and MST adaptation effects were also observed for motion direction, but unlike depth, these were not measurable throughout the visual hierarchy and could only be recorded in area V3 and beyond. This adds to neuro-physiological evidence by showing that in human MT and MST,

there is selectivity for both disparity direction and motion direction (Smith & Wall, 2008).

Rokers et al. (2009) investigated hMT+ in a series of fMRI experiments designed to establish selectivity in this area for stereo-motion perception. Bold responses were measured in retinotopic regions; V1, V2, V3, V3A/B, hV4; as well as LO1/LO2, and hMT+ (human motion complex MT/MST combined). In the first experimental condition, observers viewed a display comprising of 32 moving dot pairs which had either opposing horizontal or opposing vertical motion. When the horizontal opposing stimulus was presented dichoptically (each dot in a pair to a different eye) it stimulated 'horizontal opposing motion' which contains both CDOT and IOVD cues. fMRI bold responses to this stimulus were compared with responses to a vertical opposing motion and monoptic motion- where pairs of dots with horizontal and vertical opposing motion were presented to the same eye.

Response patterns for observers peaked when viewing horizontal dichoptically opposing motion, they were over 50% larger than those for the horizontal monocularly paired stimulus. This shows a preferential response in MT+ for 3D motion defined by horizontal disparity. There was no significant difference for vertical dichoptically opposing motion (which did not stimulate 3D-motion) and the vertical monocularly paired stimulus suggesting orientation specificity and ruling out the possibility that monocular dot density was involved in differential activation (Rokers et al., 2009).

The results suggest that hMT+ plays a role in the processing of both IOVD and CDOT

information, but as the authors conceded this does not indicate whether this area is

'important' for stereo-motion perception. To establish this, the authors investigated

directional selectivity for 3D motion. They used an adaptation paradigm in which

observers adapted to uni-directional 3D motion (towards or away from the observer).

A subsequent probe stimulus moved in either the same or opposite direction from the

adaptor (Rokers et al., 2009).

Direction selectivity for stereomotion was found in hMT+ and also in V2, V3, V3A and

LO (gradually increasing in higher processing regions), evidenced by attenuated

response in these areas for a probe stimulus moving in the same direction as the

adaptor.

This suggests that some motion and stereo information may be integrated as early in

the visual system as V2 but that the process of integration is likely to continue in later

stages as the signal moves from lower to higher areas (Rokers et al., 2009).

A further two experiments aimed to examine the role of hMT+ and associated

retinotopic regions in processing IOVD and CDOT driven stereo-motion specifically. Far

greater response to disparity based 3D motion was found in V3A, LO and most

strikingly in hMT+ in comparison to response for spatiotemporally scrambled stimuli.

In a further experiment, the CDOT signal for the stereo-motion display was degraded

using a binocularly anti-correlated stimulus (Rokers et al., 2009). Dots in the stimulus

could be fully correlated, fully anti-correlated, or correlated/anti-correlated at a ratio

of 50/50 and 25/75. There were two motion conditions: (1) Motion through depth displays (MTD) containing both IOVD and disparity information. Pairs of dots moved across the two eyes in anti-phase (displays move through one another) (2) Motion within depth displays (MWD) containing only unchanging depth information: Pairs of dots moved across the two eyes in phase (dots move in fronto-parallel planes). Correlated MTD displays resulted in oscillating motion towards and away from the observer, whilst correlated MWD gave oscillating sideways motion at various depths. Anti-correlated MTD resulted in motion towards and away through depth, MWD resulted in side-to-side motion with little or no depth, due to the degradation of the disparity signal (Rokers et al., 2008; 2009).

MT+ responses to MWD decreased as binocular correlation decreased as anticipated, due to the degraded disparity component caused by anti-correlation. MT+ responses to MTD did not decrease as binocular correlation decreased (They even may have increased slightly). This suggests a contribution of the remaining IOVD signal. The authors conclude that MT+ can be stimulated by both changing disparity and IOVD inputs to stereo-motion (Rokers et al., 2009).

In terms of integration, the results indicate that MT+ is capable of utilising IOVD and CDOT components of motion together and in isolation to drive perception. This may point to a similar architecture as in monkey MT, where both disparity and motion information is preserved in columns, but where cross-talk between motion and disparity columns is likely to occur.

Likova & Tyler (2007), identified a brain region anterior to hMT+ (hMT, MST), which was specific to cyclopean stereo-motion. This region has been labelled KO (kinetic occipital) region or V3B. The global pattern of activation for the cyclopean motion stimulus differed markedly from that of a luminance defined motion stimulus (hMT+ localiser) particularly in region V3B/KO - which was not significantly responsive to simple luminance-defined motion (Likova & Tyler, 2007). This differential pattern of activation was observed for cyclopean stereo-motion versus a disparity balanced stereo-motion control stimuli (CSMdb); which featured z-axis stereo-motion versus fixed disparity dynamic random dot displays (Likova & Tyler, 2007). This ruled out the possibility that activation was confounded by static disparities. It is therefore likely that the region V3B/KO anterior to hMT+ is responsive to dynamic disparity (Likova & Tyler, 2007). Since it is roughly the size of, as well as partially overlapping the motion complex (hMT+), this area is difficult to distinguish but well placed for processing of cyclopean stereo-motion.

The authors discuss the location of the cyclopean stereo-motion specific region of cortex, with respect to the known functional organization of the motion complex. This is hierarchical in nature and '…organized according to a posterior to anterior 'axis' of increasing functional complexity…'(Likova & Tyler, 2007). If z-axis cyclopean stereo-motion is a specialized form of 3D motion, it stands to reason that further processing of 3D motion occurs in hMT+/V5.

5.2 fMRI STUDY ON GLOBAL 3D MOTION

To investigate the integration of local 3D motion in intermediate and later stages of visual processing, we designed a stimulus that preserves local motion information but at the same time can create a percept of a full shape or form moving in depth. To this end, a stimulus was developed that extends well-known principles used in seminal experiments to investigate 2D object motion (Lorenceau & Shifrar, 1997; Lorenceau & Alais, 2001).

It has been demonstrated that the temporal and spatial characteristics of moving line stimuli define whether or not an object will be seen as a coherently moving shape or gestalt (Muckli et al., 2006; Sack et al., 2006). Psychophysical evidence has also shown that form information can greatly influence motion binding. In particular closed shapes lend themselves more readily to motion binding than open forms (Lorenceau & Alais, 2001). We use a stimulus composed of four individual moving line segments viewed through four symmetrically arranged circular apertures (see Fig. 5.1). To facilitate motion binding of the local elements into a coherent two-dimensional shape the line elements were angled at 45 degrees within the four apertures suggesting a closed diamond shape with corners occluded. This arrangement reliably produced perception of a moving diamond.

Manipulation of the phase offset of the sinusoidal motion between the left and right eyes, results in changing binocular disparity between local lines and their monocular images. To create different two-dimensional motion percepts we also varied the monocular offset between the line motion components in the four apertures (same offsets between the left and right eye). Thus motion perception depended on the inter-ocular phase offset (changing disparity) and monocular phase offset (type of two-dimensional motion). We used this stimulus in an fMRI study to investigate the response of motion area hMT+/V5 to different two-dimensional and three-dimensional motion types. For comparison we also monitored the response in primary visual area V1. Results are considered to be preliminary since they will form the basis of a body of work to investigate the integration of motion and throughout the visual cortex. These are discussed with respect to previous findings on global stereo-motion processing in hMT+/V5 and the human visual cortex.

5.2.1 Materials and Methods

Ethics

Informed written consent was obtained from all observers before participation. Experiments and experimental procedures were approved by the Faculty Ethics Committee at Glasgow University in compliance with national legislation and the Code of Ethical Principles for Medical Research Involving Human Subjects of the World Medical Association (Declaration of Helsinki).

Participants

9 naïve adult observers aged 18-36, with self-reported normal/corrected to normal visual acuity, took part in this study. Due to our opportunistic sampling of observers three observers had to be excluded from the analysis due to excessive head and eye movements. All observers were tested for stereo anomalies using the Randot Butterfly stereopsis test and achieved a score of 100 sec arc or less. This was deemed acceptable for participation in the binocular motion study (in the absence of any standard test of stereo-motion capability). Observers received no training on the task and were naïve to the experimental aims. They were paid £10 per hour.

Equipment/Scanner

A Siemens Trim-trio (3.0 Tesla) with a 12 channel head coil was used to measure fMRI BOLD signals. MRI compatible NordicNeuroLab goggles (OLED) with a resolution of 800 by 600 pixels @75Hz in each eye displayed the stimulus.

Stimuli

Stimuli were programmed in Matlab (Version 2009) using Psychtoolbox routines. The stimulus display was presented through MRI compatible binocular viewing goggles. They consisted of four circular apertures around a central fixation cross (see Figure 5.1). In each of the apertures an oblique line oscillated sinusoidally left and right through the center of the circlular aperture. The moving line elements always had the same start position and orientation to facilitate the perception of a diamond shape undergoing different motions

Left Eye   Right Eye
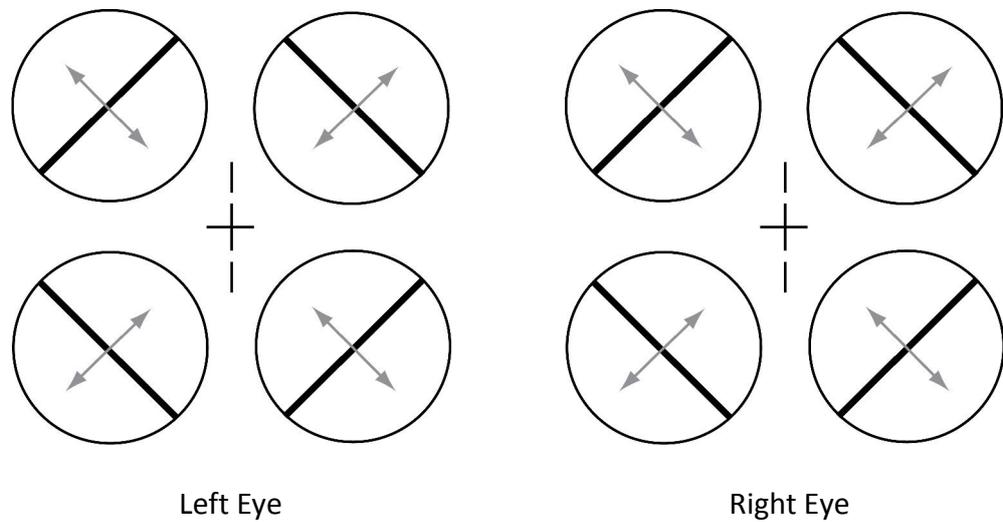
Figure 5.1 Illustration of the experimental stimulus in the left and right eye. In a stereoscopic display, four oriented lines, positioned in a diamond shape oscillated inside circular apertures. The motion of the line elements was defined by a sinusoidal function allowing manipulation of the phase offset between line elements. By choosing specific monocular offsets between pairs of lines we manipulated the global motion perception of the diamond shape; binocular phase offsets were used to create 3D motion perception.

The motion of local line elements was defined by a sinusoidal function (see Fig. 5.2). Exploiting gestalt principles, we introduced monocular phase offsets between pairs of local motion components to change the motion type of the diamond shape. Since the stimulus is presented binocularly, binocular phase shifts were introduced between monocular half images to create perception of global 3D motion (see Table 5.1)

134
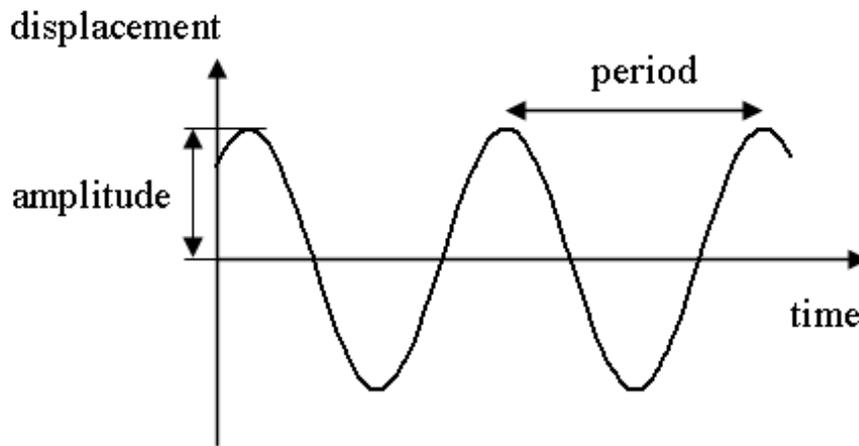
Figure 5.2 Example of sinusoidal function. Defining local line motion in this way allows manipulation of the percept through introducing monocular and binocular phase shifts between paired lined elements (i.e. shifting the sinusoid horizontally to the left or right). This enabled us to create a range of two-dimensional and three-dimensional motion types from essentially the same local oscillation.

In this experiment we manipulated phase offset of local elements to produce different global motion stimuli: (1) 2D expansion/contraction was achieved with monocular phase offset of 90° (2) 2D translation was achieved with monocular phase offset of 180° (3) 3D rotation was achieved with monocular phase offset of 180° and binocular phase offset of 90° (4) 3D looming was achieved with monocular phase offset of 180° and binocular phase offset of 180°. For the baseline condition we used the same display but removed the line elements inside the apertures (apertures only). The static control condition consisted of four motionless, oriented lines positioned inside the apertures.

		Interocular Phase Offset		
		0	90	180
Monocular	0	2D up/down translation	3D incoherent rotation	3D incoherent looming
Phase	90	**2D expa/contraction**	3D incoherent rotation	3D incoherent looming
Offset	180	**2D left/right translation**	**3D left/right rotation**	**3D looming**

Table 5.1 Here we show examples of motion types which arose during stimulus development as a result of manipulating monocular and binocular phase offset. The four moving lines vary in their monocular phase offset i.e. temporal delays between line segments in monocular half images, and their binocular phase offset. This results in different 2D, and 3D motion percepts of a diamond shape. The conditions used in the experiment are shown in bold type.

Although we labelled our motion types as "3D" and "2D" it is worth noting that the 2D stimuli also moved in depth. Unlike the 3D stimuli, they moved behind the apertures on a plane in depth with constant disparity. Static disparity in local and global motion was therefore always present in our motion stimuli (and the static control) and may explain some of the response patterns in V1.

Our stimuli also included 2 mapping conditions. These had the same circular apertures but with rectangular 'checker-board' patterns flickering between black and white at 18 Hz, The patterns were placed either inside or outside the apertures. They covered either the regions of sinusoidal line motion within the apertures or the regions of the

occluded corners of the moving diamond outside the apertures. These regions are known to facilitate motion binding and amodal completion.

Procedure

Observers completed a detailed questionnaire administered by a qualified radiographer to determine suitability for scanning. They were then tested for stereo anomalies using the Randot Butterfly stereopsis test and the results were recorded.

Inside the scanner, observers viewed the binocular display consisting of four line elements oscillating horizontally within four circular apertures (see Figure 5.1). Similar arrangements are known to facilitate amodal completion and motion binding of the lines into a diamond shape. The phase offsets of the line elements in the monocular half images were manipulated to create one of four experimental motion conditions (1) Two-dimensional expansion and contraction (2) Two-dimensional translation (sideways motion), (3) Three-dimensional rotation (around the vertical axis), (4) Three-dimensional looming in depth (towards/away from the observer). In any given trial, the observer was presented with one of the four experimental stimuli for 5 seconds, followed by 5 seconds of the static baseline condition. They were instructed to maintain fixation on a central fixation cross throughout the entire experimental block. The purpose of this was to ensure binocular stimuli were perceived, as during stimulus testing it became clear that the 3D percept was often lost when fixation was broken. With this is mind there was no reason to exclude observers who broke fixation occasionally. Attending to individual parts of the stimulus did not assist observers in

performing the task and in-fact may have hindered performance. If observers broke

fixation often they were excluded from the analysis, three observers were excluded in

total. This was verified using an eye-tracker allowing online observation of eye

movements by the experimenter, but there was no previously agreed criterion for

exclusion.

The observer's task was to indicate whether the global motion in the display was two-

dimensional or three-dimensional by pressing one of two coloured buttons, a third

button could be pressed to indicate problems with the stereoscopic display. They were

able to respond at any point during the 5 second trial and responses continued to be

registered throughout the following static baseline display.

The mapping conditions used to locate retinotopic regions of interest were built into

the experimental paradigm (see figures 5.3 and 5.4). These consisted of static versions

of the aperture display with rectangular flickering checker-board located either inside

(figure 5.3) or outside of the apertures (figure 5.4) Observers were told to respond

only when viewing one of the four experimental line conditions. To establish baseline

activity, a static display with the four circular apertures was presented for 20 seconds

at the start and end of every block.

After four blocks, observers remained in the scanner whilst a high resolution

anatomical scan was performed. They were instructed to keep their head still for the

duration of this scan but performed no additional task. High resolution T1 weighted

structural images were taken in 192 axial slices and isotropic voxels (1mm³ field of view: 256 X 256 mm², TR = 1900ms, TE = 2.92 ms, time to inversion = 900ms, FA = 90°). This high-quality brain anatomy scan was co-aligned with the intra-session anatomical data (recorded automatically during the functional scan) in order to ensure accurate alignment of functional data.

Functional images covering the whole brain (slices: 18, field of view =210x210 mm, voxel size =3X3X3 mm) were gathered, with a 2 second delay before the first stimulus onset (this was accounted for in the analysis)

Localisation

Due to the exploratory nature of this study, we limited our regions of interest to retinotopic areas hMT+/V5 in the visual hierarchy. Specifically hMT+/V5 and the anterior region V3B/KO has been implicated in stereo-motion processing (Likova & Tyler, 2007; Rokers et al. 2009; Ban et al., 2012). To investigate the differences between the processing of our global motion stimuli in lower and higher order regions of the visual cortex, we first compared activation in hMT+/V5 and primary visual cortex (V1).

The experimental design included two 'mapping' conditions. These consisted of the four circular apertures with rectangular flickering checker-boards located either inside of each aperture (covering the area of sinusoidal line motion), or outside of each aperture, covering an area where the moving corners of the diamond shape were

located (See figures 5.3 and 5.4)
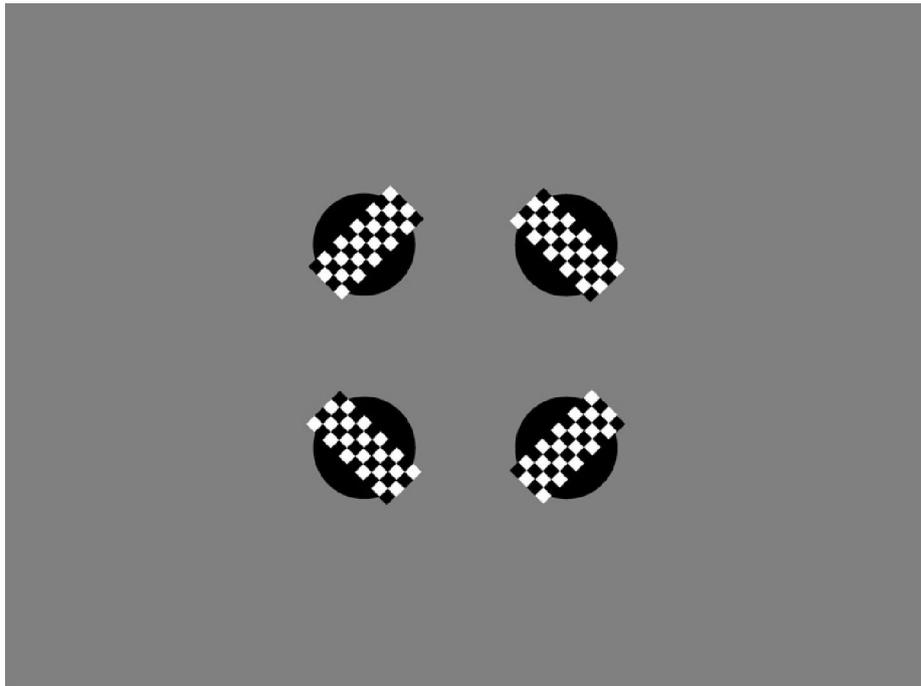
Figure 5.3 Mapping stimulus with checkerboard presented inside of the circular

apertures, used to locate retinotopic regions of interest.
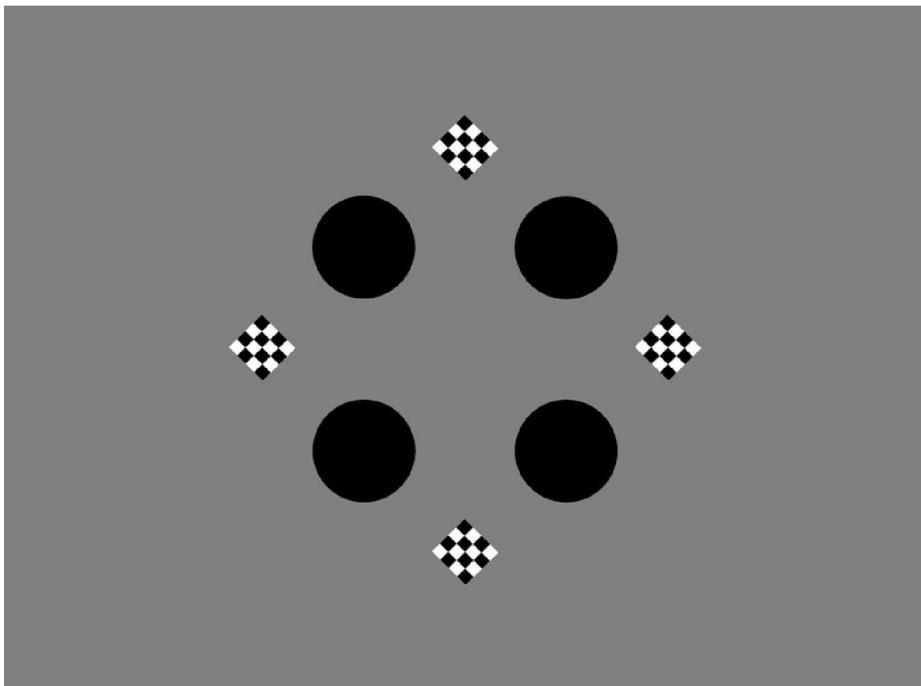
Figure 5.4 Mapping stimulus with checkerboard presented outside of the circular apertures, used to locate retinotopic regions of interest.

Since we did not carry out a detailed retinotopic mapping in this experiment, the mapping conditions inside and outside of the apertures were used as an approximation to locate the parts of early retinotopic regions that responded to activity within the circular apertures, i.e. those responsive to the locally moving line elements.

A general linear model (GLM) including all four experimental conditions as well as the static baseline condition and the two mapping conditions was conducted. In a first step a de-convolution was carried out to account for the event-related design of our study. For each individual subject, the GLM was applied to the averaged functional data across all four experimental runs to identify the region of V1 responsive to the local motions of the line elements, we included a contrast between the inside mapping condition and outside mapping condition. This highlighted that retinotopic regions are more responsive to stimulation inside of the apertures. We simply selected the most posterior active region as V1. To locate area hMT+ we ran a GLM with a contrast between the combined response to all four motion conditions and the combined response to the static and mapping conditions (inside/outside aperture). This highlighted regions more responsive to moving than static or flickering stimuli. The highlighted areas were identified as the region of the motion complex hMT+/V5 because they showed the sharpest distinction between moving and static/flickering stimuli.

Regions of interest were identified on both the two-dimensional high resolution

anatomical scans and on three-dimensional, surface based models of the individual

hemispheres. A high concordance was found between the two representations.

Since the Talairach co-ordinates for retinotopic areas are fairly well established, we

were able to verify our choices by comparing the Talairach co-ordinates of our areas to

those based on the average brain. Based on this we were reassured the selected areas

were accurately labelled.

5.2.2 Analysis

Conventional analyses on BOLD activity were carried out using Brainvoyager QX.

Anatomical data were corrected for inhomogeneity and transformed to Talairach-co-

ordinates, and the cortical surface was reconstructed. Functional MRI time series were

pre-processed using standard parameters and co-aligned to anatomical data-set.

A general linear model (GLM) including all seven conditions (including the controls) in

the experiment was applied to the activation averaged across four blocks of trials. As

mentioned previously, a de-convolution step was carried out to account for the fact

that the events in our paradigm were closely spaced. This is standard procedure when

running a GLM on data gathered using a rapid event related design, since the

haemodynamic response curves for individual events have substantial overlap.

Applying de-convolution results in the separating of the haemodynamic response

curves.

For each individual observers and region of interest (hMT+/V5, V1) we conducted a

GLM analysis, specifying contrasts and correcting for serial correlation. Although this

study was exploratory in nature, the following six predictions seemed reasonable. Each

of these contrasts was applied at the point of peak activation.

(1) Averaged 4 motion conditions (2D translation; 2D expansion/contraction; 3D

rotation; 3D looming) – averaged static 2D & 2 mapping conditions (inside;

outside). The contrast was balanced by selecting the corresponding option in

Brainvoyager. This was used both to verify our choice of motion area and to

investigate differences between lower and higher order visual regions in terms

of preference for moving over static stimuli. We anticipated that hMT+/V5

would make a significant distinction, but did not discount the possibility that V1

would also distinguish between moving and static stimuli in general.

(2) Averaged 3D motion conditions – averaged 2D motion conditions. This contrast

was used to investigate differences between lower and higher order visual

regions that may distinguish between two-dimensional and three-dimensional

motion. We anticipated that hMT+/V5 may show higher activation to 3D

moving stimuli (Likova & Tyler, 2007; Rokers et al., 2009) than lower visual

regions. A negative effect may be expected in early visual regions for this

143

contrast.

(3) 3D looming – 3D rotation. This was based on previous imaging literature, which found that hMT+/V5 (Likova & Tyler, 2007; Rokers et al, 2009) and neighbouring areas (Likova & Tyler, 2007) show a preference for cyclopean motion (in the z-dimension) and do so based on binocular information. We anticipate that hMT5+/V5 may be capable of distinguishing between different types of stereo-motion based on the global stereo component of the stimulus. This is not expected from lower order regions.

(4) 3D looming – 2D expansion/contraction. This contrast included 3D and 2D motion conditions which were designed to correspond, differing only in terms of binocular disparity. It was used to investigate whether hMT+/V5 distinguishes between motion in 3D with changing disparity or constant disparity on a 2D plane in depth.

(5) 3D rotation – 2D translation. This served the same function as the 3D loom - 2D expa/contr contrast, to investigate the role of stereo information in distinguishing between 2D and 3D motion in hMT+/V5.

(6) Mapping inside – Mapping outside. This contrast was used to identify lower order region that are more responsive to flicker inside the apertures e.g. local motion of oblique lines. It was also used to highlight differences in processing strategies between lower order and higher order visual regions. hMT+/V5 was not expected to prefer stimulation within the apertures as it is assumed to respond to more global motion information.

144

We did not to carry out a group analysis on the data due to the relatively small number of observers and the diverse set of results. Alternatively, non-conventional multivariate pattern classification analysis (MVPA) may reveal a more fine-grain representation of motion types.

5.2.3 Results

In our analysis we considered individual results from 6 observers in area hMT+/V5 and V1. Of particular interest was the BOLD response to the varying motion types in area hMT+/V5 in both left and right hemisphere (see Table 5.1). We also considered analyses on the BOLD responses in V1 in both hemispheres in order to detect differences between higher and lower order processing regions. The results are listed below by observer and region of interest

Subject LWA26: hMT+/V5 (MT)

In the left hemisphere, the location of hMT+ corresponds to Talairach co-ordinates: X: -40.5, Y: -60.6, Z: -3.8 (Fig. 5.5).
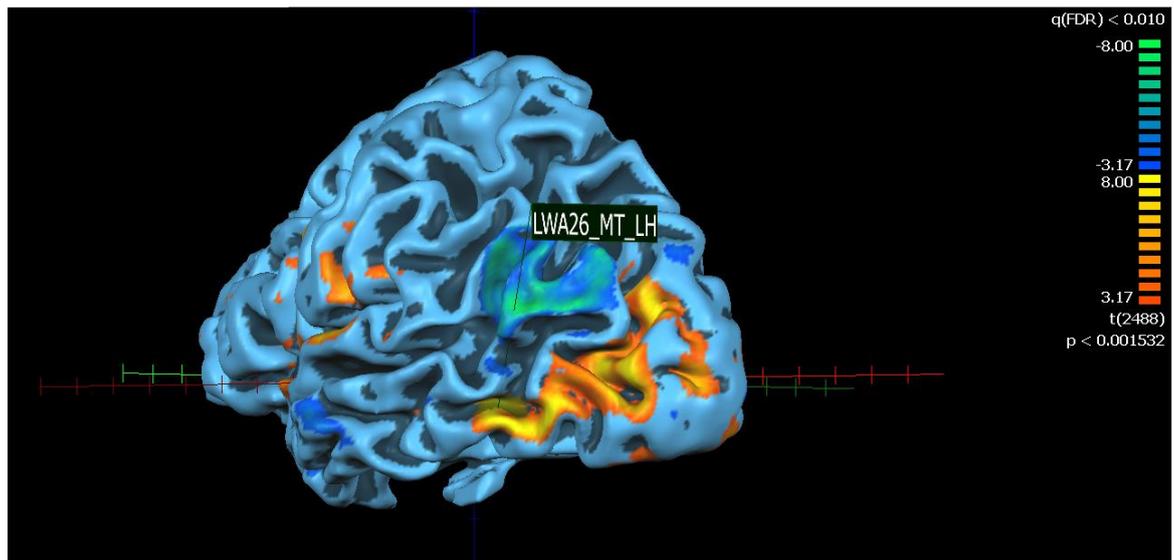
Figure 5.5: Surface view of subject LWA26 left hemisphere. The BOLD response for the left hemisphere, averaged across 4 experimental blocks is mapped onto the surface. Areas where the average BOLD activation was significantly higher for moving stimuli than static and mapping stimuli are orange/yellow in colour. The area assumed to be MT is highlighted in green and is labelled LWA26_MT_LH.

This region was significantly more responsive to moving than static images in general (t (143) = 9.94, p<0.0001) (after correction for serial correlations). As it is located further along the visual processing hierarchy, hMT+ was expected to respond preferentially to 3D moving stimuli over 2D moving stimuli. The results of the 3D>2D contrast show that the average response across 3D motion conditions is greater than that of 2D motion conditions (t (143) = 5.77, p<0.0001). In subject LWA26, hMT+ appears to show a clear preference for 3D motion.

There was no significant difference in activation for different 3D motion types (3D rotation, 3D looming) (t (143) = -0.62, p>0.05), suggesting that perceptually distinct 3D motion types cannot be distinguished in terms of hMT+ activation.

Since 2D expansion and contraction may correspond to 3D looming, a contrast

between these conditions was carried out, under the assumption that 3D looming

would engage higher order processing regions such as MT, to a greater degree than 2D

distinguished from 2D expansion/contraction in area MT (t (143) = 5.1, p<0.0001).

hMT+ also responded preferentially to 3D rotation over 2D translation (t (143) = 5.8,

p<0.0001). These two motion conditions are also corresponding because the only

difference is changing compare to constant disparity input in the context of motion. A

comparison with lower order regions would determine whether hMT+ truly

discriminates between these types of motion.

In the right hemisphere, the location hMT+ corresponds to Talairach co-ordinates: X:

41, Y: -59, Z: 8 (Figure 5.6).
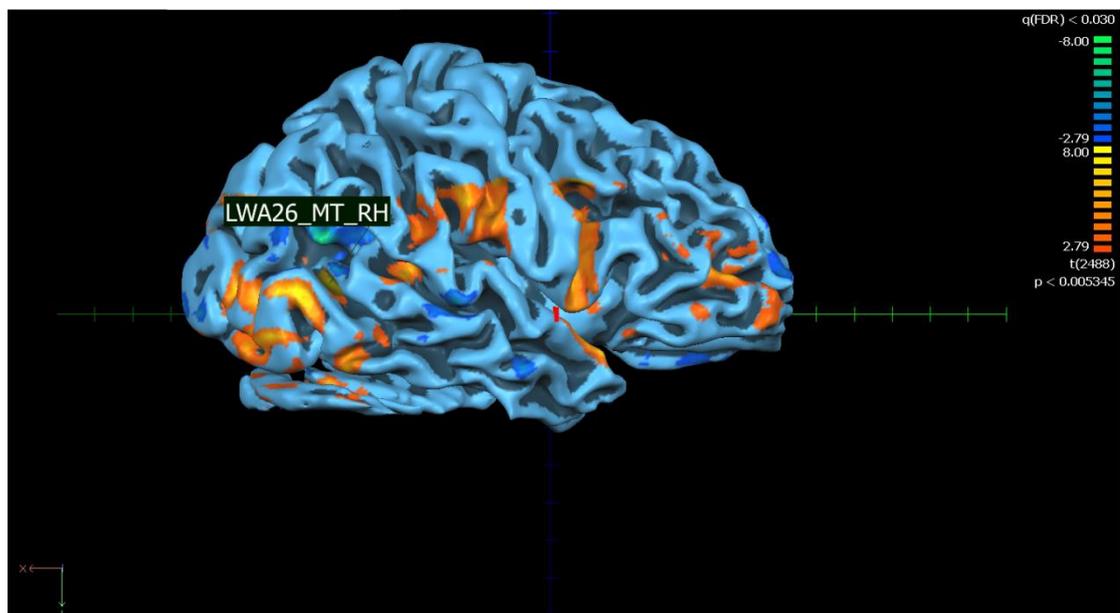
Figure 5.6: Surface view of subject LWA26 right hemisphere. The BOLD response for

the right hemisphere, averaged across 4 experimental blocks is mapped onto the

surface. Areas where the average BOLD activation was significantly higher for moving

stimuli than static and mapping stimuli are orange/yellow in colour. The area assumed to be hMT+ is highlighted in green and labelled LWA26_MT_RH.

A contrast between the four experimental motion conditions and the static and mapping conditions, confirmed that this area prefers moving to static/flicker stimuli (t (143) = 4.64, p<0.0001). As with left hemisphere hMT+, response to combined 3D conditions was greater than combined 2D conditions (t (143) = 7.2, p<0.0001). This area therefore, makes some distinction between 2D and 3D motion. There was no difference in response to 3D looming motion and 3D rotating motion (t (143) = 1.2, p>0.05).

There was, however, a significant preference in hMT+ for looming motion over 2D expansion/contraction (t (143) = 5.75, p<0.0001). This area also distinguished between 3D rotation and 2D translation, showing much greater activation to the 3D version of the stimulus (t (143) = 4.6, p<0.0001). There was no difference between response to stimulation inside and outside of the circular apertures for MT.

