Neural mechanisms of binocular motion in depth perception

Milena Kaestner

PhD

University of York

Psychology

May 2018

Abstract

Motion in depth (MID) can be cued by two binocular sources of information. These are changes in retinal disparity over time (changing disparity, CD), and binocular opponent velocity vectors (inter-ocular velocity difference, IOVD). This thesis presents a series of psychophysical and fMRI experiments investigating the neural pathways supporting the perception of CD and IOVD.

The first two experiments investigated how CD and IOVD mechanisms draw on information encoded in the magnocellular, parvocellular and koniocellular pathways. The chromaticity of CD and IOVD-isolating stimuli was manipulated to bias activity in these three pathways. Although all stimulus types and chromaticities supported a MID percept, fMRI revealed an especially dominant koniocellular contribution to the IOVD mechanism.

Because IOVD depends on eye-specific velocity signals, experiment three sought to identify an area in the brain that encodes motion direction and eye of origin information. Classification and multivariate pattern analysis techniques were applied to fMRI data, but no area where both types of information were present simultaneously was identified. Results suggested that IOVD mechanisms inherit eye-specific information from V1.

Finally, experiment four asked whether activity elicited by CD and IOVD stimuli could also be modulated by an attentional task where participants were asked to detect changes in MID or local contrast. fMRI activity was strongly modulated by attentional state, and activity in motion-selective areas was predictive of whether participants correctly identified the change in CD or IOVD MID. This suggests that these areas contain populations of neurons that are crucial for detecting, and behaviourally responding to, both types of MID.

The work presented in this thesis detail a thorough investigation of the neural pathways that underlie the computation of CD and IOVD cues to MID.

List of Contents

List of Tables

List of Figures

Acknowledgements

This project is not only mine; it also belongs to the many (brave!) people who accompanied me on this PhD journey. First and foremost, I would like to thank my supervisor Alex Wade, for the inexhaustible enthusiasm and generosity with which you have shaped my time in your lab. I have enjoyed as many challenges as I have freedoms, and it has been a dynamic and enriching experience. Thank you.

I would also like to thank the members of my thesis advisory panel, Antony Morland and Dan Baker, for the time dedicated to providing feedback that helped shape this research. Thanks also to my collaborators Julie Harris and Marina Bloj, for many stimulating conversations and for those handy six-month reviews, which kept us all on track and have turned out to be exceptionally useful in putting this thesis together.

An extra thank you to Ryan – thank you for being a great colleague and friend, and for all the adventures we had navigating the lands of binocular vision. Thank you also to the (past, present, and visiting) members of WadeLab, in particular to Lauren, Marc, Babs, Kirstie and Anna, and to my very talented friends in YNiC, the Department of Psychology, and CHyM. Working alongside you all has been at least as inspiring as it has been joyful.

Working towards a PhD also requires many a debriefing session, and I am extremely grateful for the friendships that have been a source of so much courage. Special thanks must go to Nicola, Sarah, Lucie, Ruth, Alex L, and Fran, who were there from beginning to end and who are by far the strongest, and kindest, group of people I know. Thanks also to my flatmate (and favourite lentil) Junior, for cups of tea at the end of long days.

Finally, I thank my parents, Ines and Uwe, and my little sister Louisa, for providing the foundation of support and encouragement that allowed me to reach this point.

Author's Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Conference abstracts arising from this thesis:

Kaestner, M., Maloney, R. T., Wade, A. R., Bloj, M., & Harris, J. M. (2017). Attentional state modulates responses to motion-in-depth stimuli across striate and extrastriate visual areas. Contributed talk at OSA, Fall Vision Meeting, Washington D.C., USA.

Kaestner, M., Maloney, R. T., Wade, A. R., Bloj, M., & Harris, J. M. (2017). Attending to motion-in-depth modulates fMRI responses in striate and extrastriate visual areas. Contributed talk at ECVP, Berlin, GER.

Kaestner, M., Maloney, R. T., Wade, A. R., Bloj, M., & Harris, J. M. (2017). fMRI reveals S-cone and achromatic contributions to motion-in-depth perception. Poster presented at VSS, St. Pete's Beach, Florida, USA.

Kaestner, M., Maloney, R. T., Wade, A. R., Bloj, M., & Harris, J. M. (2016). Decoding eye-of-origin signals in and beyond primary visual cortex. Poster presented at ECVP, Barcelona, ES.

Chapter 1.   Introduction

As we move about in the world, the image that falls on the retina also moves. The motion on the retina is a combination of our own motion, and the independent motion of objects around us. To make matters more complicated, the retina is a two-dimensional surface – yet our environment is inherently three-dimensional. As such, a crucial task of the mammalian visual system is to transform the 2D images that fall on the retina, into meaningful signals that convey information about 3D motion (motion in depth; MID). Despite the importance of efficient MID detection, and the saliency of MID signals in the visual world, most research to date has treated the perception of depth as separate from the perception of motion. We have only a partial understanding of how these two visual features are combined.

This thesis aims to investigate the neural pathways that underpin our ability to extract 3D motion estimates. Many cues are available to support this task, some of which are monocular, and some of which depend on a combination of signals from both eyes. Monocular cues include changes in image size (also referred to as image looming), the pattern of motion vectors on the retinal surface (optic flow), and changes in object occlusion. There are also biological cues such as lens accommodation and eye movements as we track the motion of an object. However, the perception of MID is greatly improved by binocular mechanisms that generate a vivid sensation of stereoscopic depth. Indeed, binocular visual impairments such as amblyopia typically result in difficulties perceiving MID. Despite their importance, the binocular mechanisms involved in perceiving MID are poorly understood.

This thesis will focus on two binocular cues to MID. These two cues are based on different early sources of information – binocular disparity, and monocular velocity estimates. MID can be signified by monitoring increases and decreases in binocular disparity over time (the 'changing disparity' cue, CD), or by comparing velocity vectors between the two eyes (the 'inter-ocular velocity difference' cue, IOVD). Although both cues exist in tandem in the environment, they can be dissociated mathematically and experimentally.

By breaking the constituents of binocular MID perception apart, the distinct mechanisms underlying CD and IOVD can be investigated in detail. This approach reveals differences in the neural processing and representation of both cues, relating to the different sources of

information they sample, with the eventual goal of understanding how the two interact and combine to form a generalised, binocular MID signal. Furthermore, the work presented here hopes to expose underlying computational principles that can be extended beyond the realm of binocular vision, to further our understanding of how the nervous system makes crucial inferences about its environment.

This introductory chapter will first outline visual processing and the visual pathways known to be involved in 2D motion perception, before discussing some fundamental principles of binocularity and stereopsis and describing how motion and depth mechanisms can be unified. It will then illustrate the two binocular cues to 3D motion perception that are the focus of this thesis. Psychophysical evidence for the relative contributions of both MID mechanisms, as well as computational differences between the two cues, are described. Evidence from functional magnetic resonance imaging (fMRI) and unit electrophysiology is presented, which provides the first ideas for how these two cues are represented in cortex. Outstanding questions are discussed in the summary, followed by an outline for the presentation of the chapters in this thesis.

1.1　Fundamentals of visual processing: retina to cortex

1.1.1　Retinal physiology

Visual processing begins at the retina, which transforms the 2D image that falls upon it into an efficient neural code (Field & Chichilnisky, 2007). Photons are first absorbed by rod and cone cells, of which there are three types depending on the opsin they express. Each cone is sensitive to either long (L), medium (M) or short (S) wavelengths. Rod cells express only one type of opsin (rhodopsin) and are used primarily in scotopic vision.

After phototransduction, horizontal, bipolar and amacrine cells transform and relay signals from the photoreceptor layers of the retina to the ganglion cells (see **Figure 1.1**). The functional roles of these intermediate cells are diverse (Masland, 2012), and they optimise signals from the photoreceptor cells. A surprising level of computation occurs at these early stages – for example, the dendritic trees of starburst amacrine cells are wired to some types of ganglion cell in such a way that results in direction selectivity (Briggman, Helmstaedter, &

Denk, 2011). The output cells of the retina, retinal ganglion cells (RGCs), thus project a signal to visual cortex that has already been refined in many ways.
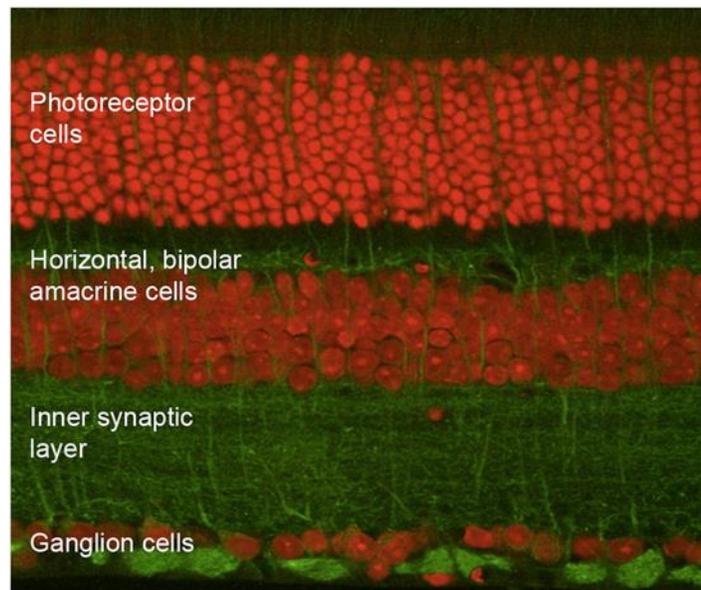
Figure 1.1 *Cross-section of the mammalian retina showing different layers of cell types. After phototransduction at the photoreceptor cells, horizontal, bipolar and amacrine cells transform and relay signals to the ganglion cells. Ganglion cells can be characterized by their projections to the lateral geniculate nucleus (LGN) and, in trichromats, signal the sums and differences of the long, medium and short-wavelength cone cells to form the basis of colour vision. Figure from Masland, 2012.*

By some estimates, there are no less than 20 different ganglion cell types in the mammalian retina (Masland, 2012). These can be classified into three broad groups, depending on the layers of the lateral geniculate nucleus (LGN) their axons project to. Because each of these three groups has characteristic response properties, it can be assumed that they carry signals optimal for different visual tasks.

Ganglion cells projecting to parvocellular layers of the LGN are termed PC cells. Their receptive fields may be colour-opponent, as the centre and surround receive inputs from different types of cone cell. Thus, a PC cell's response is a function of the wavelength and the intensity of a stimulus (De Monasterio & Gouras, 1975). Characteristics of PC cell responses include tonic (sustained) firing throughout the duration of a stimulus, as well as slower conductance due to their relatively thin axons. Signals are high in spatial resolution, due to their small receptive field centres and dendritic trees. A higher density of PC cells is required

to tessellate the retinal surface - the retina contains about eight times as many PC cells than cells projecting to magnocellular (MC) layers of the LGN (Perry, Oehler, & Cowey, 1984).

MC cells show no colour opponency in the excitatory and inhibitory parts of their receptive fields, as these receive inputs from a mixture of cone types. MC cells respond maximally to differences in luminance between the centre and surround across a broad range of wavelengths. In comparison to PC cells, MC cells respond in a phasic manner, and firing rates reduce when the stimulus is sustained. Their axons are thicker, facilitating a faster conductance. MC cells have larger receptive fields and dendritic trees, and the contrast sensitivity of MC cells is about ten times higher than that of PC cells (Kaplan & Shapley, 1986). This means whilst their signals are lower in spatial resolution, they are highly sensitive to transient changes in contrast.