Subject LWA26: V1

In the right hemisphere, V1 had Talairach co-ordinates: X: 18, Y: -83, Z: 8 (Figure 5.7)
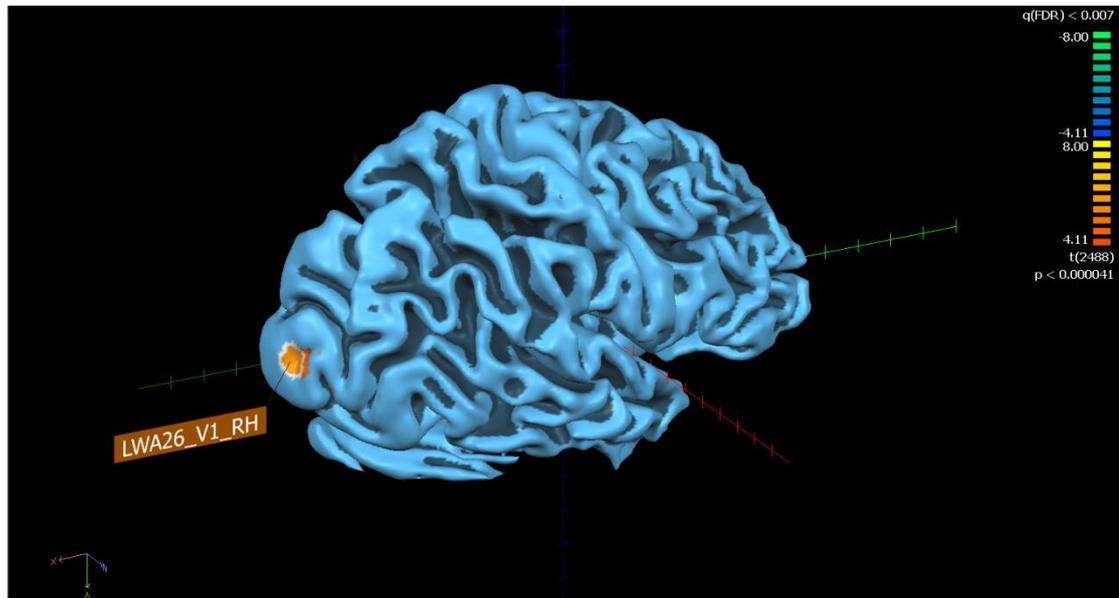
Figure 5.7: Surface view of subject LWA26 right hemisphere. Areas where the average BOLD activation was significantly higher in response to the mapping stimulus inside the apertures than to the mapping stimulus outside of the apertures are orange/yellow in colour. In this case only one distinct region was identified and so no marker colour was used. The area in V1 is highlighted in orange and labelled LWA26_V1_RH.

The motion contrast revealed a significant preference for motion stimuli in V1 (t (143) = 2.55, p < 0.02). Surprisingly, this area also showed significantly more activation for 3D motion stimuli than 2D motion stimuli (t (143) = 2.87, p<0.01).

This could be due to selective processing of local motions, however, in this observer activation in this area of V1 also distinguished between 3D looming motion and 2D expansion/contraction (t (143) = 3, p<0.01) as well as between 3D rotation and 2D

translation (t(143)= 3.1, p<0.01). This result may reflect feedback and does not necessarily indicate sensitivity to stereo-motion in early visual cortex but was contrary to our expectations.

The results of the mapping contrast may help to disambiguate these results, since the only difference between area V1 and MT seems to be that V1 responds stronger to the mapping stimulus presented inside the apertures (t(143)= 4.6, p<0.0001). The activity in V1 may therefore reflect more local stimulus features, whilst MT activity reflects more global stimulus qualities.

In the left hemisphere, V1 corresponds to Talairach co-ordinates: X: -9, Y: -99; Z: -5 (Figure 5.8). In this region, we observed the same response pattern as in the right hemisphere V1. There was significant preference for moving over static stimuli (t (143) = 2.2, p<0.05) and for 3D moving stimuli over 2D moving stimuli (t (143) = 4.9, p<0.0001). When comparing analogous 3D and 2D stimulus pairs, again we found that V1 distinguished between 3D looming and 2D expansion/contraction (t (143) = 4, p<0.0001) and between 3D rotation and 2D translation (t (143) = 4.3, p<0.0001). There was significantly higher activation in response to the inside mapping condition (t (143) = 3.6, p<0.01)
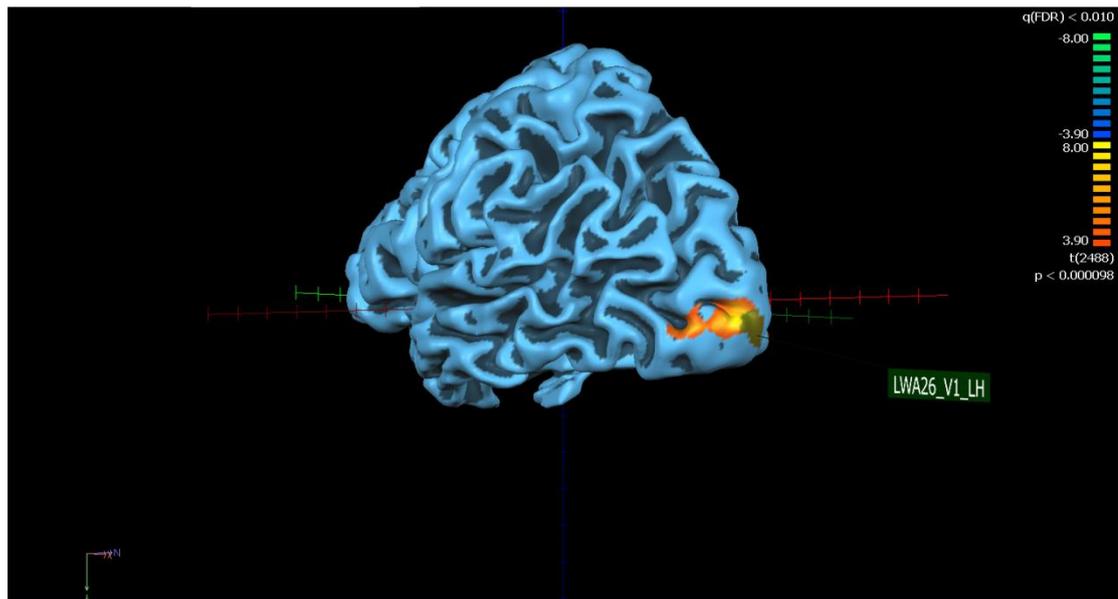
Figure 5.8: Surface view of subject LWA26 right hemisphere. The BOLD response for the right hemisphere, averaged across 4 experimental blocks is mapped onto the surface. Areas where the average BOLD activation was significantly higher in response to the mapping stimulus inside the apertures than to the mapping stimulus outside of the apertures are orange/yellow in colour The activated area in V1 is highlighted in green and labelled LWA26_V1_LH.

Subject FFN20: MT Right Hemisphere

In subject FFN20, right hemisphere hMT+ corresponds to Talairach co-ordinates: X: -59, Y: -38, Z: 20 (Figure 5.9) As expected, hMT+ was significantly more responsive to moving stimuli than to the static or flicker stimuli (t (143) = 11.2, p<0.0001). hMT+ also showed significant preference for 3D motion over 2D motion (t (143) = 6.1, p<0.0001), but contrary to previous findings, also distinguished between 3D looming motion and 3D rotation (t (143) = 2.7, p = 0.02).

For the corresponding 3D and 2D motion stimuli, this area was significantly more active in response to 3D looming motion versus 2D expansion/contraction (t (143) = 5.3, p<0.0001). It preferred 3D rotation over 2D translation (t (143) = 3.3, p<0.0001). There was no significant difference in response to inside and outside mapping conditions (t (143) = 0.5, p>0.05).

In the left hemisphere, hMT+ corresponds to Talairach co-ordinates; X: -59, Y: -38, Z: 20 (Figure 5.10). There was a significantly greater response to moving than static/flicker stimuli   (t (143) = 3.9, p<0.0001). There was also a significant preference for 3D motion stimuli over 2D motion stimuli (t (143) = 2.1, p<0.05) but to a lesser degree, perhaps suggesting some lateralisation in this subject.

hMT+ was not selective to sub-types of 3D motion (t (143) = 1.4, p>0.05) but did distinguish between 3D looming motion and 2D expansion/contraction (t (143) = 2.3, p<0.03). There was no distinction between 3D rotation and 2D translation (t (143) = 0.67, p>0.05). Again there was no significant difference in activation for mapping conditions inside and outside of the aperture, possibly suggesting a more global processing strategy in hMT+.
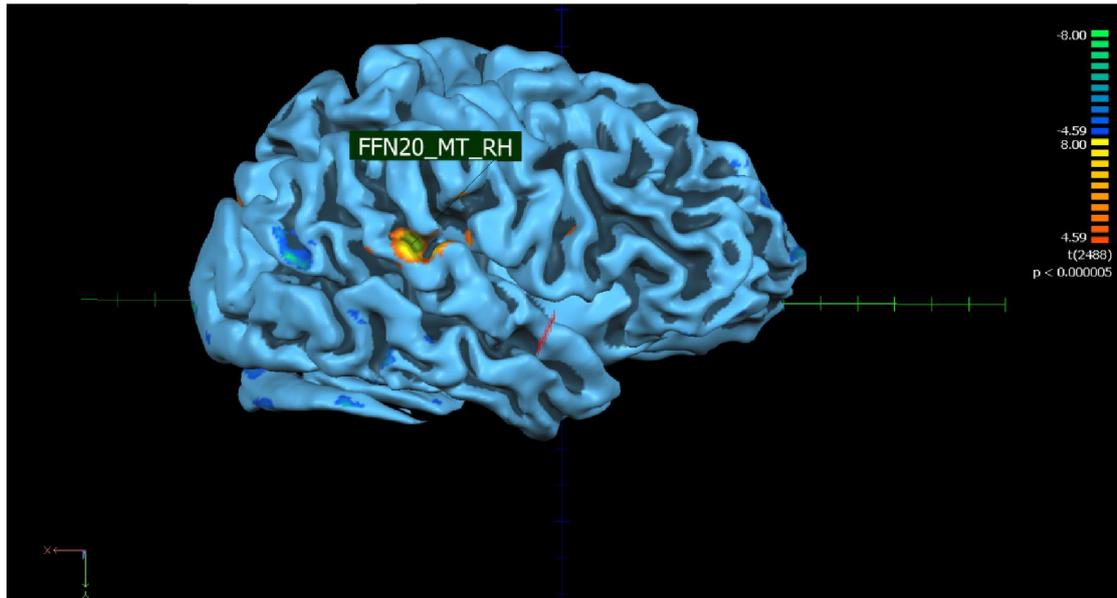
Figure 5.9: Surface view of subject FFN20 right hemisphere. The BOLD response for the right hemisphere, averaged across 4 experimental blocks is mapped onto the surface. Areas where the average BOLD activation was significantly higher for moving stimuli than static and mapping stimuli are orange/yellow in colour. The area assumed to be hMT+ is highlighted in green and is labelled FFN20_MT_RH.
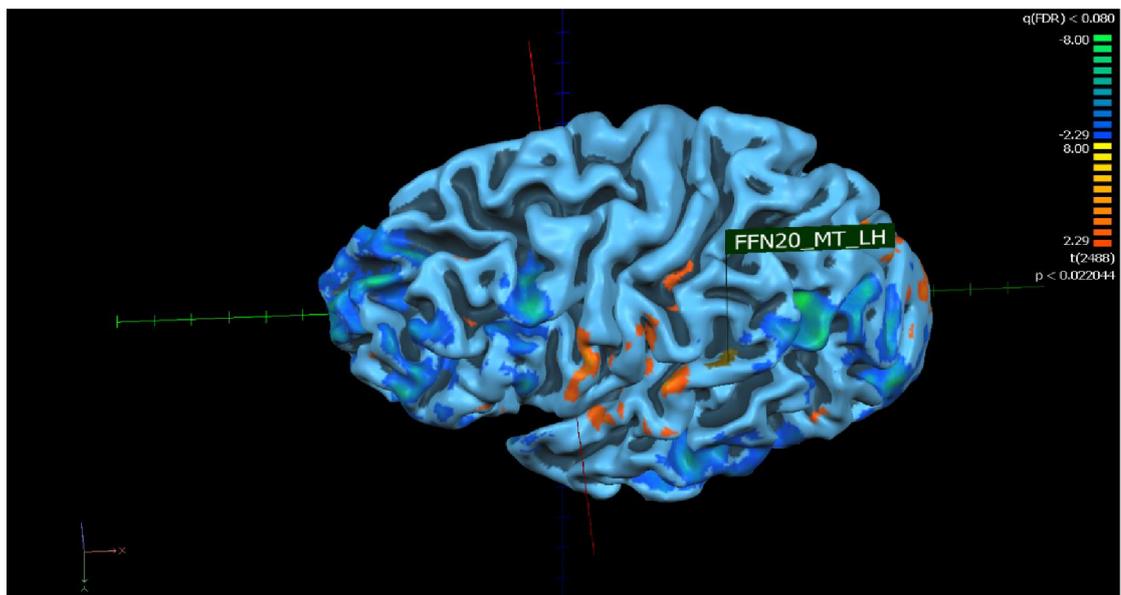
Figure 5.10: Surface view of subject FFN20 left hemisphere. The BOLD response for the left hemisphere, averaged across 4 experimental blocks is mapped onto the surface.

Areas where the average BOLD activation was significantly higher for moving stimuli

than static and mapping stimuli are orange/yellow in colour. The area assumed to be in

hMT+ is highlighted in green and is labelled FFN20_MT_LH.

Subject FFN20 V1 Right Hemisphere

In this observer, V1 was located at Talairach co-ordinates; X: 13, Y: -96, Z: 12 (Figure

5.11). The results of the region of interest (ROI) contrast, confirmed that it was more

responsive to a mapping stimulus presented inside of the circular apertures (t (143) =

6, p<0.0001).
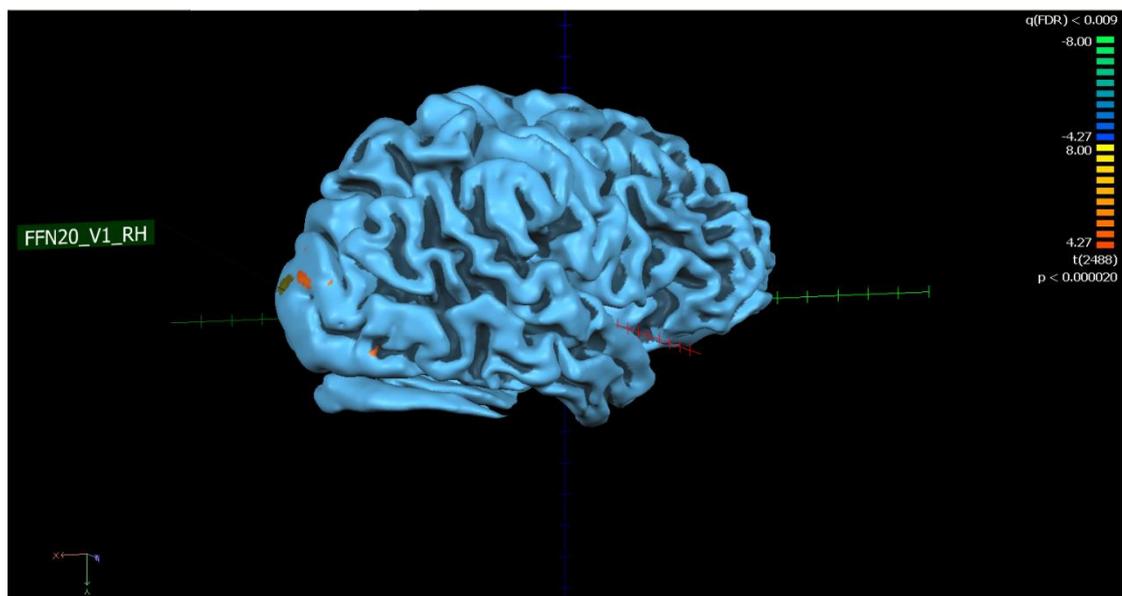
Figure 5.11: Surface view of subject FFN20 right hemisphere. The BOLD response for

the right hemisphere, averaged across 4 experimental blocks is mapped onto the

surface. Areas where the average BOLD activation was significantly higher in response

to the mapping stimulus inside the apertures than to the mapping stimulus outside of

the apertures are orange/yellow in colour The area activated in V1 is highlighted in green and is labelled FFN20_V1_RH.

V1 showed a significant preference for motion over the static/flicker conditions (t (143) = 4.1, P<0.0001), almost to the same extent as the area in hMT+. It also showed a moderately significant preference for 3D moving stimuli over 2D moving stimuli (t (143) = 2.5, p = 0.01). Contrary to the findings in the right hemisphere MT, activation in V1 did not distinguish between the two sub-types of 3D motion (t (143) = -0.8, P>0.05). This suggests that the distinction made in MT and V1 may result from different underlying processes.

Activation in V1 did not distinguish between 3D looming and 2D expansion/contraction (t (143) = 1.7, p>0.05), nor was there a distinction between 3D rotation and 2D translation (t (143) = 1.8, p>0.05).

In the left hemisphere the activated area in V1 corresponds to Talairach co-ordinates; X: -10, Y: -98, Z: 8 (Figure 5.12). A mapping contrast performed in this region of interest, confirmed it to be more responsive to stimuli presented inside of the circular apertures (t (143) = 6.2, p<0.0001).
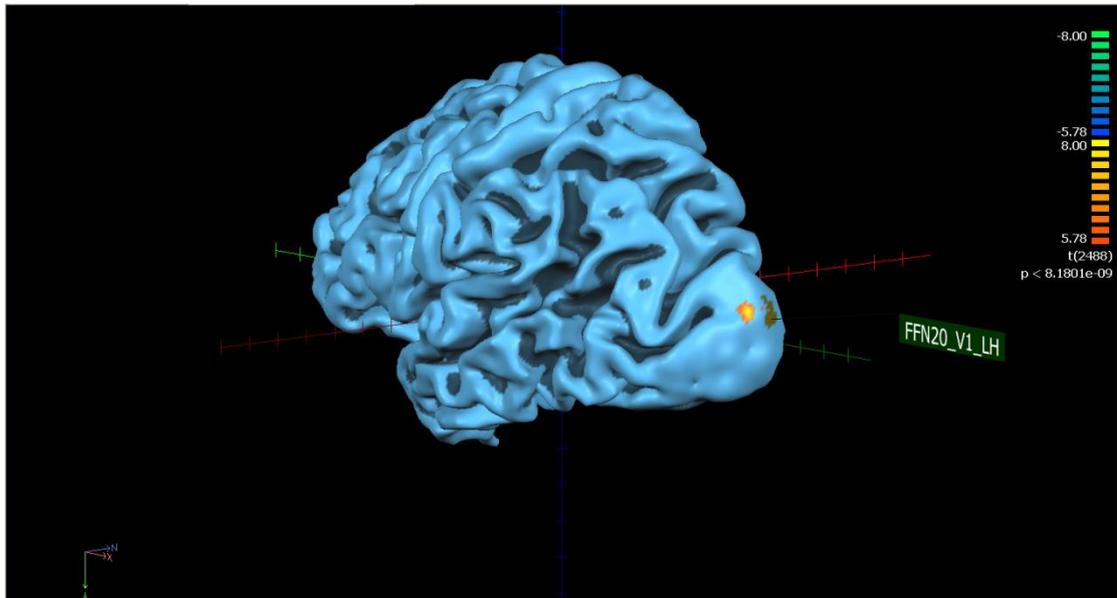
Figure 5.12: Surface view of subject FFN20 left hemisphere. The BOLD response for the left hemisphere, averaged across 4 experimental blocks is mapped onto the surface. Areas where the average BOLD activation was significantly higher in response to the mapping stimulus inside the apertures than to the mapping stimulus outside of the apertures are orange/yellow in colour The activated area in V1 is highlighted in green and is labelled FFN20_V1_LH.

This region preferred moving over static stimuli (t (143) = 3.7, p<0.01) but in this observer there was no distinction between 3D and 2D motion conditions (t (143) = 0.3, p>0.05).

There was also no difference in response between sub-types of 3D motion (t (143) = 0.04, p>0.05), between 3D looming and 2D expansion/contraction (t (143) = 0.03, p>0.05), and between 3D rotation and 2D translation (t (143) = 0.37, p>0.05).

It seems that in subject FFN20, there is convincing evidence for lateralisation of both feature level and more advanced global motion processing. The ROIs in the right hemisphere distinguish better between the different sub-types of motion with area hMT+ specifically responding to the global stereo-motion in the 3D stimuli.

Subject AME14: hMT+ Right Hemisphere

In observer AME14 right hemisphere hMT+ corresponds to Talairach co-ordinates; X: 39, Y: -60, Z: 0 (Figure 5.13). A contrast between motion conditions and static/flicker conditions revealed significant preference for moving stimuli in this area (t (143) = 10.7, p<0.0001).
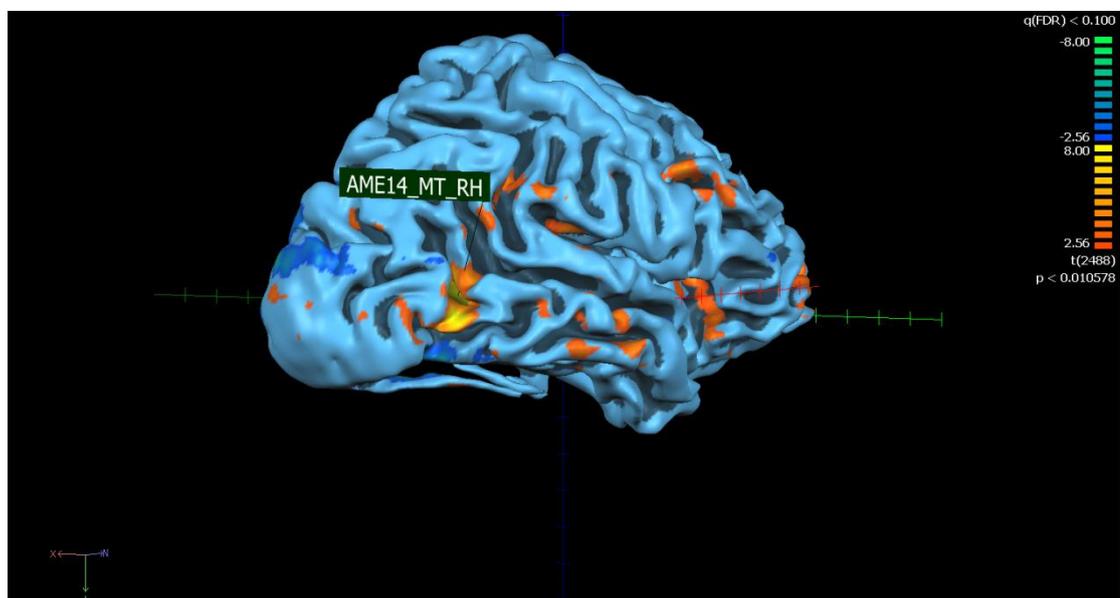
Figure 5.13: Surface view of subject AME14 hMT+ right hemisphere. The BOLD response for the right hemisphere, averaged across 4 experimental blocks is mapped onto the surface. Areas where the average BOLD activation was significantly higher for

moving stimuli than static and mapping stimuli are orange/yellow in colour. The area assumed to be hMT+ is highlighted in green and is labelled AME14_MT_RH.

Contrary to findings in other observers, hMT+ showed no preference for 3D motion (t(143)= 0.26, p>0.05). A contrast between the two sub-types of stereo-motion was also non-significant (t (143)<1, p>0.05), as were the two contrasts comparing analogous 3D and 2D motion conditions, looming versus 2D expansion and rotation versus 2D translation (p>0.05). The contrast between mapping conditions inside and outside was non-significant (t (143) =-0.6, P>0.05).

Left hemisphere hMT+ was located at Talairach co-ordinates; X: -44, Y:-63, Z: 3 (Figure 5.14). This area showed exactly the same responses as right hemisphere hMT+. There was a significant preference for moving over static stimuli (t (143) = 9.9, p<0.0001) but no distinction between 3D and 2D motion (t (143) = 1.3, p>0.05) or between sub-types of 3D motion (t (143) = 0.23, p>0.05). There was also no significant difference in activation when viewing 3D looming versus 2D expansion/contraction (t (143) = 1.5, p>0.05) or when viewing 3D rotation versus 2D translation (t (143) = 0.3, p>0.05).

The only anticipated result was a non-significant contrast between the inside and outside mapping conditions (t (143) = 0.09, p>0.05).
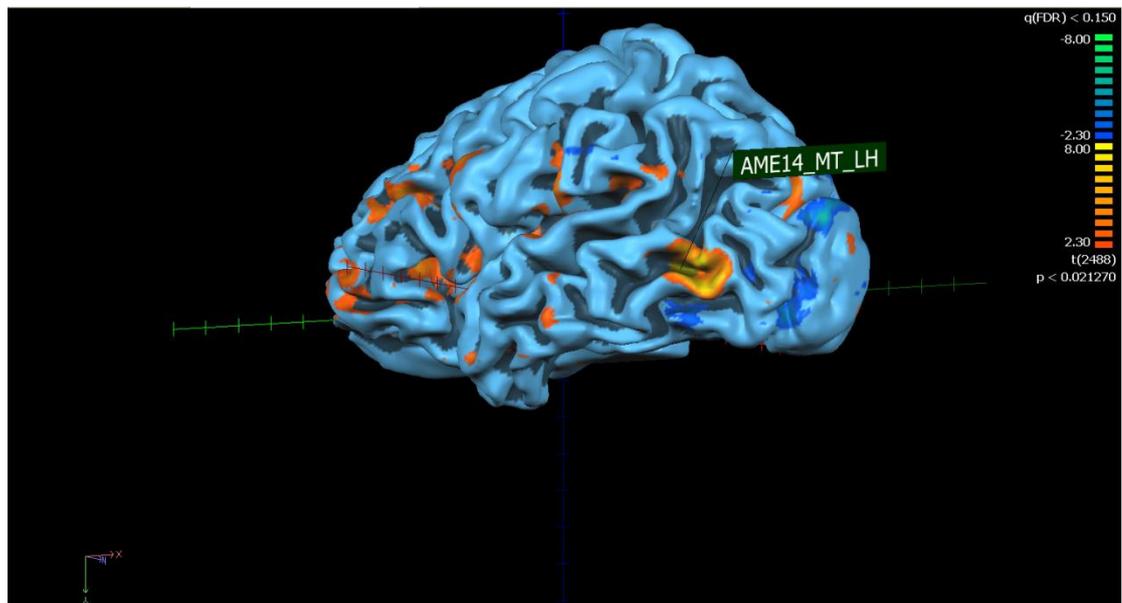
Figure 5.14: Surface view of subject AME14 MT left hemisphere. The BOLD response for the left hemisphere, averaged across 4 experimental blocks is mapped onto the surface. Areas where the average BOLD activation was significantly higher for moving stimuli than static and mapping stimuli are orange/yellow in colour. The area assumed to be MT is highlighted in green and is labelled AME14_MT_LH.

Subject AME14: V1 Right Hemisphere

In this observer right hemisphere V1 corresponds to Talairach co-ordinates; X: 11, Y: -99, Z: 1 (Figure 5.15). An ROI contrast confirmed that this area was more responsive to stimuli presented inside the circular aperture (t (143) = 7.3, p<0.05).
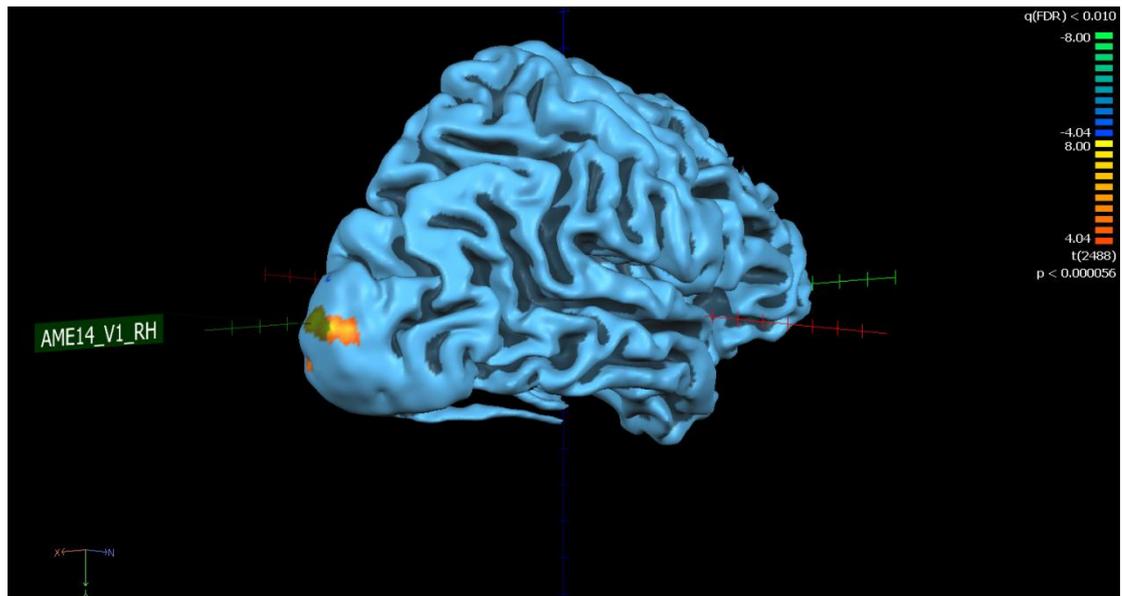
Figure 5.15: Surface view of subject AME14 V1 right hemisphere. The BOLD response for the right hemisphere, averaged across 4 experimental blocks is mapped onto the surface. Areas where the average BOLD activation was significantly higher in response to the mapping stimulus inside the apertures than to the mapping stimulus outside of the apertures are orange/yellow in colour. The area in V1 is highlighted in green and is labelled AME14_V1_RH.

As with previous observers, V1 shows significant preference for moving stimuli compared to static stimuli (t (143) = 4.3, p<0.0001). We did not observe significantly higher activation for 3D moving stimuli than 2D moving stimuli (t (143) = 0.9, p>0.05), suggesting that activation in this area does not distinguish between 3D and 2D motion.

There was a significant negative result for the contrast between 3D looming motion and 3D rotation (t (143) = -2.1, p<0.05), suggesting that processing in this region distinguishes between sub-types of 3D motion, although this may be explained by differences in local features. The even-related de-convolution plot for this subject suggests that there was little or no activation for the looming condition in V1 over the

160

time-points measured (Figure 5.16). This may indicate that looming motion is detected by higher order processing (perhaps even top-down processes) whilst 3D rotation may be based on disparity in a local line component.

No significant difference was found between 3D looming versus 2D expansion/contraction (t (143) = 0.1, $p > 0.05$) or between 3D rotation and 2D translation (t (143) = 1.2, $p > 0.05$). This supports the theory that the distinction between looming and rotation is supported by local differences.

In observer AME14, the area in left hemisphere V1 corresponds to Talairach co-ordinates; X: -13, Y: -97, Z: 11 (Figure 5.17). A contrast between the 2 mapping conditions confirmed that this area was more responsive to stimuli presented inside the circular apertures ($t(143)$ = 6.6, $p < 0.0001$. Again, this area shows much higher activation in response to moving than static stimuli (t (143) = 4.5, $p < 0.0001$) but does not respond stronger to 3D moving stimuli than 2D moving stimuli (t (143) = 0.6, $p > 0.05$). In keeping with the results of right hemisphere V1 in this observer, the contrast between 3D looming motion and 3D rotation was moderately significant (t (143) = 2.1, $p < 0.05$), suggesting this area distinguished between different 3D motion sub-types but did so based on low level stimulus features rather than global motion direction in depth.