In comparison to PC and MC cells, ganglion cells that project to the koniocellular (KC) layers of the LGN form a much more heterogeneous group (Hendry & Reid, 2000). KC cells show no concentric receptive fields, but instead respond to the onset or offset of light in any part of their receptive field, or to moving spots of light (De Monasterio & Gouras, 1975). Others do show colour opponent responses, but in the absence of an inhibitory surround as centre-only ON or OFF cells (De Monasterio, 1978). KC cells also have lower spatial resolution than MC or PC cells (White, Solomon, & Martin, 2001).

Signals from the PC, MC and KC ganglion cells are strictly segregated in the LGN, and remain so until primary visual cortex (V1). These three primary pathways are discussed below.

1.1.2 *The primary visual pathways: retina to V1*

Pathways from the RGCs to V1 are schematised in **Figure 1.2**. After phototransduction at the retina, visual signals are fed along the optic nerve. At the optic chiasm, fibres in the optic nerve cross over into the optic tract, such that information from the left and right visual hemifields in each eye is projected to the contralateral hemisphere. Throughout this, MC, PC and KC cell outputs remain segregated, and project to separate layers of an intermediate structure, known as the LGN.

In old world primates, the LGN is divided into 6 major layers, 3 of which receive input from the ipsilateral eye, and 3 of which receive input from the contralateral eye. Layers 1 and 2 receive input from MC cells, whereas layers 3-6 are driven by PC cells (Leventhal, Rodieck, & Dreher, 1981). Interspersed between these major layers are thinner stripes of KC cell output (Hendry & Reid, 2000).
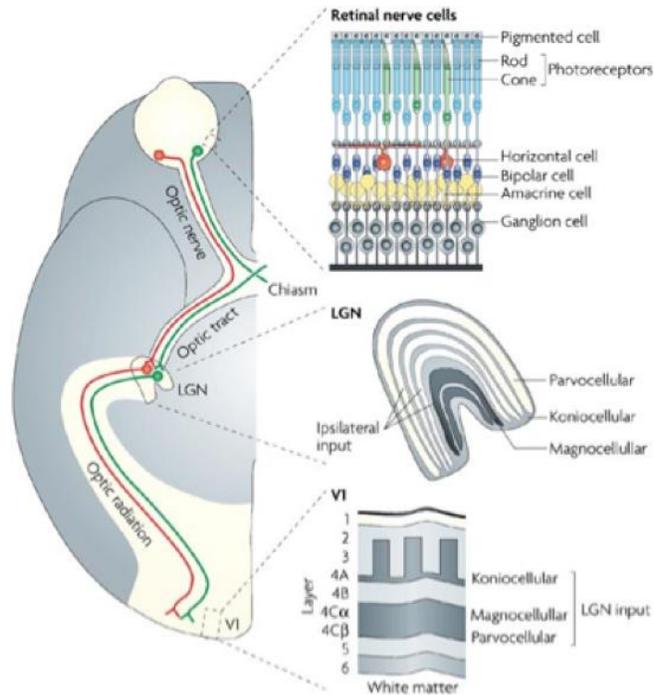
Figure 1.2 *An overview of visual processing from retina to cortex. After phototransduction in the rod and cone cells of the retina, ganglion cells axons form the optic nerve. Outputs from different ganglion cell types are segregated in different layers of the lateral geniculate nucleus (LGN). The magnocellular layers of the LGN project to layer 4Cα of V1, parvocellular (PC) layers of the LGN project into layer 4Cβ of V1, and koniocellular layers project to the blobs in layer 2/3 of V1. This arrangement ensures that the outputs of different ganglion cell classes remain strictly segregated from the retina to early visual cortex. Figure from Solomon & Lennie, 2007.*

Because the characteristics of MC and PC layers in the LGN are inherited from their ganglion cell outputs, activity in the MC and PC pathways can be biased by different stimulus chromaticities (Derrington & Lennie, 1984). As MC cells exhibit no colour opponency, their responses are strongly modulated by achromatic contrast given by the sum of L+M cone outputs. On the other hand, the colour opponent responses of PC cells are most commonly driven by differences between the L and M cone outputs; thus, cells in the PC pathway are strongly driven by red/green contrast (Derrington, Krauskopf, & Lennie, 1984). A minority of PC cells show a more complex configuration where different parts of their receptive fields are

driven by both L and M cone outputs, and a further subset are driven by the S cones – though these are far less common than L-M cone opponent cells (Derrington, Krauskopf, & Lennie, 1984).

Historically, very little attention has been given to the KC pathway. Indeed, the landmark wiring diagram by Felleman and Van Essen of the primate visual system – marvellous though it is – omits its contribution entirely (see section 1.8: Supplementary information, **Figure 1.1**, first published by Felleman & Van Essen, 1991 and reproduced in Silvanto, 2015). This is in part due to the technical difficulty of recording directly from the small, 'dustlike' cells (Hendry & Reid, 2000). However, more recently, these technical difficulties have been overcome to reveal some of the characteristics of this evolutionarily ancient pathway.

As previously stated, KC ganglion cells form a far more heterogeneous group, and this heterogeneity is further represented by the characteristics of the LGN KC layers. Dorsal KC layers have larger receptive field sizes even than MC cells at the same eccentricity (Xu Xiangmin et al., 2004), and the characteristic response of the majority of KC cells is spatially broadband. Cells in the central two layers of macaque LGN are driven especially by signals from the S cones, and show mostly blue-ON or yellow-OFF responses (Hendry & Reid, 2000). Thus, KC pathways can be driven by an S cone stimulus, though some KC layers in the LGN will also respond to achromatic contrast.

Some additional tuning properties of KC cells in the LGN are noteworthy. Whilst MC and PC cells receive monocular input only, some neurons in the KC layers of the LGN signal binocular responses (Zeater, Cheong, Solomon, Dreher, & Martin, 2015). Others also show orientation tuning, although these cells receive negligible input from the S cones (Cheong, Tailby, Solomon, & Martin, 2013). Selectivity for directional drift in monochromatic gratings has been observed in S-cone driven KC cells, though this is unlikely to drive directional responses of neurons in extrastriate areas (Tailby et al., 2010). These additional properties indicate that the role of KC pathways in more specialised processing streams may be underappreciated.

The segregation of MC, PC and KC cells through the LGN is maintained to V1. MC cells project primarily to layer 4Cα, whereas PC cells project to layer 4Cβ. The KC pathway projects to layers 4A and to the blobs in layer 2/3 (see **Figure 1.3**). Within V1, these signals interleave at the first and second intracortical synapses, meaning that segregation of the three early pathways becomes less clear beyond V1. For example, pyramidal neurons in layer 4B receive inputs from both layers 4Cα and 4Cβ (Sincich & Horton, 2005).
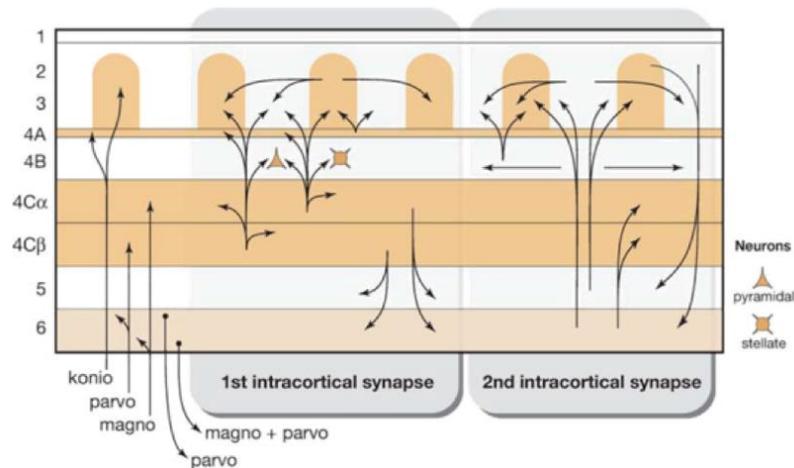
Figure 1.3 Magnocellular (MC), parvocellular (PC), and koniocellular (KC) signals in V1. The initial inputs from the MC, PC and KC layers of the lateral geniculate nucleus (LGN) are shown at the left of the figure. The KC (konio) pathway projects into layer 4A and the blobs in layer 2/3. The PC (parvo) pathways projects mainly into layer 4Cβ, whereas the MC (magno) pathway projects mainly into layer 4Cα. At the 1st and 2nd intracortical synapse, these signals continue to interleave and diffuse though several layers. Examples of pyramidal and stellate neurons in layer 4B show how signals from different pathways combine within the same type of neural population. Figure from Sincich & Horton, 2005.

In general, the different receptive field properties of MC, PC and KC cells in the LGN imply several parallel pathways that may be suited to performing different computations. A stereoscopic depth mechanism that relies on fine spatial resolution would benefit from signals carried in the PC pathway particularly. The transient, fast signalling of the MC pathway suits motion and velocity computation. The KC pathway may also contribute to these with additional properties of binocularity, and selectivity for direction and orientation.

1.2 Binocularity and stereopsis

A fundamental task of the visual system is to transform the two-dimensional images that fall onto our retinae into a seamless, three-dimensional view of the world. In doing so, it faces the challenge that our two eyes see two slightly different views of the world. The image that falls on the left retina is horizontally shifted in comparison to the one that falls on the right retina. This horizontal shift is known as binocular disparity, and provides one of the key cues to depth that allows the transformation to 3D perception to take place.
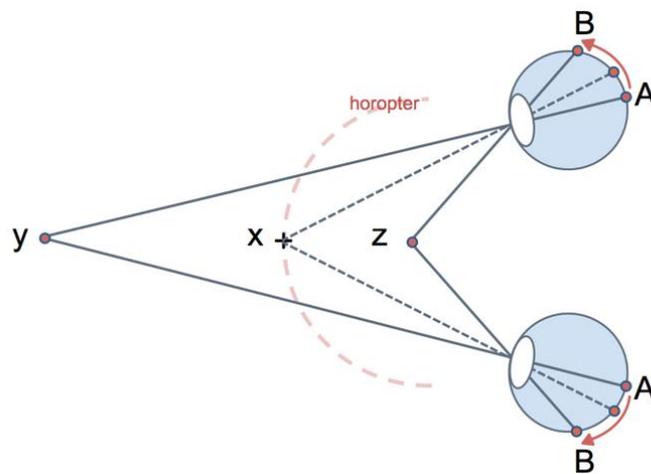
Figure 1.4 Basic principles of binocular vision. When fixating at point x, the images of point y and point z fall on different points on the retina in the left and right eyes, at points A and B. The horizontal offset between point A across both eyes, and point B across both eyes, is known as the binocular disparity. As an object moves from position y to position z, its binocular disparity increases. Furthermore, the image that falls on the retina within one eye travels from point A to point B. These two cues form the basis of binocular MID mechanisms.

Figure 1.4 illustrates this basic principle, referring to the two sources of information that binocular MID perception depends on: binocular disparity, and motion signals on the retinae. When we fixate on a particular object, at point *x*, the image of that object is focused on the fovea of the retina – thus, it stimulates corresponding points in the left and right eye. The plane of fixation around point *x* is known as the horopter, and is the point of zero disparity between the left and right images on the retina.

Point *y*, however, is beyond the horopter and thus stimulates different parts of the left and right retinae. This is a point at uncrossed disparity. Point *z* is within the horopter, at

crossed disparity. Because the relative disparity of images on the retinae depends on their distance from the horopter, this disparity implies the depth of an object relative to the observer. This is the fundamental mechanism of stereoscopic depth perception. The visual system performs a low-level cross-correlation between retinal inputs to match corresponding points in the left and right eye images (Nishihara, 1984), allowing it to estimate depth from the disparity cue.

If static binocular disparity allows us to perceive depth, dynamic disparity allows us to monitor the motion of objects towards and away from us. Referring back to **Figure 1.4**, it is clear that the binocular disparity of an object at point *y* increases as it moves towards point *z*. These changes in the retinal disparity of an object as it moves through space provide binocular cues to perceiving MID. One of these cues depends on the change in binocular disparity over time (disparity between points A and B in **Figure 1.4**).

Another possible cue depends on comparing the monocular motion over time of an image between the eyes (the arrows between points A and B in **Figure 1.4**). Note that, in this example, because motion is directly towards and away from the eyes, the motion vectors on the retinae are pointing outwards (away from the nose) in opponent directions. This is the basis of the velocity cue to MID.

This diagram illustrates a simplified version of how disparity and retinal motion cues occur in tandem. In the real world, the situation is more complex as the motion vectors on the retina differ across the whole retinal surface. For example, consider the instance when motion is moving directly towards one eye. Here, the direction of the motion vectors would be outwards (away from the nose) on the far, temporal portion of the retina, but it would be inwards (towards the nose) on the nasal portion of the retina. There would also be motion upwards and downwards in the top and bottom halves of the retina. Furthermore, the magnitude of the motion vector would also depend on the distance from the fovea. This is illustrated in **Figure 1.5**, and was recently described as the 'binoptic flow field' (BFF; Cormack, Czuba, Knöll, & Huk, 2017) – an extension of the well-known optic flow cue to include binocular mechanisms. Note that in stimuli based on this naturalistic representation, it is impossible to tease apart the combined cues of binocular disparity, retinal motion, and optic flow.
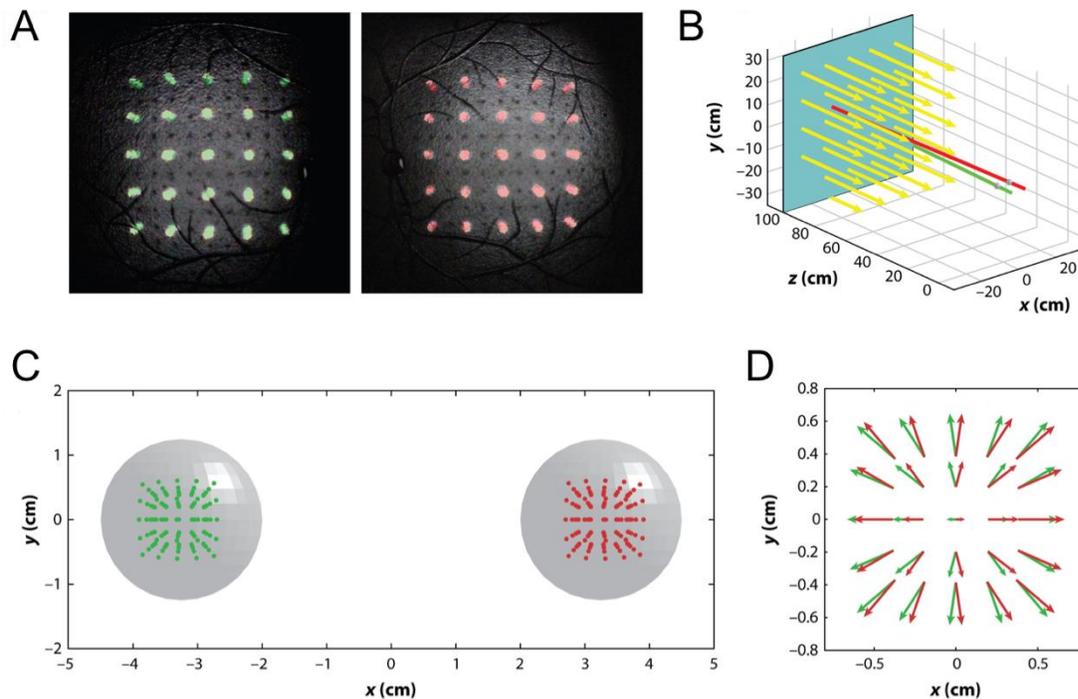
Figure 1.5 *A computational model of the binocular flow field (BFF), illustrating realistic motion vectors generated in the left (green arrows) and right (red arrows) eyes when MID is directly towards the observer. Panel A shows a depiction of the BFF as stimulated points in the back of the left and right eyes, assuming fixation on the centre of the stimulus. Panel B shows the position of the eyes within the field of motion modelled in A, C and D. This type of motion could be induced by forward self motion. Panel C shows the point in A projected on two eyeballs, set a realistic distance apart. The visual system faces the task of combining motion vectors in the left and right eyes into a single binocular percept, by aligning corresponding retinal points. This is depicted in panel D, showing slight differences in motion vectors on corresponding points of the left and right retinae. These differences generate a cue to MID. Motion, disparity and optic flow cues are combined in this simulation. Figure adapted from Cormack et al., 2017.*

1.3 Stereograms and kinematograms: the separate study of depth and motion

In the past, the perception of disparity has largely been considered separately from the perception of motion, perhaps due to the clever isolation of stereoscopic depth and motion cues in random-dot stimuli. Our understanding of the stereoscopic depth mechanism, and the correspondence problem for how the visual system combines information from the two eyes, was revolutionised by the development of the random dot stereogram (RDS) by Bela Julesz in 1971. An example of the RDS is shown in **Figure 1.6**.

Figure 1.6 *An example of one of Bela Julesz's random dot stereograms. Images containing white noise are shown to the left and right eye. No structure is discernible in each monocular half-image, but after binocular fusion these types of stimuli convey shape information. Try for yourself by fixating on the central cross and slightly crossing the eyes, fusing the two images into one. The white noise patterns are correlated between the eyes – and there is a small subset of elements in the centre where the dots are shifted horizontally with respect to one another. Binocular disparities can be extracted by the visual system on the basis of such low-level cross-correlations, yielding a depth percept. Image adapted from Julesz, 1971.*

The discovery of this type of stimulus was remarkable, as it contains no apparent structure, and yet is able to convey a striking and convincing sense of depth. By introducing small horizontal offsets within a white noise pattern shown to the left and right eyes, Julesz demonstrated that disparity processing likely occurs early on in the visual system at the site of binocular combination, before the extraction of identifiable edges or contours.

This led to the assumption that, as disparity mechanisms can operate on low-level texture primitives, they likely *only* depend on this kind of information. As such, disparity processing was considered a module in its own right, and was considered independently from other visual functions such as motion processing (Cormack et al., 2017).

Likewise, motion perception has been considered as an independent module with the introduction of another elegant stimulus – the random dot kinematogram (RDK). This was essentially one of Julesz's RDSs presented monocularly, where a portion of the dots was displaced over time (Anstis, 1970). Later, sparse dot fields were introduced where parameters such as dot lifetime and the proportion of moving dots to static dots (signal to noise) could be manipulated independently to test the perceptual limits of motion perception (Braddick, 1974; Morgan & Ward, 1980; Williams & Sekuler, 1984). These manipulations allowed crucial inferences to be made about motion processing in the brain, for example by correlating the signal to noise ratio in a random-dot stimulus against the firing rate of

individual neurons (Britten, Shadlen, Newsome, & Movshon, 1992, 1993; Newsome & Pare, 1988).

These approaches were essential in developing an understanding of motion and depth perception. However, the treatment of these two systems as separate from one another naturally lead to the viewpoint that motion processing was inherently two-dimensional, taking place in various disparity planes so that motion signals across discrete depth planes could be combined to form 3D motion percepts. This viewpoint neglects the manner in which depth and motion signals are inherently linked in the natural world, and seems to propose a rather clunky and inefficient way of extracting the crucial MID signal. Furthermore, it misses the contribution from other binocular MID mechanisms. These are discussed below.

1.4   Binocular cues to motion in depth perception

1.4.1   Changing disparity

One cue to binocular MID perception relies on fine changes in binocular disparity, introduced in section 1.2: Binocularity and stereopsis. As an object moves towards and away from an observer, the image that falls on the left retina is horizontally displaced in comparison to the one that falls on the right retina. If the motion is coming towards the observer, the retinal disparity increases. If the motion is receding away from the observer, retinal disparity decreases. The absolute magnitude of retinal disparity provides a strong cue to how far the object is away from the observer. Thus, monitoring the rate and extent of retinal disparity changes generates an estimate of the trajectory of object MID (see **Figure 1.7**).
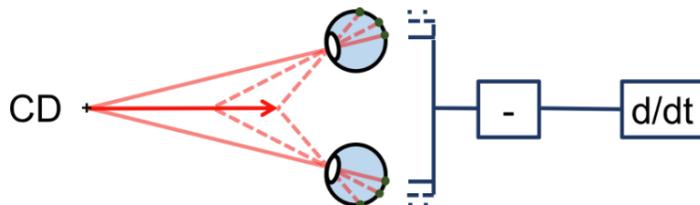
Figure 1.7 *The CD cue to MID. When an object moves towards an observer, the binocular disparity between stimulated points on the retinae increases systematically, providing an estimate of MID direction over time.*

This binocular cue is known as the 'changing disparity' (CD), and its computation is expressed as

$$\frac{d(r-l)}{dt}$$

where *r* and *l* are the retinal locations of each image in the left and right eyes, *dt* is the temporal derivative, and *(r – l)* is the horizontal disparity between left and right retinal images (Rushbass & Westheimer, 1961).

1.4.2   Inter-ocular velocity difference

Simply rearranging the above equation results in another potential cue to binocular MID perception. First proposed by Rushbass and Westheimer in 1961, this cue is mathematically equivalent to the CD cue and can be expressed as

$$\frac{dr}{dt} - \frac{dl}{dt}$$

where *r* and *l* are the retinal locations of each image in the left and right eyes, and *dt* is the temporal derivative (Rushbass & Westheimer, 1961).