Contrasts between analogous 3D and 2D motion stimuli gave the same results as in the right hemisphere, with no significant preference for the 3D conditions ($p > 0.05$).
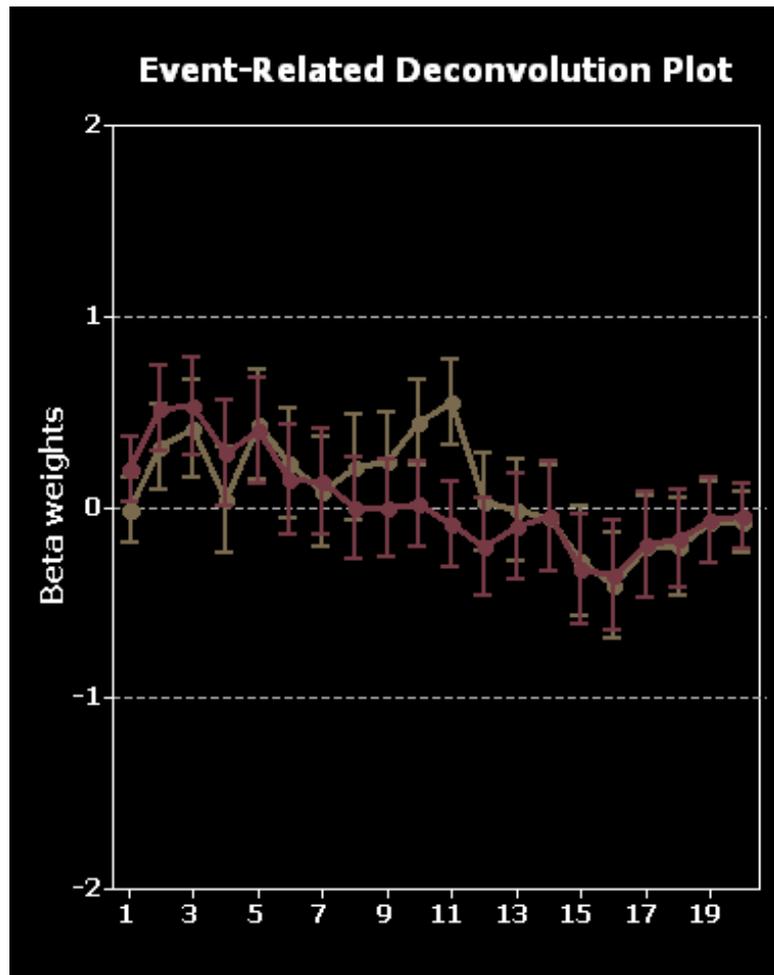
Figure 5.16: Event related de-convolution plot (after correction for serial correlations) for subject AME14 in right hemisphere primary visual cortex V1. The time-point of interest here, are 8-11. Activation is shown across time for the 3D looming (red) and 3D rotation (olive). At the relevant time-points the response to looming is close to 0, suggesting that low level visual areas are not involved in detecting 3D looming motion in this subject, whereas the detection of rotational motion may be based on the processing of local disparity changes.
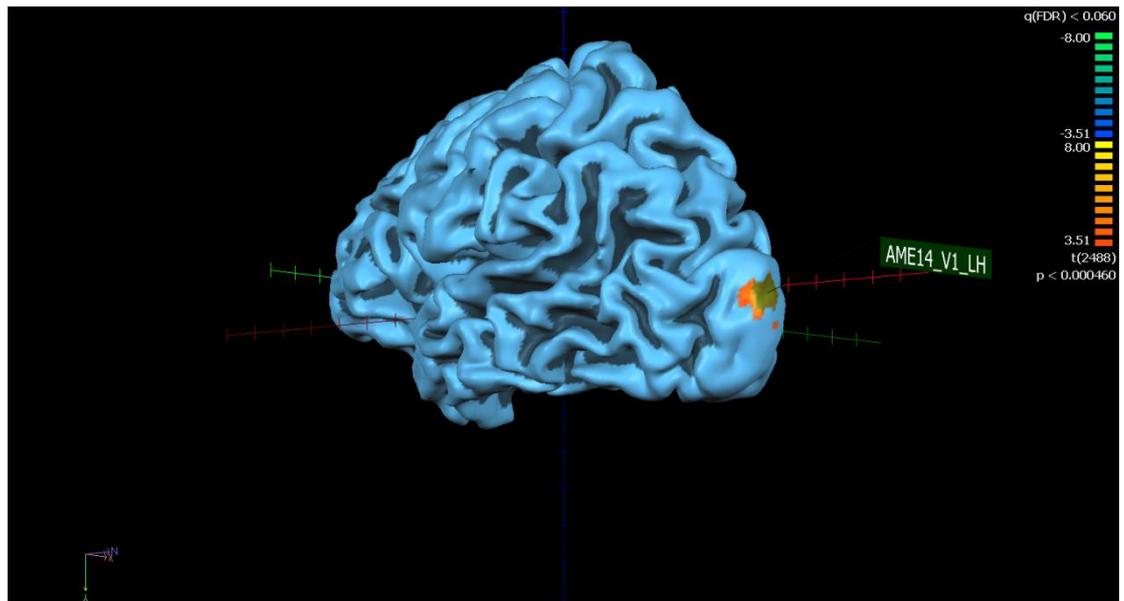
Figure 5.17: Surface view of subject AME14 V1 left hemisphere. The BOLD response for the left hemisphere, averaged across 4 experimental blocks is mapped onto the surface. Areas where the average BOLD activation was significantly higher in response to the mapping stimulus inside the apertures than to the mapping stimulus outside of the apertures are orange/yellow in colour. The area in V1 is highlighted in green and is labelled AME14_V1_LH.

Subject KDS11 hMT+ Right Hemisphere

In observer KDS11 area hMT+ on the right hemisphere was located at Talairach co-ordinates; X: 39, Y: -61, Z: 5 (Figure 5.18). This region was significantly more responsive to moving than static stimuli (t (143) = 11.4, p<0.0001). A contrast between 3D and 2D motion conditions failed to reach significance (t (143) = 1.4, p>0.05).
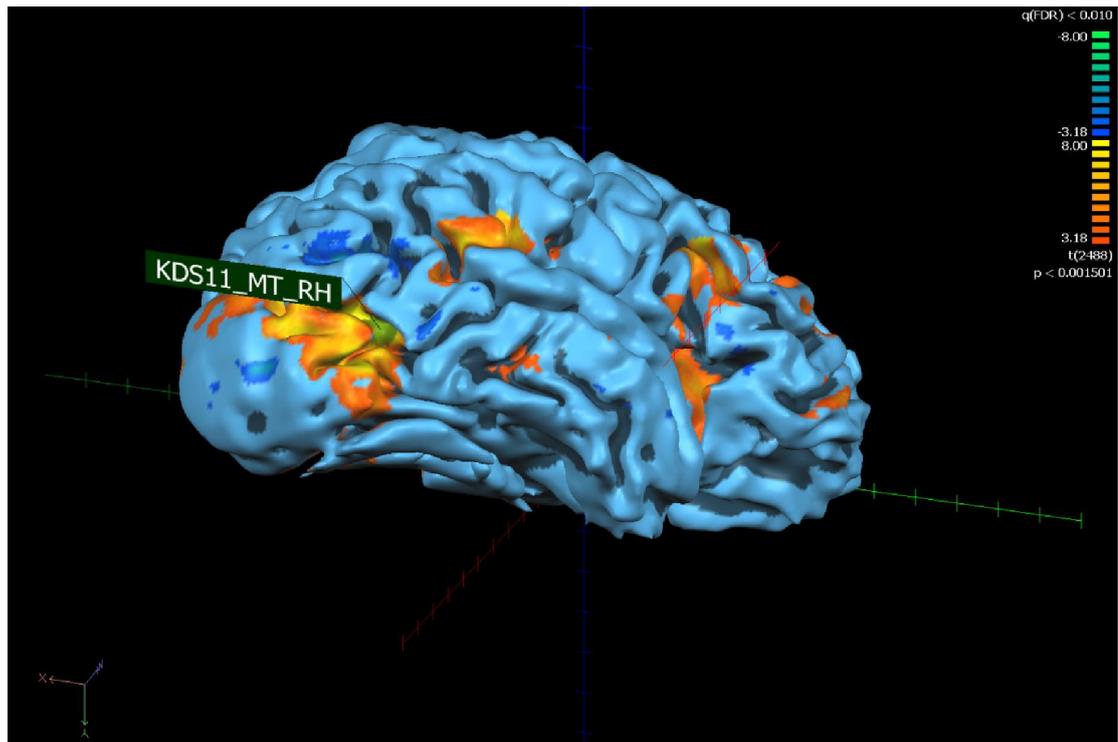
Figure 5.18: Surface view of subject KDS11 hMT+ right hemisphere. The BOLD

response for the right hemisphere, averaged across 4 experimental blocks is mapped

onto the surface. Areas where the average BOLD activation was significantly higher for

moving stimuli than static and mapping stimuli are orange/yellow in colour. The area

assumed to be hMT+ is highlighted in green and is labelled KDS11_MT_RH.

This area did not show preference for looming motion over rotation (t (143) = 1.8,

p>0.05).It did show a significantly greater response to 3D looming motion over 2D

expansion/contraction (t (143) = 2.4, p< 0.05) but not between 3D rotation and 2D

translation (t (143) = -0.3, p> 0.05). The contrast between mapping conditions inside

and outside was non-significant (t (143) = 0.01, p>0.05).

Subject KDS11: hMT+ Left Hemisphere

In the left hemisphere the location of the region labelled MT corresponds to Talairach co-ordinates; X: 40, Y: -62, Z: -1 (Figure 5.19). This region showed preference for moving over static stimuli (t (143) = 8.4, p<0.0001) but as we observed in right hemisphere hMT+ , did not discriminate between 3D and 2D motion conditions (t (143) = -0.4, p> 0.05).
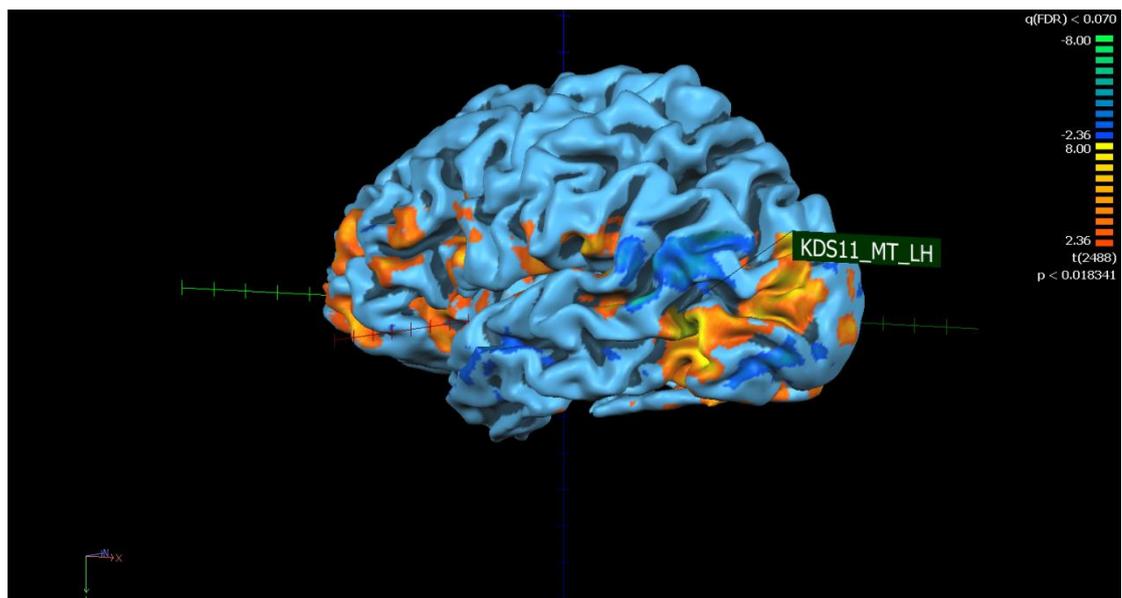

Figure 5.19: Surface view of subject KDS11 MT left hemisphere. The BOLD response for the left hemisphere, averaged across 4 experimental blocks is mapped onto the surface. Areas where the average BOLD activation was significantly higher for moving stimuli than static and mapping stimuli are orange/yellow in colour. The area assumed to be hMT+ is highlighted in green and is labelled KDS11_MT_LH.

In this region we found a significantly higher response to looming motion than to rotation   (t (143) = 2.6, p<0.01). Unlike right hemisphere hMT+, there was no significant difference in response between looming motion and 2D expansion/contraction (t (143) = 1.3, p>0.05) nor between 3D rotation and 2D translation (t (143) = -1.9, p>0.05). However, in the latter contrast, a negative difference almost reached significance, perhaps indicating a slight preference for 2D translation.

No difference was found in response to mapping conditions inside and outside apertures (t (143) = 0.6, p>0.05).

Subject KDS11: V1 Right Hemisphere

V1 in this observer was located at Talairach co-ordinates X: 9, Y: -96, Z: -3 (Figure 5.20) (t (143) = 3.2, p< 0.01).

The area in V1 right hemisphere of KDS11 did not prefer moving over static/flicker stimuli (t (143) = 1.2, p>0.05). It also shows no preference for 3D moving stimuli over 2D moving stimuli (t (143) = 0.7, p>0.05) but does respond stronger to 3D looming motion compared to 3D rotation (t (143) = 2.4, p <0.05). A contrast between 3D looming and 2D expansion/contraction failed to reach significance (t (143) = 1.9, p>0.05). There was also no significant difference in response between 3D rotation and the analogous 2D translation condition (t (143) = -0.9, p>0.05).
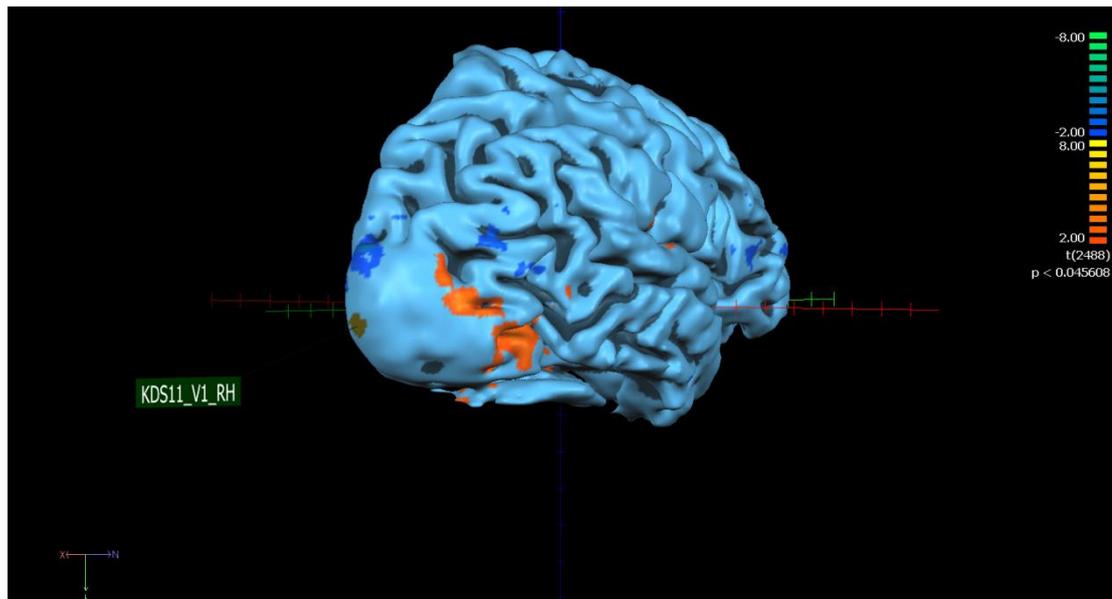
Figure 5.20: Surface view of subject KDS11 V1 right hemisphere. The BOLD response for the right hemisphere, averaged across 4 experimental blocks is mapped onto the surface. Areas where the average BOLD activation was significantly higher in response to the mapping stimulus inside the apertures than to the mapping stimulus outside of the apertures are orange/yellow in colour. The area in V1 is highlighted in green and is labelled KDS11_V1_RH.

Subject KDS11: V1 Left Hemisphere

In the left hemisphere, the location of V1 corresponds to Talairach co-ordinates; X: -16, Y: -96, Z: -1, (Figure 5.21). A contrast between mapping conditions confirmed that this area is more responsive to stimulation inside the circular apertures ($t(143) = 3.6$, $p<0.001$). Unlike right hemisphere V1 in this subject, this area is significantly more responsive to moving stimuli than static/ flicker stimuli ($t(143) = 5.9$, $p<0.0001$) but it was not significantly more responsive to 3D moving stimuli over 2D moving stimuli (t

(143) = 0.8, p>0.05). There was no distinction between response for looming versus

rotation (t (143) = 0.6, p>0.05) and no distinction between either, looming motion and

2D expansion/contraction (t (143) = 1.1, p>0.05) or rotation and 2D translation (t (143)

= 0.003, p>0.05). This area responds well to motion but does not seem to process any

3D features of the stimuli.
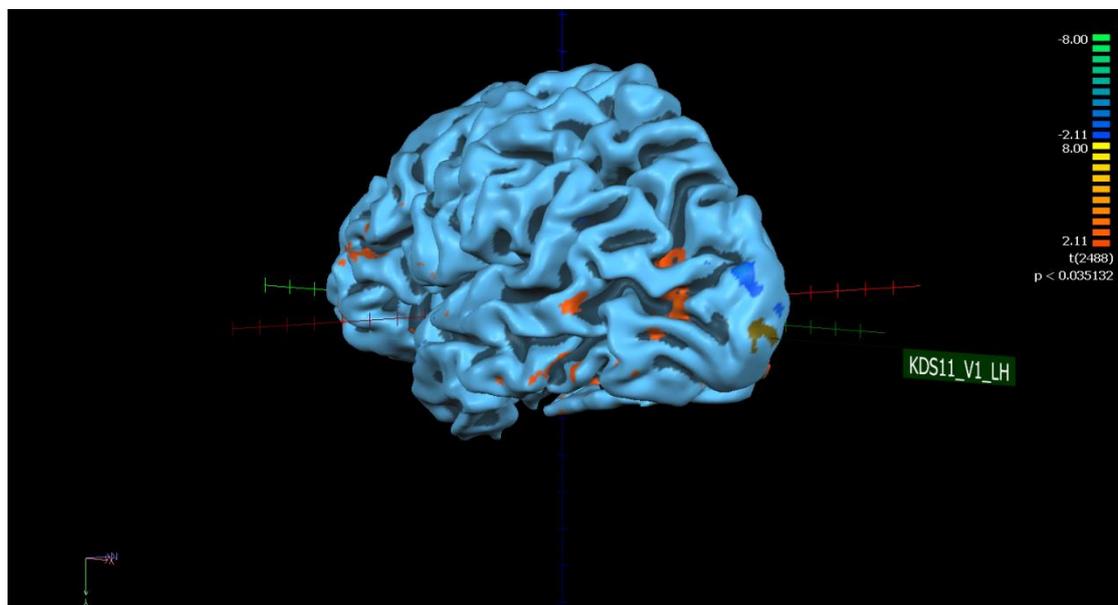
Figure 5.21: Surface view of subject KDS11 V1 left hemisphere. The BOLD response for

the left hemisphere, averaged across 4 experimental blocks is mapped onto the

surface. Areas where the average BOLD activation was significantly higher in response

to the mapping stimulus inside the apertures than to the mapping stimulus outside of

the apertures are orange/yellow in colour. The area assumed to be V1 is highlighted in

green and is labelled KDS11_V1_LH.

Subject BRY25: hMT+ Right Hemisphere

In observer BRY25 hMT+ was located in the right hemisphere in a region corresponding to Talairach co-ordinates; X: 40, Y: -58, Z: 1 (Figure 5.22). This area is significantly more sensitive to motion stimuli than static/flicker stimuli (t (143) = 4.3, p<0.0001). It was significantly more responsive to 3D motion than 2D motion (t (143) = 3.2, p<0.01) but did not distinguish between looming motion and rotation (t (143) = 1.6, p>0.05). This area showed significant preference for looming motion over 2D contraction/expansion (t (143) = 3.8, p<0.01). There was no difference in response between 3D rotation and 2D translation (t (143) = 0.8, p>0.05).
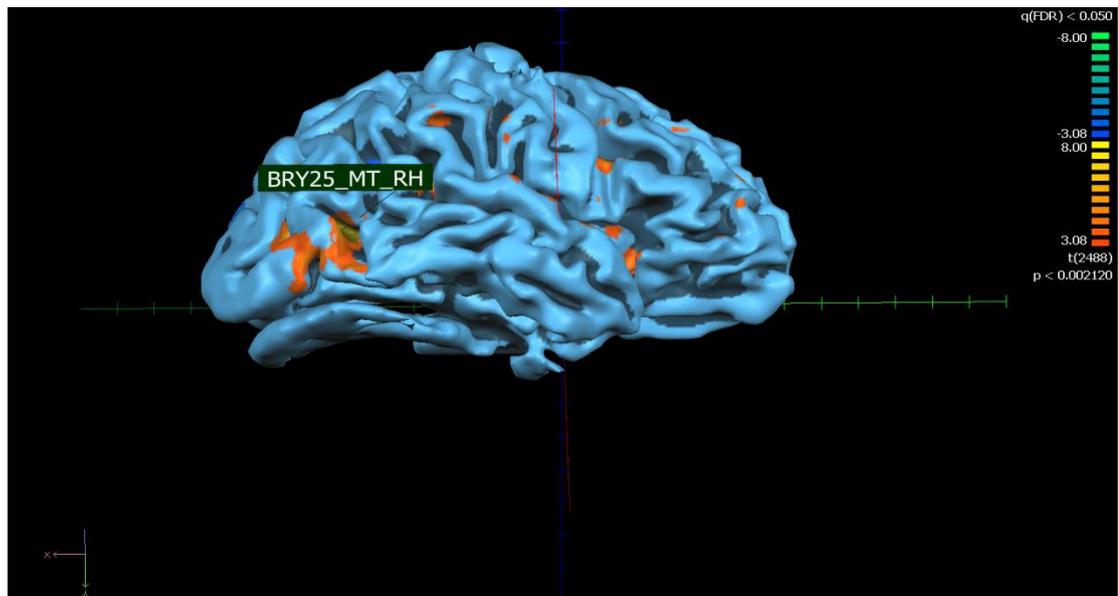
Figure 5.22: Surface view of subject BRY25 MT right hemisphere. The BOLD response for the right hemisphere, averaged across 4 experimental blocks is mapped onto the surface. Areas where the average BOLD activation was significantly higher for moving

stimuli than static and mapping stimuli are orange/yellow in colour. The area assumed to be hMT+ is highlighted in green and is labelled BRY25_MT_RH.

Subject BRY25 hMT+ Left Hemisphere

hMT+ was located in the left hemisphere at Talairach co-ordinates; X: -38, Y: -59, Z: -1 (Figure 5.23). This area was more responsive to moving than static/flicker stimuli (t (143) = 6, p<0.0001). It also showed a preference for 3D motion over 2D motion (t (143) = 4.3, p<0.001). There was no significant difference in activation to 3D looming motion and 3D rotation (t (143) = 0.09, p>0.05) so activation levels in hMT+ did not distinguish between sub-types of stereo-motion.
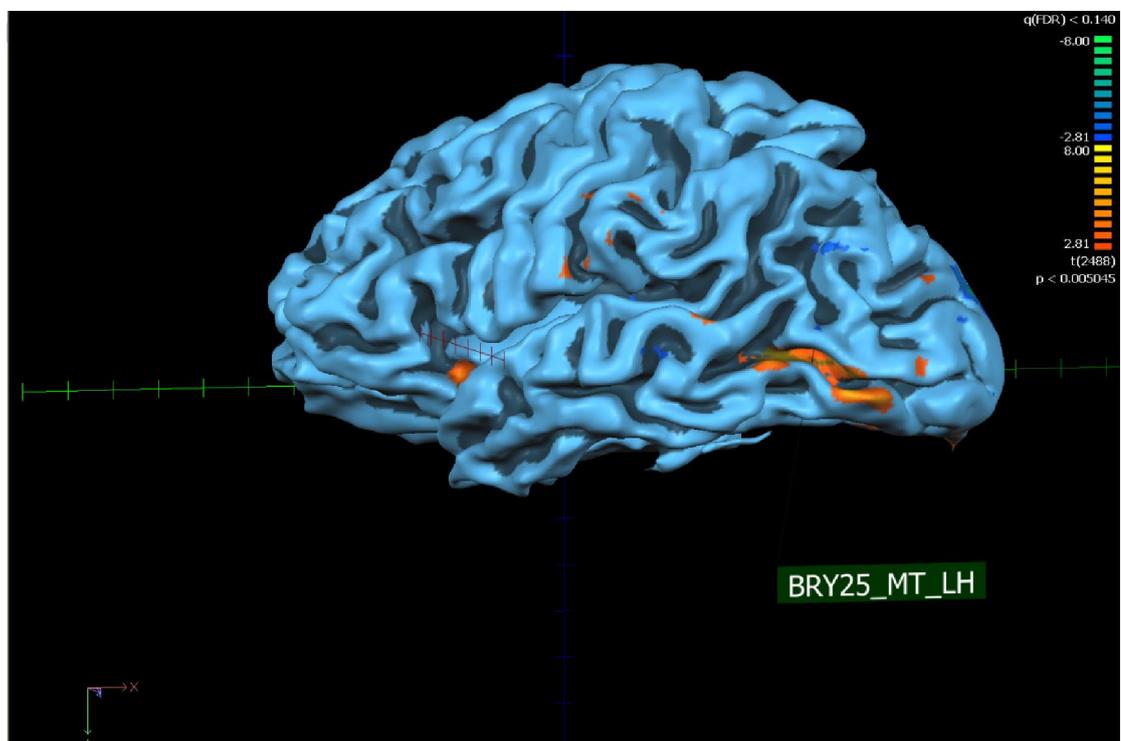
Figure 5.23: Surface view of subject BRY25 MT left hemisphere. The BOLD response for the left hemisphere, averaged across 4 experimental blocks is mapped onto the surface. Areas where the average BOLD activation was significantly higher for moving stimuli than static and mapping stimuli are orange/yellow in colour. The area assumed to be hMT+ is highlighted in green and is labelled BRY25_MT_LH.

A contrast between 3D looming and 2D expansion/contraction yielded a significant result   ($t(143) = 4.3$, $p<0.0001$). hMT+ was also equally sensitive to stimulation inside and outside of the aperture ($t(143) = -0.4$, $p>0.05$) and is therefore more responsive to global stimulus features. This distinction appears to be specific for 3D looming motion as there is no difference in response to 3D rotation and 2D translation in hMT+ ($t(143) = 1.7$, $p>0.05$).

Subject BRY25: V1 Right Hemisphere

In the right hemisphere, V1 was located at Talairach co-ordinates; X: 24, Y: -94, Z: 0, (Figure 5.24)
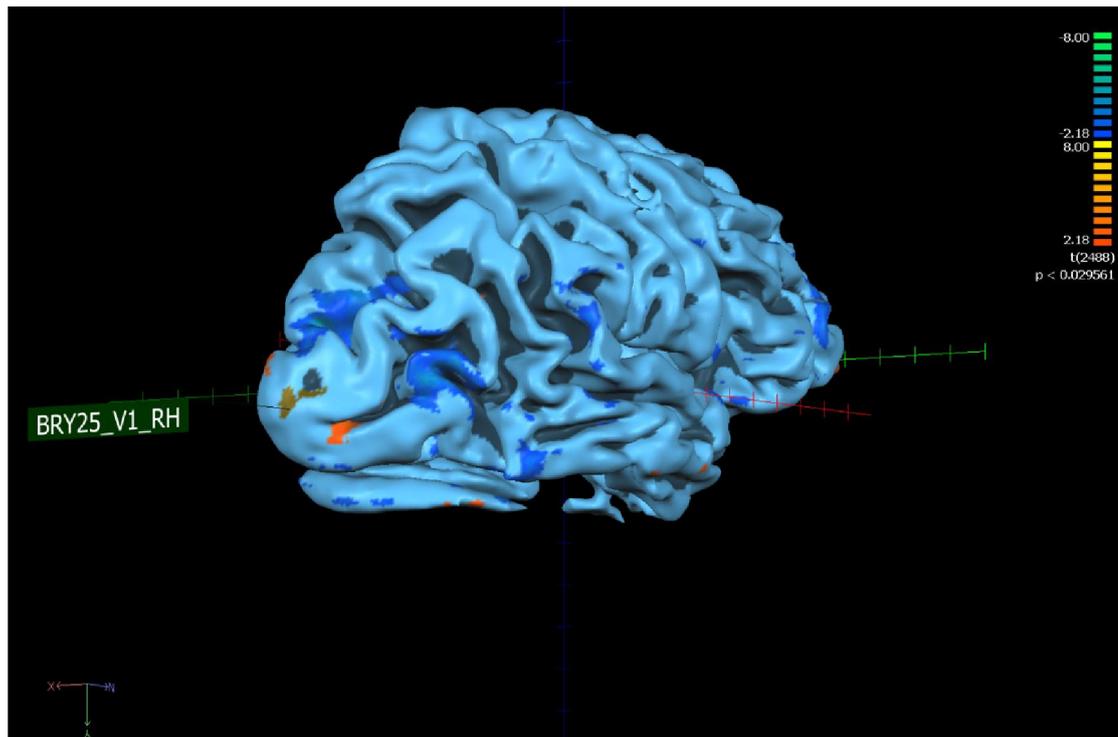
Figure 5.24: Surface view of subject BRY25 V1 right hemisphere. The BOLD response

for the right hemisphere, averaged across 4 experimental blocks is mapped onto the

surface. Areas where the average BOLD activation was significantly higher in response

to the mapping stimulus inside the apertures than to the mapping stimulus outside of

the apertures are orange/yellow in colour. The area in V1 is highlighted in green and is

labelled BRY25_V1_RH.

The most posterior region responded preferentially to stimulation inside of the

apertures (t (143) = 2.1, p<0.05). There was no significant preference for moving over

static stimuli (t (143) = -1.4, p>0.05); and no distinction between 3D and 2D moving

stimuli (t (143) = -0.1, p>0.05) or between sub-types of 3D motion (t (143) = 0.25,

p>0.05). In this observer V1 showed no difference in response magnitude to 3D

looming over 2D expansion/contraction (t (143) =0.02, p>0.05) or to 3D rotation over

2D translation (t (143) = -0.2, p>0.05). This area also showed no particular preference

for motion in general and between sub-types of motion. In right hemisphere hMT+ we found a preference for 3D motion over 2D motion and specifically a strong distinction between looming motion and 2D expansion/contraction (t-test). Since these distinctions are not made in lower processing regions in this subject, it seems that MT may be responding to the binocular component of the stimulus.

Subject BRY25: V1 Left Hemisphere

In this observer V1 was located at Talairach co-ordinates; X: -15, Y: -100, Z: 1, (Figure 5.25) and was confirmed to be selective for activity inside the apertures ($t(143) = 2.4$, $p < 0.05$). This area was not found to respond stronger to moving over static stimuli ($t(143) = -1.2$, $p > 0.05$), made no distinction between 3D and 2D motion ($t(143) = -0.9$, $p > 0.05$) or between sub-types of 3D motion ($t(143) = 0.9$, $p > 0.05$). There was also no difference in activity for looming versus 2D expansion/contraction ($t(143) = -0.5$, $p > 0.05$) or for 3D rotation versus 2D translation ($t(143) = -0.8$, $p > 0.05$). Activation in the left hemisphere V1, in this subject, was not particularly responsive to motion and did not distinguish between motion types.
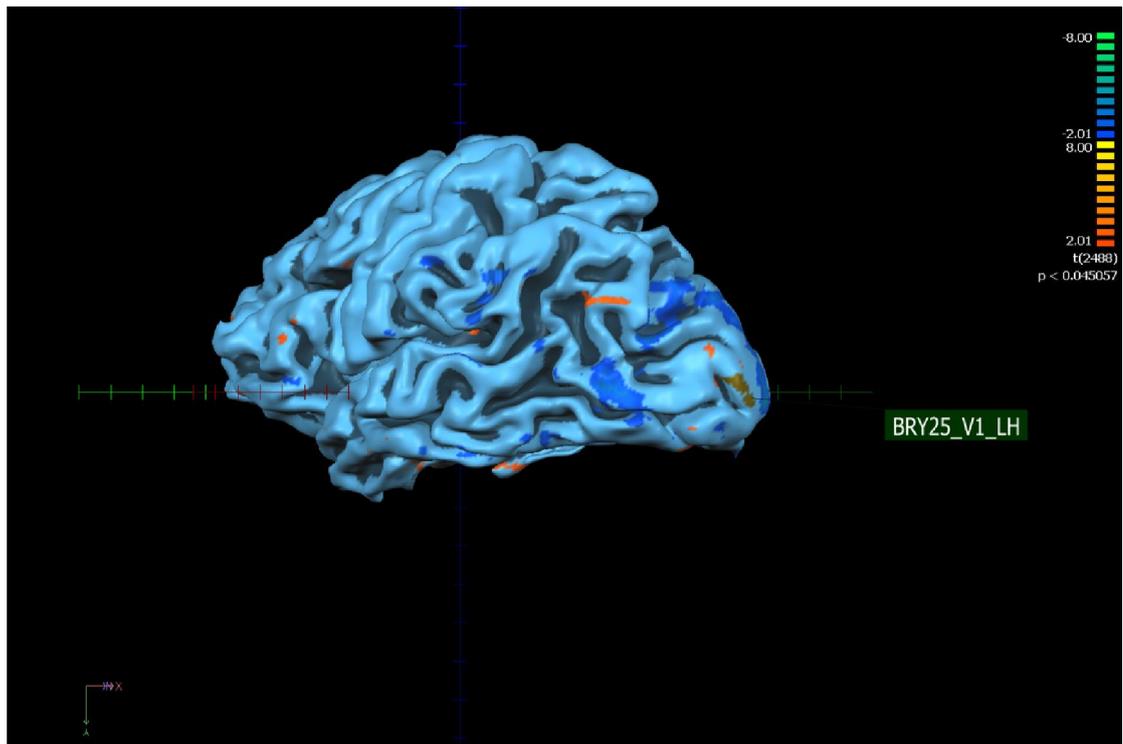
Figure 5.25: Surface view of subject BRY25 V1 left hemisphere. The BOLD response for the left hemisphere, averaged across 4 experimental blocks is mapped onto the surface. Areas where the average BOLD activation was significantly higher in response to the mapping stimulus inside the apertures than to the mapping stimulus outside of the apertures are orange/yellow in colour. The area in V1 is highlighted in green and is labelled BRY25_V1_LH.

Subject MKE22: hMT+ Right Hemisphere

hMT+ was located in a region corresponding to Talairach co-ordinates; X: 46, Y:-64 , Z: 3, (Figure 5.26) This showed significantly higher response to moving than static stimuli in general (t (143) = 7.8, p<0.0001). There was a significant preference for 3D motion over 2D motion (t (143) = 3.4, p<0.001). This preference seemed to be driven mainly by a preference for looming motion, since it also distinguished between looming motion and 3D rotation (t (143) = 2.3, p<0.05) and between looming and the analogous 2D expansion/contraction (t (143) = 3.6, p<0.001). There was no difference in response to 3D rotation and 2D translation (t (143) = 1.2, p>0.05). There was difference in response magnitude to mapping conditions inside and outside (t (143) = 1.2, p>0.05).
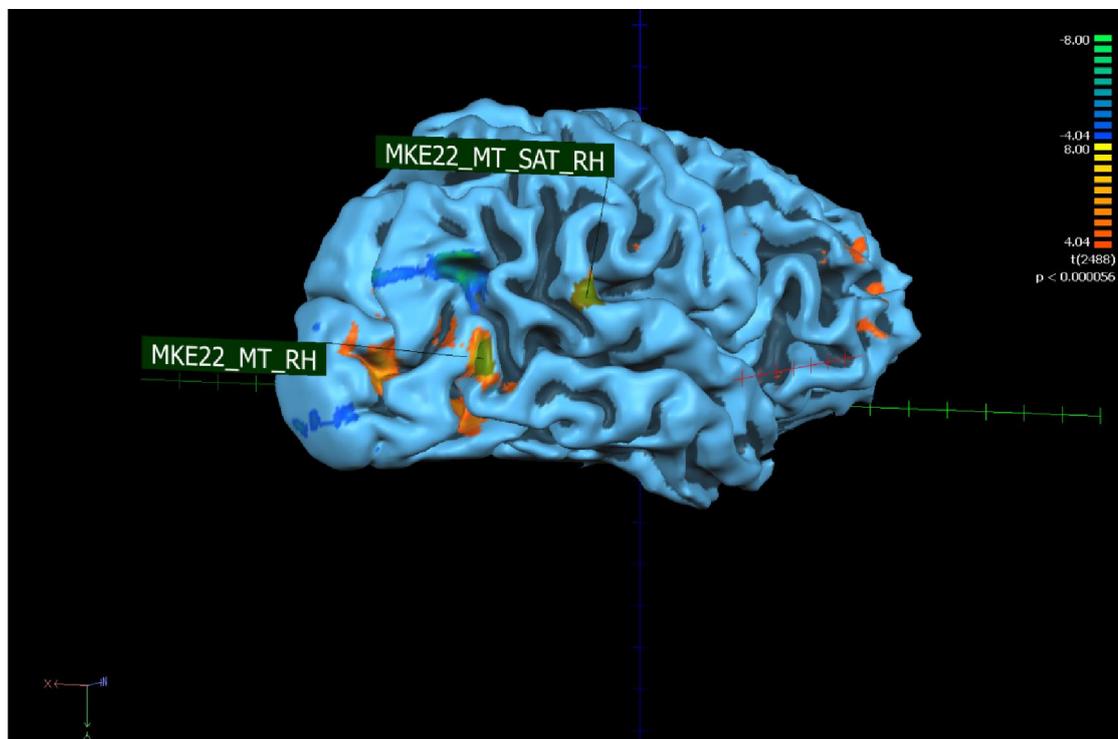
Figure 5.26: Surface view of subject MKE22 MT right hemisphere. The BOLD response for the right hemisphere, averaged across 4 experimental blocks is mapped onto the

surface. Areas where the average BOLD activation was significantly higher for moving stimuli than static and mapping stimuli are orange/yellow in colour. The area assumed to be hMT+ is highlighted in green and labelled MKE22_MT_RH. Also shown on this surface model is a region further along the visual hierarchy appears to show a very significant preference for looming motion, labelled MKE22_MT_SAT_RH.