This expression implies that MID can be estimated by comparing image velocities on the left and right retinae. As an object moves directly towards and away from an observer, the image that falls on the left retina moves in equal and opposite direction to that on the right retina (see **Figure 1.8**). Comparing these monocular velocity cues can also generate an estimate of MID. This cue is known as the interocular velocity difference (IOVD).
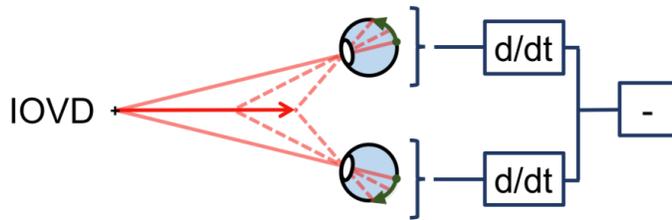
Figure 1.8 *The IOVD cue to MID. When an object moves towards an observer, it generates a monocular motion vector on each retina that is in the opposite direction between the eyes. Comparing these monocular motion vectors generates an estimate of the direction of MID.*

The difference in order of processing that these equations imply is illustrated in **Figure 1.9**. CD requires the instantaneous binocular disparity to be monitored over time, and thus requires a high-resolution comparison of left and right eye positions before any temporal derivative is taken. IOVD, on the other hand, requires a monocular velocity cue to be computed first while the differencing happens second. High spatial resolution is not required. Because the order of operations is reversed in the IOVD computation compared to CD, different neuronal machinery might be used to compute the two percepts.
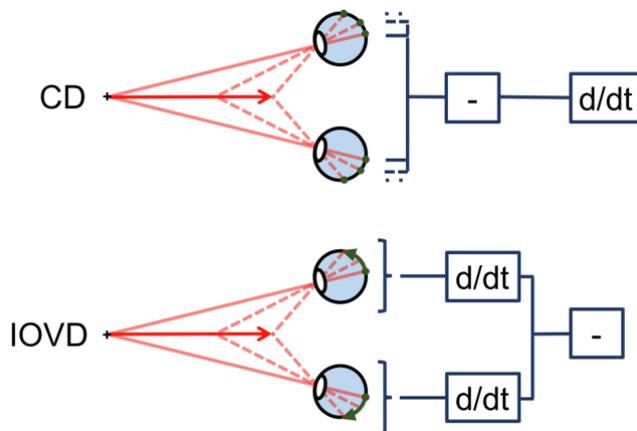
Figure 1.9 *Differences in order of processing between the CD and IOVD cues. For CD, outputs from the left and right eyes are compared first to estimate binocular disparity. Binocular disparity is then monitored over time to estimate MID. For IOVD these operations are reversed, and monocular disparity over time is computed first to generate a velocity vector in each eye. These vectors are combined at a later stage in the computational hierarchy to form a binocular signal for MID.*

1.4.3 Isolating CD and IOVD cues: stimulus design

To show that CD and IOVD provide independent cues to MID, researchers have attempted to generate stimuli that isolate both cue types. Isolating the CD pathway is relatively straightforward using stereoscopic random dot stimuli. Spatial correlations between elements in the left and right eyes are maintained to generate a depth cue, whilst horizontal binocular disparity between element pairs is systematically increased or decreased to generate MID (Julesz, 1971).

Importantly, no depth information is present in the monocular images and the dot pattern is refreshed at each video frame, thus eliminating any possible monocular motion cues. Closing one eye results in the perception of random flicker. Viewed stereoscopically, the subjective percept is one of a plane of dots moving smoothly through depth. A cartoon of typical CD stimuli is shown in **Figure 1.10**, panel A.
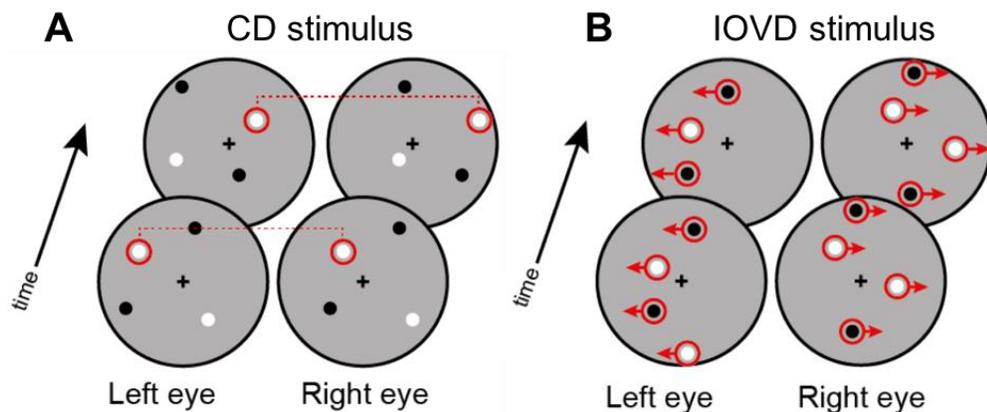
Figure 1.10 *Isolating CD (panel A) and IOVD (panel B) cues in stereoscopic random-dot stimuli. For the CD stimulus, the retinal disparity of binocular dot pairs is increased or decreased systematically over time. Because the dot pattern is regenerated at the onset of each video frame, there are no monocular motion cues. Monocular, temporal correlations are disrupted, whilst binocular correlations are maintained. For the IOVD stimulus, patterns in the left and the right eyes move in opposite directions, generating horizontally opposing motion vectors. Binocular correlations between dot pairs can be disrupted by a number of methods to reduce disparity leakage. In this example, dot patterns are generated independently for each eye's view. The dots persist over several video frames to generate monocular, temporal correlations that convey the lateral motion.*

Generating a 'pure' IOVD stimulus is considerably more challenging, with researchers relying on 'devilish trickery' (Cormack et al., 2017, p. 300) to isolate the cue. In general, the aim is to create dot fields that move in antiphase in the left and right eyes. Such a stimulus generates a pattern of velocities in each eye that can be compared to produce an IOVD cue (see **Figure 1.10**, panel B).

It is theoretically possible that random spatial correlations between the dots in the left and right eyes can give rise to spurious CD cues. The 'pure' IOVD stimulus can therefore become contaminated with another weak but consistent cue to MID (Peng & Shi, 2014). In the work presented in this thesis, we mitigated this effect in several ways that are described in detail in the second and third chapters (see sections 2.3.4, 2.3.5, 3.3.4, and 3.3.6). Other groups have attempted to minimise this problem using anticorrelated, decorrelated, or 'banded' display types. These methods are illustrated in **Figure 1.11**.
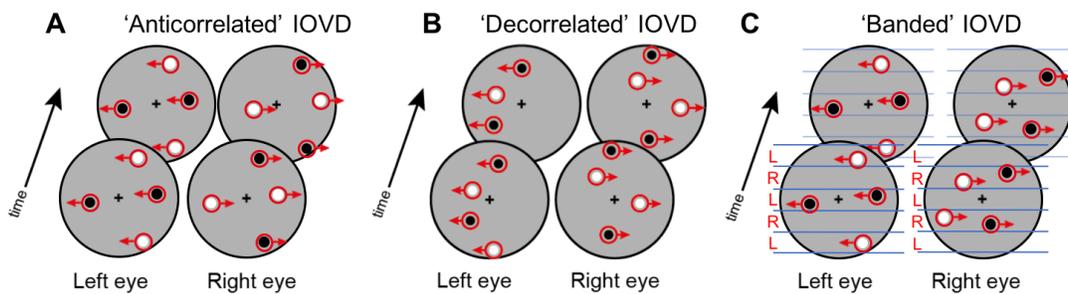
Figure 1.11 Different methods for isolating the IOVD cue to MID. Binocular disparity cues can be degraded, thereby minimising CD cue leakage, by anticorrelating the contrast polarity of binocular dot pairs (panel A), decorrelating the dot patterns shown to the left and right eyes (panel B), or by presenting dots in alternating bands in the stimulus between the two eyes (panel C).

Anticorrelating the contrast polarity of left and right eye dot pairs (**Figure 1.11**, panel A) has been shown to severely degrade the stereo depth mechanism (Harris & Rushton, 2003; Rokers, Cormack, & Huk, 2008) and thus has been used to minimise CD leakage (Czuba, Rokers, Huk, & Cormack, 2010). This method cannot wholly eliminate the CD cue – neurons in V1 do still respond to anticorrelated disparity, though these responses are absent in extrastriate areas (Bridge & Parker, 2007).

Alternatively, decorrelated dot patterns (**Figure 1.11**, panel B) can be shown to the left and right eyes, where the pattern of dots in each half-image is random (Nefs, O'Hare, & Harris, 2010; Shioiri, Nakajima, Kakehi, & Yaguchi, 2008; Shioiri, Saisho, & Yaguchi, 2000). In the latter case, spurious binocular matches are difficult to quantify. A 'banded' stimulus (**Figure 1.11**, panel C), where spatially alternating stripes of elements are shown to the left and right eyes, may minimise this problem (Shioiri et al., 2000).

Because the IOVD cue is based on a monocular motion trajectory, elements persist over several video frames and move in coherent, horizontally opposing directions between the two eyes. Closing one eye results in the perception of unidirectional lateral motion. When viewed with both eyes, the IOVD stimulus generates a sensation of MID but, due to the elimination of binocular disparity, does not convey a strong perception of depth relative to the observer (Rokers et al., 2008).

1.4.4   Evidence for the CD mechanism

Early reports showing that a MID percept could be generated based on changing disparities alone (Norcia & Tyler, 1984; Tyler, 1971) contributed to a zeitgeist where binocular MID mechanisms were thought to depend almost wholly on CD information. Detection thresholds for MID stimuli were found to be well-predicted by stereoacuity thresholds, but not by monocular (2D) motion thresholds (Cumming, 1995). Furthermore, introducing IOVD signals into a pure CD stimulus did not appear to improve stereomotion detection thresholds (Cumming & Parker, 1994). These findings led to the assertion that to include an IOVD mechanism into models of binocular MID perception would be 'superfluous' (Cumming, 1995, p. 113).

Despite the apparent reliance of MID mechanisms on CD information, it was unclear whether these were built on static disparity detectors, or whether they were processed independently. In support of a 'two streams' model for static disparity and disparity-based MID pathways, no cross-adaptation effects between static stereo and dynamic MID stimuli were observed (Beverley & Regan, 1973). In addition, many subjects were unable to detect stereoscopic MID in parts of the visual field where stereoacuity was nonetheless normal (Hong & Regan, 1989; Regan, Erkelens, & Collewijn, 1986; Richards & Regan, 1973).

In contrast to this view, Cumming and Parker proposed that CD mechanisms are built on the same circuitry that underpins static disparity computations (Cumming & Parker, 1994). Both stereoacuity and stereomotion thresholds showed similar dependencies on mean disparity pedestals and visual eccentricity, and participants with poor stereoacuity thresholds also showed poor stereomotion detection thresholds (Cumming, 1995). More recently, it has been shown that psychophysical responses to disparity modulation, such as the sensitivity limit for temporal frequency, are mirrored by response properties and receptive field sizes of disparity-selective neurons in V1 (Nienborg, Bridge, Parker, & Cumming, 2004, 2005). These similarities imply that MID perception based on the CD mechanism is likely to draw on the same computations that support the perception of static disparity.

The existence of neural populations tuned for CD-defined MID is further suggested by results from motion adaptation paradigms. Despite the fact that CD stimuli do not contain any coherent monocular motion, they are nonetheless capable of producing 3D motion after-effects (Czuba, Rokers, Huk, & Cormack, 2012; Rokers, Czuba, Cormack, & Huk, 2011). Recently, a physiologically plausible model for CD processing has been proposed (Peng & Shi, 2010, 2014) that combines existing disparity energy models (Ohzawa, DeAngelis, & Freeman, 1990) with 2D motion energy models (Adelson & Bergen, 1985; Watson & Ahumada, 1985). The model proposes that MID tuning is more easily developed from disparity than from velocity signals, and its outputs are consistent with speed discrimination thresholds measured empirically using dynamic random dot stereograms (Brooks & Stone, 2004, 2006).

1.4.5   Evidence for the IOVD mechanism

The historical lack of evidence for IOVD inputs to MID mechanisms stem, in part, from the technical difficulty in generating a 'pure' IOVD stimulus (Cormack et al., 2017). Early studies were limited to drawing comparisons between a CD-isolating stimulus and a stimulus that contained both CD and IOVD cues (Portfors-Yeomans & Regan, 1996), or by attempting to degrade available disparity cues in mixed stimuli where MID was beyond the temporal limits of stereopsis (Cumming & Parker, 1994). However, as techniques for isolating the IOVD cue have improved, the contribution of velocity signals to MID perception has become clearer (Fernandez & Farell, 2005; Rokers et al., 2008, 2011; Shioiri et al., 2000).