Subject MKE22: hMT+ SATELLITE Right Hemisphere

An area beyond MT (see Figure 5.26), highlighted in the same contrast used to localize MT showed response properties similar to hMT+ but is placed at a later stage in the visual hierarchy. As well as being significantly more responsive to moving than static stimuli ($t (143) = 7.3$, $p<0.0001$) this area showed a convincing preference for 3D motion stimuli over 2D motion stimuli ($t (143) = 5.6$, $p<0.0001$) and an even stronger preference for looming motion over 3D rotation than we observed in right hemisphere MT ($t (143) = 3.1$, $p<0.002$).

In addition, activation in this MT satellite area clearly distinguished between looming motion and the analogous 2D expansion/contraction ($t (143) = 5.9$, $p<0.0001$). A small but significant distinction between 3D rotation and 2D translation was also noted ($t (143) = 2$, $p<0.05$). Since it is a higher order processing region and makes no significant distinction between mapping conditions ($t (143) = 1.9$, $p>0.05$), we suggest that this area may be specifically responsive to the global 3D looming motion.

Subject MKE22: hMT+ Left Hemisphere

176

Left hemisphere hMT+ was localised in a region corresponding to Talairach co-ordinates: X: -38,Y: -72, Z: -2 (Figure 5.27). This area shows significant preference for moving over static/flicker stimuli (t(143) = 4, p<0.0001). Activation in left hemisphere hMT+ also distinguished between 3D and 2D moving stimuli (t (143) = 4.1, p<0.0001), suggesting a strong preference for 3D motion. Activation levels in this area did not distinguish between 3D looming motion and 3D rotation (t (143) = 0.56, p>0.05) but responded preferentially to the 3D motion conditions compared to the corresponding 2D conditions. The response to 3D looming motion was significantly higher than the response to 2D expansion/contraction (t(143) = 2.9, p<0.01). In adition the response to 3D rotation was significantly higher than to 2D translation (t (143) = 2.9, p<0.01). Again, there was no difference in response to mapping stimuli inside and outside.
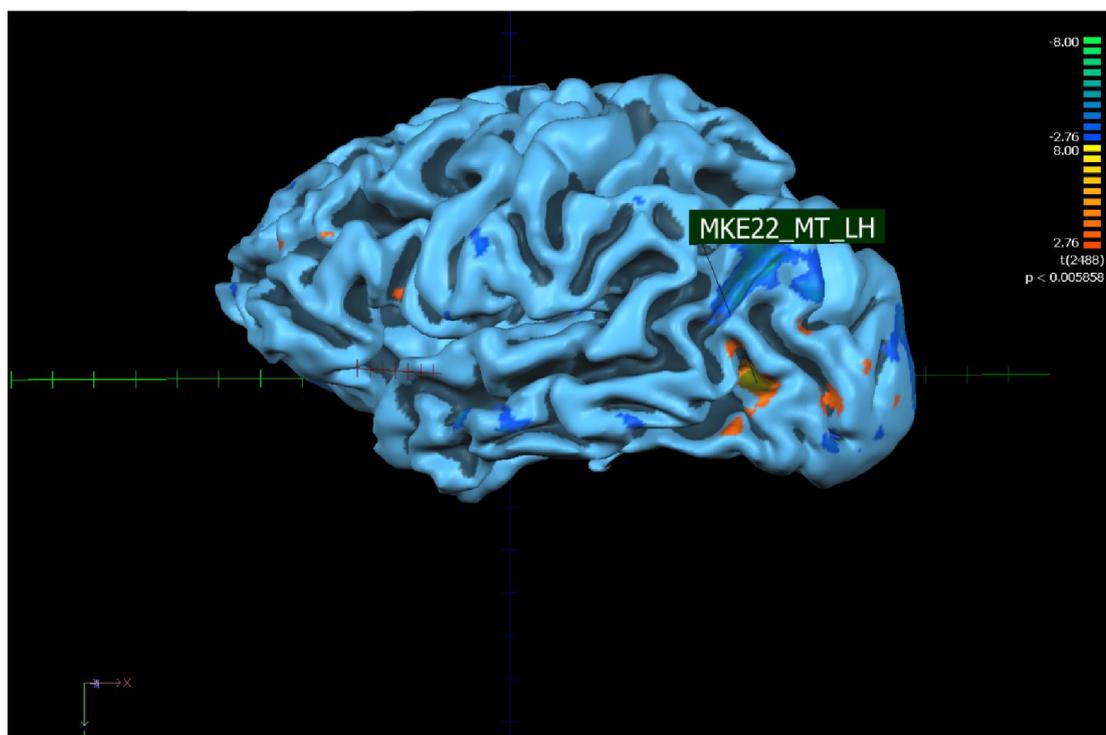
Figure 5.27: Surface view of subject MKE22 MT left hemisphere. The BOLD response for the left hemisphere, averaged across 4 experimental blocks is mapped onto the surface. Areas where the average BOLD activation was significantly higher for moving stimuli than static and mapping stimuli are orange/yellow in colour. The area assumed to be hMT+ is highlighted in green and is labelled MKE22_MT_LH.

Subject MKE22: V1 Right Hemisphere

V1 in the right hemisphere was located at Talairach co-ordinates; X: 13, Y: -98, Z: -3 (Figure 5.28) and was significantly more responsive to stimulation inside the apertures ($t (143) = 6.2$, $p<0.0001$). In line with our original hypothesis, but contrary to activation in the majority of previous subjects, V1 did not respond preferentially to moving over static stimuli ($t (143) = -1.5$, $p>0.05$). There was no significant preference for 3D moving stimuli over 2D moving stimuli ($t (143) = -2.8$, $p>0.05$).

Interestingly, activation levels in right hemisphere V1, distinguished between 3D looming motion and 2D rotation, but in the opposite direction as observed in hMT+ for other subjects ($t (143) = -4.1$, $p<0.0001$). Activation in V1 in this observer appeared to prefer 3D rotation over 3D looming. This may reflect disparity modulation in line components, as there was also a significant negative result in the contrast between 3D looming motion and the corresponding 2D expansion/contraction condition ($t(143) = -3.5$, $p<0.001$). We did not observe the same distinction between 3D rotation and 2D translation ($t (143) = -0.3$, $p>0.05$), suggesting that the distinction between looming and rotation may reflect a large negative response to looming.
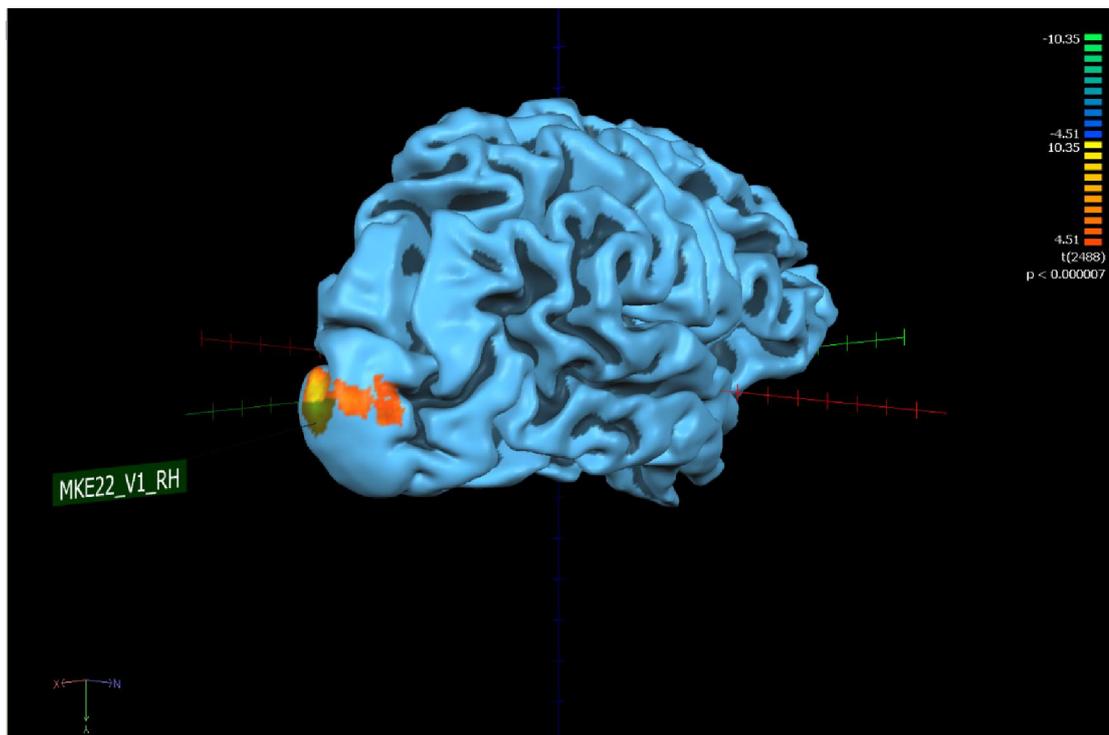
Figure 5.28: Surface view of subject MKE22 right hemisphere. The BOLD response for the

right hemisphere, averaged across 4 experimental blocks is mapped onto the surface.

Areas where the average BOLD activation was significantly higher in response to the

mapping stimulus inside the apertures than to the mapping stimulus outside of the

apertures are orange/yellow in colour. The area in V1 is highlighted in green and is labelled

MKE22_V1_RH
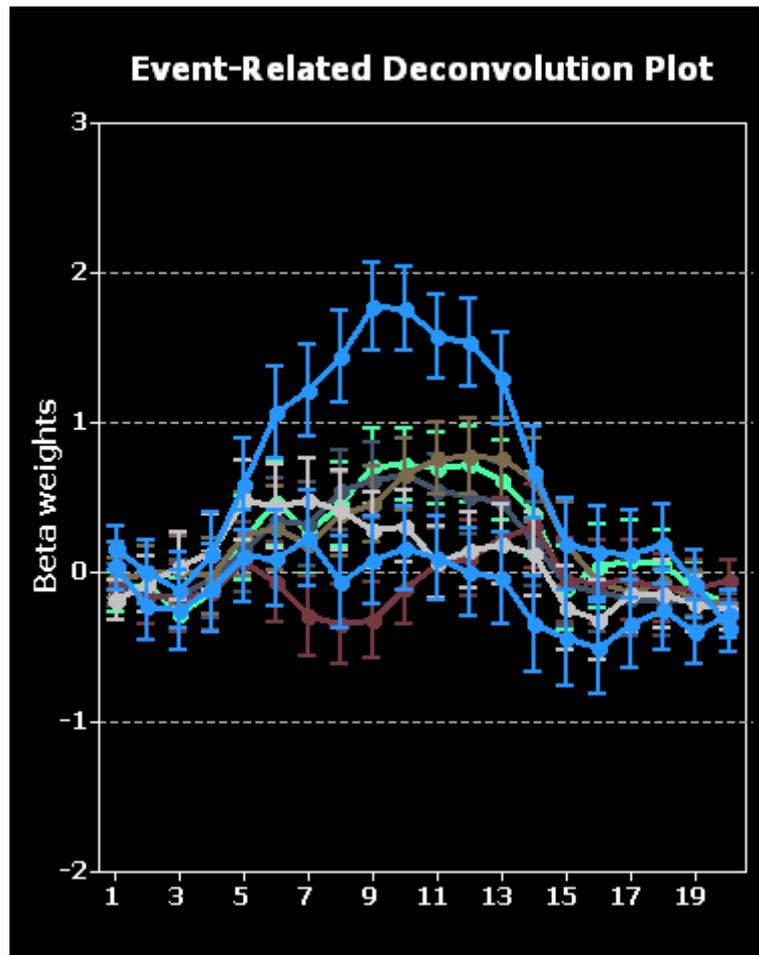
Figure 5.29: The event related de-convolution plot for subject MKE22 V1 right hemisphere. The response for 3D looming motion against time are plotted in red. At the time points of interest (8-11) we see a negative response in this area to 3D looming stimuli. This may account for the significant difference observed between peak responses (averaged over time-points 8-11) for 3D looming motion versus 3D rotation. This suggests that, rather than a preference for 3D rotation, this area appears to treat 3D looming motion differently from all other conditions, perhaps representing a role of higher order feedback in inhibiting response to the looming stimulus.

Subject MKE22: V1 Left Hemisphere

V1 was located in the left hemisphere in the region corresponding to Talairach co-ordinates; X: -14, Y: -97, Z: -11 (Figure 5.30). This area showed a significantly stronger response to stimulation inside of the apertures (t (143) = 5.9, p<0.0001). The motion contrast gave a significant negative t-value (t(143) = -2.4, p<0.05), also showing no preference for moving over static stimuli. This area showed a strong response in the opposite direction, appearing to prefer static over moving stimuli. Again, there is no significant difference in response magnitude between 3D and 2D moving stimuli (t (143) = -1.2, p>0.05).

Consistent with right hemisphere V1 in this subject, a distinction was made between 3D looming motion and 3D rotation (t (143) = -3, p<0.01), so this area showed a significant preference for 3D rotation. Again, looking at the de-convolution plot (Figure 5.26) we see that this is due to a large negative response to 3D looming motion rather than a large positive response to 3D rotation.

This area also distinguished between 3D looming and 2D expansion/contraction (t(143) = -2.2, p<0.05). Again there was no distinction between 3D rotation and 2D translation (t(143)= 0.5, p>0.05) and the t-value here was positive, indicating that there wasn't the same negative response to 3D rotation as for 3D looming motion.
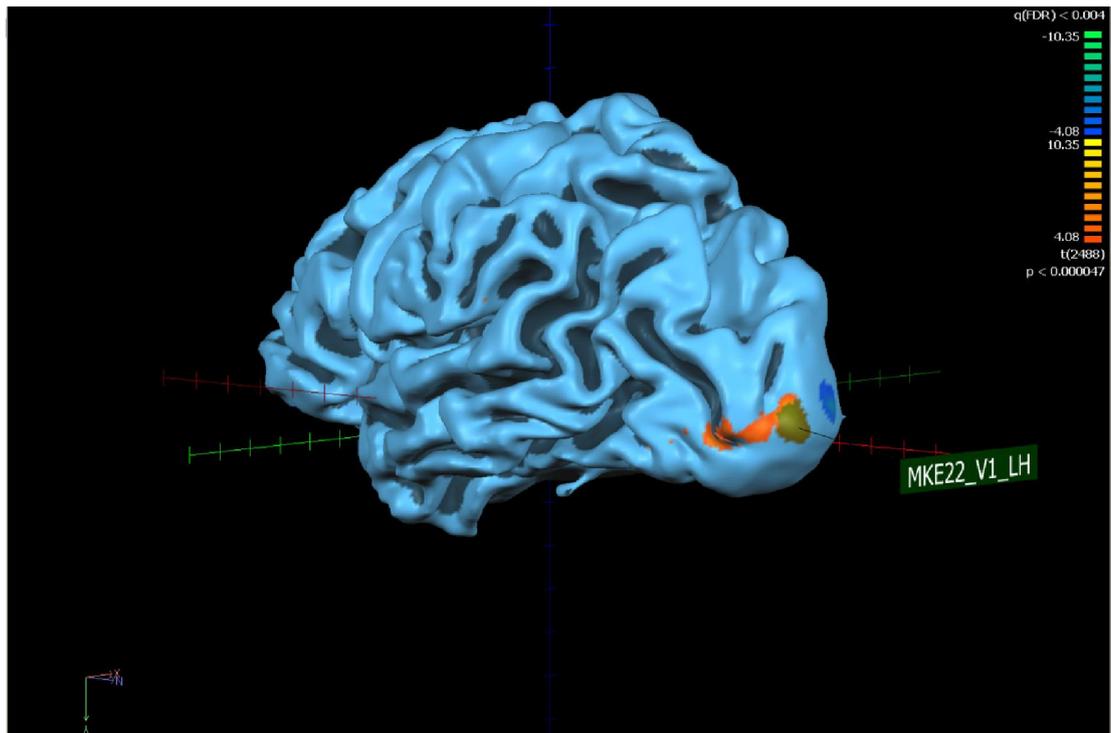
Figure 5.30: Surface view of subject MKE22 V1 left hemisphere. The BOLD response for the

left hemisphere, averaged across 4 experimental blocks is mapped onto the surface.

Areas where the average BOLD activation was significantly higher in response to the

mapping stimulus inside the apertures than to the mapping stimulus outside of the

apertures are orange/yellow in colour. The activated area in V1 is highlighted in green and

is labelled MKE22_V1_LH.

Results Summary hMT+/V5

Subject	Motion vs Static (p-value)	3D vs 2D	Looming vs Rotation	3D Loom vs 2D Expa	3D Rot vs 2D Trans	Mapin vs Mapout	Hemisphere
LWA26	<0.0001	<0.0001	>0.05	<0.0001	<0.0001	>0.05	R
LWA26	<0.0001	<0.0001	>0.05	<0.0001	<0.0001	>0.05	L
FFN20	<0.0001	<0.0001	<0.02	<0.0001	<0.0001	>0.05	R
FFN20	<0.0001	<0.05	>0.05	<0.03	>0.05	>0.05	L
AME14	<0.0001	>0.05	>0.05	>0.05	>0.05	>0.05	R
AME14	<0.0001	>0.05	>0.05	>0.05	>0.05	>0.05	L
KDS11	<0.0001	>0.05	>0.05	<0.05	>0.05	>0.05	R
KDS11	<0.0001	>0.05	<0.01	>0.05	>0.05	>0.05	L
BRY25	<0.0001	<0.01	>0.05	<0.01	>0.05	>0.05	R
BRY25	<0.0001	<0.001	>0.05	<0.0001	>0.05	>0.05	L
MKE22	<0.0001	<0.001	<0.05	<0.001	>0.05	>0.05	R
MKE22	<0.0001	<0.0001	>0.05	<0.01	<0.01	>0.05	L
MKE22	<0.0001	<0.0001	<0.002	<0.0001	<0.05	>0.05	SAT_RH

Table 5.2: Summary of individual results, area hMT+/V5 left and right hemisphere.

Results are grouped by subject and show the p-value associated with each of the 6

GLM contrasts. Shown in yellow are results from subjects which can be grouped in

terms of significant preference for motion, 3D motion and preference for at least one

of either: 3D Looming-2D expansion/contraction, or, 3D rotation-2D translation. In

green are results grouped based on significant preference for motion but no

preference for any of the 3D conditions. In red are the results from the satellite region beyond hMT+/V5 located in subject MKE22 right hemisphere, since these are of theoretical interest.

Table 5.2 summarises the individual results of 6 observers, for hMT+/V5 in the left and right hemisphere. As expected, across both hemispheres and in all observers, there was significantly higher response to moving stimuli than static stimuli and no significant difference in response to the two mapping conditions.

In 4 out of 6 observers (LWA26, BRY25, MKE22, FFN20) MT was significantly more responsive to 3D motion than 2D motion. For subjects LWA26, BRY25 and MKE22, this effect was observed in both hemispheres to a similar extent, subject FFN20 shows a degree of lateralisation favouring the right hemisphere. To clarify, by 'lateralisation' here, and in the discussion following, we refer to a stronger significance level for the effect in one hemisphere over another.

Of these 4 observers, only FFN20 and MKE22 showed greater response to 3D looming over 3D rotation, this effect was lateralised to the right hemisphere. Not surprisingly general activation does not suggest a distinction between sub-types of stereo-motion in MT. All 4 observers showed significantly higher activation in MT (both hemispheres) in response to 3D looming motion compared to 2D expansion/contraction condition. However, only LWA26 showed a significantly higher response to 3D rotation over 2D translation in both hemispheres. In subjects FFN20 and MKE22 the effect was lateralised, to the right hemisphere for FFN20 and to the left hemisphere in MKE22. In subject BRY25 no significant difference was observed for 3D rotation>2D translation in either hemisphere. The similarity in response pattern in observers LWA26, FFN20,

BRY25 and MKE22 seems to be consistent enough to group them together for discussion.

The response pattern in hMT+/V5 for the remaining 2 observers (AME14 and KDS11) is also strikingly similar. There was no significant preference for 3D motion over 2D motion in this area. Subject KDS11 showed significant preference for looming motion over rotation, but this effect was lateralised to the left hemisphere. There was no significant difference in activation for looming motion versus 2D expansion/contraction in either KDS11 or AME14 (both hemispheres) and no difference between 3D rotation and 2D translation.

Results Summary V1

Subject	Motion vs Static (p-Value)	3D vs 2D	Looming vs Rotation	3D Loom vs 2D Expa	3D Rot vs 2D Trans	Mapin vs Mapout	Hemisphere
LWA26	<0.02	<0.01	>0.05	<0.01	<0.01	<0.0001	R
LWA26	<0.05	<0.0001	>0.05	<0.0001	<0.0001	<0.01	L
FFN20	<0.0001	<0.01	>0.05	>0.05	>0.05	<0.0001	R
FFN20	<0.01	>0.05	>0.05	>0.05	>0.05	<0.0001	L
AME14	<0.0001	>0.05	<0.05 (-ve)	>0.05	>0.05	<0.05	R
AME14	<0.0001	>0.05	<0.05 (-ve)	>0.05	>0.05	<0.0001	L
KDS11	>0.05	>0.05	<0.05	>0.05	>0.05	<0.01	R
KDS11	<0.0001	>0.05	>0.05	>0.05	>0.05	<0.001	L
BRY25	>0.05	>0.05	>0.05	>0.05	>0.05	<0.05	R
BRY25	>0.05	>0.05	>0.05	>0.05	>0.05	<0.05	L
MKE22	>0.05	>0.05	<0.0001 (-ve)	<0.001 (-ve)	>0.05	<0.0001	R
MKE22	<0.05 (-ve)	>0.05	<0.01 (-ve)	<0.05 (-ve)	>0.05	<0.0001	L

Table 5.3 : Results summary for V1. Results are grouped by subject and hemisphere.
Right hemisphere results are shown across the rows in dark grey. The associated p-
value for each of the 6 GLM contrasts are given. For the left hemisphere p-values for
each of the 6 GLM contrasts are shown across the rows in light grey.

Since the pattern of response in V1 showed far greater individual variation, we refrained from grouping any of the observers together based on response pattern. In 4 out of 6 subjects we observed greater activation in response to moving over static stimuli. In subject KDS11 this was lateralised in the left hemisphere but the effect was highly significant ($p < 0.0001$). In subject BRY25 there was no significant difference in response to moving and static stimuli in V1, whilst subject MKE22 showed a small but significant preference for static stimuli in the left hemisphere and no difference in the right.

A significantly larger response to 3D over 2D motion was observed in only 2 observers, LWA26 and FFN20, in the latter this was lateralised in the right hemisphere V1. The remaining 4 observers showed no significant difference in response to 3D over 2D moving stimuli. In four observers (LWA26, FFN20, KDS11, BRY25) there was no difference in activation to 3D looming and 3D rotation motion. Interestingly in the other 2 observers (MKE22, AME14) there was significant negative activity (below baseline) associated with viewing looming motion, this resulted in a significant negative result, since in these subjects there was little response to 3D rotation. Perhaps, indicating that the perception of looming motion is governed by higher order processing in these observers.

Only in observer LWA26, there was an increased response to looming motion compared to 2D expansion/contraction. In MKE22, again we observed a significant negative response to looming motion that gave a significant contrast. For the remaining 4 observers there was no significant difference in response. LWA26 was also the only case where a significant difference in activation was observed for 3D rotation

compared to the analogous 2D translation condition (both hemispheres). In all other observers there was no significant difference in response to these stimuli. In all observers there was significantly higher response to stimulation inside the mapping condition than outside, so any differences in response are most likely to be caused by more peripheral location of the outside stimulation.

5.3 DISCUSSION

Data from six naïve observers were analysed. In an fMRI experiment observers viewed four types of diamond motion in a display with four apertures (see stimuli): (1) 2D expanding/contracting motion (2) 2D translational motion (3) 3D rotational motion (4) 3D Looming motion. The purpose of this experiment was to investigate differences in early visual processing regions (V1) and later, more specialised, visual processing regions (hMT+/V5) in response to 2D and 3D global motion displays. An important feature of the motion conditions was that local monocular motion of the line segments remained the same when different types of a globally moving diamond were generated. This means that monocular motion processing in early visual regions should be the same but integration into global object motion and disparity input differed.

Role of hMT+/V5 in Stereo-motion Perception

This study was exploratory in nature, as to our knowledge no-one has yet studied global stereo-motion perception in this way. Nevertheless, there were a few noteworthy findings and a strong basis for continuation of this paradigm in the future. One finding was a general trend towards preference for three-dimensional over two-

dimensional motion in hMT+/V5. This was seen in 4 out of 6 observers. There was also significant distinction between corresponding 3D-2D motion pairs. This suggests that preference for stereo-motion observed in hMT+/V5 is based specifically on the binocular component in the stimuli.

Likova & Tyler (2007) identified a region of cortex anterior to hMT+ which is specialised for stereo-motion even when controlled for static disparity, i.e. activation is due to dynamic disparity rather than static disparity. The region is of roughly equal in size to the entire hMT+ region and located anterior to hMT+ as well as being adjacent to/partially overlapping with the motion complex. In our study we identified selectivity for stereo-motion in hMT+ itself, although we cannot discount the role of static disparities. In a future version of this experiment a static stimulus should be included, since the motion complex MT (V5) MST is known to contain both motion and disparity selective cells, with many of these cells sensitive even to static disparities (Palanca & DeAngelis, 2003).

Disparity selective populations within hMT+ have the potential to provide input to a mechanism calculating disparity modulation over time (CDOT) and, although quite out-with the scope of this pilot study, there is potential for exploring the contribution of IOVD and CDOT in stereo-motion selectivity in hMT+.

This has been investigated previously by Bas Rokers and colleagues who found a unique preferential response in hMT+ to horizontally opponent, dichoptic motion - the only stimulus that simulated three-dimensional motion - using stimuli containing both

IOVD and CDOT information. They carefully speculated that 'MT+ carries signals which are specific to stereomotion' (Rokers et al., 2009).

To investigate the role of IOVD and CDOT in MT, they used dynamic random dot stimuli and isolated the changing disparity cue from interocular velocity. They found a significant preference on MT for coherent CDOT defined stereo-motion versus a spatio-temporally scrambled version of the stimulus which did not result in perception of stereo-motion. The contribution of the IOVD signal was isolated using binocularly anti-correlated stimuli, it was found that IOVD signals contributed to stereomotion selectivity but to a lesser extent than CDOT signals (Rokers et al., 2009). Again this points to the importance of investigating the role of disparity-based signals in our display.

Most importantly, Rokers et al asked whether MT was simply a carrier of stereo-motion information or whether it was integral in stereo-motion perception. The presence of direction selective adaptation in MT in response to stereomotion input, was taken as evidence that MT makes a robust contribution to stereo-motion perception.

Applied to our findings, it seems we have good reason to believe that the selectivity for 3D motion observed in hMT+ reflects a strong contribution of this region in global stereo-motion perception. In future, we may consider using an adaptation paradigm to test whether there is directional adaptation for our 3D motion stimuli in hMT+/V5. If

this can be shown we have strong evidence that MT plays more than a passive role in global stereo-motion perception.

Another possible research direction we may consider is to investigate the contribution of CDOT and IOVD information to stereomotion perception in hMT+. This would provide an important link with our investigation of trajectory judgements in low level stereomotion perception so that we could begin to trace the interaction of these cues throughout the visual hierarchy. This approach could also be taken at stages in between V1 and hMT+, and indeed, we have some preliminary results for subject MKE22 from intermediate retinotopic regions using the current basic paradigm.

Individual Differences in hMT+

In two of our observers, there was no apparent distinction between three-dimensional and two-dimensional motion in hMT+/V5. One observer showed preference for looming motion over rotational motion in this region but did not distinguish between looming and the analogous 2D condition expansion/contraction. There are various possibilities to consider here, the first being that upon failing to carry out a robust test of stereomotion perception, we have included observers who may be perceptually ill-suited for this study.

As we mentioned in our chapter on stereo testing, stereomotion specific deficits have been identified in 'normal' observers (Richards, 1971; Richards & Regan, 1973; Regan, Erkelens, & Collewijn, 1986; Miles, 1998), therefore testing only static stereopsis is a

potential problem. Since we concluded in our previous work (see Chapter 2- 4) that

CDOT and IOVD cues are likely to be processed in parallel streams, it is possible that a

deficit in either or both streams may interrupt perception of a 3D motion stimulus.

There is a growing body of evidence in support of specific low-level stereomotion

deficits (Watanabe et al, 2008; Nefs, O'Hare, & Harris, 2010).

Observers with stereomotion deficits may rely upon top down/global processing of the

stimulus and therefore may show a different response in hMT+. This would not

necessarily affect the ability to perceive our global displays, but may be reflected in

brain activity. It is important not to make the mistake of considering visual processing

and visual perception to be one and the same thing.

In this case the behavioural data would provide an unsatisfactory way of discounting

this. The two-alternative forced choice task (2D/3D motion) used in this experiment

does not provide a way to assess individual differences in underlying processing

strategies, so we cannot draw firm conclusions on individual differences in our

observers.

Role of V1 in Stereo-motion Processing

In terms of V1 activity in response to the various levels of independent variable, we

observed mixed results. Four out of the six observers showed a significant preference

for moving over static stimuli in this area. This was unanticipated, since early visual

regions are not known to be particularly functionally specific. However, upon looking

at our approach to analysis, it is possible we selected regions of V1 dedicated to

processing of local motion inputs - since we selected those regions which responded favourably to stimulation inside of the apertures.

In subsequent versions of this experiment, the use of retinotopic mapping to identify regions of interest, rather than the coarse method used here, is a definite requirement.

In the majority of observers the results of the remaining contrasts were as expected, there was little evidence for selectivity to three-dimensional moving stimuli over two-dimensional moving stimuli or for any contrasts between 2D and 3D motion conditions. There are two notable exceptions to this, observers LWA26 and MKE22, who showed distinct response patterns and they will be discussed individually in the following.

Subject LWA26

The response pattern of observer LWA26 was similar in hMT+/V5 and in V1. In V1 there was substantial motion selectivity, a very significant preference for three-dimensional motion and also for the 3D versions of corresponding 2D/3D motion types. There are two possible approaches to explain this. First, since the only difference between response in hMT+/V5 and V1 in this subject was a preference in V1 for stimulation within the circular apertures compared to outside the apertures, any distinction between three-dimensional and two-dimensional stimuli in V1 may be

based on local components. If this is the case, the activation that we see could reflect feedback information. Why this is case for observer LWA26 only, is an outstanding question.

The second possibility is that individual observers may have taken different approaches to carrying out the task. We encouraged observers to maintain fixation throughout the trials so as to prevent them from deliberately tracking differences in local motions. It is possible that some individuals may have used only local cues to discriminate between the different types of motion. Since we deliberately preserved local motion cues (by relying on completion to obtain global motion percepts) this was always a potential pitfall.

Subject MKE22

Of all of the observers, MKE22 showed a response pattern that was in line with our predictions. In area V1 there was a significant negative result in the left hemisphere for the motion contrast. This suggests a preference for static/flickering stimuli in V1. There was no preference for three-dimensional motion over two-dimensional motion in general and most importantly, a strong negative response to looming motion, both in the looming-rotation contrast and in the contrast between looming and the analogous 2D expansion/contraction. This was not observed for the rotation versus 2D translation contrast and so the effect seems to be specific to looming motion. Suggesting that for this subject the perception of looming motion is a higher-order function.

Additional evidence for this comes from the identification of a 'satellite' region to MT in this observer. This area is located anterior to hMT+/V5 in the right hemisphere and shows a significant preference for looming motion over 3D rotation ($p<0.002$) as well as selectivity for looming over the analogous 2D expansion/contraction condition. This suggests specialisation for looming motion in this region that might explain the strong negative response in V1. It is unclear whether this area is likely to be present in all observers.

Future Research Directions

Firstly, we have evidence that hMT+/V5 is involved in the processing of global stereo-motion perception which corroborates findings by the main studies in this area (Likova & Tyler, 2007; Rokers et al. 2009; Ban et al., 2012).

hMT+ is involved in processing stereo-motion information. In order to assess the importance of the hMT+ contribution we can use MVPA (multivariate pattern analysis) to distinguish between types of 3D motion in different regions of interest.

In an additional study we could use an adaption paradigm in conjunction with MRI to study selective adaption in hMT+/V5 to our 3D stimuli. Repeated exposure to the same stimulus causes an increasingly attenuated response in the activated neuronal population. With respect to our stimulus, this can be exploited in at least two ways. 1) To compare activation for 3D stimuli and their associated 2D pairings (i.e. Looming versus 2D expansion/contraction; Rotation versus 2D translation). In order to show

more robustly that these stimulus pairings are treated as distinct in hMT+/V5 but may not be in earlier visual regions such as V1. For example, when several exposures to looming motion is followed by presentation of 2D expansion/contraction we may expect to see a boost in activation in MT but not in lower regions, where a 3D/2D distinction is not made. 2) To show directional adaptation to 3D motion stimuli in hMT+/V5. To do this we would create two rotating stimuli, one where rotation occurs to the left around the vertical axis and one in which rotation is to the right around the vertical axis. We would expect an attenuated response to rotation in the same direction, but no adaptation for motion in the opposite direction. We could also use looming and rotating stimuli, to show that adaptation to looming motion is reversed when shown a rotating stimulus. This would show that hMT+/V5 distinguishes between 3D motion types in terms of adaptation even if activation alone does not show this

In terms of being able to understand the processing of stereo-motion from basic low level visual processing to higher order perceptual processing, we must also provide information about intermediate levels of the visual hierarchy. It has become clear that the best way to conceptualise stereo-motion processing is not as a series of unique processing stages, but as a constantly changing and interactive loop of processes.

Another consideration is to link our MRI findings with earlier findings on parallel processing streams for IOVD and CDOT driven stereo-motion. Our stimuli contain both CDOT and IOVD cues to motion-in-depth, however, there is some indication that the stereo-component of the binocular displays is particularly appealing to hMT+. Rokers

and colleagues found that the main contribution to cyclopean motion selectivity came from the CDOT mechanism, but that the role of IOVD cues is by no means redundant (Rokers et al., 2009). To investigate this further we must derive a way to isolate CDOT and IOVD components in our displays. This is not a trivial task but would provide an important link to our previous research.

One simple improvement that can be made to this paradigm is the use of an explicit retinotopic mapping procedure to identify regions of interest. The method used here was coarse and time consuming and may have had implications for the analysis, particularly in V1. It would be equally important to carry out behavioural experiments to train observers and to carefully assess their 3D motion performance and skills. As mentioned in the discussion there is evidence of individual differences in task approach and potentially in the underlying processing strategy in our experimental stimuli. Knowledge of these differences would contribute greatly to our understanding of the brain-imaging results. Whilst fMRI provides valuable insight into the location of stereo-motion processing it is perhaps too temporally limited to shed light on the complex interactions which take place when the visual system process stimuli of an intermediate level of complexity.This problem may be better addressed in an EEG/MEG study to complement fMRI results.

CHAPTER 6: STEREODEFICIENCIES

Abstract

Binocular deficits are relatively common within a typical sample of observers. This has implications for research on binocular vision, as a variety of stereo deficits can affect performance. Despite this, there is no agreed standard for testing stereo capabilities in observers and many studies do not report visual abilities at all. Within the stereo literature, failure to report screening and sampling has the potential to undermine the results of otherwise strictly controlled research. We reviewed research articles on binocular vision published in three journals between 2000-2008 to illustrate how screening for binocular deficits and sampling of participants is approached. Our results reveal that 44% of the studies do not mention screening for stereo deficits and 91% do not report selection of participants. The percentage of participants excluded from studies that report stereo screening amounts to 3.9% and 0.7% for studies that do not report stereo screening. These low numbers contrast with the exclusion of 17.6% of participants in studies that report screening for binocular deficits as well as selection of participants. We discuss various options for stereo testing and the need for stereo-motion testing with reference to recent research on binocular perception.

6.1. Introduction

The process of interpreting depth information veridically is anything but trivial and yet, the computations involved are well masked from our conscious experience. If information about depth in the environment is weak or ambiguous, the visual system nevertheless arrives at an interpretation, so expertly that we usually do not notice ambiguities or missing information. Similarly, if depth information in a visual scene is abundant we are usually not aware of different cues contributing to our depth perception. For this reason studies that investigate stereopsis and depth perception must consider whether performance in a binocular task is based on impaired and/or different input for individual observers.

The aim of this review is to document screening and sampling procedures in psychophysical and neuroscientific studies of binocular vision and to discuss implications. We call for greater attention to screening for stereo deficits and sampling of observers and draw on literature from clinical ophthalmology and binocular vision research. In addition, we suggest that stereo-motion capabilities must be considered when assessing stereo perception.

6.1.1 Stereo Deficiencies

The ability to see in depth is facilitated by the basic anatomical structure of the primate visual system (Howard & Rogers, 2002). The presence of two forward facing eyes in combination with vergence (and cyclotorsion) of the eyes creates a large degree of binocular overlap or crossover between the visual fields. Animals with a large degree of binocular overlap have developed stereopsis; they can exploit differences between retinal images in the left and right eye to perceive depth in the environment.

During stereopsis the disparate, 2D retinal images are fused into a single 3D percept. Through identifying corresponding points, 'matching primitives' in the slightly offset retinal images of the left and right eye, the images can be aligned in such a way as to

reveal depth information. There are various monocular and binocular depth cues available in a typical 3D scene and depth perception can be achieved by exploiting stereo correspondences in static as well as dynamic images. It was reported, for example, that dynamic depth cues enhance depth perception (Bradshaw & Cumming, 1997), especially when disparity cues are weak or ambiguous (van Ee & Anderson, 2001).

Stereo deficiency, or, the inability to correctly perceive depth, occurs in various ocular conditions. Amblyopia ('lazy eye') is a non-disease related reduction in visual acuity that has a worldwide prevalence rate of 2% to 2.5% (Parker, 2007; Farvardin & Afarid, 2007).

Stereopsis is also affected in strabismic (cross-eyed) patients, who suffer from varying degrees of misalignment between the two eyes. This condition has different forms including esotropia, where one eye deviates inwards, exotropia, where one eye deviates outwards, and hypertropia, where one eye deviates upwards. [Strabismus can be further categorised as constant when it occurs under all viewing conditions or as intermittent when it alternates between the eyes. Patients with intermittent strabismus are likely to develop normal binocular functioning, whilst those with constant strabismus are unlikely to develop normal binocular functioning unless they are identified and treated early.] Patients with anisometropia, a condition in which the ability to focus is degraded in one eye relative to the other, are also known to show deficits in stereopsis and depth perception.

Most of the clinical conditions associated with poor stereopsis affect one eye only. A condition that usually involves both eyes and affects between 2.3% and 13% of the U.S. population (Convergence Insufficiency Treatment Trial Study Group, 2008), known as convergence insufficiency, is a binocular disorder resulting in a range of visual impairments especially at near-point distances. This condition often leads to diplopia or double vision.

6.1.2 Screening and Prevalence

Prevalence estimates of stereo deficits vary widely between reports. Using a Keystone stereo test and a hexagonal stereo test Coutant and Westheimer (1993) reported that 97.3% of a convenience sample of N=188 biology students could detect horizontal disparities of 138 arcsec (2.3 arcmin) or less. On the other hand, 20% of the sample could not detect depth at less than 30 arcsec.

Ament, et al. (2008) tested visual abilities at near point distances in 200 college students with self-reported 'normal' vision. The visual capabilities of participants were measured using the Keystone Visual Skills Series[1]. This series of tests provides a thorough assessment of the visual capabilities of the observer, including measures of binocular functioning and stereopsis. It is mainly used in clinical practice because it requires a telebinocular. The authors were particularly concerned with binocular convergence at near (16 in or 0.4 m) and far-point distances (20 ft or 6 m) and resting lateral posture of the eyes, the direction of the line of sight for each eye at rest. In this study they also measured fusion ability, the level of binocular co-ordination that is present, under the assumption that maximum processing of stimuli occurs only in cases where normal lateral posture and fusion are present.

Near point binocular convergence dysfunction in terms of lateral posture was found in 65% of the sample, 40% were exophoric whilst 25% were esophoric. In terms of fusion, they reported that 23.5% of the students displayed some kind of binocular dysfunction, with 10.5% exophoric and 13% esophoric. Reduced fusion is likely to affect the ability of the observer to adequately combine monocular half images in a binocular display.

In a review of 10 selected studies mostly from clinical populations and children Cacho-Martinez, Garcia-Munoz, & Ruiz-Cantero (2010) tried to establish prevalence of accommodative and non-strabismic binocular dysfunctions. They found a wide range

---

[1] Keystone View, Nevada Capital Group Inc., 2200 Dickerson Road, Reno, NV 89503.

of accommodative insufficiencies (2%-61.7%) as well as convergence insufficiencies (2.25%-33%) across studies.

In their literature review on screening practices in research Ament et al. (2008) sampled all studies using visual stimuli published between 1997-2004 in four psychology journals: *Journal of Experimental Psychology: Learning Memory and Cognition, Journal of Experimental Psychology: Human Perception and Performance, Perception and Psychophysics and the American Journal of Psychology*.
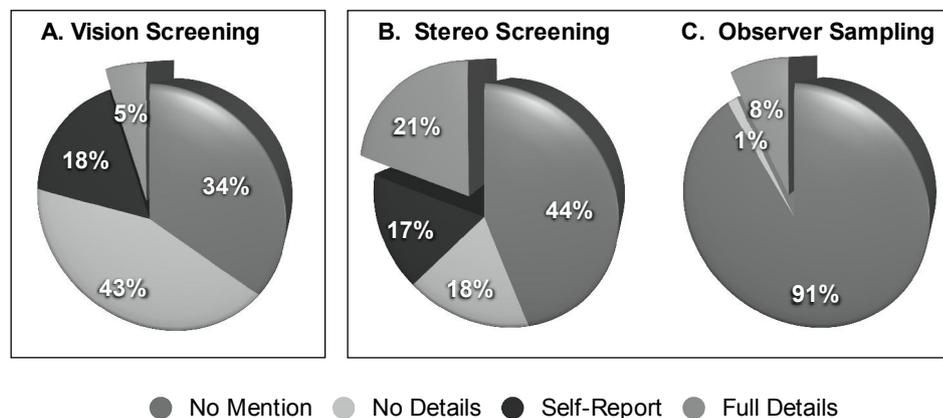
Figure 6.1 Pie-charts with percentages of articles from journals that make no mention (mid-gray), give no details (light-gray), provide self-report/in-task testing (dark-gray), and give full details (separate slice) for A. vision screening as reported by Ament et al. (2008), B. screening for stereo deficits, and C. selective sampling of participants from a literature review of studies published between 2000-2008 (see text for details).

A total of 34% of the articles reviewed made no mention of the visual abilities of research participants, a further 43% mentioned a test but gave no specific details, 18% used self-report measures and only 5% used a standardised visual test (see Figure 1A).

6.2 Survey of Stereo Literature

As illustrated by the literature review prevalence as well as the degree of binocular dysfunction and stereo deficits is relatively unknown in the general population. This coupled with the gross underreporting of visual abilities in typical research, indicates a potential sampling bias in studies of binocular vision. If participants are selected from a subpopulation without stereo deficits then results may not be representative of the general population. On the other hand, if observers with stereo deficits are included then performance in a specific binocular task may be biased.

In order to document screening and sampling in the field of binocular perception, we conducted a literature survey on studies published in three journals dedicated to research on visual perception: *Vision Research*, *Perception,* and *Journal of Vision.*

We surveyed stereo testing and sampling of observers in recent publications devoted to the study of binocular vision. Thereto, we conducted a Web of Science search on the key term 'binocular vision', refining our search to the years 2000-2008 inclusive and the three journals. Our criteria for inclusion were that the article (a) included only human observers (b) concerned only non-clinical samples of observers, and (c) used a binocular stimulus display. To satisfy the last point we excluded studies on binocular rivalry or ocular dominance. This gave us a sample of 266 studies with a total of 2660 observers. Although our main focus was screening and sampling of observers it is worth mentioning that 70% of the studies employed static stereo stimuli, 19% moving stereo stimuli, and 11% both.

First, each study was coded according to the type (if any) of stereo test used: 'Not Mentioned' (0), if the depth capabilities of the observers were mentioned only briefly/not at all; 'No Details' (1), if capabilities of depth perception for observers was explicitly mentioned but the testing procedure was not; 'Self-Report' (2) if a test was mentioned which is not a standard method of testing stereopsis, e.g. a self-report, an

in-task method designed by the experimenters, or a test which is not commercially available/scientifically validated; and finally 'Full Details' (3) if the article mentioned a specific test which is considered a standard test of stereopsis.

Based on this categorization, we found that 44% of the studies did not mention any stereo screening and a further 18% mentioned screening but no specific stereo test. Of those studies that specifically mentioned a test (38%), 17% used a non-standard self-report or in-task test. Approximately 21% of the studies used a standard stereo test (see Figure 1B). Given that all studies are concerned with binocular vision, the results illustrate widespread uncertainty among researchers whether or not screening for stereo deficits is feasible or necessary and whether or not to report it.

In terms of sampling of participants, 91% of the studies fail to report selection or exclusion of participants, 1.5% mention exclusion but provide no numbers or details, and 7.5% provide a detailed report (see Figure 1C).

Studies that report screening for binocular deficits excluded 3.9% (56 of 1454) of participants and studies that fail to report stereo screening excluded 0.7% (9 of 1,206); although it should be mentioned that exclusion of participants due to inability to carry out the task is probably underreported. Whilst the percentage of excluded participants seems rather low these results suggest that screening for stereo deficits increased the number of reported exclusions more than five times. In addition, the percentage of observers excluded from studies that report detailed screening for stereo deficits as well as sampling amounts to 17.6% (56 of 318). As the report on sampling and exclusion of participants is likely to be confounded, detection of stereo deficits severe enough to merit exclusion probably spans 3.9% to 17.6%. The wide range is likely to reflect uncertainty about the severity of stereo deficits across individual observers as well as the exclusion criterion across studies.

The low number of exclusions (0.7%) in studies without stereo screening may be explained as follows. (1) Observers with binocular dysfunctions were not identified as a consequence of omitting screening for stereo deficits. (2) Selective sampling may have

occurred so that results from these studies are not representative of stereo capabilities in the general population. (3) As mentioned above, it is possible that stereo screening as well as sampling is only reported when participants were excluded.

We also coded the type of participant involved in each study: 75% of all participants were naïve observers, that is participants who were unaware of the aims of the experiment, 10% were trained or experienced observers also unaware of the aims, 3.5% were classified as authors, if participants of the studies were members of the research group involved in the study, and 11.5% of participants were mixed, if participants were naïve as well as authors.

In 89% of the studies at least one participant was naïve to the experimental aims and procedures and in 51% of the studies, comprising 75% of all participants, observers were naïve and without experience. Naive participants are unlikely to be guided by the experimental hypotheses but they are also unfamiliar with stereoscopic displays so that screening for stereo deficits may be more important. Of those studies with at least one naïve participant, about 41% report screening with a specific stereo test whereas 41% failed to report screening. The remaining 18% mention non-specific stereo screening. Not surprisingly, no exclusion of participants was reported for studies with authors only (0 of 94), and relatively few for mixed samples of authors and naïve observers (2 of 303).

In 41% of the studies, accounting for 25% of all participants, either experienced observers or authors took part. These samples may fall short of measuring the average observer. This was recently illustrated in a study where naïve participants (all of whom were psychology students) performed similarly to infants and children rather than the experienced observers' normally included in this type of research (Horwood & Riddell, 2010). Repeated employment of the same observers across studies and a lack of reporting selective sampling may exacerbate the bias in average performance because observers stereo capabilities are simply better than average.

We conclude that in the vast majority of studies screening and sampling procedures are not sufficiently documented, almost certainly leading to biased reports of stereo capabilities.

6.3 Measuring Stereopsis

Stereo tests should provide a quick and effective measure of stereo acuity and a number of stereo tests are used to screen for deficits in stereo perception, on their own or in clinical studies in addition to tests of monocular and binocular eye function. The most common stereopsis tests (see Appendix), are random dot based tests such as the TNO test, the Lang I and II tests and the Random-Dot E test (Kriegbaum-Stehberger, Jiang, & Mojon, 2008; Reinecke & Simons, 1974). Traditional ways of administering these tests include the use of anaglyph or polarized targets and glasses to stimulate retinal disparity between the left and right eye. Two panels with random dot stimuli are presented side by side in front of the observer at increasing viewing distance.

Repeated ability to correctly identify the panel with target (butterfly or letter E) defined in depth is used to determine a detection threshold for stereopsis.

Stereopsis can be measured using anaglyph cancellation to present disparate images in the left and right eyes. Targets are typically viewed through anaglyph glasses and these tests are also available with polarized glasses (Yamada, Scheiman, & Mitchell, 2008). Comparison between the polarized and anaglyph versions of the RDE stereo-test, the Random Dot Butterfly, the stereo numbers test, the stereo circles test and the stereo animals test in a sample of 60 children with no strabismus, amblyopia or high refractive error and normal ocular health, gave mixed results.

Both of the Random dot based tests showed a high level of agreement in terms of the number of children who correctly identified the letter 'E' or the 'butterfly' on four out

of four trials. The Random Dot Butterfly test showed very strong agreement with its polarized counterpart, a Wilcoxon signed rank test indicated virtually no difference between the two RDE tests (p=0.99). The stereo animals test also showed a high level agreement: Of the 58 observers who achieved 100 arcsec using the anaglyph test 57 also had 100 arcsec on the polarized version. The overall agreement was 0.95 ($CI_{95\%}$ 0.895-1.0). However, agreement was less than 0.6 for the Stereo Numbers Test and less than 0.35 for the Stereo Circles test (Yamada et al., 2008).

More recently, methods have been introduced that use a special prismatic printing process so that disparate images can be presented to each eye without the need for anaglyph or polarized glasses. Several versions of these tests are available, including: the Titmus Stereo Test, the Stereo Reindeer Test, the Random Dot Butterfly, the Random Dot Figures, and the Random E, Circle, Square (Hatch & Richman, 1994). Within-subjects performance on the five individual tests against their traditional counterparts showed a high correlation (r =0.997, r =0.998, r =0.997, r =1.0, and r =1.0, respectively) with regard to levels of stereopsis identified by traditional and non-polarized versions of each test (Hatch & Richman, 1994). There are advantages to non-polarized versions of tests for researchers, in terms of cost effectiveness and ease of use. Since there is no quantifiable difference in performance these versions of traditional tests may be considered as reliable substitutes.

RDS tests of binocular functioning have the advantage of minimizing the presence of monocular cues. Contour based tests such as the Titmus/Stereo Circles test are known to contain monocular cues (e.g. Holmes et al., 1999) suggesting unimpaired stereopsis even in individuals with previously identified binocular deficits (Fawcett, 2005). Fawcett (2005) compared four tests; the Titmus Circles test, the Randot (Version 2) Circles test, the Preschool Randot Stereo-acuity test and the Titmus Fly test (non-contour based) in 91 patients (age 5-85 years) with a history of binocular deficits and 54 normal controls (age 6-72 years). She found that the Preschool Randot test consistently identified stereo-acuity scores with a lower disparity than the Stereo Circles test, more so at the highest disparity levels. Of the 22 patients identified as being 'stereo-blind' by both the Randot Stereo-acuity test and the Titmus Fly test (non-contour based), 19 were identified as having stereoscopic vision by the Titmus Circles

test (Fawcett, 2005) and 12 by the Randot Circles test. This suggests that contour based tests tend to misclassify patients with known binocular deficits as having normal stereopsis. Thus, contour based tests are unreliable and should be avoided.

In terms of test-retest reliability the literature is scarce. A recent study concerning the reliability of the Random Dot E test examined between-tester agreement in a sample of 1,195 preschool children selected after initial screening of visual acuity and refractive error (Schmidt, et al., 2006). All children that had failed in the initial screening were included in the study (58%; target conditions) whereas the remaining children (42%; no condition) had passed initial screening. A battery of tests, including the Random Dot E test, was administered on two separate occasions by licensed eye care professionals. The second tester was blind to the results of the previous screening. The results suggest that only 59% of children showed the same sensitivity to disparity on both testing occasions. The inter-tester agreement measured by Cohen's weighted kappa was κ_w=0.43. Although the proportion of children showing identical scores at the two testing sessions increased significantly with age (identical scores were achieved by 54% at 3 years, 59% at 4 years, and 63% at 5 years) inter-tester agreement κ_w did not change significantly across age (p=0.49). Agreement between test results was higher among children with target conditions (κ_w=0.44) than for children who did not have a condition (κ_w =0.33, p=0.02) but the percentage of children with identical scores was higher among children with no condition (66%) than among children with targeted conditions (42%; p<0.001). This suggests that the moderate test-retest reliability of the Random Dot E test varies to some extent with age and level of deficit. The developmental component makes it difficult to generalize to a population of older children/adults who show higher test-retest reliability. There is some disagreement as to how important developmental constraints are in such tests (Heron, Dholakia, Collins, & McLaughlan, 1985; Fox, Patterson, & Francis, 1986; Ohlsson et al., 2001).

Whilst some studies have found adult-like stereopsis in 7-year-old children for Randot and TNO (Heron et al., 1985), others argue that stereo acuity is already fully developed

at an age of 3-5 years (Fox, Patterson, & Francis, 1986). However, developmental patterns seem to differ between testing methods. Some researchers therefore suggested that the conceptual knowledge of younger participants rather than their stereo acuity may explain inconsistencies (Ohlsson et al., 2001). Ohlsson and colleagues tested 1035 12-13 year olds (454 male, 581 female) using five binocular screening tools (Lang II, Frisby, Randot, Titmus, and TNO). They found that of a total of 60 children who had either strabismus and/or amblyopia, only 8 were identified by all five tests. 26 of the children were not identified by any of the tests although 23 were amblyopic. Their results led the authors to conclude that none of the 5 tests are feasible as visual screening and selection tools as they found no significant cut-off point between visually healthy and impaired participants (Ohlsson et al., 2001).

Instead of visual acuity and stereopsis tests, some studies either use self-report measures or simply ask participants about their visual abilities (17% in our stereo survey and 18% according to Ament et al., 2008). This approach appears to be far less robust than using a direct measure. However, Coren & Hakstian (1996) developed a self-report screening inventory for stereopsis, demonstrating a high degree of classification accuracy with combined stereopsis measures using a keystone telebinocular (Coren & Hakstian, 1996; N=1115). Similarly, self reported skill in seeing magic eye™ stereograms, predicted performance on the TNO stereo-test (Wilmer & Backus, 2008; N=194). Due to their subjective nature however, the use of self-report measures by itself is not recommended.

Stereo tests have also been combined with eye tracking in order to provide a more objective measurement of stimulus detection (Kriegbaum-Stehberger et al., 2008). Breyer, Jiang, Rutsche and Mojon (2006) applied a monitor-based random-dot stereo test for use in young children coupled with infrared photo-oculography (eye tracking). Although this arrangement provides an objective measure of stereopsis under natural viewing conditions and seems to be superior to the Lang I stereo test, the authors concede that such a combined test would be unlikely to replace current tests which are less expensive and easier to administer.

It seems that anaglyph or non-polarized versions of the RDS based tests such as the Random Dot E/ Butterfly represent the least expensive, most diagnostic tools for measuring stereopsis in the adult population. In terms of non-verbal methods for use with very young children or disabled adults, the Infant Random-dot stereoacuity cards (Birch & Salomao, 1998) or the Preschool Randot Stereoacuity test (Fawcett, 2005) provide acceptable solutions.

6.4 Stereopsis and Stereo-Motion

In addition to monocular motion cues to depth (e.g., motion parallax, looming) there are two basic binocular cues available in stereo-motion stimuli. The first cue, binocular disparity (Julesz, 1971), is usually addressed in the context of static stereo displays. If however, the visual system tracks binocular disparity over time then such a mechanism describes stereo-motion perception, known as changing disparity over time (CDOT; Cumming & Parker, 1994). Depending on the binocular viewing geometry, both eyes also receive motion signals of different magnitude and direction. This difference, which may be computed by the subtraction or division of horizontal motion signals in each eye (Regan, 1993), is known as interocular velocity difference (IOVD). Both binocular cues, CDOT and IOVD, are confounded in natural scenes but rely on different physiological encoding. Stereo-motion may be triggered not only by stimulus motion in the scene but also by observer movements such as eye, head, and body movements (e.g., Miles, 1998; Harris, 2006).

There is some contention in the literature whether the visual system extracts disparity information first (CDOT); motion information first (IOVD) or whether motion and disparity is jointly encoded (JEMD) to establish binocular 3D motion perception (Brooks & Stone, 2004; Lages, Mamassian & Graf, 2003; Lages, 2006; Lages, Dolia, & Graf, 1997; Nefs & Harris, 2010). The answer to this question has proven difficult, despite the wealth of evidence in favour of any one of these accounts (for reviews see Harris, Nefs & Grafton, 2008; Regan & Gray, 2009). At least for local encoding of arbitrary 3D motion trajectories, the inverse problem of binocular 3D motion

perception can only be solved when disparity and motion input are combined (Lages & Heron, 2008; 2010; see also chapters 1, 2, 3 and 4), possibly at a late processing stage (Rokers, Cormack, & Huk, 2009). As discussed in chapter 5, neuroscientific evidence in monkeys suggests parallel processing of disparity and motion information in the ventral and dorsal stream (Ponce, Lomber, & Born, 2008). Motion and disparity constraints may be integrated along a 'complexity axis', with projections from motion area MT to the adjacent region MST (dorsal and ventral) increasing in functional complexity from posterior to anterior regions (Born, 2000). Likova and Tyler (2007) using fMRI BOLD signals identified dedicated processing areas, selective for depth driven (cyclopean) stereo-motion in the anterior region adjacent to/partially overlapping the motion complex hMT+ (Rokers, et al., 2009). If depth and motion cues are integrated late then it seems possible that deficits in either stereo or motion processing can have differential effects on performance in binocular tasks.

There is a need in binocular vision research to measure stereo-motion specific deficits, due to evidence for stereo-motion specific scotoma without obvious deficits in static stereopsis (Richards, 1971; Richards & Regan, 1973; Regan, Erkelens, & Collewijn, 1986; Miles, 1998). Similarly, it is possible that stereo deficits in static stimulus displays do not necessarily indicate impaired stereo-motion perception.

Since patients with strabismus can interact rather well with a dynamic 3D visual environment and even enjoy 3D films differences between the processing of dynamic and static stereo stimuli have been explored (Fujikado et al., 1998). Fujikado et al. (1998) carried out individual comparisons between the Titmus/Lang stereo test, measuring fine grain depth only, and dynamic random dot stereograms (DRDS). DRDS stimuli present a succession of random dot displays, showing changing disparity of a global form or surface over time without coherent local motion signals. Success in the DRDS stimulus conditions was determined by the ability of observers to correctly locate a dynamic pattern (circle or square) moving in depth. According to Cumming & Parker (1994) most observers consistently perceive stereo-motion in such displays. A total of 52 patients with various forms of strabismus were tested and the results show that a significant number of patients who failed to demonstrate stereopsis in the

Titmus fly and Lang test displayed coarse dynamic stereopsis as measured by DRDS (58.3% for Titmus fly and 56.5% for Lang test). The DRDS test had a statistically significant higher detection rate than the two static stereo tests leading Fujikado et al. (1998) to the conclusion that the use of motion in tests of stereopsis is useful for revealing lower levels of intact stereopsis.

Watanabe et al. (2008) determined detection/discrimination thresholds for stereo-motion perception in 52 strabismic patients using four types of computerised dynamic random dot stereograms. The first stimulus contained both binocular disparity and interocular velocity cues (disparity+velocity condition; RDS), the observer was asked to identify the direction of rotation (clockwise versus counter-clockwise) of two parallel planes rotating in depth around a vertical axis. The second stimulus was a temporally correlated but binocularly uncorrelated version of the first such that local interocular velocity was available but no disparity (velocity condition, TCRDS), again the observer indicated direction of rotation. In the third condition, the RD stimulus was a binocular correlated but temporally uncorrelated version of the first RD stimulus, promoting disparity but eliminating local velocity cues (disparity condition, DRDS). Observers indicated the presence/absence of motion in depth. In the fourth, and final condition the random dot stereogram was a rotating cylinder in which the upper and lower halves rotated in opposite directions. The task of the observer was to identify the border between these two parts.

For some patients discrimination thresholds for motion in depth were too high in either or both velocity and disparity conditions. It is therefore difficult to say whether strabismic patients have different velocity and disparity thresholds. However, the data did allow comparing static stereopsis (as measured by the Titmus stereo test) with dynamic stereopsis. Interestingly, some degree of dissociation was found within their sample: six patients failed to detect depth at 1,200 arcsec on the Titmus stereo test out of a total of 18 patients who could see motion in depth at 500 arcsec or smaller in the RDS display. On the other hand four patients showed deficits for dynamic stereopsis at 1,200 arcsec out of a total of seventeen patients who could detect static depth at 500 arcsec or less.

213

Recent evidence suggests that individual observers without strabismus and average visual acuity use IOVD (TCRDS) and changing disparity information (DRDS) differently (Nefs, O'Hare, & Harris, 2010). Out of a sample of 62 naïve participants eight did not reach threshold performance for perceiving motion in depth in any of the three stimuli (RDS, DRDS, and TCRDS). 47 observers (76%) reached acceptable thresholds levels for the RDS, 48 (77%) for the DRDS and 33 (53%) for the TCRDS. Interestingly, they found a substantial number of observers who could reliably detect motion in depth in RDS and either DRDS or TCRDS stimuli, but not in both (14 and 1, respectively). In total 29 people achieved acceptable threshold performance (75% correct at 100% signal), or better, in all three conditions.

In summary, these studies add to neuroscientific evidence that functionally distinct pathways for stereo and stereo-motion processing are likely to exist and that deficits in either stream of processing may have differential effects on stereopsis and depth perception.

6.5 Discussion

The lack of screening and sampling reports in studies of binocular vision is problematic. Standard stereo tests, despite their insufficiencies, provide a tool to assess individual stereo capabilities. This is better than no screening at all.

It follows from our review of the stereo literature that there should be no predetermined inclusion/exclusion criterion for the participation in studies of binocular vision. Our results and the results by Ament et al. (2008) as well as others suggest that stereo deficiencies are likely to occur in a significant proportion of the population. This raises the question, why stereo deficits are underreported in the majority of binocular vision research and to which extent this is due to insufficient screening and selective sampling. Poor documentation of stereo anomalies and unreported selection of participants poses a serious problem if we want to generalize results of binocular vision research to a wider population.

Thorough control of stimulus characteristics and experimental design is only one side of the coin. Binocular vision research also needs to assess and report the stereo capabilities of observers. Every report should inform how observers were screened and sampled and which exclusion criterion, if any, was applied before experimental testing. It seems reasonable to screen observers using standardized stereo tests so that basic stereo acuity can be assessed and documented. In addition, we suggest developing a simple stereo-motion test that can measure dynamic stereo capabilities.

Selection and exclusion of participants, however, should always be reported and limited to observers who cannot perform in the main experimental task under investigation. If stereo screening and sampling procedures are reported in detail it might emerge, for example, that observers with severe stereo deficits can perform in a given binocular task whereas observers without obvious stereo deficits cannot. These dissociations would be highly informative.

One possibility to improve performance in stereo tasks is to train observers before testing. Observers may show significant improvement in performance through practice (Lu, Chu, Dosher, & Lee, 2005). Perceptual learning has been demonstrated in a wide range of visual tasks, including depth perception (Sowden, Davies, Rose, & Kayne, 1996) and figure detection (O'Toole & Kersten, 1992) in random dot stereograms. The effects of training are thought to be persistent as some observers demonstrated retention of training effects for months or even years (Gantz, Patel, Chung, & Harwerth, 2007; Fahle, 2005). Less is known about the mechanisms underlying perceptual learning and visual long-term memory (Lages & Paul, 2006; Lages & Treisman, 1998; 2010). There is always the danger that observers with stereo deficits develop intelligent coping strategies.

Relatively simple, cost effective tests for assessing stereopsis are available. Of these, RDS based tests such as the Randot E/Butterfly, are commonly used and recommended. They are quick and easy to administer and minimize monocular cues to depth. However, they provide only a coarse measure of stereo acuity because they rely on the detection of a global form in depth and movements of stimulus and observer are difficult to control.

Throughout this review we have highlighted the importance of a detailed report on screening for stereo deficits as well as sampling. We discussed stereo-motion perception as a critical example since performance in standard stereo tests may be confounded by dynamic stereo cues, especially when stimulus, eye, head, and body movement are not controlled. Stereopsis based on dynamic cues appears to be different from stereopsis based on static cues, not only in terms of low-level encoding. With this in mind, we suggest that observers should be screened for stereo deficiencies in both.

CHAPTER 7. CONCLUSION

In this thesis the role of motion and disparity cues in computing local motion trajectories was investigated. We used an onscreen adjustment paradigm, in which observers adjusted and online depth probe to show the perceived direction of motion of a single line stimulus oscillating in depth behind a circular aperture. The inherent ambiguity in the display allowed us to study perceptual bias for local motion trajectory, defined by orientation disparity and interocular velocity difference (IOVD). The empirical results showed that observer judgments did not follow geometric predictions but varied systematically with the orientation disparity and IOVD of the stimuli. Response patterns were very similar across four observers, thus prompting the use of geometric statistical models to explain the data.

We first compared results to two computational models (details given in chapter 2); the Vector Normal model (VN) assumed that local image velocities remained ambiguous, defining a constraint plane in 3D space. A vector normal was used to derive a unique solution. The Cyclopean Average model (CA) assumed early motion averaging, so that a unique 2D solution was derived first and combine with disparity information to determine a unique solution in 3D. In chapter 4, we show that the VN model provides a better solution, despite failing to account for the data. This model was therefore used as a basis to develop Bayesian statistical models, provide a more flexible framework to understand perceived 3D motion direction.

Bayesian models of binocular motion perception for dots or local features moving on a horizontal depth plane (Lages, 2006; see chapter 3) can be extended to 3D motion perception of lines or edges, predicting perceived azimuth and elevation under ambiguity (Lages & Heron, 2010; Heron & Lages, 2009). In chapter 4 we tested the utility of two Bayesian models in predicting observer judgments of motion trajectory. The one parameter BVN model was a direct extension of the VN model and assumed perceptual judgments under uncertainty were influenced by a motion prior (preference for slow motion).

The two parameter BDVN model was also based on the VN model but assumed a weak contribution of a motion prior and a stronger contribution of a disparity prior (perception was largely driven by a preference for motion towards the fronto-parallel fixation plane). We found that the BDVN model gave a reasonable fit to our empirical data.

It was concluded that using a generalized Bayesian motion model (Lages & Heron, 2009; Wang, Heron & Lages, 2013), noisy velocity constraint planes define velocity likelihoods that, combined with a 3D motion prior, can explain perceptual bias under uncertainty and motion perception under ambiguity (see chapter 4). We suggest that the visual system integrates velocity constraints with feature tracking from disparity processing to arrive at velocity estimates of moving features and objects.

In chapter 5 we used fMRI, to investigate the role of hMT+ in global 3D motion integration. We used a stimulus in which monocular and binocular phase offsets defined the perceived motion direction of a diamond shape (see chapter 5 for description of stimuli and methods). Observers viewed four types of motion display inside the scanner: 1) 2D expansion/contraction 2) 2D translation 3) 3D rotation 4) 3D looming. A forced choice decision task (2D versus 3D motion) was used to focus attention on the displays.