Studies investigating speed and velocity perception in particular emphasize the utility of the IOVD cue. IOVD information improves speed discrimination thresholds above those measured for a CD-only stimulus (Brooks & Stone, 2004), and at high speeds the facilitatory effect of the IOVD cue for detecting 3D motion is especially pronounced (Wardle & Alais, 2013). In addition, deficits in stereomotion perception in parts of the visual field – taken earlier as evidence for separate static disparity and stereomotion pathways – are correlated with observers' sensitivity to IOVD, but not to CD (Barendregt, Dumoulin, & Rokers, 2016).

Adapting to a stimulus moving in depth has also revealed a central role for IOVD, where motion after-effects were more readily induced using stimuli that contained IOVD cues rather than the CD cue alone (Brooks, 2002; Czuba et al., 2012; Sakano, Allison, & Howard, 2012; Sakano & Allison, 2014). Because these 3D motion after-effects could not be explained by adaptation to component 2D motion, they suggest the existence of neural populations that are tuned specifically to a direction of motion in depth (Czuba, Rokers, Guillet, Huk, & Cormack, 2011).

In contrast to 2D motion after-effects, 3D motion after-effects based on IOVD stimuli are independent of spatial frequency, suggesting that IOVD is computed at a relatively late stage of motion processing (Shioiri, Kakehi, Tashiro, & Yaguchi, 2009). Because the IOVD cue depends on a comparison of binocular opponent motion signals, these results further imply the preservation of eye of origin information to relatively late stages of visual computation (Fernandez & Farell, 2006; Shioiri et al., 2009). Indeed, very sparse IOVD stimuli can generate a MID percept, even when motion is integrated across areas larger than V1 receptive fields (Rokers et al., 2011), and when monocular velocity cues occur in non-corresponding parts of the far visual field (Greer, Bonnen, Huk, & Cormack, 2016). This late stage of binocular integration stands in direct contrast to the early extraction of disparity information that the CD cue depends on, suggesting that both cues may provide dissociable contributions to MID perception.

1.4.6   Relative contributions of CD and IOVD to 3D motion perception

Despite the fact that both disparity-based and velocity-based signals have been shown to contribute to MID perception, it is unclear how the two cues interact and supplement one another, or which cue provides the primary input to MID perception. Earlier papers claimed that sensitivity to CD was higher, whereas the IOVD cue appeared weak and was thought to add little information to MID estimates (Cumming & Parker, 1994; Gray & Regan, 1996; Lages, Mamassian, & Graf, 2003).

In addition, there is substantial variation across observers – in a study measuring dot coherence thresholds to MID stimuli across a large range of naïve participants, Nefs et al. found that only 53% of all participants were able to perform above chance level on the IOVD task. In comparison, 77% of participants tested reached acceptable threshold performance for the CD task. Whilst most observers showed a preference (higher sensitivity) to the CD cue, a minority did show the opposite pattern (Nefs et al., 2010).

More recently, other groups have shown that IOVD provides a significant contribution to MID mechanisms given a particular set of stimulus conditions (Czuba et al., 2012; Fernandez & Farell, 2005; Sakano et al., 2012), and may even dominate over the CD cue under those conditions (Czuba et al., 2010; Shioiri et al., 2008). Thus, task selection is a critical factor in assessing the relative contribution of IOVD – as mentioned previously, the IOVD mechanism appears particularly advantageous in tasks involving speed perception (Brooks & Stone, 2004, 2006; Fernandez & Farell, 2005; Harris & Watamaniuk, 1995), and experiments using motion after-effects have found evidence primarily for IOVD-based MID mechanisms (Czuba et al., 2012; Sakano et al., 2012). These findings suggest that the relative dominance of CD and IOVD may be situation-dependent, with both cues supporting MID extraction under different conditions.

Some of this confusion has arisen because few papers have attempted to map the parameter space over which CD and IOVD mechanisms are optimally sensitive. In many cases, stimuli have been matched on physical properties such as element size, speed, or stimulus field of view (Nefs et al., 2010). Whilst this balances the low-level information present between MID stimulus types, it may produce conditions that favour one cue over the other. When spatial and temporal parameters have been manipulated and sensitivity to CD

and IOVD stimuli has been measured, substantial differences between the cue types have been revealed (Czuba et al., 2010; Shioiri et al., 2008). Differences in spatiotemporal tuning of the two cues is consistent with the relatively poor temporal and spatial resolution of stereopsis, in comparison to that for lateral motion (Norcia & Tyler, 1984; Regan & Beverley, 1973; Tyler, 1971).

1.4.7   Spatiotemporal tuning of binocular MID mechanisms

The spatiotemporal tuning differences between the CD and the IOVD cues were tested extensively in a 2010 paper by Czuba and colleagues (Czuba et al., 2010). Using a technique similar to those commonly used to test frontoparallel motion, the authors titrated sensitivity to CD, IOVD or CD+IOVD by measuring coherence thresholds in random dot displays.

Stimuli were tested over a range of eccentricities (3-7, 7-11, 11-15 degrees from fixation) and speeds (0.3, 0.6, 0.9, 1.8 and 2.7 degrees per second (°/s); monocular motion). At lower speeds, increasing eccentricity degraded the MID percept at a similar rate irrespective of stimulus type, although overall sensitivity was higher to CD than IOVD. At higher speeds, sensitivity to the CD cue broke down, whereas mixed and IOVD cues showed a peak in sensitivity at around 2°/s that was maintained even at far eccentricities.
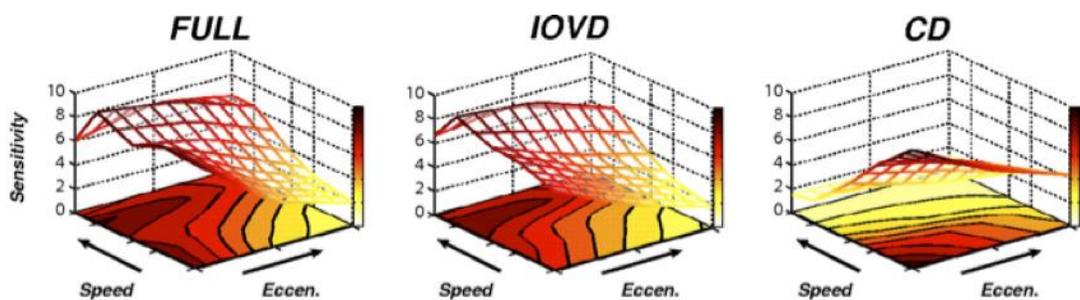
Figure 1.12 *Spatiotemporal tuning of FULL (CD+IOVD), IOVD and CD mechanisms. Sensitivity, measured by dot coherence thresholds and here normalised to vary between 0-10, is plotted as a function of speed and eccentricity. The peak in sensitivity for an IOVD cue occurs at higher speeds and is maintained across a range of eccentricities. For CD, the peak in sensitivity occurs at lower speeds and low eccentricities. Note also that the FULL surface combines sensitivity profiles from both cue types, indicating that both are used in tandem to perceive MID across different spatiotemporal parameters. Figure from Czuba et al., 2010.*

Czuba et al. showed that CD and IOVD cues operate optimally over distinct parts of the parameter space (see **Figure 1.12**). Together, these cues may be used to perceive MID across wider ranges of eccentricities and stimulus speeds. CD is optimal at lower speeds (Gheorghiu & Erkelens, 2005; Shioiri et al., 2008) and closer to the fovea, whereas IOVD is more useful across a wider range of speeds and across larger portions of the visual field. When sensitivity to CD and IOVD is mapped across a fuller spatiotemporal parameter space, it is easy to see how previous studies focusing on only one portion of this space may have underappreciated the contribution of the IOVD cue to MID mechanisms. In fact, Czuba et al. argue that because the sensitivity map for the IOVD cue predicts the sensitivity map for the FULL (CD+IOVD) cue more closely than the CD cue alone does, IOVD provides the more substantive contribution to binocular MID mechanisms. Finally, the distinctive spatiotemporal characteristics of the CD and IOVD mechanisms also suggest that they are processed by dissociable neural pathways.

1.5   Neural encoding of 3D motion

1.5.1   Motion and depth encoding

Because the MID mechanisms discussed in the previous section depend on the same sources of information as do 2D motion and depth perception, it is likely that they share a common architecture with mechanisms that extract frontoparallel motion and depth from disparity (Huk, 2012). Indeed, early approaches suggested that MID mechanisms depend on integrating 2D motion processing across various static disparity planes (Cormack et al., 2017). Since then, psychophysical evidence from adaptation studies suggest the existence of neural populations tuned specifically to 3D motion direction. Current models of MID processing make explicit reference to pathways encoding CD- and IOVD-defined MID that sit parallel to those pathways involved in the extraction of frontoparallel motion (Baker & Bair, 2016; Peng & Shi, 2014).

A candidate region for the extraction of MID information is area MT, which is heavily implicated in 2D motion processing in primates (Born & Bradley, 2005; Maunsell & Newsome, 1987). As many as 90% of the cells in MT are tuned to 2D motion direction (Baker, Petersen, Newsome, & Allman, 1981; Felleman & Kaas, 1984; Maunsell & Van Essen, 1983a, 1983b; Van Essen, Maunsell, & Bixby, 1981; Zeki, 1980; Zeki, 1974), with a typical tuning bandwidth of roughly 90° around its preferred direction (Britten & Wezel, 1998). In addition, neurons in

MT are tuned to a range of different speeds, and their tuning curves can be approximated by a logarithmic function whose limits correspond to perceptual speed discrimination thresholds measured psychophysically (Nover, Anderson, & DeAngelis, 2005).

Many of these motion tuning characteristics are inherited from V1, which raises the question of how MT contributes to motion processing beyond that already carried out in early visual areas (Born & Bradley, 2005). Whilst V1 extracts component motion, MT may be more involved in pooling these signals for pattern motion computations (Movshon, Adelson, Gizzi, & Newsome, 1985). Additionally, MT may be involved in noise reduction, velocity computations, and estimating structure-from-motion (Born & Bradley, 2005).

Many neurons in MT are also disparity selective (DeAngelis, Cumming, & Newsome, 1998; DeAngelis & Newsome, 1999, 2004; DeAngelis & Uka, 2003), with disparity selectivity arising from the indirect pathway via V2 and V3 (Ponce, Hunter, Pack, Lomber, & Born, 2011; Ponce, Lomber, & Born, 2008), rather than directly from V1. Up to two-thirds of neurons in MT are selective for horizontal and vertical binocular disparity (Maunsell & Van Essen, 1983). Populations of disparity selective cells are arranged in columns (DeAngelis & Newsome, 1999), interspersed with columns of directionally selective cells (Albright, Desimone, & Gross, 1984; Geesaman, Born, Andersen, & Tootell, 1997). The combined encoding of depth and motion information in MT highlights its potential as a site for the extraction of 3D motion vectors.

1.5.2   Cortical and subcortical routes into MT

As already mentioned, motion selectivity in MT likely arises through its direct pathway from V1 layer 4, which receives its inputs primarily from the magnocellular (MC) layers of the lateral geniculate nucleus (LGN). Reversible deactivation of LGN MC layers results in a substantial loss of visual responsiveness in MT/MST (Maunsell, Nealey, & DePriest, 1990), indicating the relative importance of MC cell projections over signals from the parvocellular (PC) and koniocellular (KC) pathways. These findings contributed to the viewpoint that motion perception is largely achromatic, and that there is a functional division between motion and colour processing (Livingstone & Hubel, 1988; Zeki, 1993).

Disparity selectivity is mediated primarily via the indirect pathway through V2 and V3. This pathway includes a small projection from the PC layers of the LGN (Yabuta, Sawatari, & Callaway, 2001), in part due to the interleaving of MC and PC signals in the intracortical synapses of V1 (Sincich & Horton, 2005).

Although the majority of neural connections into MT originate from cortical inputs, lesion studies have shown that inactivation of V1 does not abolish all visual function in this region (Girard, Salin, & Bullier, 1992; Rodman, Gross, & Albright, 1989). Subcortical pathways from the superior colliculus (SC) and the inferior pulvinar (IP) may mediate this residual function; however, lesions to the SC alone cause no significant changes in MT responses (Rodman et al., 1989).

Alternatively, there is another direct, subcortical pathway from the lateral geniculate nucleus (LGN) to MT. This pathway appears to originate exclusively from the koniocellular (KC) cell layers in the LGN and its projections equal about 10% of the V1 input to MT (Sincich, Park, Wohlgemuth, & Horton, 2004; Yoshida & Benevento, 1981). It draws on cell populations across all KC layers of the LGN and has been proposed to mediate residual visual function in patients with blindsight (Ajina & Bridge, 2018; Ajina, Pestilli, Rokem, Kennard, & Bridge, 2015; Bridge et al., 2010). Thus, the KC inputs to MT represent a curious addition to the function of MT, and the role of these KC inputs is, as yet, not fully understood.

A Gestalt map of cortical and subcortical routes from the retinal ganglion cells (RGC) to MT is shown in **Figure 1.13**, where the thickness of each line denotes the relative strength of the connection. This is based on a meta-analysis of the anatomical literature, combining the number of projecting neurons with characteristics of their axon terminals (Born & Bradley, 2005).
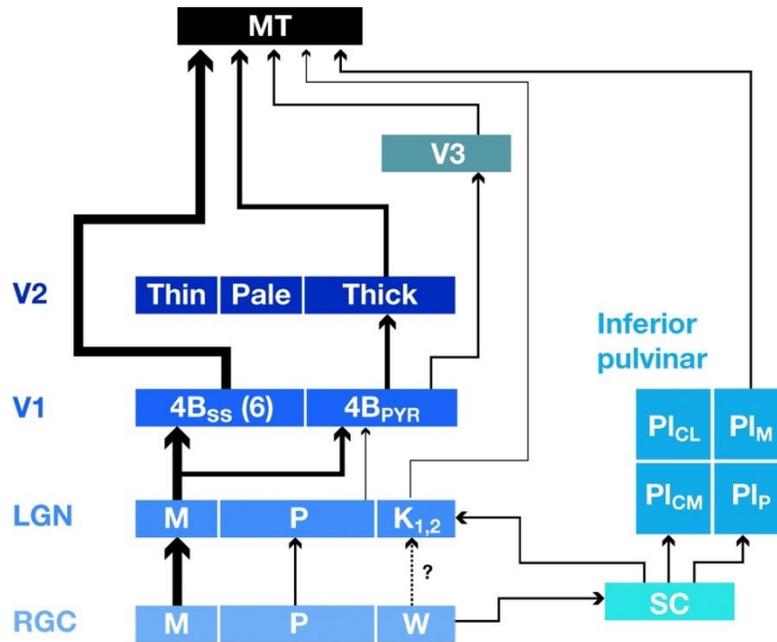
Figure 1.13 *Schematic of some known feedforward inputs from the retinal ganglion cells, through to MT. The line weights translate roughly to the magnitude of the inputs. The dominant projection is from the magnocellular layers of the LGN to MT via V1. Note also the subcortical projection from the koniocellular layers that by-pass V1 and project directly to MT from the LGN. Abbreviations: RGC retinal ganglion cell; LGN lateral geniculate nucleus; M magnocellular pathway; P parvocellular pathway; W retinal inputs to the K layers, the precise nature of which is not known but whose responses are W-like in the galago; K koniocellular pathway; SC superior colliculus; PI_{CL} central lateral nucleus of the inferior pulvinar; PI_M medial nucleus of the inferior pulvinar; PI_{CM} central medial nucleus of the inferior pulvinar; PI_P posterior nucleus of the inferior pulvinar; $4B_{SS}$ spiny stellate neurons in layer 4B of V1, $4B_{PYR}$ pyramidal neurons in layer 4B of V1. Figure from Born & Bradley, 2005.*

1.5.3   Evidence for MID processing in MT

In line with the idea that MID processing may be multiplexed in pathways known to underlie 2D motion and stereoscopic depth perception, two separate groups have recently published evidence for 3D motion tuning in MT neurons. Czuba et al., and Sanada and DeAngelis, conducted extracellular recordings in macaque MT. Using random dot stimuli containing CD and IOVD cues, both groups found that around 50% of all cells in MT selectively coded for motion directly towards or away from the observer (Czuba, Huk, Cormack, & Kohn, 2014; Sanada & DeAngelis, 2014), with a further subset of neurons showing a MID bias (Czuba et al., 2014). Thus in total, around 70% of neurons in MT could represent information about 3D motion. The majority of these were selective for IOVD cues, with relatively little contribution from CD (Sanada & DeAngelis, 2014). Additionally, MID

tuning was found to arise from different monocular motion preferences, as well as from nonlinear interactions of these signals when both eyes were stimulated (Czuba et al., 2014).

These results contradict earlier electrophysiological work (Maunsell & Van Essen, 1983) that found little evidence for neurons tuned to MID in this region. Maunsell and Van Essen found that their MID responses could be fully accounted for by separable contributions from 2D motion direction and static disparity. However, their recordings were focused on extending tuning for frontoparallel motion to the third dimension – in other words, they defined their MID vectors by the preferred lateral motion component of individual units. This means that the majority of their stimuli were not moving directly towards and away from the observer. Czuba et al. found that there was an overrepresentation of trajectories directly approaching or receding the observer, which may explain why Maunsell and Van Essen did not record a significant number of MID tuned units.

Direct evidence from extracellular recordings in macaque MT dovetails with indirect evidence using functional magnetic resonance imaging (fMRI) in humans. To date, three fMRI papers have shown that responses in cortical regions in or around human area hMT+ are modulated by MID stimuli containing CD and IOVD cues (Joo, Czuba, Cormack, & Huk, 2016; Rokers, Cormack, & Huk, 2009), or CD cues in isolation (Likova & Tyler, 2007).

Likova and Tyler identified an area anterior to hMT+ that seems particularly responsive to CD MID – the putative 'cyclopean stereo motion' area (Likova & Tyler, 2007; **Figure 1.14**). They used a random-dot CD stimulus containing no monocular motion cues. Responses to the CD stimulus were compared to responses to two different control stimuli – a static depth plane at zero disparity, and a disparity-matched control stimulus. This stimulus contained no coherent motion in depth but did contain static disparity over the same depth range as in their CD stimulus. Responses comparing the CD stimulus against either control stimulus were highly correlated. Thus, observed responses to the CD stimulus could not be accounted for by static disparity tuning. The authors suggest that disparity sensitive neurons identified in macaque MT may provide input to this anterior region for the processing of CD MID.
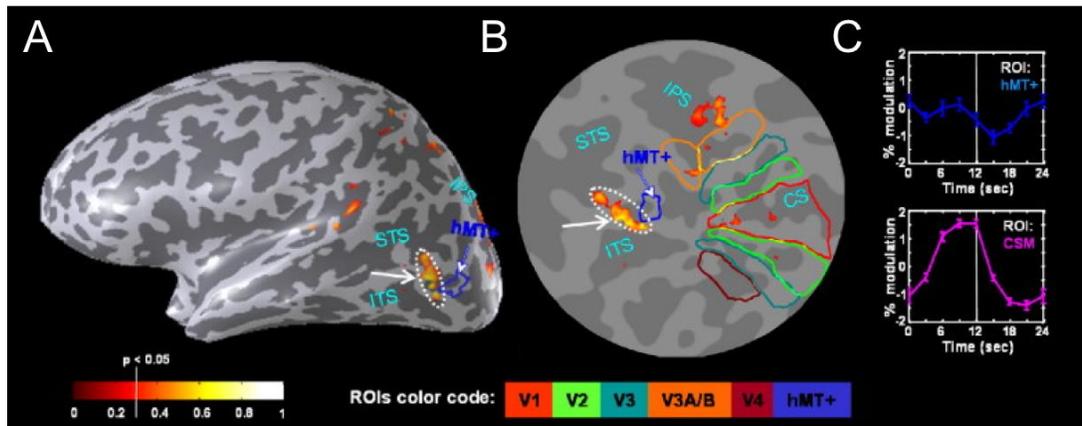
Figure 1.14 *Responses to a CD motion stimulus, compared against a control stimulus. In this example, the control stimulus consisted of a static depth plane at zero disparity. Responses are shown on an inflated brain (panel A) of an example subject, and on a flattened representation showing the location of the peak activation (in the CSM ROI, dotted white line) relative to earlier visual areas. Note that the CSM is situated anterior to hMT+, and that there is no activation in hMT+ itself. In addition, there is a robust activation in the intraparietal parietal sulcus (IPS) but this is not discussed at length by the authors. Panel C shows the timecourse from within the hMT+ ROI and the CSM area during the CD stimulus condition. The CSM responds twice as strongly. Figure adapted from Likova & Tyler, 2007.*

Rokers, Cormack and Huk also used fMRI to investigate the cortical locus of binocular MID processing (Rokers et al., 2009). Their paper outlined four separate experiments. Experiment 1 used a mixed CD+IOVD stimulus to identify areas involved in MID processing. In experiments 2 and 3, they investigated the representation of CD and IOVD separately (see **Figure 1.15**). Finally, an adaptation paradigm revealed areas showing directionally selective response patterns to CD+IOVD cues. Their principal finding is that hMT+ is involved in the processing of all these stimulus types. In addition, their final experiment showed that, in hMT+, the fMRI signal was depressed when the test pattern was moving in the same direction in depth as the adaptor. This adaptation effect suggests that hMT+ contains neurons tuned to a specific direction of MID.
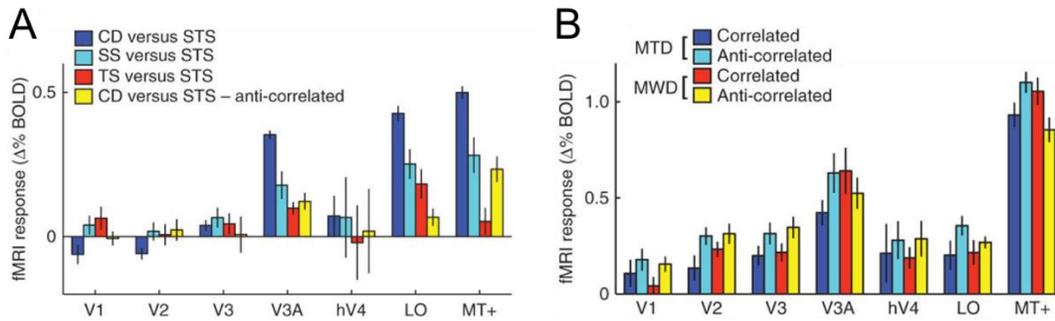
Figure 1.15 *fMRI responses measured across visual areas to CD (panel A) and IOVD (panel B) stimuli. In panel A, the dark blue bar represents the response difference between a CD stimulus and a spatio-temporally scrambled (STS) control stimulus. Other bars show control conditions, where responses to spatially scrambled (SS) CD motion and temporally scrambled (TS) CD motion were compared against the STS stimulus. The final comparison was between anti-correlated CD motion against an anti-correlated STS stimulus. In panel B, the dark blue bar shows the response to a motion through depth (MTD) stimulus where binocular dot pairs were correlated (containing both CD and IOVD). The light blue bar shows an anticorrelated version of the stimulus, which biases it towards IOVD motion. The motion within depth (MWD) conditions provide controls, where the same monocular motion energy is preserved but there is no MID signal to extract because binocular dot pairs are moving in the same direction. Error bars represent ±1 SEM. Figure adapted from Rokers et al., 2009.*