Results showed a strong general preference for global 3D motion displays over 2D motion displays in hMT+. The precise role of hMT+ in perception of global 3D motion is unclear from this study, but in most observers this area also distinguished between analogous 3D/2D pairings (which differed only in binocular phase offset), preferring the 3D version of stimuli; and so there is evidence for a distinction based on the disparity component of global 3D motion stimuli in hMT+. In comparison, V1 did not show preference for 3D stimuli in most observers and did not respond preferentially to 3D versions of analogous pairings. Our findings corroborated those in the two existing

studies in the area (Likova & Tyler, 2007; Rokers et al. 2009) adding to growing evidence that suggests an important role for hMT+ in processing stereo-motion.

7.1 FUTURE RESEARCH DIRECTIONS

How the visual system establishes binocular 3D motion perception from image-based local motion and disparity input remains a difficult and unresolved issue. It is hoped, however, that the present line of research improves understanding of local constraints in binocular 3D motion perception. The results should be of interest to researchers in computer vision, neuroscience, and psychology and may inform developments of 3D technology in applied areas.

Our results suggest that a geometric-statistical approach as exemplified by the Bayesian models provides a powerful framework to model binocular 3D motion perception. However, more empirical data are needed to evaluate perceived 3D trajectories and systematic distortions.

More specifically, in order to validate the Bayesian approach it would be important to verify the motion prior for slow motion through real-world measurements of scene flow as well as experimental data from discrimination tasks adapted to binocular 3D motion. Note that due to environmental constraints the motion prior may not be isotropic and Gaussian. There is empirical evidence that the motion prior in 2D velocity space has heavier tails than a Gaussian (Stocker & Simoncelli, 2006). Also, if experience shapes the motion prior it may not only reflect slow motion but also horizontal motion along a ground plane and possibly downward motion aligned with the pull of gravity. As a consequence, the motion prior may not be entirely isotropic. On the other hand, any of these effects is likely to depend on multisensory integration and top-down processing, mainly affecting global object rather than local motion perception (Lages, Jenkins & Hillis, 2008).

We have only considered translation in 3D space but a moving line or edge may also change orientation over time. It is immediately clear that the present approach, which

relies on translational motion, has a difficulty with encoding rotational movement of a stimulus.

It seems likely that encoding of rotation of a local line stimulus requires tracking of features together with frequent updates of binocular depth, whereas translational motion is the result of instantaneous spatio-temporal encoding by local motion filters. According to recent neuroscientific evidence (Ponce et al. 2008; Rokers, et al., 2009) local input from motion and disparity processing is integrated late and disparity of features may be used to disambiguate velocity constraints. It seems possible that integration of velocity constraints and features in depth follows a characteristic time course similar to 2D motion perception (Osborn et al., 2004; Treue, Hol & Rauber, 2000; Montagnini et al., 2007).

In our fMRI study, we observed preferential response in hMT+ to global 3D motion versus global 2D motion (defined by phase disparities). The 3D motion stimuli contained both monocular and binocular phase disparities and appeared to translate in depth. 2D stimuli contained only monocular phase offsets and appeared to move in a single depth plane.

To further understand the significance of our findings it would be advantageous to use multi-variate pattern analysis (MVPA) on the existing data. MVPA is used to relate a pattern of brain activity (across multiple voxels) to one cognitive state versus another. This allows us to relate a pattern of brain activity to a particular mental representation i.e. we would be able to associate the pattern of activation in the visual cortex with perception of looming motion or rotation respectively.

The analysis used, GLM, attempts to find individual voxels which are activated significantly in response to the stimulus, spatially averaging across voxels which respond to a specific experimental condition. This discounts voxels which may carry information about a given stimulus, on the basis that their response is below the significant level and does not consider fine-grain spatial patterns which may discriminate between different stimulus conditions (due to spatial averaging).

MVPA uses pattern classification techniques to detect patterns of activity over multiple voxels and thus does not employ spatial averaging. It is also suited to our purposes in that it focuses on idiosyncratic patterns of response (unique to the individual observer), this is advantageous, given that our data appears to be best understood on an individual level (Norman et al, 2006). This would allow us to detect differences in patterns of activation between the two types of 3D motion (looming, rotation) which were not shown in hMT+/V5 in our original analysis.

In future studies using this stimulus, the aim would be to show directional adaptation to global 3D motion stimuli in hMT+. In fMRI, repeated exposure to the same stimulus causes a gradual attenuation of BOLD response to that stimulus.
As mentioned previously in the discussion section of chapter 5, showing that adaptation to one 3D motion stimulus does not generalize to a different type of 3D motion would provide evidence that the two are recognized as being distinct in hMT+/V5. This would suggest an important role for perception of global 3D motion, as opposed to merely processing (Rokers et al.2009; Ban et al.2012).

In the current study we focused only on area hMT+/V5 and as a comparison region, area V1. In future, we would extend this to include intermediate areas, in order to compare response to our stimuli throughout the visual hierarchy. Stereo-motion processing is best understood as a complex interactive loop of processes and to capture this fully we must explore the activity of other regions in the loop and also (using MVPA) attempt to understand how these regions interact with one another to achieve the perceptual goal of global 3D motion perception.

To this end, it would also be advantageous to gather behavioural data on observers, this would enable us to understand more about the mechanisms involved in processing our stimuli and would make results obtained from brain imaging studies more meaningful. In chapter 6 we highlight the importance of considering individual differences in stereo-motion processing, the behavioural results from the fMRI task

indicate that the six observers may not have approached the 3D/2D classification task in the same way.

It is unclear whether differences were due to perception of the stimuli or a misunderstanding of the task since we did not test or train observers prior to participation. In future we will make it a priority to do so.

Our fMRI results would also be complemented using a brain imaging technique with higher temporal resolution such as EEG to delve deeper into the complexity of interactions which result in a global percept of 3D object motion. This stimuli used here would be ideal for exploration in an EEG paradigm and in using such a technique, which enjoys superior temporal resolution in comparison to fMRI, we could gather information on the time-course of integration of the local features in our display into a full global percept of motion-in-depth.

Finally, to link this strand of research to the previous strand, where we studied the role of IOVD and CDOT driven mechanisms in computing local motion trajectories, we could attempt to isolate IOVD and CDOT components of our displays. This would allow us to investigate the driving mechanisms of global 3D motion perception in hMT+.

References

Adams, D.L., Sincich, L.C., Horton, J.C. (2007) Complete pattern of ocular dominance columns in human primary visual cortex. *Journal of Neuroscience*, 27 (39), 10391-10403.

Adelson, E.H. (1984) Binocular disparity and the computation of two-dimensional motion. *Journal of the Optical Society of America* A 1: 1266 (Abstract).

Adelson, E.H., Bergen, J.R. (1985) Spatio-temporal energy models for the perception of motion. *J Optical Soc America A 2*: 284-299.

Adelson, E.H., & Movshon, J.A. (1982) Phenomenal coherence of moving visual patterns. *Nature* 300, 523-525.

Albright, T.D. (1984) Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology*. 52 (6), 1106-1130.

Albright, T.D., Desimone, R., Gross, C.G. (1984) Columnar organization of directionally selective cells in visual area MT of the macaque. *Journal of Neurophysiology*. 51(1), 16-31.

Ament, P.A., Robbins, D.K., Brockman, C.J., Hawkins, R., Lake. S., & Bazin, M.B. (2008) Prevalence and reporting of near-point binocular convergence and acuity vision deficits in potential research participants. *Behaviour Research Methods*, 40 (1), 309-314.

Anderson, B.L. (1999) Stereoscopic occlusion and the aperture problem for motion: a new solution. Vision Research 39(7), 1273-1284.

Anderson, J.C., & Martin, K.A. 2002. Connection from cortical area V2 to MT in macaque monkey. *Journal of Computational Neurology*. 443(1), 56-70.

Anderson, J.C., Martin, K.A. C. (2009) The synaptic connections between cortical areas V1 and V2 in macaque monkey. Journal of Neuroscience, 29 (36), 11283-11293.

Anzai, A., Ohzawa, I., Freeman, R.D. (2001) Joint encoding of motion and depth by visual cortical neurons: neural basis of the Pulfrich effect. *Nature Neuroscience*, 4(5), 513-518.

Ascher, D., & Grzywacz, N.N. (2000) A Bayesian model for the measurement of visual velocity. *Vision Research*, *40(24)*, 3427-3434.

Bakin J. S., Nakayama K., & Gilbert, C. D. (2000) Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *Journal of Neuroscience*, 20, 8188-

8198.

Ban, H., Preston, T.J., Meeson, A., & Welchman, A.E. (2012) The integration of motion and disparity cues to depth in dorsal visual cortex. *Nature Neuroscience*, 15 (4), 636-643.

Banks, M.S., Backus, B.T. (1998) Extra-retinal and perspective cues cause the small range of the induced effect. *Vision Res*, *38*: 187-194.

Barlow, H.B., Blakemore, C., & Pettigrew, J.D. (1967) The neural mechanism of binocular depth discrimination. *J. Physiol. Lond*, 193: 327-342.

Berkeley,G. (1709/1975) *Philosophical Works; Including the Works on Vision*. In: M Ayers M, editor. London, Dent.

Beverley, K.I., Regan, D., 1973. Evidence for the existence of neural mechanisms selectively sensitive to the direction of movement in space. *J. Physiol*. 235, 17-29.

Beverley, K.I., Regan, D., 1974. Visual sensitivity to disparity pulses: evidence for directional selectivity. *Vision Res*. 14, 357-361.

Beverley, K.I., Regan, D. (1974) Temporal integration of disparity information in stereoscopic perception. *Exp Brain Res 19*: 228–232.

Beverley, K.I., Regan, D. (1975) The relation between discrimination and sensitivity in the perception of motion in depth. *J Physiology 249*: 387–398.

Birch, E.E., & Salomao, S. (1998) Infant random dot stereoacuity cards. *J Pediatr Opthalmol Strabismus*, 35, 86-90.

Blasdel, G. G. & Lund, J.S. (1983). "Termination of afferent axons in macaque striate cortex." *J. Neurosci*. 3, 1389-1413.

Born, R.T. (2000). Center-surround interactions in the middle temporal visual area of the owl monkey. *Journal of Neurophysiology*, 84, 2658-2669.

Born, R.T., & Bradely, D.C. (2005). Structure and Function of Visual Area MT. *Annual Review of Neuroscience*, 28, 157-189.

Braddick, O.J. (1974) A short-range process in apparent motion. *Vision Res 14*: 519–527.

Braddick, O.J. (1980) Low-level and high-level processes in apparent motion. *Philosoph Trans R Soc 290B*: 137–151.

Bradely, D.C., Chang, G.C.,Andersen, R.A. (1998) Encoding of three-dimensional structure from motion by primate area MT neurons. *Nature* 392(6516), 609-611.

Bradley, D.C., Qian, N., Andersen, R.A. (1995) Integration of motion and stereopsis in

middle temporal cortical area of macaques. *Nature*, 373(6515): 609-611.

Bradshaw, M.F., & Cumming, B.G. (1997) The direction of retinal motion facilitates binocular stereopsis. *Proc R Soc London B 264*: 1421-1427.

Breyer, A., Jiang,X., Rutsche, A., & Mojon, D. (2006) A new 3D monitor-based random-dot stereotest for children. *Investigative Ophthalmology and Visual Science*, 47 (11), 4842-4846.

Britten, K.H. (2003). The middle temporal area: motion processing and the link to perception. In *The visual neurosciences*, ed. LM Chalupa, JF Werner, pp.1203-1216. Cambridge, MA: MIT Press.

Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S., Movshon, J.A. (1996). A relationship between behavioural choice and the visual responses of neurons in macaque MT. *Visual Neuroscience*. 13(1), 87-100.

Brooks, K.R. (2002) Interocular velocity difference contributes to stereomotion speed perception. *J Vision 2*: 218–231.

Brooks, K.R., & Stone, L.S. (2004). Stereomotion speed perception: Contributions from both changing disparity and interocular velocity. *Journal of Vision*, 4, 1061-1079.

Bruhn, A., Weickert, J., & Schnörr, C. (2005). Lucas/Kanade meets Horn/Schunck: combining local and global optic flow methods. *Int J of Computer Vision*, *61(3)*, 211-231.

Burt, P., & Sperling, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review* 88: 171195.

Cacho-Martinez, P., Garcia-Munoz, A., & Ruiz-Cantero, M.T. (2010). Do we really know the prevalence of accommodative and nonstrabismic binocular dysfunction? *Journal of Optometry*, *3(4)*, 185-197.

Callaway, E.M., & Katz, L.C. (1999). Photostimulation using caged glutamate reveals functional circuitry in living brain slices. *Proc. Natl. Acad. Sci. USA*. 90(16), 7661-7665.

Callaway, E. M. & Wiser, A. K. Contributions of individual layer 2–5 spiny neurons to local circuits in macaque primary visual cortex. *Vis. Neurosci.* **13**, 907–922 (1996).

Carceroni, R.L., & Kutulakos, K.N. (2002). Multi-view scene capture by surfel sampling: from video streams to non-rigid 3D motion, shape and reflectance. *International Journal of Computer Vision*, 49(2), 175-214.

Carney, T., Paradiso, M.A., & Freeman, R.D. (1989). A physiological correlate of the Pulfrich effect in cortical neurons of the cat. *Vision Research*, *29*, 155-165.

Cagnello, R., & Rogers, B.J. (1993) Anisotropies in the perception of stereoscopic surfaces: the role of orientation disparity. *Vision Research* 33, 16:2189-2201.

Chatterjee, S. & Callaway, E.M. (2003) "Parallel colour-opponent pathways to primary visual cortex." *Nature* 426, 668-671.

Convergence Insufficiency Treatment Trial Study Group. (2008) Randomized clinical trial of treatments for symptomatic convergence insufficiency in children. *Archives of Ophthalmology*, Oct, 126 (10), 1336-1349.

Coren, S., & Hakstian, A.R. (1996) Screening for stereopsis without the use of technical equipment: Scale development and cross validation. *International Journal of Epidemiology*, 25(1), 146-152.

Coutant, B.E., & Westheimer, G. (1993). Population distribution of stereoscopic ability. *Ophthal. Physiol. Opt.*, *13*, 3-7.

Cumming, B.G., & Parker, A.J.(1994) Binocular mechanisms for detecting motion in depth. *Vision Res 34*, 483–495.

Czuba,T.B., & Rokers, B., Huk, A.C., Cormack, L.K. (2010) Speed and eccentricity tuning reveal a central role for the velocity-based cue to 3D visual motion. *J Neurophysiology* 104, 2886-2899.

Cynader, M., & Regan, D. (1978). Neurones in cat parastriate cortex sensitive to the direction of motion in three-dimensional space. *Journal of Physiology*, 274, 549-569.

Cynader, M., & Regan, D. (1982). Neurons in cat visual cortex tuned to the direction of motion in depth: E ect of positional disparity. *Vision Research*, 22(8), 967-982.

Dacey, D. M., & Lee, B.B.(1994) "The 'blue-on' opponent pathway in primate retina originates from a distinct bistratified ganglion cell type." *Nature* 367, 731-735.

DeAngelis, G.C., Cumming, G.C., Newsome, W.T. (1998) Cortical area MT and the perception of stereoscopic depth. *Nature* 394(6694), 677-680.

DeAngelis, G.C., Ohzawa, I., Freeman, R.D. (1993) Spatiotemporal organization of simple-cell receptive fields in the cats striate cortex. 1. General characteristics and postnatal development. *Journal of Neurophysiology*, 69(4), 1091-1117.

DeAngelis, G.C., Newsome, W.T. (1999) Organization of disparity-selective neurons in macaque area MT. *Journal of Neuroscience*, 19: 13981415.

DeAngelis, D.C. & Newsome, W.T. (2004) Perceptual "read-out" of conjoined direction and disparity maps in extrastriate area MT. *PLoS Biology* 2, 0349.

Decartes, R. (1633/1664). *Traité de l'Homme (De Homine, 1662)*.

DeYoe, E. A. & Van Essen, D. C. (1988) Concurrent processing streams in monkey visual cortex. *Trends Neurosci.* **11**, 219–226.

Dodd, J.V., Krug, K., Cumming, B.G., Parker, A.J. (2001) Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *Journal of Neuroscience*. 21(13), 4809-4821.

Domini, F., Caudek, C., & Tassinari, H. (2006) Stereo and motion information are not independently processed by the visual system. *Vision Research* 46, 1707-1723.

Dubner, R., Zeki, S.M. (1971) Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Research*. 35(2), 528-532.

Edwards, M., & Schor, C.M. (1999) Depth aliasing by the transient stereo-system. *Vision Research*, *39*, 4333-4340.

Fahle, M. (2005) Perceptual learning: Specificity versus generalisation. *Current Opinion in Neurobiology*, 15, 154-160.

Farell, B. (1998) Two-dimensional matches from one-dimensional stimulus components in human stereopsis. *Nature*, 395, 689-693.

Farvardin, M., & Afarid, M. (2007). Evaluation of stereo tests for screening of amblyopia. *Iranian Res Crescent Medical Journal*, 9(2), 80-85.

Faugeras, O.D. (1992). What can be seen in three dimensions with an uncalibrated stereo rig? In: Proceedings of ECCV 1992, 563-578.

Fawcett, F.L. (2005) An evaluation of the agreement between contour-based circles and random dot-based near stereoacuity tests. *Journal of AAPOS*, 9(6), 572-578.

Fernandez J.M., & Farell, B. (2005) Seeing motion-in-depth using inter-ocular velocity differences. *Vision Res 45*: 2786–2798.

Ferster, D.A. (1981) A comparison of binocular depth mechanisms in areas 17 and 18 of the cat visual cortex. *J. Physiol. Lond*, 311, 623-655.

Fox, R., Patterson. R., & Francis, E.L. (1986) Stereoacuity in young children. *Invest Opthalmol Vis Sci*, 27, 598-600.

Friedman, H.S., Zhou, H., von der Heydt, R. (2003) The coding of uniform colour figures in monkey visual cortex. *Journal of Physiology*. 548(Pt.2), 593-613.

Fujikado, T., Hosohata , J., Ohmi, G., Asonuma, S., Yamada, T., Maeda, N., & Tano, Y. (1998) Use of dynamic and coloured stereogram to measure stereopsis in strabismic patients. *Japanese Journal of Ophthalmology*, 42(2), 101-107.

Gantz, L., Patel, S.S., Chung, S.T.L., & Harwerth, R.S. (2007) Mechanisms of perceptual learning of depth discrimination in random dot stereograms. *Vision Research*, 47, 2170-2178.

Gellert, W., Gottwald, S., Hellwich, M., Kästner, H., Künstner, H. (Eds) (1989) *Plane*. In *VNR Concise encyclopedia of mathematics* (2nd ed). New York: Van Nostrand Reinhold.

Goodale, M.A., Jakobson, L.S., Keillor, J.M. (1994b) Differences in the control of pantomimend and natural grasping movements. Neuropsychologia. 32(10), 1159-1178.

Goodale, M. A. and & Milner, A.D.(1992). "Separate visual pathways for perception and action." *Trends Neurosci.* 15: 20-25.

Goodale, M.A., Milner, A.D., Jakobson, L.S., Carey, D.P. (1991) A neurological dissociation between perceiving objects and grasping them. *Nature*. 349(6305), 154-156.

Goodale, M. A., G. Króliczak, Milner, A.D. (2005). Dual routes to action: contributions of the dorsal and ventral streams to adaptive behavior. *Progress in Brain Research*. R. W. G. V.A. Casagrande and S. M. Sherman, Elsevier. Volume 149: 269-283.

Graf, E.W., Adams, W.J., & Lages, M. (2004). Prior monocular information can bias motion perception. *Journal of Vision*, *4*, 427-433.

Greenwald, H.S., & Knill, D.C. (2009) Orientation disparity: a cue for 3D orientation? *Neural Computation*, 21, 2581-2604.

Hansard, M., & Haroud, R. (2008). Cyclopean geometry of binocular vision. *J Opt Soc Am A, 25(9)*, 2357-2369.

Harris, J.M. (2006). The interaction of eye movements and retinal signals during the perception of 3-D motion direction. *J. Vision*, *6*, 777-790.

Harris, J.M., & Dean, P.J. (2003) Accuracy and precision of binocular 3-D motion perception. *Journal of Exp Psychol: Hum Percept Perform*, 29: 869-881.

Harris, J.M., Drga (2005) Using visual direction in three-dimensional motion perception. *Nature Neuroscience*, 8(2): 229-233.

Harris, J.M., Nefs, H.R., & Grafton, C.E. (2008). Binocular vision and motion-in-depth. *Spatial Vision*, 21(6), 531-547.

Harris JM, Rushton SK (2003) Poor visibility of motion-in-depth is due to early motion averaging. *Vision Res*, *43*, 385–392.

Hartley, R., & Zisserman, A. (2004). Multiple view geometry in computer vision (2nd ed.). Cambridge University Press.

Hatch, S.W. & Richman, J.E. (1994) Stereopsis testing without polarized glasses: A comparison study on five new stereoacuity tests. *J Am Optom Assoc*, 65, 637-641.

von Helmholtz, H. (1910/1962) In: Southall, JP editor. *Helmholtz's Treatise on Physiological Optics*, Vol 1. Dover: New York, USA. pp. 312–313.

Hendrickson, A. E., J. R. Wilson, et al. (1978) "The neuroanatomical organization of pathways between the dorsal lateral geniculate nucleus and visual cortex in Old World and New World primates." *J. Comp. Neurol*. 182: 123-136.

Hendry, S. H., & Reid, R.C. (2000). "The koniocellular pathway in primate vision." *Annu. Rev. Neurosci*. 23: 127-153.

Hendry, S. H., & Yoshioka, T. (1994) "A neurochemically distinct third channel in the macaque dorsal lateral geniculate nucleus." Science 264, 575-577.

Heron, G., Dholakia, S., Collins, D.E., & McLaughlan, H. (1985) Stereoscopic thresholds in children and adults. *AM J Optom Physiol Opt*, 62, 505-515.

Heron, S., & Lages, M. (2009) Measuring azimuth and elevation of binocular 3D motion direction [Abstract]. *J Vision 9*: 637a.

Hibbard, P.B. (2007) A statistical model of binocular disparity. *Visual Cognition* 15(2), 149-165.

Hildreth, E.C. & Koch, C. (1987) The analysis of visual motion from computational theory to neuronal mechanisms. *Annual Review of Neuroscience* 10, 477-533.

Hildreth, E.C. (1984) The computation of the velocity field. *Proc of the Roy Soc of London B 221*, 189 – 220.

Hinkle, D.A., Connor, C.E. (2002) Three-dimensional orientation tuning in macaque area V4. *Nat Neurosci 5*, 665-670.

Hogervorst, M.A., & Eagle, R.A. (1998) Biases in three-dimensional structure-from-motion arise from noise in the early visual system. *Proceedings of the Royal Society: Biological Sciences*, 265, 1587-1593.

Hogervorst, M.A. & Eagle, R.A. (2000) The role of perspective e ects and accelerations in perceived three-dimensional structure-from-motion. J of Exp Psych Hum Percept Perform 26: 934-955.

Holmes, J.A. & Leske, D.A. Monocular clues in tests of stereoacuity. In: C. Pritchard, Editor, *Transactions IX International Orthoptic Congress*, Berufsverband der Orthoptistinnen Deutschlands eV, Nurnberg Germany (1999), pp. 103–106.

Horn, B.K.P., & Schunck, B.G. (1981) Determining optical flow. Artificial Intelligence 17:

185-203.

Horton, J. C., & Hubel, D.H. (1981) "Regular patchy distribution of cytochrome oxidase staining in primary visual cortex of macaque monkey." *Nature* 292: 762-764.

Horton, J.C. (1984). Cytochrome oxidase patches: a new cytoarchitectonic feature of monkey visual cortex. Philos. Trans. R. Soc. London. Ser. B 304(1119), 199-253.

Horwitz, G.D., Chichilnisky, E.J., Albright, T.D. (2004) Spatial opponency and color tuning dynamics in macaque V1. *Soc. Neurosci. Abstr*. 34, 370-379.

Horwood, A.M., & Riddell, P.M. (2010) Differences between naïve and expert observers' vergence and accommodative responses to a range of targets. *Ophthal.Physiol.Opt*, 30, 152-159.

Howard, I.P. & Rogers, B.J. (2002). Seeing in depth (Vol 2). Toronto: I. Porteous, Toronto University Press.

Hubel, D.H. & Wiesel, T.N. (1968) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology* 195, 215-243.

Hubel, D.H., & Wiesel, T.N. (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiology (London) 160*, 106-154.

Hubel, D.H., Wiesel, T.N. (1970) Stereoscopic vision in macaque monkey. Cells sensitive to binocular depth in area 18 of the macaque monkey cortex. *Nature* 225, 41-42.

Hurlimann, F., Kiper, D.C., & Carandini, M. (2002) Testing the Bayesian model of perceived speed. *Vision Research* 42, 2253-2257.

Ito, H. (2003) The aperture problems in the Pulfrich effect. *Perception* 32, 367-375.

Jeffreys, H., Jeffreys, B.S. (1988) *Methods of Mathematical Physics 3rd ed*. Cambridge, England: Cambridge University Press.

Ji, H., Fermüller, C. (2006) Noise causes slant underestimation in stereo and motion. *Vision Res 46*, 3105-3120.

Johnson, E. N., M. J. Hawken, et al. (2001) The spatial transformation of color in the primary visual cortex of the macaque monkey. *Nature Neurosci.* 4, 409-416.

Johnston, E.B., Cumming, B.G., & Landy, M.S. (1994) Integration of stereopsis and motion shape cues. *Vision Research*, 34: 2259-2275.

Johnstone, A., McOwan, P.W., & Benton, C. (1999) Robust velocity computation from a biologically motivated model of motion perception. *Proc. R. Soc. Lond. B 266*, 509-518.

Julesz, B. (1971) Foundations of Cyclopean Perception. University of Chicago Press:

Chicago.

Kass, H.H., Huerta, M.F., Weber, J.T., Harting, J.K. (1978) Patterns of retinal terminations and laminar organization of the lateral geniculate nucleus of primates. *Journal of Computational Neurology*. 182(3), 517-553.

Knill, D.C., Kersten, D., & Yuille, A.L. (1996) Introduction: A Bayesian formulation of visual perception. In D.C. Knill, & W. Richards (eds.), Perception as Bayesian Inference, pp. Cambridge, UK: Cambridge University Press.

Knill, D.C., & Richards, W. (1996). *Perception as Bayesian inference*. Cambridge U. Press: Cambridge, UK.

Koenderink, J.J., van Doorn, A.J. (1991) Affine structure from motion. *J Optical Soc of America* 8, 377–385.

Koenderink JJ, van Doorn AJ, & Kapers AML (1992) Surface perception in pictures. *Percept Psychophysics* 52:487-496.

Koka, K. (1935/1963) Principles of Gestalt psychology. New York: A Harbinger Book, Harcourt, Brace & World, Inc.

Korte, A. (1915) Kinematoskopische Untersuchungen [Kinematoscopic investigations]. *Zeitschrift fur Psychologie* 72, 194296.

Kriegbaum-Stehberger, B., Jiang, X., & Mojon, D.S. (2008) Performance of a new, 3D-montior based random-dot stereotest for children under 4 years of age. *Graefes Arch Clin Exp Opthalmol*, 246, 1-7.

Lachica, E. A., P. D. Beck, et al. (1992). "Parallel pathways in macaque monkey striate cortex: anatomically defined columns in layer III." *Proc. Natl Acad. Sci. USA* 89, 3566-3570.

Lages, M. (2006) Bayesian models of binocular 3-D motion perception. *Journal of Vision* 6(4), 508-522.

Lages, M., Dolia, A., Graf, E.W. (2007) Dichoptic motion perception limited to depth of fixation? *Vision Res 47*, 244-252.

Lages, M., & Heron, S. (2008) Motion and disparity processing informs Bayesian 3D motion estimation. *Proc Nat Acad Sci USA 105*, E117.

Lages, M., & Heron, S. (2009) Testing generalized models of binocular 3D motion perception [Abstract]. *J Vision 9*, 636a.

Lages, M., & Heron, S. (2010) On the inverse problem of binocular 3D motion

perception. PLoS Computational Biology, 6(11): e1000999.

Lages, M., Jenkins, R., & Hillis, J.M. (2008) Anticipation of gravity alters perception of average speed. *Perception (Suppl.)*, *37*, 28.

Lages, M., Mamassian, P. & Graf, E.W. (2003) Spatial and temporal tuning of motion-in-depth. *Vision Research*, 43(27), 2861-2873.

Lages, M., & Paul, A. (2006) Visual long-term memory for spatial frequency? *Psychonomic Bulletin & Review*, *13(3)*, 486-492.

Lages, M., & Treisman, M. (1998) Spatial frequency discrimination: Visual long-term memory or criterion setting? *Vision Research*, *38*, 557-572.

 Lages, M., & Tresiman, M. (2010) A criterion setting theory of discrimination learning that accounts for anisotropies and context effects. *Seeing & Perceiving*, *23*, 401-434.

Ledgeway, T., & Smith, A.T. (1994) Evidence for separate motion-detecting mechanisms for first-order and 2nd-order motion in human vision. *Vision Res 34*, 2727-2740.

Le Vay, s., & Voigt, T. (1988) Ocular dominance and disparity coding in cat visual cortex. *Visual Neusrosci*, 1, 395-414.

Likova, L.T., Tyler, C.W. (2007) Stereomotion processing in the human occipital cortex. *Neuroimage 38*, 293–305.

Liu, J., & Newsome, W.T. (2003b) Functional organization of speed tuned neurons in visual area MT. *Journal of Neurophysiology*. 89(1), 246-256.

Livingstone, M., & Hubel, D. (1988) "Segregation of form, color, movement, and depth: anatomy, physiology, and perception." *Science* 240, 740-749.

Loffler, G., Orbach, H.S. (2001) Anisotropy in judging the absolute direction of motion. *Vision Research*, 41(27), 3677-3692.

Lorenceau, J., & Alais, D. (2001) Form constraints in motion binding. *Nature Neuroscience*, *4(7)*, 745-751.

Lorenceau J., Shifrar M., Wells N., Castet E. (1993) Different motion sensitive units are involved in recovering the direction of moving lines. *Vision Research*, 33, 1207-1217.

Lu, Z.L., Chu, W., Dosher, B.A., & Lee, S. (2005) Independent perceptual learning in monocular and binocular motion systems. *Proceedings of the National Academy of Sciences*, 102, 5624-5629.

Lu, H., Lin, T., Lee, A.L.F., Vese, L., & Yuille, A. (2010) Functional form of motion priors in human motion perception. *Proceedings of Neural Information Processing Systems (NIPS)*, 1-9.

233

Lu, Z.L., & Sperling, G. (1995) The functional architecture of human visual motion perception. *Vision Research*, *35*, 2697-2722.

Lu, Z.L., & Sperling, G. (2001) Three systems theory of human visual motion perception: review and update. *J Optical Soc of America A 18*, 2331-2370.

Lucas, B.D. & Kanade, T. (1981) An iterative image registration technique with an application to stereo vision, DARPA Image Understanding Workshop, pp. 121130 (see also IJCAI81, pp. 674-679).

Lüneburg, R.K. (1947) *Mathematical analysis of binocular vision*. Princeton, NJ: Princeton University Press.

Majaj, N., Carandini, M., & Movshon, J.A. (2007) Motion integration by neurons in macaque MT is local not global. *J Neurosci*, 27, 366-370.

Malach, R.,Schirman, T.D., Harel, M., Tootell, R.B.H., Malonek, D. (1997) Organization of connections in owl monkey area MT. *Cerebral Cortex*. 7(4), 386-393.

Maloney, L.T., & Landy, M.S. (1989) A statistical framework for robust fusion of depth information. In W. A. Pearlman (Ed.), Visual communications and image processing IV (vol. 1199, pp. 1154 1163). Proceedings of SPIE.

Mamassian, P., Landy, M.S., & Maloney, L.T. (2002) Bayesian modelling of visual perception. In: R. Rao, B Olshausen, M Lewicki (Eds.) probabilistic models of the brain: perception and neural function (pp. 13-36). MIT Press, Cambridge, MA.

Marr, D. (1982) *Vision*. W.H. Freeman & Co: New York, US.

Marr, D., & Ullman, S. (1981) Directional sensitivity and its use in early visual processing. *Proceedings of the Royal Society of London*, Series B 211: 151 -180

Masson, G.S., & Castet, E. (2002) Parallel motion processing for the initiation of short-latency ocular following in humans. *J Neurosci 22*, 5149-5163.

Maunsell, J.H., & van Essen, D.C. (1983a) Functional properties of neurons in middle temporal visual area of the macaque monkey: I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology*, 49, 1127-1147.

Maunsell, J.H., & van Essen, D.C. (1983b) The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. Journal *of Neuroscience*, 3(12), 2563-86.

Maunsell, J.H., & Van Essen, D.C. (1983c) The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *Journal of Neuroscience*.3(12), 2563-2586.

Maunsell, J.H., & Newsome, W.T. (1987) Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*. 10, 363-401.

Mayhew JEW, Longuet-Higgins HC (1982) A computational model of binocular depth perception. *Nature*, 297: 376-378.

McDermott, J., Weiss, Y., & Adelson, E. H. (2001) Beyond junctions: Non-local form constraints on motion interpretation. *Perception*, 30, 905923.

Miles, F.A., (1998) The neural processing of 3-D visual information: evidence from eye movements. *European Journal of Neuroscience*, *10(3)*, 811-822.

Milner, A.D., & Goodale, M.A. (2008) Two visual systems re-viewed. *Neuropsychologia*. 46, 774-785.

Milner, A.D., Perret, D.I., Johnston, R.S., Benson, P.J., Jordan, T.R., Heelet, D.W., et al (1991) Perception and action in 'visual form agnosia'. *Brain*, 114 (Pt.1B), 405-428.

Mitchison, G., & McKee, S. (1990) Mechanisms underlying the anisotropy of stereoscopic tilt perception. *Vision Research* 30, 1781-1791.

Min, D., & Sohn, K. (2006) Edge-preserving simultaneous joint motion-disparity estimation. *Proceedings of the 18th International Conference on Pattern Recognition* (ICPR'06) Vol 2, 74-77.

Montagnini, A., Mamassian, P., Perrinet, L., Castet, E., & Masson, G.S. (2007) *Journal of Physiology*, *101*, 64-77.

Morgan, M.J., & Castet, E. (1997) The aperture problem in stereopsis. *Vision Research*, 37(19), 2737-2744.

Morgan, M.J., & Fahle, M. (2000) Motion-stereo mechanisms sensitive to interocular phase. *Vision Research*, *40*, 1667-1675.

Movshon, J.A., Adelson, E.H., Gizzi, M.S., Newsome, W.T. (1985) The analysis of moving visual patterns. In Pattern Recognition Mechanisms, ed. C Chagas, RGattass, C Gross, pp. 117-151. Rome: Vatican.

Muckli, L., T. Schicke, A. L. Beer, W. Singer, R. Goebel, F. Roesler & B. Roeder. (2006) Tight covariation of BOLD signal changes and slow ERPs in the parietal cortex in a parametric spatial imagery task with haptic acquisition*. European Journal of Neuroscience*. Apr, 23(7), 1910-1918.

Nakayama, K. (1985) Biological image motion processing: A review. *Vision Research* 25, 625-660.

Nakayama, K., & Silverman, G.H. (1988) The aperture problem II: Spatial integration of velocity information along contours. *Vision Research,* 28, 747-753.

Nassi, J.J., Callaway, E.M. (2009) Parallel processing strategies of the primate visual system. *Nature Reviews Neuroscience*. 10(5), 360-372.

Nefs, H., & Harris, J.M. (2010). What visual information is used for stereoscopic depth displacement discrimination. *Perception*, *39*, 727-744.

Nefs, H., O'Hare, L., & Harris, J. (2010). Two independent mechanisms for motion-in-depth perception: evidence from individual differences. *Frontiers in Psychology - Perception Science*, *1*, 155.

Neri, P. (2005) A stereoscopic look at visual cortex. *Journal of Neurophysiology*, 93 (4), 1823-1826.

Nguyenkim, J.D., DeAngelis, G.C. 2004. Macaque MT neurons are selective for 3D surface orientation defined by multiple cues. *Soc. Neurosci.Abstr*. 30, 368-412.

Nikara, T., Bishop, P.O., & Pettigrew, J.D. (1968) Analysis of retinal correspondence by studying receptive fields of binocular single units in cat striate cortex. *Exp, Brain Res*, 6, 353-372.

Norman, K. A., S. M. Polyn, et al. (2006) Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10(9), 424-430.

Nover, H., Anserson, CH,&DeAngelis,G.C.(2005) A logarithmic, scale-invariant representation of speed in macaque middle temporal area accounts for speed discrimination performance, *J of Neuroscience* 25(43): 10049-10060.

Ogle, K.N. (1940) Induced size effect with the eyes in asymmetric convergence. *Archives of Ophthalmology, 23*, 1023-1028.

Ohlsson, J., Villarreal, G., Abrahamsson, M., Cavasos, H., Sjostrom, A., & Sjostrand, J. (2001) Screening merits of the Lang II, Frisby, Randot, Titmus and TNO stereo tests. *Journal of AAPOS*, 5(5), 316-322.

Ohzawa, I., DeAngelis, G.C., Freeman, R.D. (1990) Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. *Science*, 249(4972), 1037-1041.

Orban, G.A. 1997. Visual processing in macaque area MT/V5 and its satellites (MSTd and MSTv). In Extrastriate cortex in Primates, ed. KS Rockland, JF Kaas, A Peters, pp. 359-434. New York: Plenum.

Orban, G.A. (2003) Similarites and differences in motion processing between the human and macaque brain: evidence from fMRI. *Neuropsychologia*, 41, 1157-1768.

Orban, G.A. (2008) Higher order visual processing in macaque extrastriate cortex. *Physio. Rev* 88: 59-89.

Osborne, L.C., Bialek, W., & Lisberger, S.G. (2004). Time course of information about motion direction in visual area MT of macaque monkeys. *Journal of Neuroscience*, *24(13)*, 3210-3222.

O'Toole, A. & Kersten, D. (1992) Learning to see random dot stereograms. *Perception*, 21, 227-243.

Pack, C.C., Born, R.T., Livingstone, M.S. (2003) Two-dimensional substructure of stereo and motion interactions in macaque visual cortex. *Neuron*, *37*: 525-535.

Palanca, B.J., & DeAngelis, G.C. (2003) Macaque middle temporal neurons signal depth in the absence of motion. *Journal of Neuroscience*. 23, 7647-7658.

Parker, A.J. (2007) Binocular depth perception and the cerebral cortex. *Nature*, 8, 379-391.

Ponce C.R., Lomber S.G., Born R.T. (2008) Integrating motion and depth via parallel pathways*. Nature Neuroscience*, *11*, 216–223.

Peng, Q & Shi, B.E. (2010) The changing disparity energy model. *Vision Research* 50, 181-192.

Perry, E.H., Oehler, R., Cowey, A. (1984) Retinal ganglion cells that project to the dorsal lateral geniculate nucleus in the macaque monkey. *Neuroscience* 12(4), 1101-1123.

Pettigrew, J.D., Nikara, T., & Bishop, P.O. (1968) Binocular interaction on single units in cat striate cortex: simultaneous stimulation by single moving slit with receptive fields in correspondence. *Exp. Brain Res*, 6, 391-410.

Pizlo, Z. (2001) Perception viewed as an inverse problem. *Vision Res*, *41*, 3145-3161.

Poggio, G.F., & Fischer, B. (1977) Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving Rhesus monkey. *J. Neurophysiol*. 40, 1392-1405.

Poggio, G. F., & Talbot, W. H. (1981) Mechanisms of static and dynamic stereopsis in foveal cortex of the rhesus monkey. *Journal of Physiology*, 315, 469-492.

Poggio, T., Torre, V., Koch, C. (1985) Computational vision and regularization theory. *Nature*, *317*: 314-319.

Ponce, C.R., Lomber, S.G., & Born, R.T. (2008) Integrating motion and depth via parallel pathways. *Nature Neuroscience*, 11, 216-223.

Portfors-Yeomans, C.V., & Regan, D. (1996) Cyclopean discrimination thresholds for the direction and speed of motion in depth. *Vision Research*, *36*, 3265-3279.

Prince, S.J.D. & Eagle, R.A. (2000) Weighted directional energy model of human stereo correspondence. *Vision Research,* 40(9), 1143-1155.

Qian, N. (1994) Computing stereo disparity and motion with known binocular cell properties. *Neural Comp 6*, 390-404.

Qian, N., & Andersen, R.A. (1997) A physiological model for motion-stereo integration and a unified explanation of Pulfrich-like phenomena. *Vision Res 37*, 1683–1698.

Raftery, A.E. (1995) Bayesian model selection in social research. Im PV Marsden (Ed.) Sociological method-ology (pp. 111-196). Cambridge, MA: Blackwell.

Raftery, A.E. (1999) Bayes factors and BIC. *Sociological Methods & Research*, 27, 411-427.

Rashbass, C., & Westheimer, G. (1961) Disjunctive eye movements. *J Physiology 159*, 339-360.

Read, J.C.A. (2002a) A Bayesian approach to the stereo correspondence problem. *Neural Computation*, *14*, 1371-1392.

Read, J.C.A. (2002b). A Bayesian model of stereopsis depth and motion direction discrimination. *Biological Cybernetics*, *86(2)*, 117-136.

Read, J.C.A., & Cumming, B.G. (2005a) Effect of interocular delay on disparity-selective V1 neurons: Relationship to stereoacuity and the Pulfrich effect. *J Neurophys 94*: 1541-1553.

Read, J.C.A., & Cumming, B.G. (2005b) The stroboscopic Pulfrich effect is not evidence for the joint encoding of motion and depth. *J Vision*, *5*, 417-434.

Read, J.C.A., Phillipson, G.P., Glennerster, A. (2009) Latitude and longitude vertical disparities. *J Vision*, *9*: 1-37.

Regan, D. (1993) Binocular correlates of the direction of motion in depth. *Vision Research*, 33(16): 2359-2360.

Regan, D., & Beverley, K.I. (1973) Some dynamic features of depth perception, *Vision Res 13*, 2369–2379.

Regan, D., & Beverley, K.I. (1980) Binocular and monocular stimuli for motion in depth: Changing- disparity and changing-size feed the same motion-in-depth stage. *Vision Research 19*, 1331-1342.

Regan, D., Beverley, K.I., Cynader, M., Lennie, P. (1979) Stereoscopic subsystems for position in depth and for motion in depth. *Proc R Soc Lon B* 42(204), 485-501.

Regan, D., Erkelens, C.J., & Collewijn, H. (1986) Visual field defects for vwergence eye movements and for stereomotion perception. *Invest Ophthalmol Vis Sci*, *27*, 806-819.

Regan, D., & Gray, R. (2009) Binocular processing of motion; some unresolved problems. *Spatial Vision 22*, 1-43.

Richards, W. (1971) Anomalous stereoscopic depth perception. *Journal of the Optical Society of America*, 61(3), 410-414.

Richards, W. (1985) Structure from stereo and motion. Journal of the *Optical Society of America*, A 2343-349

Richards, W. & Regan, D. (1973) A stereo field map with implications for disparity processing, *Invest. Ophthalmol*, **12**, 904–909.

Rizzolatti, G. & Matelli, M. (2003) Two different streams form the dorsal visual system: anatomy and functions. *Experimental Brain Research*. 153, 146-157.

Reinecke, R.D, & Simons, K. (1974). A new stereoscopic test for amblyopia. *American Journal of Ophthalmology*, 78, 714-721.

Rokers, B., Cormack, L.K., Huk, A.C. (2008) Strong percepts of motion through depth without strong percepts of position in depth. *J Vision 8*, 1–10.

Rokers, B., Cormack, L.K., Huk, A.C. (2009) Disparity- and velocity-based signals for three-dimensional motion perception in human MT+. *Nat Neurosci*, *12*, 1050–1055.

Rokers, B., Czuba, T.B., Cormack, L.K., & Huk, A.C. (2011) Motion processing with two eyes in three dimensions. *Journal of Vision*, 11(2):10, 1-9.

Rushton, S.K., & Duke, P.A. (2007). The use of direction and distance information in the perception of approach trajectory. *Vision Research*, *47*, 899-912.

Rushton, S.K., & Warren, P.A. (2005) Moving observers, relative retinal motion, and the detection of object movement. *Current Biology*, *15(14)*, R542.

Sacks, A. T., Kohler, a., Linden, D.E.J., Goebel, R., & Muckli, L. (2006). The temporal characteristics of visual motion processing in hMTV5+. *NeuroImage*. Feb 15, 29(4), 1326-35

Sakai, K., Ogiya, M., & Hirai, Y. (2011) Decoding of depth and motion in ambiguous binocular perception. *J Opt Soc Am* A 28(7), 1445-1452.

Scarfe, P., & Hibbard, P.B. (2011) Statistically optimal integration of biased sensory estimates. *Journal of Vision* 11(7), 12, 1-17.

Scharr, H., & Küsters, R. (2002) A linear model for simultaneous estimation of 3D motion and depth. *IEEE Workshop on Motion and Video Computing*, Orlando FL, 1-6.

Schmidt, P., Maguire, M., Kulp, M.T., Dobson, V., & Quinn, G. (2006) Random dot E stereotest: Testability and reliability in 3- to 5-year-old children. *Journal of AAPOS*, 10(6), 507-514.

Schreiber, K.M., Hillis, J.M., Filippini, H.R., Schor, C.M., Banks, M.S. (2008) The surface of the empirical horopter. *Journal of Vision 8*, 1-20.

Shimojo, J., Silverman, G.H., Nakayama, K. (1988) An occlusion-related mechanism of depth perception based on motion and interocular sequence. *Nature* (London) 222, 265-268.

Shimojo, J., Silverman, G.H., Nakayama, K. (1989) Occlusion and the solution to the aperture problem. *Vision Research* 29, 619-626

Shioiri, S., Saisho, H., Yaguchi, H. (2000) Motion in depth based on inter-ocular velocity differences, *Vision Res 40*, 2565–2572.

Shipp, S., & Zeki. (1985) Segregation of pathways leading from area V2 to areas V4 and V5 of macaque monkey visual cortex. *Nature* 315(6017), 322-325.

Shipp, S., & Zeki. (1989b) The organization of connections between areas V5 and V2 in macaque monkey visual cortex. *European Journal of Neuroscience*. 1(4), 333-354.

Shipp, S. & Zeki, S. (2002) The functional organization of area V2, I: specialization across stripes and layers. *Vis. Neurosci.* **19**, 187–210.

Sinich, L.C., & Horton, J.C. (2002) Divided by cytochrome oxidase: a map of the projections from V1 to V2 in macaques. *Science*. 295(5560), 1734-1737.

Sinich, L.C., & Horton, J.C. (2005) The circuitry of V1 and V2: Integration of color, form and motion. *Annual Review of Neuroscience*. 28, 303-326.

Smith, A.T. & Wall, M.B. (2008) Sensitivity of human visual cortical areas to the stereoscopic depth of a moving stimulus. *Journal of Vision* 8(10), 1-12.

Snowden, R.J., Treue, S, Erickson, R.G., Andersen, R.A. (1991) The response of area MT and V1 neurons to transparent motion. *Journal of Neuroscience*. 11(9), 2768-2785.

Sowden, P., Davies, I., Rose, D., & Kayne, M. (1996) Perceptual learning of stereoacuity. *Perception*, 25, 1043-1052.

Spies, H., Jähne, B.J. & Barron, J.L. (2002) Range flow estimation. *Computer Vision Image Understanding (CVIU2002)* 85, 209-231.

Stocker, A.A., & Simoncelli, E.P. (2006) Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience*, *9(4)*, 578-585.

Sung, K., Wojtach, W.T., Purves, D. (2009) An empirical explanation of aperture effects. *Proc Nat Acad Sci USA*, *106*, 298-303.

Thomas, O.M., Cumming, B.G., Parker, A.J. (2002) A specialization for relative disparity in V2. *Nat Neurosci*, *5*, 472-478.

Tootell, R.B., Silverman, M.S., DeValois, R.L., Jacobs, G.H. (1983) Functional organization of the second cortical visual area in primates. *Science* 220(4598), 737-739.

Treue, S., Hol, K., & Rauber, H.J. (2000) Seeing multiple directions of motion – physiology and psychophysics. *Nature Neuroscience*, *3*, 270-276.

Tyler, C.W. (1971) Stereoscopic depth movement: Two eyes less sensitive than one. *Science*, *174*, 958–961.

Ugolini, G. (1995) Specificity of rabies virus as a transneuronal tracer of motor networks: transfer from hypoglossal motorneurons to connected second order and higher order central nervous system cell groups. *Journal of Computational Neurology*. 356(3), 457-480.

Uka, T., & DeAngelis, G.C. 2004. Contribution of area MT to stereoscopic depth perception: choice related response modulations reflect task strategy. *Neuron* 42(2), 297-310.

Ullman, S., & Yuille, A. (1989) Rigidity and smoothness of motion. In: S. Ullman & W. Richards, *Image understanding*. Norwood, NJ: Ablex Publishing Corporation.

Ungerleider, L.G., & Mishkin, M. Two cortical visual systems. In The Analysis of Visual Behaviour, ed. DJ Ingle, RJW Mansfield, MS Goodale, pp. 549-586. Cambridge, MA: MIT Press.

Vanduffel, W., Tootell, R.B., Schoups, A.A., Orban, G.A. (2002). The organization of orientation selectivity throughout macaque visual cortex. *Cerebral Cortex*. 12(6), 647-662.

van Dam, L.C.J., & van Ee, R. (2004) Stereoscopic matching and the aperture problem. *Perception* 33: 769-787.

van Ee, R., & Anderson, B.L. (2001) Motion direction, speed and orientation in binocular matching. *Nature 410*, 690-694.

van Ee, R., & Schor, C.M. (2000) Unconstrained stereoscopic matching of lines. *Vision Res 40*, 151-162.

van Essen, D.C. & Maunsell, J.H., Bixby, J.L. (1981) The middle temporal visual area in the macaque: myloarchitecture, connections, functional properties and topographic organization. *Journal of Computational Neurology*. 199(3), 293-326.

Vedula, S., Baker, S., Rander, P., Collins, R., & Kanade, T. (2005) Three-dimensional scene-flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 475-480.

von Helmholtz, H. (1910/1962) In: Southall, JP editor. *Helmholtz's Treatise on Physiological Optics*, Vol 1. Dover: New York, USA. pp. 312–313.

Wallach, H. (1935) Über visuell wahrgenommene Bewegungsrichtung. *Psychol Forschung*, *20*: 325-380.

Wang, H., & Lages, M. (2011) A biologically-inspired Bayesian model of 3D velocity estimation. *AVA Meeting, Cardiff*.

Watanabe, Y., Kezuka, T., Harasawa, K., Usui, M., Yaguchi, H., & Shioiri, S. (2008) A new method for assessing motion-in-depth perception in strabismic patients. *Br J Opthalmol*, 92, 47-50.

Watt, S.J., Akeley, K., & Banks, M.S. (2005) Focus cues affect perceived depth. *J Vision*, *5(10)*, 834-862.

Weiss, Y., & Fleet, D.J. (2001) In: Probabilistic models of the brain: perception and neural function R.P.N. Rao, B. Olshausen, & M.S. Lewicki (Eds.), Velocity likelihoods in biological and machine vision, pp.81-100. Cambridge, MA: MIT Press.

Weiss, Y., Simoncelli, E.P., Adelson, E.H. (2002) Motion illusions as optimal percepts. *Nat Neurosci 5*, 598-604.

Welchman, A.E., Lam, J.M., Bülthoff, H.H. (2008) Bayesian motion estimation accounts for a surprising bias in 3D vision. *Proc Nat Acad Sci USA*, *105*, 12087-92.

Welchman, A.E., Tuck, V.L., & Harris, J.M. (2004) Human observers are biased in judging the angular approach of a projectile. *Vision Research*, *44*, 2027-2042.

Wilmer, J.B., & Backus, B.T. (2008) Self reported magic eye ™ stereogram skill predicts stereoacuity. *Perception*, 37(8), 1297-1300.

Wilson, H.R., Ferrera, V.P., Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Vis Neurosci 9(1)*, 79-97.

Xu, X., Bosking, W., Sary, G., Stefansic, J., Shima, D., Casagrande, V.A. (2004) Functional organisation of visual cortex in the owl monkey. *Journal of Neuroscience*. 24(28), 6237-6247.

Yabuta, N. H. and E. M. Callaway (1998) "Functional streams and local connections of layer 4C neurons in primary visual cortex of the macaque monkey." *J. Neurosci*. 18, 9489-9499.

Yabuta, N.H., Sawatari, A., Callaway, E.M. (2001) Two functional channels from primary visual cortex to dorsal visual cortical areas. *Science*. 292(5515), 297-300

Yamada, T., Scheiman, M., & Mitchell, G.L. (2008) A comparison of stereopsis testing between red/green targets and polarized targets in children with normal binocular vision. *Optometry*, 79, 138-142.

APPENDIX

Chapter 2

Appendix A2: Analytic Geometry

In the following we assume a fixed binocular viewing geometry with the cyclopean origin $O = (0,0,0)$ centered $\pm i/2$ between the nodal points of the left and right eye and the eyes verged on a fixation point straight ahead at viewing distance D (see Fig. 2.1). More complicated geometries arise if we take into account version, cyclovergence, and cyclotorsion of the eyes (Read, Phillipson & Glernnerster, 2009; Schreiber et al. 2008). For the sake of simplicity we ignore the non-linear aspects of visual space (Lüneburg, 1947) and represent perceived 3D motion as a linear vector in a three-dimensional Euclidean space where the fixation point is also the starting point of the motion stimulus.

Since we are not concerned about particular algorithms and their implementation, results are given in terms of analytic geometry (Jeffrey & Jeffreys, 1988; Gellert et al., 1989).

Intersection of Constraint Lines

If the eyes remain verged on a fixation point in a binocular viewing geometry then the constraint line in the left and right eye can be defined by pairs of points $(\mathbf{a}, \mathbf{b})$ and $(\mathbf{c}, \mathbf{d})$, respectively. The nodal point in the left eye $\mathbf{a} = (-i/2, 0, 0)$ and a projection point $\mathbf{b} = (x_L, y_L, z_L)$ of the motion vector on the left retina define a constraint line for the left eye.

Similarly, points $\mathbf{c} = (+i/2,0,0)$ and $\mathbf{d} = (x_R, y_R, z_R)$ determine a constraint line in the right eye. The corresponding vector directions are given by

$$
\begin{aligned}
(\mathbf{a} - \mathbf{c}) &= (-i/2,0,0) - (+i/2,0,0) = (-i,0,0) \\
(\mathbf{b} - \mathbf{a}) &= (x_L, y_L, z_L) - (-i/2,0,0) = (x_L + i/2, y_L, z_L) \\
(\mathbf{d} - \mathbf{c}) &= (x_R, y_R, z_R) - (+i/2,0,0) = (x_R - i/2, y_R, z_R)
\end{aligned}
\tag{2.1}
$$

Each constraint line can expressed by a pair of points $(\mathbf{a}, \mathbf{b})$ and $(\mathbf{c}, \mathbf{d})$ together with scalar t:

$$
\begin{aligned}
\mathbf{x}_L &= \mathbf{a} + (\mathbf{b} - \mathbf{a})t \\
\mathbf{x}_R &= \mathbf{c} + (\mathbf{d} - \mathbf{c})t
\end{aligned}
\tag{2.2}
$$

The two lines intersect for

$$
t = \frac{[(\mathbf{c} - \mathbf{a}) \times (\mathbf{d} - \mathbf{c})] \cdot [(\mathbf{b} - \mathbf{a}) \times (\mathbf{d} - \mathbf{c})]}{\|(\mathbf{b} - \mathbf{a}) \times (\mathbf{d} - \mathbf{c})\|^2}
\tag{2.3}
$$

if and only if

$$
(\mathbf{a} - \mathbf{c}) \cdot [(\mathbf{b} - \mathbf{a}) \times (\mathbf{d} - \mathbf{c})] = 0
\tag{2.4}
$$

where $\cdot$ is the scalar product also called the dot product, $\times$ denotes the cross product, and $\|.\|$ the norm of a vector. Otherwise, the two lines are skew, and the inverse problem is ill posed.

We can exclude the trivial case $(\mathbf{a} - \mathbf{c}) = \mathbf{0}$ because the two eyes are separated by $i > 0$. We also exclude the special case where the cross product is zero because the motion vectors in the left and right eye are identical or opposite.

The cross product in (2.4) can be written as

$$(\mathbf{b} - \mathbf{a}) \times (\mathbf{d} - \mathbf{c}) =$$
$$(y_L z_R - z_L y_R, z_L(x_R - i/2) - (x_L + i/2)z_R, (x_L + i/2)y_R - y_L(x_R - i/2)) \tag{2.5}$$

Since $(\mathbf{a} - \mathbf{c}) = (-i, 0, 0)$ in Eq. (2.4) we are only concerned with the product $-i(y_L z_R - z_L y_R)$ which equals zero if and only if

$$y_L z_R = y_R z_L \quad or \quad \frac{y_L}{y_R} - \frac{z_L}{z_R} = 0, \tag{2.6}$$

The ratio of z co-ordinates on the right-hand side may be different from 1 as a result of eye vergence and the left-hand side reflects the corresponding ratio of vertical displacements.

In the following we consider the simpler case of projections onto a fronto-parallel screen (coplanar retinae) at a fixed viewing distance D (see Fig. 2.2). In this case epipolar lines are horizontal with equivalent co-ordinates $z_L = z_R = z_C$ on the z-axis.

Again, since $(\mathbf{a} - \mathbf{c}) = (-i, 0, 0)$ in (2.4) we only have to evaluate $-i z_C(y_L - y_R)$ which is zero if and only if:

$$y_L - y_R = 0 \tag{2.7}$$

For an intersection to exist the left and right eye motion vector must have equivalent horizontal *y* co-ordinates or zero vertical disparity.

Intersection of Constraint Planes

Monocular line motion defines a constraint plane with three points: the nodal point of an eye and two points defining the end position of the projected line (see Fig. 3). In order to find the intersection of the left and right eye constraint plane we use the plane normal in the left and right eye. If the two planes are specified in Hessian normal form

$$\mathbf{n}_L \cdot \mathbf{p} = d_L,$$
$$\mathbf{n}_R \cdot \mathbf{p} = d_R$$

(2.8)

where $.$ is again the dot product, $\mathbf{n} = (a, b, c)$ is a vector describing the surface normal to a plane, $\mathbf{p} = (x, y, z)$ is a vector representing all points on the plane, and *d* is a scalar.

We need to check whether the constraint planes are parallel or coincident, that is if

$$\mathbf{n}_L \times \mathbf{n}_R = \mathbf{0}$$

(2.9)

before we can determine their intersection. The equation for the intersection of the two constraint planes is a line here written as

$$\mathbf{p} = c_L \mathbf{n}_L + c_R \mathbf{n}_R + u(\mathbf{n}_L \times \mathbf{n}_R)$$

(2.10)

where u is a free parameter. Taking the dot product of the above with each plane normal gives two equations with unknown scalars c_L and c_R.

$$\mathbf{n}_L \cdot \mathbf{p} = d_L = c_L (\mathbf{n}_L \cdot \mathbf{n}_L) + c_R (\mathbf{n}_L \cdot \mathbf{n}_R)$$
$$\mathbf{n}_R \cdot \mathbf{p} = d_R = c_L (\mathbf{n}_L \cdot \mathbf{n}_R) + c_R (\mathbf{n}_R \cdot \mathbf{n}_R)$$

(2.11)

Solving the two equations for c_L and c_R gives

$$c_L = [d_L (\mathbf{n}_R \cdot \mathbf{n}_R) - d_R (\mathbf{n}_L \cdot \mathbf{n}_R)]/\Delta,$$
$$c_R = [d_R (\mathbf{n}_L \cdot \mathbf{n}_L) - d_R (\mathbf{n}_L \cdot \mathbf{n}_R)]/\Delta$$

(2.12)

where $\Delta = (\mathbf{n}_L \cdot \mathbf{n}_L)(\mathbf{n}_R \cdot \mathbf{n}_R) - (\mathbf{n}_L \cdot \mathbf{n}_R)^2$.

Inserting c_L and c_R in (10) determines the intersection of constraints or constraint line $\mathbf{p}$.

In analogy to the 2D aperture problem and the intersection of constraints we can now define two plausible strategies for solving the 3D aperture problem:

Vector Normal (VN)

The shortest distance in 3-D (velocity) space between the starting point $\mathbf{p}_0 = (0,0,D)$ of the stimulus line and the constraint line $\mathbf{p}$ is the line or vector normal through point $\mathbf{p}_0$. In order

to determine the intersection point of the vector normal with the constraint line we pick two arbitrary points $\mathbf{p}_1$ and $\mathbf{p}_2$ on intersection constraint line $\mathbf{p}$ by choosing a scalar u (e.g., 0.5).

$$
\begin{aligned}
\mathbf{p}_1 &= c_L \mathbf{n}_L + c_R \mathbf{n}_R - u(\mathbf{n}_L \times \mathbf{n}_R) \\
\mathbf{p}_2 &= c_L \mathbf{n}_L + c_R \mathbf{n}_R + u(\mathbf{n}_L \times \mathbf{n}_R)
\end{aligned}
\tag{2.13}
$$

Together with point $\mathbf{p}_0$ we can compute scalar t_n as

$$
t_n = -\frac{(\mathbf{p}_1 - \mathbf{p}_0) \cdot (\mathbf{p}_2 - \mathbf{p}_1)}{\|\mathbf{p}_2 - \mathbf{p}_1\|^2}
\tag{2.14}
$$

which determines the closest intersection point $\mathbf{x}$ on the constraint line:

$$
\mathbf{x} = \mathbf{p}_1 + (\mathbf{p}_2 - \mathbf{p}_1)t_n
\tag{2.15}
$$

Cyclopean Average (CA)

We can define a cyclopean constraint line in terms of the cyclopean origin $O = (0,0,0)$ and projection point $\mathbf{p}_C = (x_C, y_C, z_C)$ on a fronto-parallel screen where $x_C = (x_L + x_R)/2$ and $y_C = (y_L + y_R)/2$ are the averages of the 2D normal co-ordinates for the left and right eye projections.

If we measure disparity δ at the same retinal coordinates as the horizontal offset between the left and right eye anchored at position $\mathbf{p}_C$ then we can define new points $\mathbf{b}$ with $x_L' = x_C - \delta/2$ and $\mathbf{d}$ with $x_R' = x_C + \delta/2$. (Alternatively, we may establish an epipolar or more sophisticated disparity constraint.) The resulting two points together with the

248

corresponding nodal points **a** and **c** define two constraint lines as in (2.2), one for the left and the other for the right eye. By inserting the new co-ordinates from above into (2.4) it is easy to see that condition (2.6) holds and the scalar for the intersection of lines can be found as in (2.3).

Transformation into spherical co-ordinates

The intersection $\mathbf{x} = (x, y, z)$ in cartesian co-ordinates can be transformed into spherical co-ordinates $(\alpha, \beta, \|\mathbf{s}\|)$ using vectors $\mathbf{q} = (x, 0, z - D)$ and $\mathbf{r} = (x, 0, D)$ to determine azimuth α in the horizontal plane

$$\alpha = \arccos\left(\frac{\mathbf{q} \cdot \mathbf{r}}{\|\mathbf{q}\| \|\mathbf{r}\|}\right) \tag{2.16}$$

Similarly, for base vectors $\mathbf{s} = (x, y, z - D)$ and $\mathbf{q} = (x, 0, z - D)$ elevation β is given by

$$\beta = \arccos\left(\frac{\mathbf{s} \cdot \mathbf{q}}{\|\mathbf{s}\| \|\mathbf{q}\|}\right) \tag{2.17}$$

Speed in 3D space is equivalent to the norm of vector **s** written as $\|\mathbf{s}\|$.

Chapter 4

Appendix A3: Generalized Bayesian Binocular 3D Motion Model

Two points on the image plane

$$\mathbf{p}_1 = (t_R, 0, 0)$$
$$\mathbf{p}_2 = (t_R - \cos(\theta_R), \sin(\theta_R), 0) \tag{3.13}$$

define a constraint line (2D intensity gradient) for the right eye where t_R is the horizontal translation and θ_R measures the line orientation from horizontal in the image plane. Similarly t_L is the horizontal translation and line orientation θ_L for the left eye. We define interocular velocity difference (iovd) as $(t_L\text{-}t_R)$ and orientation disparity as $(\theta_L\text{-}\theta_R)$ in the image plane. The nodal point of the right eye is given by

$$\mathbf{p}_0 = (+i/2, 0, -D) \tag{3.14}$$

where i denotes internodal distance and D is the distance to the image plane and fixation point. Then the 3D plane normal which describes the right constraint plane can be expressed as

$$\mathbf{n}_R = \frac{(\mathbf{p}_2 - \mathbf{p}_0) \times (\mathbf{p}_1 - \mathbf{p}_0)}{\|(\mathbf{p}_2 - \mathbf{p}_0) \times (\mathbf{p}_1 - \mathbf{p}_0)\|^2} = \frac{\nabla R}{\|\nabla R\|} \tag{3.15}$$

where $\perp$ denotes the cross product and $\| . \|$ the norm. The three components of the normal may be understood as (first-order Taylor approximations of) intensity gradients for the left ∇L and right eye ∇R, respectively. In a strict sense, intensity gradients are only defined on a surface, so the gradient in z is undefined. However, we can extend the 2D definition to 3D by setting 3D intensity gradients constant for all x,y, and z that project to the same 2D image gradients. The resulting constraint planes describe all possible 3D positions of the line for the left and right eye, respectively. Alternatively, the constraint planes and their intersection (IOC) may be understood as the end position in 3D when the moving line is sampled over time. This implies that disparity rather than motion processing determines line positions in 3D before they are integrated by a 3D motion system.

In analogy to 2D motion gradients in the image plane (Weiss & Fleet, 2001), 3D intensity gradients may be approximated by first-order Taylor series expansion using brightness constraints in the left and right eye.

$$\begin{aligned}
\mathbf{v}^T \nabla L + L_t &= 0, \\
\mathbf{v}^T \nabla R + R_t &= 0
\end{aligned}$$
(3.16)

where $\nabla L = (L_x, L_y, L_z)$ and $\nabla R = (R_x, R_y, R_z)$ denote the spatial gradients associated with the corresponding constraint planes in the left and right eye, respectively.

Bayesian Vector Normal (BVN) Model

The velocity constraint planes may be noisy due to microsaccades and neural encoding. If we make the simplifying assumption that spatial derivatives of the constraint planes are precise but temporal derivatives have additive noise then

$$\begin{aligned}
L_t &= L_t(x, y, z, t) + \eta(x, y, z, t), \\
R_t &= R_t(x, y, z, t) + \eta(x, y, z, t)
\end{aligned}$$
(3.17)

where $\eta(x, y, z, t)$ has Gaussian density with zero mean and variance $\sigma_v^2 I$, or $\mathcal{N}(0, \sigma_v^2 I)$ for short (**I** is the 3 x 3 identity matrix). Given the gradient constraint equation holds, that is the intensity of the line or edge does not change with position in 3D it follows that

$$\begin{aligned}
\tilde{L}_t &\sim \mathcal{N}(-\mathbf{v}^T \nabla L, \sigma_v^2 I), \\
\tilde{R}_t &\sim \mathcal{N}(-\mathbf{v}^T \nabla R, \sigma_v^2 I)
\end{aligned}$$
(3.18)

If temporal noise is small and viewing distance large then adding noise to the temporal gradient approximates uncertainty of 3D line motion inside a local aperture. If 3D velocity $\mathbf{v}$ is known then the probability of observing $(\nabla L, \tilde{L}_t)$ for the left and $(\nabla R, \tilde{R}_t)$ for the right eye can be expressed as

$$
\begin{aligned}
p(\nabla L, \tilde{L}_t \mid \mathbf{v}) &= \frac{1}{\sqrt{2\pi}\sigma_v} \exp\!\left(\frac{-(\mathbf{v}^T \nabla L + \tilde{L}_t)^T (\mathbf{v}^T \nabla L + \tilde{L}_t)}{2\sigma_v^2} \right), \\
p(\nabla R, \tilde{R}_t \mid \mathbf{v}) &= \frac{1}{\sqrt{2\pi}\sigma_v} \exp\!\left(\frac{-(\mathbf{v}^T \nabla R + \tilde{R}_t)^T (\mathbf{v}^T \nabla R + \tilde{R}_t)}{2\sigma_v^2} \right)
\end{aligned}
\tag{3.19}
$$

If the constraint planes through the left and right eye are not coincident or parallel their intersection constrains 3D velocities but does not provide a unique solution or local velocity estimate. As a consequence additional constraints are needed to disambiguate local velocity estimates. It seems plausible that disparity information from feature tracking, together with other depth cues, helps to disambiguate 3D motion perception (Lages & Heron, 2008; 2010). However, if disparity features and similar depth information is unavailable a weak prior for 3D motion resolves ambiguity and provides a local default estimate.