There are some inconsistencies between the findings of these two papers. Likova and Tyler found no tuning to CD in hMT+. Instead, they identified CD responses in an adjacent, anterior area. Rokers, Cormack and Huk did not report results from this adjacent area, but did find CD tuning in hMT+. Differences in stimuli may have contributed to this – both groups used stimuli that were similar in MID speed, dot size and dot density, but the Rokers et al. stimulus was divided into four quadrants, where the MID trajectory was phase shifted by 180º in alternate quadrants. Thus, the stimulus contained both relative and absolute disparity cues. The Likova and Tyler stimulus consisted of a plane of dots oscillating in depth, thus containing absolute disparity, but no relative disparity cues. Absolute and relative disparity appear to be processed via dissociable neural mechanisms, with absolute disparity engaging more dorsal visual areas (Neri, Bridge, & Heeger, 2004). There were also differences in the control stimulus that was used to subtract the signal from static disparity mechanisms – Rokers et al. used a control stimulus that was a non-structured cloud of dots containing a range of disparities, whereas Likova and Tyler used or a near or far disparity 'surface' of dots. These factors may contribute to the differences in cortical responses measured in these papers.

Joo et al. extended the findings from the fMRI-adaptation study described by Rokers et al., by measuring separate adaptation effects to CD and IOVD isolating stimuli. In hMT, the adaptation effect was stronger for IOVD than for CD stimuli, but no cross-cue adaptation effects were observed (Joo et al., 2016; see **Figure 1.16**). The adaptation effects in V1, V2, V3 and V3A were weak, suggesting that whilst these areas may provide inputs to hMT that are pertinent to the extraction of MID, they are unlikely to contain neurons tuned to 3D motion direction directly. In addition, the lack of cross-cue adaptation implies that whilst hMT encodes both CD and IOVD-defined MID, different sub-populations of neurons in this area process either cue type. The authors suggest that this functional distinction may reflect the relative utility of CD and IOVD mechanisms for different perceptual tasks, consistent with the tuning properties of both cues (reviewed above in sections 1.4.6: Relative contributions of CD and IOVD to 3D motion processing, and 1.4.7: Spatiotemporal tuning of binocular MID mechanisms).
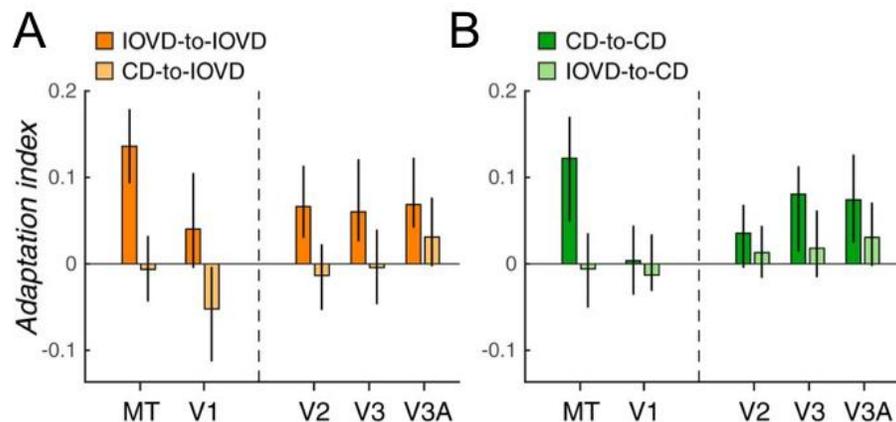
Figure 1.16 Within-cue and cross-cue adaptation effects for MID stimuli. Panel A shows within-cue adaptation for IOVD in orange, whilst cross-cue from CD to IOVD is shown in yellow. Panel B shows within-cue adaptation for CD in dark green, and cross-cue adaptation from IOVD to CD in light green. Effects were measured using fMRI across a range of visual areas. The adaptation index was calculated by normalising the difference between the response amplitude to MID in the same direction between test and adapter, and the response amplitude to MID in the opposite direction between test and adapter, by its sum. Error bars indicate 68% confidence intervals. Figure adapted from Joo et al., 2016.

1.6   Summary and outstanding questions

There is now considerable evidence for the use of both CD and IOVD cues in the perception of MID. The spatiotemporal tuning profiles of both cues are sufficiently distinct to

suggest that both cues may support different perceptual tasks, and that they are also dissociated in cortical and, potentially, subcortical pathways. The neural mechanisms that underpin CD and IOVD have begun to be explored. However, there are some crucial questions that remain to be addressed. First, there is some discrepancy in the literature regarding the neural locus of the CD cue. Secondly, fMRI studies to date have localised possible sites for CD and IOVD computations, but they have only considered feedforward mechanisms and have not linked fMRI activation with behaviour explicitly. Thirdly, it is unclear how the fundamental sources of information that the two cues depend on – such as eye of origin information – are maintained in areas that compute MID. Finally, the relative contributions of precortical MC, PC and KC pathways, whose tuning characteristics suggest different contributions to either mechanism, has not yet been explored. As such, it is unclear how exactly signals are transformed throughout the visual hierarchy to arrive at a coherent MID percept. It is important to address these questions before we can arrive at a complete explanation for how CD and IOVD cues are combined to form a general, binocular MID signal.

1.7   Outline of this thesis

This thesis aims to present a detailed and thorough exploration of the outstanding questions issued above. The first chapter investigates the contribution of MC, PC and KC signals to CD and IOVD, by measuring psychophysical motion direction discrimination thresholds in random dot stimuli. In the second chapter, fMRI is used to investigate the neural representation of CD and IOVD, whilst also distinguishing between achromatic and S-cone isolating inputs. In the third chapter, multivariate pattern analysis and classification techniques are applied to fMRI data, to investigate how eye-specific signals are carried to regions in the brain, revealed in the previous chapter, that support the computation of IOVD. The final chapter investigates how top-down attentional demands impact the fMRI signal in pathways involved in MID perception. This approach reveals neural populations that are critical for solving a motion detection task based on CD or IOVD-isolating cues. Together, this work describes an exploration of both feedforward and feedback mechanisms, and details how neural pathways involved in the perception of MID draw on fundamental sources of visual information.

1.8 Supplementary information
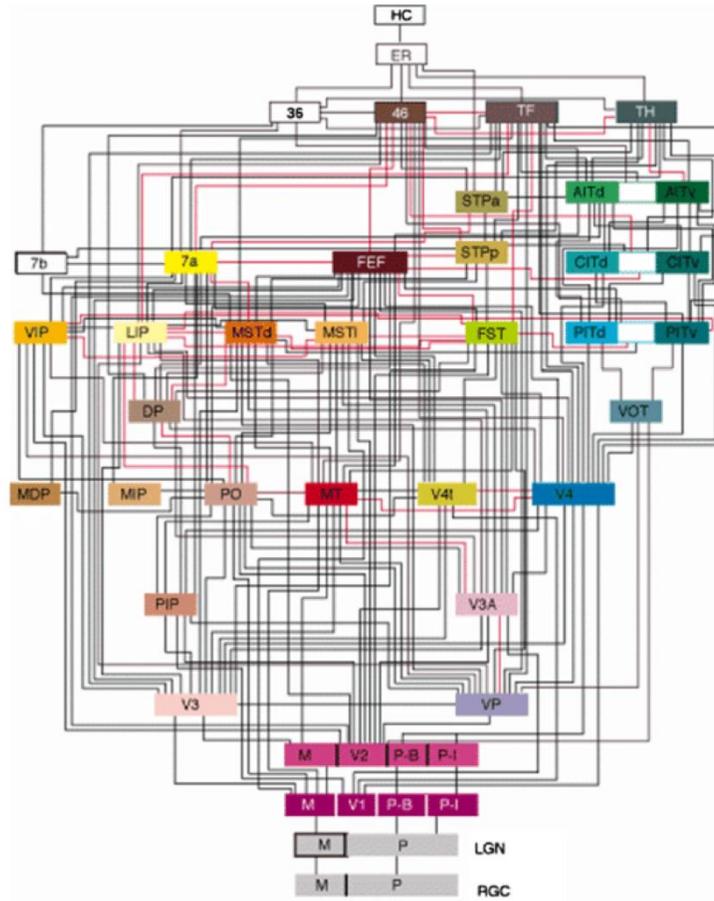
Figure 1.17 *Anatomical connections in the primate visual system, demonstrating the hierarchical nature of visual processing and feed-forward connections. Despite its already fabulous complexity, this diagram misses the contribution from the bistratified retinal ganglion cells (RGCs), which project to the koniocellular layers of the lateral geniculate nucleus (LGN) – partly because of the historical lack of research on this evolutionarily ancient pathway. Figure adapted from Felleman & Van Essen, 1991 and reproduced in Silvanto, 2015.*

Chapter 2.　Chromatic disparity and velocity cues contribute to motion in depth perception

2.1　Abstract

The visual system draws on at least two binocular cues to monitor an object's motion in depth (MID). These are changing disparity (CD), the temporal derivative of retinal disparity, and inter-ocular velocity differences (IOVD), the disparity of monocular motion cues. To potentially dissociate the pathways involved in MID computation, we asked whether CD and IOVD could also be computed from L+M, L-M or S cone isolating stimuli. Given adequate levels of stimulus contrast, chromatic pathways contribute to 2D motion perception. Therefore, we used contrast-scaled L+M, L-M and S isolating random-dot stimuli to measure motion discrimination performance for CD or IOVD-type MID.

We found that approximately equal L-M and L+M+S cone contrasts (2-3%) were required to achieve the same performance threshold, while S-cone stimuli required approximately ten times more contrast to reach the same performance. Once these contrasts were set, there were no differences between IOVD and CD motion discrimination performance. Adding noise to the stimuli degraded performance as expected with larger performance decrements in IOVD than CD conditions, consistent with the two mechanisms having different integrative field sizes. All three chromatic channels were affected in the same way from reduction in coherence.

Finally, we asked whether detection mechanisms pooled across different chromatic channels using signal and mask dots of different colours. We found no evidence of independent detection mechanisms: chromatic signals were masked as effectively by luminance noise as by noise from within their own channel.

We conclude that both types of MID computation draw on signals from all three precortical pathways, in a manner analogous to 2D motion detectors.