In their influential paper on 2D motion illusions Weiss et al. (2002) suggested an 2D Gaussian motion prior for slow motion perception in x-y space. Similarly, Lages (2006) introduced a bivariate Gaussian motion prior on the x-z axes to explain bias in perceived azimuth and speed of a target moving on a horizontal plane. If we assume that most features and objects in a scene are stationary or tend to move slowly on an arbitrary trajectory in 3D space then a symmetric 3D Gaussian provides a plausible world prior for binocular 3D motion perception.

Here we propose the 3D Gaussian as conjugate motion prior

$$
p(\mathbf{v}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\!\left(\frac{-\mathbf{v}^T \mathbf{v}}{2\sigma^2} \right)
\tag{3.20}
$$

253

This world prior simply reflects a preference for slow motion in every direction. This is a plausible assumption as most features in natural scenes remain static and moving objects tend to move slowly.

Similar to Equation 3.1, the posterior distribution is the result of combining likelihood constraints and prior using Bayes' Rule where $L(\mathbf{x}, t)$ and $R(\mathbf{x}, t)$ describe intensities in world co-ordinates associated with the left and right eye. The denominator can be dropped because the expression is independent of $\mathbf{v}$ and scales the posterior by a constant factor.

$$p(\mathbf{v} \mid L(\mathbf{x}, t), R(\mathbf{x}, t)) \propto p(L(\mathbf{x}, t) \mid \mathbf{v}) p(R(\mathbf{x}, t) \mid \mathbf{v}) p(\mathbf{v}) \tag{3.21}$$

We can then approximate the posterior by replacing the intensities through gradients.

$$p(\mathbf{v} \mid (\nabla L, \tilde{L}_t), (\nabla R, \tilde{R}_t)) \propto p(\nabla L, \tilde{L}_t \mid \mathbf{v}) p(\nabla R, \tilde{R}_t \mid \mathbf{v}) p(\mathbf{v}) \tag{3.22}$$

The posterior distribution gives a random variable as an estimate. In order to find the most probable velocity or MAP estimate, we take the negative logarithm of the posterior, differentiate it with respect to $\mathbf{v}$ and set the derivative equal to zero.

The logarithm of the posterior is quadratic in $\mathbf{v}$ so that the solution can be written in closed form using standard linear algebra.

$$-\hat{\mathbf{v}} = \left(\mathbf{M} + \frac{1}{\sigma^2} \mathbf{I} \right)^{-1} \mathbf{b} \tag{3.23}$$

where $\mathbf{I}$ is again the 3 by 3 identity matrix, $\mathbf{M} = 1/\sigma_v^2 \left(\nabla L \nabla L^T + \nabla R \nabla R^T \right)$, and $\mathbf{b} = 1/\sigma_v^2 \left(\nabla L \tilde{L}_t + \nabla R \tilde{R}_t \right)$. If the matrix in Eq (3.23) has full rank and is invertible then the solution is unique and a velocity estimate can be determined.

Chapter 5

Talairach Co-ordinates of Regions of Interest

Observer MKE22

Right Hemisphere hMT+/V5 : X: 48, Y: -64, Z: -1

Left Hemisphere hMT+/V5: X: -38, Y: -72, Z: -2

Right Hemisphere Mt Satellite: X: 59, Y: -41, Z: 18

Right Hemisphere V1: X: 13, Y: -98, Z: -3

Left Hemisphere V1: X: -14, Y: -97, Z: -11

Observer KDS11

Right Hemisphere hMT+/V5: X: 39, Y: -61, Z: 5

Left Hemisphere hMT+/V5: X: -40, Y:-61, Z: -1

Right Hemisphere V1: X: 9, Y: -96, Z: -3

Left Hemisphere V1: X: -16, Y: -96, Z: -1

Observer BRY25

Right Hemisphere hMT+/V5: X: 40, Y: -58, Z: 1

Left Hemisphere hMT+/V5: X: -39, Y: -59, Z: -1

Right Hemisphere V1: X: 24, Y: -94, Z: 0

Left Hemisphere V1: X: -15, Y: -100, Z: 1

Observer LWA26

Right Hemisphere hMT+/V5: X: X: 41, Y: -59, Z: 8

Left Hemisphere hMT+/V5: X: -40.5, Y: -60.6, Z: -3.8

Right Hemisphere V1: X: 18, Y: -83, Z: 8

Left Hemisphere V1: X: -9, Y:-99; Z:-5

Observer AME14

Right Hemisphere hMT+/V5: X: 39, Y: -60, Z, 0

Left Hemisphere hMT+/V5: X: -44, Y:-63, Z: 3

Right Hemisphere V1: X: 11, Y: -99, Z: 1

Left Hemisphere V1: X: -13, Y: -97, Z: 11

Observer FFN20

Right Hemisphere hMT+/V5: X: -59, Y: -38, Z: 20

Left Hemisphere hMT+/V5: X: X: -59, Y: -38, Z: 20

Right Hemisphere V1: X: 13, Y: -96, Z: 12

Left Hemisphere V1: X: -10, Y: -98, Z: 8

fMRI Output hMT+/V5 Left and Right Hemisphere (6 Subjects)

Contrasts in order: 3D>2D, Looming>rot, Looming>expa, Looming>translation, mapin>mapout, 4motion>2map/static

Observer KDS11 hMT+ Right Hemisphere

Contrast / LF	value	se	*t*	*p*
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 - 1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0	1.283	0.896	1.431	0.152427

Matrix				
0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0				
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1.137	0.640	1.775	0.075969
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 - 1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1.502	0.633	2.373	0.017711
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-0.219	0.638	-0.343	0.731455
0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0	0.011	0.892	0.012	0.990289
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 - 4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0	55.518	4.854	11.438	0.000000

Observer KDS11 hMT+ Left Hemisphere

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-0.369	0.937	-0.394	0.693417
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1.731	0.664	2.605	0.009245
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.886	0.663	1.336	0.181582
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-1.256	0.664	-1.892	0.058640
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0	0.569	0.937	0.607	0.543840
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0	42.649	5.052	8.441	0.000000

Observer FFN20 hMT+ Right Hemisphere

Contrast / LF	value	se	*t*	*p*
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 - 1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	4.711	0.771	6.108	0.000000
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1.248	0.551	2.267	0.023480
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 - 1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2.912	0.547	5.321	0.000000
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1.799	0.546	3.295	0.000997
0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0	0.382	0.781	0.490	0.624520
0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 - 4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0	46.562	4.151	11.216	0.000000

260

0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0				

Observer FFN20 hMT+ Left Hemisphere

Contrast / LF	value	se	*t*	*p*
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1.892	0.898	2.106	0.035267
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.865	0.641	1.350	0.177006
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1.474	0.637	2.313	0.020823
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.418	0.636	0.658	0.510630
0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0	-0.635	0.908	-0.699	0.484867

	value	se	t	p
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0	19.082	4.834	3.947	0.000081

Observer AME14 MT Right Hemisphere

Contrast / LF	value	se	*t*	*p*
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.202	0.779	0.259	0.796031
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-0.000	0.559	-0.000	1.000000
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.523	0.553	0.945	0.344814
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-0.321	0.553	-0.580	0.561790

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0				
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0	-0.697	0.790	-0.881	0.378149
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0	44.980	4.220	10.659	0.000000

Observer AME14 MT Left Hemisphere

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1.114	0.857	1.300	0.193790
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.141	0.613	0.230	0.818361
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.930	0.608	1.530	0.126109

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.184	0.608	0.303	0.762207
0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0	0.075	0.867	0.086	0.931414
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0	46.103	4.636	9.944	0.000000

Observer BRY25 MT Right Hemisphere

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2.748	0.856	3.209	0.001349
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.978	0.608	1.607	0.108101
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2.281	0.607	3.760	0.000174

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.467	0.605	0.772	0.440423
0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0	-1.545	0.858	-1.801	0.071788
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0	19.142	4.612	4.150	0.000034

Observer BRY25 MT Left Hemisphere

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	3.497	0.824	4.243	0.000023
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.052	0.587	0.089	0.928730
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2.529	0.584	4.330	0.000016

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.968	0.582	1.663	0.096413
0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0	-0.339	0.826	-0.410	0.681626
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0	26.460	4.435	5.966	0.000000

Observer MKE22 hMT+ Right Hemisphere

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2.921	0.869	3.360	0.000791
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1.409	0.621	2.269	0.023347
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2.212	0.616	3.591	0.000336

266

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.709	0.615	1.152	0.249364
0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0	1.058	0.879	1.203	0.229140
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0	36.505	4.685	7.793	0.000000

Observer MKE22 MT SAT Right Hemisphere

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	5.268	0.938	5.615	0.000000
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2.084	0.668	3.119	0.001833
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	3.943	0.665	5.931	0.000000

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1.325	0.664	1.997	0.045945
0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0	1.796	0.946	1.899	0.057700
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0	36.757	5.060	7.265	0.000000

fMRI Contrasts V1 Left and Right Hemisphere (6 Subjects)

Observer KDS11 V1 Right Hemisphere

Contrast / LF	value	se	*t*	*p*
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.679	0.970	0.700	0.483742
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1.635	0.683	2.394	0.016734
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1.284	0.687	1.868	0.061907

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0				
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-0.604	0.684	-0.884	0.376878
0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0	3.084	0.974	3.167	0.001561
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0	6.242	5.213	1.197	0.231300

Observer KDS11 V1 Left Hemisphere

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.758	0.935	0.810	0.418051
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.399	0.659	0.607	0.544164

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0				
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.755	0.663	1.139	0.254640
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.002	0.659	0.003	1.000000
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0	3.346	0.940	3.561	0.000377
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0	29.881	5.028	5.943	0.000000

Observer BRY25 V1 Right Hemisphere

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-0.130	1.016	-0.128	0.897962

270

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0				
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.179	0.721	0.249	0.803707
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.012	0.719	0.017	0.986266
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-0.142	0.717	-0.198	0.842936
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0	2.139	1.016	2.105	0.035408
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0	-7.543	5.473	-1.378	0.168264

Observer BRY25 V1 Left Hemisphere

Contrast / LF	value	se	t	p

0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-0.968	1.019	-0.950	0.342144
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.682	0.723	0.943	0.345898
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-0.362	0.722	-0.502	0.615662
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-0.606	0.720	-0.841	0.400224
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0	2.482	1.019	2.435	0.014982
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0	-6.412	5.490	-1.168	0.242910

Observer AME14 V1 Right Hemisphere

272

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.839	0.912	0.919	0.357921
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-1.366	0.651	-2.098	0.035967
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.064	0.647	0.098	0.921829
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.775	0.647	1.199	0.230577
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0	6.762	0.921	7.340	0.000000
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0	21.391	4.934	4.336	0.000015

Observer AME14 V1 Left Hemisphere

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.581	0.937	0.620	0.535456
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-1.393	0.668	-2.087	0.036986
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-0.190	0.664	-0.286	0.774703
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.771	0.664	1.161	0.245591
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0	6.235	0.945	6.598	0.000000
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0	22.606	5.067	4.461	0.000009

274

Observer FFN20 V1 Right Hemisphere

Contrast / LF	value	se	*t*	*p*
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	2.188	0.875	2.502	0.012428
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	-0.519	0.624	-0.832	0.405303
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1.035	0.620	1.669	0.095330
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	1.152	0.619	1.862	0.062717
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0	5.279	0.884	5.969	0.000000
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0	19.475	4.708	4.137	0.000036

275

Contrast / LF	value	se	t	p
0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0				

Observer FFN20 V1 Left Hemisphere

Contrast / LF	value	se	*t*	*p*
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.247	0.886	0.279	0.780535
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.027	0.632	0.042	0.966368
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.017	0.628	0.027	0.978287
0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0.230	0.627	0.366	0.714242

276

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0			
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1 -1 -1 -1 0 0 0 0 0 0 0 0	5.564	0.896	6.210	0.000000
0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 3 3 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -4 -4 -4 -4 0 0 0 0 0 0 0 0	17.674	4.768	3.707	0.000215

Chapter 6

Web of knowledge/ web of science: Search Strategy

Search term: binocular vision

Year: 2000>>>2008

Journals articles: Vision Research, Journal of Vision, Perception, Visual Neuroscience

Number of articles found: 442

Excluding: any journal concerning: non human subjects, clinical populations, binocular rivalry and ocular dominance

Total Number of Articles: 279

Articles Dealt With (Endnote List)

{Banks, 2001 #429}; {van Ee, 2002 #388}; {Krauskopf, 2002 #389}; {Brooks, 2002 #390}; {Backus, 2003 #298}; {Berends, 2003 #299}; {Both, 2003 #300}; {Boyaci, 2003 #301}; {Hol, 2003 #316}; {Likova, 2003 #322}; {Sheliga, 2003 #336}; {Trommershauser, 2003 #339}; {Welchman, 2003 #346}; {Yang, 2003 #349}; {Adams, 2004 #248};{Boyaci, 2004 #250}; {Brooks, 2004 #251}; {Farell, 2004 #254}; {Farell, 2004 #255}; {Farell, 2004 #256}; {Glennerster, 2004 #260}; {Graf, 2004 #262}; {Hillis, 2004 #269}; {McKee, 2004 #277}; {Obein, 2004 #283}; {Berends, 2005 #196}; {Brascamp, 2005 #197}; {Burge, 2005 #198}; {Coubard, 2005 #199}; {Hartung, 2005 #212}; {Knill, 2005 #216}; {Liu, 2005 #217}; {Mamassian, 2005 #218}; {McKee, 2005 #222}; {Norcia, 2005 #224}; {Norcia, 2005 #225}; {Paffen, 2005 #229}; {Read, 2005 #234}; {Read, 2005 #235}; {Watt, 2005 #245}; {Brooks, 2006 #117}; {Brooks, 2006 #118}; {Brooks, 2006 #119}; {Harris, 2006 #140}; {Lages, 2006 #155}; {Meese, 2006 #162}; {Petrov, 2006 #175}; {Rauschecker, 2006 #178}; {Saunders, 2006 #180}; {Sohn, 2006 #186}; {Erkelens, 2007 #61}; {Muller, 2007 #76}; {McKee, 2007 #105}; {Brooks, 2007 #106}; {Schreiber, 2008 #17}; {Hoffman, 2008 #20}; {Wendt, 2008 #21}; {Tsirlin, 2008 #25}; {Sayim, 2008 #39}; {Blohm, 2008 #53}; {He, 2000 #439}; {Ninio, 2000 #444}; {Norman, 2000 #445}; {Ono, 2000 #446}; {Bravo, 2001 #395}; {Brooks, 2001 #397}; {Grove, 2001 #404}; {Wang, 2001 #427}; {Bacon, 2002 #353}; {Berends, 2002 #355}; {Morikawa, 2002 #370}; {Palmisano, 2002 #373}; {Poom, 2002 #376}; {Viswanathan, 2002 #384}; {Ebenholtz, 2003 #305}; {Grove, 2003 #311}; {Ichikawa, 2003 #318}; {Petrov, 2003 #330}; {Rose, 2003 #333}; {van Ee, 2003 #343}; {Hogervorst, 2004 #270}; {Kennedy, 2004 #273}; {Lewist, 2004 #274}; {Papathomas, 2004 #285}; {van Dam, 2004 #291}; {Gheorghiu, 2005 #204}; {Gonzalez, 2005 #207}; {Grove, 2005 #210}; {Mitsudo, 2005 #223}; {Norman, 2005 #226}; {Ono, 2005 #227}; {Poom, 2005 #233}; {Simmons, 2005 #240}; {Wilcox, 2005 #246}; {Calabro, 2006 #123}; {Lee, 2006 #157}; {Norman, 2006 #171}; {Vuong, 2006 #193}; {Gillam, 2007 #63}; {Matsumiya, 2007 #72}; {O'Kane, 2007 #78}; {Taya, 2007 #88}; {Wilcox, 2007 #94}; {Mapp, 2007 #98}; {Ono, 2007 #103}; {Sachtler, 2007 #109}; {Tassinari, 2008 #12}; {Bertamini, 2008 #23}; {Hammad, 2008 #24}; {Grove, 2008 #43}; {van Bogaert, 2008 #44}; {Wilmer, 2008 #45}; {Allison, 2000 #432}; {Bowd, 2000 #433}; {Bradshaw, 2000 #434}; {Brooks, 2000 #436}; {Gray, 2000 #438}; {Hogervorst, 2000 #440}; {Morgan, 2000 #443}; {Patterson, 2000 #447}; {Popple, 2000 #449}; {Prince, 2000 #450}; {Read, 2000 #452}; {Reich, 2000 #453}; {Schrauf, 2000 #455}; {Shioiri, 2000 #456}; {Stevenson, 2000 #458}; {Taroyan, 2000 #459}; {Tsai, 2000 #460}; {van Ee, 2000 #461}; {Wilcox, 2000 #463}; {Yin, 2000 #464}; {Ziegler, 2000 #465}; {Ziegler, 2000 #466}; {Berends, 2001 #393}; {Berends, 2001 #394}; {Brenner, 2001 #396}; {Durgin, 2001 #398}; {Erkelens, 2001 #399}; {Fredenburg, 2001 #400}; {Fukusima, 2001 #402}; {Gorea, 2001 #403}; {Hakkinen, 2001 #405}; {Harris, 2001 #406}; {Hillis, 2001 #407}; {Jaschinski, 2001 #408}; {Masson, 2001 #410}; {Maxwell, 2001 #411}; {Poom, 2001 #417}; {Sato, 2001 #420}; {Sato, 2001 #421}; {Schor, 2001 #422}; {Shorter, 2001 #424}; {Tyler, 2001 #426}; {Zlatkova, 2001 #428}; {Adams, 2002 #352}; {Banks, 2002 #354}; {Chen, 2002 #358};{Forte, 2002 #363}; {Grove, 2002 #365}; {Masson, 2002 #368}; {McKee, 2002 #369}; {Ono, 2002 #372}; {Petrov, 2002 #374}; {Regan, 2002 #377}; {Simmons, 2002 #380}; {Sumnall, 2002 #381}; {Vreven, 2002 #386}; {Yang, 2002 #387}; {Allison, 2003 #297}; {Duke, 2003 #304}; {Farell, 2003 #306}; {Harris, 2003 #312}; {Hayashi, 2003 #314}; {Hess, 2003 #315}; {Howard, 2003 #317}; {Ichikawa, 2003 #319}; {Lages, 2003 #321}; {Likova, 2003 #323}; {Nishina, 2003

#326}; {Oruc, 2003 #327}; {Pardhan, 2003 #328}; {Patel, 2003 #329}; {Pianta, 2003 #331}; {Pianta, 2003 #332}; {Ukwade, 2003 #341}; {Ukwade, 2003 #342}; {van Ee, 2003 #344}; {Yang, 2003 #348}; {Yang, 2003 #350}; {Gheorghiu, 2004 #258}; {Gillam, 2004 #259}; {Gosselin, 2004 #261}; {Gray, 2004 #263}; {Ishii, 2004 #271}; {Maxwell, 2004 #275}; {Muller, 2004 #281}; {Petrov, 2004 #286}; {Petrov, 2004 #287}; {Schlerf, 2004 #288}; {Wallace, 2004 #293}; {Welchman, 2004 #294}; {Yazdanbakhsh, 2004 #295}; {Zhang, 2004 #296}; {Alvarez, 2005 #194}; {Arnold, 2005 #195}; {Duke, 2005 #201}; {Fernandez, 2005 #202}; {Gheorghiu, 2005 #205}; {Gillam, 2005 #206}; {Greenwald, 2005 #208}; {Hibbard, 2005 #213}; {Khuu, 2005 #215}; {Pizlo, 2005 #231}; {Sakai, 2005 #237}; {Bradshaw, 2006 #114}; {Brooks, 2006 #116}; {Buckthought, 2006 #121}; {Domini, 2006 #126}; {Duke, 2006 #127}; {Fernandez, 2006 #129}; {Fischmeister, 2006 #130}; {Freeman, 2006 #131}; {Fukuda, 2006 #132}; {Goutcher, 2006 #133}; {Gray, 2006 #134}; {Greenwood, 2006 #135}; {Griffiths, 2006 #136}; {Grove, 2006 #137}; {Grove, 2006 #139}; {Harris, 2006 #141}; {Hess, 2006 #142}; {Huang, 2006 #145}; {Jaschinski, 2006 #148}; {Khuu, 2006 #151}; {Knapen, 2006 #153}; {Kuroki, 2006 #154}; {Li, 2006 #158}; {Makino, 2006 #160}; {Mitsudo, 2006 #165}; {Nakamura, 2006 #166}; {Nieman, 2006 #169}; {Norman, 2006 #170}; {Okada, 2006 #172}; {Palmisano, 2006 #173}; {Patel, 2006 #174}; {Petrov, 2006 #177}; {Scarfe, 2006 #181}; {Sheliga, 2006 #185}; {Vreven, 2006 #192}; {Backus, 2007 #57}; {Buckthought, 2007 #59}; {Di Luca, 2007 #60}; {Gonzalez, 2007 #64}; {Hess, 2007 #67}; {Lee, 2007 #70}; {Mitsudo, 2007 #75}; {Ono, 2007 #79}; {Poom, 2007 #80}; {Rushton, 2007 #81}; {Sheliga, 2007 #83}; {Shimono, 2007 #84}; {Suryakumar, 2007 #87}; {Treder, 2007 #89}; {Tsui, 2007 #90}; {Wilcox, 2007 #95}; {Jaschinski, 2008 #4}; {Buckthought, 2008 #6}; {Devisme, 2008 #7}; {Bock, 2008 #8}; {Harris, 2008 #9}; {Hess, 2008 #13}; {Takase, 2008 #34}; {Rambold, 2008 #37}; {Georgeson, 2008 #42}; {Norman, 2008 #47};

Article Coding

Articles were coded as follows:

Stimulus Type:

 0 = Static

1= Moving

2= both

Participant Code:

0= Naïve

1= experienced/naïve

 2= Trained

3= Authors (not naïve and highly trained) nb:// if equal numbers of authors and participants then code as mixed (more other participants>code for other participants)

4= Mixed

Stereopsis Test Code:

 0= Not Mentioned

 1= Non-specific

2= Specific

Test Used:

0=Task Specific

1= Standardised Test (commercially available/clinical)

Exclusion Code:

0= Not Mentioned

 1= Not Specified

 2= Specified

Raw Data From Article Coding

Year	Num Part	Stim Type	Part Code	Stereo Test	Test Used (if specific)	Excl Code	Num Exc	Stereo Exc
2000	2	0	1	0	0	0		
2005	5	0	0	0	0	0		
2000	2	1	4	0		0		
2000	3	0	1	0		0		
2000	20	0	0	0		0		
2000	5	2	4	0		0		
2000	20	1	2	0		0		
2000	2	2	0	0		0		
2000	3	0	0	0		0		
2001	3	0	3	0		0		
2001	3	0	0	0		0		
2001	16	0	0	0		0		
2001	2	1	3	0		0		
2001	2	0	0	0		0		
2001	5	0	2	0		0		
2001	2	0	3	0		0		
2001	3	0	3	0		0		
2002	3	0	0	0		0		
2002	4	0	4	0		0		
2002	9	0	4	0		0		
2002	5	2	4	0		0		

2002	3	0	3	0		0	
2002	12	0	0	0		0	
2002	4	1	3	0		0	
2002	12	0	0	0		0	
2002	3	0	0	0		0	
2002	4	2	0	0		0	
2002	3	2	0	0		0	
2002	6	0	0	0		0	
2002	2	0	3	0		0	
2003	3	0	3	0		0	
2003	6	0	1	0		0	
2003	24	0	0	0		0	
2003	4	1	1	0		0	
2003	2	0	3	0		0	
2003	5	1	0	0		0	
2003	4	0	4	0		0	
2003	6	0	0	0		0	
2003	8	0	4	0		0	
2003	3	0	0	0		0	
2003	6	2	4	0		0	
2003	3	0	3	0		0	
2003	5	0	1	0		0	
2003	4	0	4	0		0	
2004	6	0	0	0		0	
2004	6	2	4	0		0	
2004	10	1	0	0		0	

282

2004	5	2	4	0		0	
2004	4	1	2	0		0	
2004	6	0	0	0		0	
2004	3	0	4	0		0	
2004	3	0	2	0		0	
2004	15	0	0	0		0	
2004	12	0	0	0		0	
2005	4	2	0	0		0	
2005	3	0	0	0		0	
2005	6	1	0	0		0	
2005	3	0	4	0		0	
2005	4	0	1	0		0	
2005	4	0	4	0		0	
2005	6	0	3	0		0	
2005	3	0	0	0		0	
2005	4	0	3	0		0	
2005	3	1	3	0		0	
2005	3	2	3	0		0	
2005	3	2	0	0		0	
2006	4	0	2	0		0	
2006	4	1	2	0		0	
2006	8	1	0	0		0	
2006	6	2	0	0		0	
2006	20	1	0	0		0	
2006	8	0	0	0		0	
2006	3	1	0	0		0	

2006	9	1	1	0		0	
2006	3	0	1	0		0	
2006	34	0	0	0		0	
2006	3	2	3	0		0	
2006	4	0	0	0		0	
2006	11	0	0	0		0	
2006	7	1	0	0		0	
2006	5	0	4	0		0	
2006	4	0	0	0		0	
2006	5	0	4	0		0	
2006	4	0	0	0		0	
2006	5	0	0	0		0	
2006	4	1	4	0		0	
2006	3	0	0	0		0	
2006	100	0	0	0		0	
2006	3	0	3	0		0	
2006	5	0	2	0		0	
2006	3	0	2	0		0	
2006	5	0	4	0		0	
2006	4	0	0	0		0	
2006	3	0	3	0		0	
2006	5	1	4	0		0	
2006	4	0	4	0		0	
2007	12	2	0	0		0	
2007	2	1	4	0		0	
2007	6	0	4	0		0	

284

2007	9	0	0	0		0	
2007	25	2	0	0		0	
2007	3	0	3	0		0	
2007	24	0	0	0		0	
2007	6	0	0	0		0	
2007	12	0	0	0		0	
2007	3	2	3	0		0	
2007	17	2	0	0		0	
2007	4	0	2	0		0	
2008	12	2	0	0		0	
2008	6	1	0	0		0	
2008	3	0	0	0		0	
2008	11	0	0	0		0	
2008	4	0	4	0		0	
2008	3	0	4	0		0	
2008	5	0	0	0		0	
2004	4	0	1	0		2	1	0
2000	27	2	4	1		0	
2000	5	0	4	1		0	
2000	5	0	0	1		0	
2000	2	0	1	1		0	
2000	3	0	3	1		0	
2000	3	0	3	1		0	
2001	4	0	4	1		0	
2001	4	0	0	1		0	
2001	8	0	0	1		0	

2001	3	0	3	1	0		
2001	6	0	0	1	0		
2001	3	0	3	1	0		
2002	4	0	4	1	0		
2002	10	2	0	1	0		
2002	2	0	3	1	0		
2003	5	0	4	1	0		
2003	10	0	0	1	0		
2003	5	2	2	1	0		
2003	4	1	4	1	0		
2003	12	0	2	1	0		
2003	9	0	0	1	0		
2003	4	2	4	1	0		
2003	3	0	0	1	0		
2004	5	0	1	1	0		
2004	6	1	2	1	0		
2004	5	0	4	1	0		
2004	3	0	4	1	0		
2004	2	0	1	1	0		
2004	4	0	4	1	0		
2004	5	1	2	1	0		
2004	5	0	4	1	0		
2004	3	0	1	1	0		
2005	13	1	4	1	0		
2005	17	0	0	1	0		
2005	6	0	4	1	0		

Year								
2005	4	0	2	1		0		
2005	4	0	1	1		0		
2006	20	0	4	1		0		
2006	30	0	0	1		0		
2006	5	0	0	1		0		
2006	7	1	0	1		0		
2007	10	1	0	1		0		
2007	3	1	1	1		0		
2007	5	0	3	1		0		
2008	3	0	0	1		0		
2008	6	1	1	1		0		
2004	22	0	0	1		2	1	1
2005	29	0	0	1		2	1	0
2006	19	0	0	1		2	6	6
2000	5	1	4	2	0	0		
2000	4	0	1	2	0	0		
2000	5	2	4	2	0	0		
2000	4	1	4	2	0	0		
2001	4	1	0	2	0	0		
2001	3	1	0	2	0	0		
2001	9	0	0	2	0	0		
2001	3	0	3	2	0	0		
2002	3	1	4	2	0	0		
2002	24	2	0	2	0	0		
2002	7	0	0	2	0	0		
2003	3	0	3	2	0	0		

2003	4	0	4	2	0	0		
2003	4	0	0	2	0	0		
2003	18	0	4	2	0	0		
2003	11	1	4	2	0	0		
2003	30	0	0	2	0	0		
2003	5	0	0	2	0	0		
2004	5	1	4	2	0	0		
2004	4	2	2	2	0	0		
2004	16	0	0	2	0	0		
2004	5	0	0	2	0	0		
2005	3	1	0	2	0	0		
2005	3	1	0	2	0	0		
2005	34	0	0	2	0	0		
2006	6	1	0	2	0	0		
2006	4	1	0	2	0	0		
2006	4	0	0	2	0	0		
2006	11	1	0	2	0	0		
2006	4	0	0	2	0	0		
2007	16	0	2	2	0	0		
2007	6	0	0	2	0	0		
2007	9	0	0	2	0	0		
2007	74	0	0	2	0	0		
2000	59	0	0	2	0	1		
2001	4	0	0	2	0	1		
2002	4	1	1	2	0	1		
2000	67	0	0	2	0	2	10	10

2002	4	0	1	2	0	2	2	2
2002	20	2	0	2	0	2	5	4
2003	15	0	0	2	0	2	7	4
2004	36	1	0	2	0	2	2	0
2005	7	0	0	2	0	2	8	8
2007	5	0	4	2	0	2	2	2
2007	8	0	0	2	0	2	1	1
2008	65	0	0	2	0	2	2	2
2000	6	0	2	2	1	0		
2000	41	0	0	2	1	0		
2000	12	0	0	2	1	0		
2000	11	0	4	2	1	0		
2001	7	1	1	2	1	0		
2001	5	1	3	2	1	0		
2001	3	0	0	2	1	0		
2001	3	0	3	2	1	0		
2002	4	0	4	2	1	0		
2003	5	0	0	2	1	0		
2003	17	1	0	2	1	0		
2003	4	0	4	2	1	0		
2003	3	0	0	2	1	0		
2003	4	0	4	2	1	0		
2003	24	0	0	2	1	0		
2003	4	2	1	2	1	0		
2004	8	0	0	2	1	0		
2004	20	1	1	2	1	0		

2004	3	0	3	2	1	0	
2005	16	0	0	2	1	0	
2005	8	0	0	2	1	0	
2005	6	0	1	2	1	0	
2005	28	0	0	2	1	0	
2005	24	0	0	2	1	0	
2006	6	1	0	2	1	0	
2006	5	0	0	2	1	0	
2006	4	0	0	2	1	0	
2006	8	1	0	2	1	0	
2006	5	0	0	2	1	0	
2006	4	0	0	2	1	0	
2006	5	0	2	2	1	0	
2006	3	0	1	2	1	0	
2007	4	0	0	2	1	0	
2007	37	0	0	2	1	0	
2007	5	0	1	2	1	0	
2007	6	0	0	2	1	0	
2007	22	0	0	2	1	0	
2007	4	1	0	2	1	0	
2008	16	0	0	2	1	0	
2008	7	0	0	2	1	0	
2008	8	0	0	2	1	0	
2008	12	1	1	2	1	0	
2008	2	0	1	2	1	0	
2008	3	0	2	2	1	0	

2008	48	0	0	2	1	0		
2008	5	0	0	2	1	0		
2008	194	0	0	2	1	0		
2001	40	0	0	2	1	1		
2000	6	0	1	2	1	2	2	2
2001	6	1	0	2	1	2	3	0
2003	12	2	2	2	1	2	1	1
2005	10	2	0	2	1	2	1	0
2006	14	0	0	2	1	2	1	1
2008	20	0	0	2	1	2	5	3
2008	23	0	0	2	1	2	4	1

Table of Stereopsis Tests

This table is designed to give a brief overview of the most commonly used tests in terms of the type and range of stereopsis tested, whether they are contour or RDS based and how they are administered. Other tests available tend to be variations on the tests included here and differ only in the image presented.

Name of Test	Contour/Non contour	Type of Stereo	Test Components	Scale Type	Range (seconds of arc)
Titmus Fly test	Contour	Analglyph, polarized	House fly test, Graded circles test (800-40 seconds of arc), animal testing for young children (400-100)	Coarse and fine grain	3500-40
Stereo Butterfly test	Contour	Anaglyph, polarized	Butterfly test (2500-1200 secs of arc), graded circles test (800-400), animal testing for young children	Coarse and fine grain	2500-400
Frisby test	Contour	Natural	Three test plates (6mm, 3mm, 1.5mm thick) featuring a hidden shape which is revealed binocularly	Fine grain	600-15
Lang I stereotest	Non contour: random dots, cylindrical gratings	Natural	Cat, car, star images	Coarse stereopsis in children	1200-550
Lang II stereotest	Non contour: random dots cylindrical gratings	Natural	Moon and star (200 secs of arc), car (400), elephant(600): random dots smaller/sparser than Lang I	Fine grain stereopsis in children	600-200

			293		
Random Dot E	*Non contour*	*Anaglyph, polarized*	*Raised 'E' target, non stereo target which the patient must differenciate between. Distance manipulated to vary secs of arc*	*Coarse and fine grain, most suitable test for children*	*Varied in terms of distance*
Randot Stereo test (inventory	*Non contour and contour based tests*	*Anaglyph, polarized*	*Randot test with 6 geometric shapes and 2 levels (500 and 250 secs of arc), graded circles test (400-20), animal testing for young children (400-100)*	*Coarse and fine grain*	*500-20*
TNO Stereo test	*Non contour*	*Anaglyph*	*A series of random dot based test plates with retinal disparities ranging from 480-15 secs of arc*	*Coarse and fine grain*	*480-15*