2.2   Introduction

The perception of motion in depth (MID) – relative motion towards and away from an observer – is a fundamental prerequisite for interacting with a three-dimensional world. Binocular mechanisms provide at least two such cues (Harris, Nefs, & Grafton, 2008; Regan, 1993; Rushbass & Westheimer, 1961). Despite a recent increase in attention given to studying these two cues (Czuba, Rokers, Huk, & Cormack, 2012), the neural mechanisms that underpin them are still unclear. Functional magnetic resonance imaging (fMRI; Joo, Czuba, Cormack, & Huk, 2016; Likova & Tyler, 2007; Rokers, Cormack, & Huk, 2009) and unit electrophysiology (Czuba, Huk, Cormack, & Kohn, 2014; Sanada & DeAngelis, 2014), studies indicate that these 3D cues may be multiplexed in areas classically associated with 2D motion perception (Huk, 2012). In this study, we ask how signals from three precortical pathways contribute to the perception of MID based on two binocular cues.

2.2.1   Binocular cues to motion in depth

The trajectory of an object moving through space generates at least two dissociable binocular cues (**Figure 2.1**). The first cue, called changing disparity (CD), relies on increases and decreases in the relative horizontal offset, or binocular disparity, between stimulated areas on the left and right retina. As an object moves towards an observer, binocular disparity increases. As the object moves away, binocular disparity decreases.

The second cue, called inter-ocular velocity difference (IOVD), relies on a comparison of monocular velocity signals as an image is displaced on each retina separately. Motion directly towards and away from an observer generates monocular motion signals that are equal in amplitude, but opposite in direction. Motion towards an observer generates motion vectors in the temporal direction, whereas motion away generates vectors in the nasal direction. Comparing these signals between the eyes generates an estimate of MID.
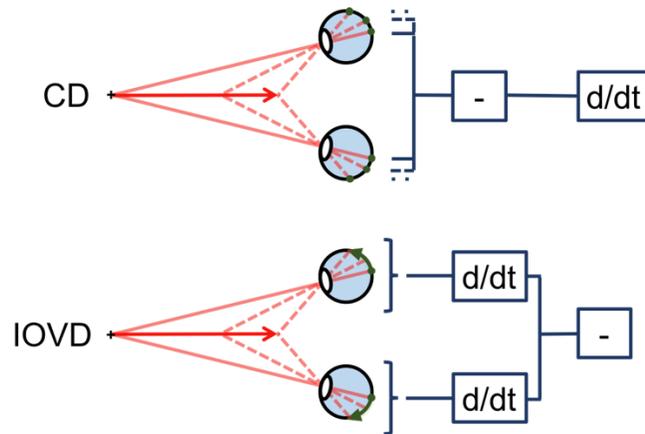
Figure 2.1 *The two binocular cues to motion in depth. CD (top) is computed by estimating the binocular disparity first, and monitoring how that disparity changes over time. IOVD (bottom) is computed by generating a monocular motion vector first, before comparing those vectors between the eyes.*

Early psychophysical work suggested that the CD cue alone is sufficient to detect MID, because detection thresholds for a CD-isolating stimulus were as low as for a stimulus containing both CD and IOVD information (Cumming & Parker, 1994; Gray & Regan, 1996). However, several groups have now shown that the IOVD cue does provide a valuable contribution, and that the visual system uses both cue types to determine MID (Allen, Haun, Hanley, Green, & Rokers, 2015; Brooks & Stone, 2004; Czuba, Rokers, Huk, & Cormack, 2010; Fernandez & Farell, 2005; Nefs, O'Hare, & Harris, 2010; Rokers, Czuba, Cormack, & Huk, 2011; S Shioiri, Saisho, & Yaguchi, 2000). Differences in their spatial and temporal sensitivities imply that the two mechanisms may be used in tandem to extract MID information across wider ranges of stimulation (Satoshi Shioiri, Nakajima, Kakehi, & Yaguchi, 2008). IOVD may be used at higher speeds when the CD cue breaks down (Czuba et al., 2010), and at times when there are contrast mismatches between left and right retinal inputs (Rokers, Cormack, & Huk, 2008). Because the disparity mechanism is temporally sluggish (Beverley & Regan, 1973), MID speed judgements may rely more on IOVD than on CD cues (Harris & Watamaniuk, 1995). CD may be more useful for static depth perception and the extraction of form from motion (Czuba et al., 2010).

Because both cues rely on different mechanisms, and because they show such different characteristics, they are also likely to be dissociated at a neural level. Recent neuroimaging and neurophysiological findings implicate motion area MT in the processing of both cue types (Czuba et al., 2014; Joo et al., 2016; Likova & Tyler, 2007; Rokers et al., 2009; Sanada & DeAngelis, 2014). This area contains populations of neurons tuned for

motion direction (Albright, Desimone, & Gross, 1984; Geesaman, Born, Andersen, & Tootell, 1997) and static disparity (DeAngelis & Newsome, 1999), and thus may encode IOVD and CD directly, or provide cues for their computation in other regions. One challenge therefore is to understand how signals arrive at MT, and how different neural pathways contribute to these computations.

2.2.2 Chromatic cues to motion perception

One way to isolate the different pathways contributing to MID perception is to take advantage of an existing segregation of cone outputs present in the precortical pathways. Activity in the magnocellular (MC) pathway is driven almost exclusively by low spatial resolution, temporally band-pass achromatic (L+M+S) signals (Lee, 2010). The parvocellular (PC) pathway appears to carry both summed- and differenced-signals from the L and M cones with a preference for lower temporal resolutions, and the koniocellular (KC) pathway is driven optimally by signals generated by the opponent S-(L+M) cone mechanisms (Tailby, Szmajda, Buzás, Lee, & Martin, 2008).

Early models of functional specialization posit a segregation of motion and colour (Livingstone & Hubel, 1988; Zeki, 1993), where chromatic pathways were thought to contribute little to motion processing. However, this strong functional segregation may not persist past the input layers of V1 (Sincich & Horton, 2005). Neuroimaging and neurophysiological research has also demonstrated sensitivity in extrastriate motion areas to inputs from all three pathways, including areas that have been implicated in MID perception (Gegenfurtner et al., 1994; Saito, Tanaka, Isono, Yasuda, & Mikami, 1989; Thiele, Dobkins, & Albright, 1999, 2001; Wandell et al., 1999).

Psychophysically, it has also been shown that isoluminant L-M and S cone stimuli can contribute towards lateral motion perception (Cavanagh & Anstis, 1991; Cavanagh & Anstis, 1986; Cropper & Derrington, 1996), though their speed appears slower than that of achromatic stimuli (Dougherty, Press, & Wandell, 1999; Lee & Stromeyer, 1989; Stone & Thompson, 1992). Motion nulling experiments have demonstrated that 2D motion detectors pool across the contrast of achromatic as well as isoluminant stimuli, integrating signals from all three precortical pathways (Chichilnisky, Heeger, & Wandell, 1993). Finally, L-M grating stimuli have been found to convey the percept of stereoscopic MID (Tyler & Cavanagh,

1991). Phase thresholds for chromatic stimuli were higher than for achromatic stimuli, but they could be predicted based on the contrast sensitivity thresholds of each stimulus type. The contribution of S-cone signals was not investigated in this study, and no distinction was made between CD and IOVD cues.

An important point is that the three pre-cortical pathways do not contribute equally to motion perception. Despite being the most visible stimuli when scaled in units of cone contrast, L-M cone isolating stimuli require several times more contrast than luminance stimuli to convey the same motion percept and S cone stimuli require a ten-fold increase in contrast to match the luminance target (Dougherty et al., 1999).

If MID mechanisms use the same neural circuitry as that used for 2D motion perception, we would expect to find similarities in the way that isoluminant chromatic signals generate a percept of MID.

In this study, we first measured the contrast thresholds required to elicit a robust MID percept. After contrast-scaling our stimuli, we compared CD and IOVD mechanisms based on a common metric of sensitivity (3D motion coherence thresholds, determined by varying the ratio of signal and noise dots in the stimulus). Finally, we masked chromatic signal elements with luminance noise. Our results provide evidence for MID mechanisms that pool signals across all three precortical pathways to estimate the direction of MID, in a manner analogous to 2D motion detectors.

2.3   Methods

2.3.1   Participants

Five experienced psychophysical observers (including one author, MK) with normal or corrected-to-normal vision performed the experiment. Participants were screened for normal colour vision using the Ishihara colour vision test (24 plate edition). Stereoacuity for all participants was normal (below 120 arcsec) as measured by the TNO test (19th edition, Laméris Ootech, Ede, The Netherlands). All participants were able to reliably detect MID from

both CD and IOVD stimuli. This study was approved by the departmental ethics committee at the University of York and was designed in compliance with the Declaration of Helsinki.

2.3.2   Materials

Experiments were run using an Apple OSX computer (Intel I5) running Psykinematix software (KyberVision, Montreal, Canada). Stimuli were presented on a CRT monitor (NEC Multisync, 1124 x 768 pixels, 75Hz refresh rate) viewed at an effective distance of 165cm. Luminance, phosphor colour coordinates and gamma functions were measured using a Spyder4 colorimeter (Datacolor, Lawrenceville, NJ, USA) and gamma correction was performed in software using a lookup table. Participants viewed stimuli through a mirror stereoscope using a chin rest to aid fusion of side-by-side 2D images. A PS3-type gamepad (Afterglow AP.1, PDP, Los Angeles, CA, USA) was used as a response device.

2.3.3   Stimulus design

A percept of MID can be elicited in dynamic, random-dot stereograms. Our stereograms were two side-by-side circular patches (radius = 1.7 degrees of visual angle, °) containing pseudorandomly distributed dots (N = 50 per eye, density = 2 dots/deg^2 covering around 12% of the stimulus at any moment). Each element had a Gaussian profile (sigma = 0.06°) to reduce chromatic aberration, and was defined by either a positive or negative contrast polarity displayed on a mean grey background (luminance = 142.75 cd/m^2).

To aid stable binocular fusion, a white central fixation cross (100% contrast) was displayed throughout each trial (size = 0.15°). An additional four white fusion markers at peak luminance were displayed at the corners of a square around the stimulus edges (distance from fixation = 2.6°).

We specified our stimuli in MacLeod-Boynton cone contrast space (MacLeod & Boynton, 1979). To account for variability in the isoluminant L:M cone contrast ratio across participants, we determined the point of equiluminance for each participant using a motion nulling task (Anstis & Cavanagh, 1983). Participants adjusted the L:M ratio of a 1 cpd grating

until motion appeared minimal. This ratio was given by the theta value (θ), which represents the angle in cone contrast space that sets the point of equiluminance in L:M contrast. The mean θ across five repeats was taken as the point of isoluminance and was used to scale the L and M properties in L-M stimuli. LMS vectors used to define signal and noise dots are listed in **Table 2.1**

Table 2.1 LMS vectors in MacLeod-Boynton cone contrast space used to define the chromatic properties of experimental stimuli.

	L	**M**	**S**
Achromatic	1	1	1
Isoluminant red/green	*-sinθ*	*cosθ*	0
Isoluminant blue/yellow	0	0	1

Additional stimulus properties depended on whether the stimulus contained CD or IOVD MID, but for coherence thresholding the same pattern of noise elements was used for both cue types. Noise elements were physically identical to signal elements, but were pseudorandomly positioned in each eye, and were refreshed each monitor frame. Noise elements gave the perception of a static, flickering plane of dots at zero disparity. The motion coherence of the stimulus was determined by systematically varying the proportion of signal to noise elements between 0-100% coherence. **Figure 2.2** shows examples of our stimuli.
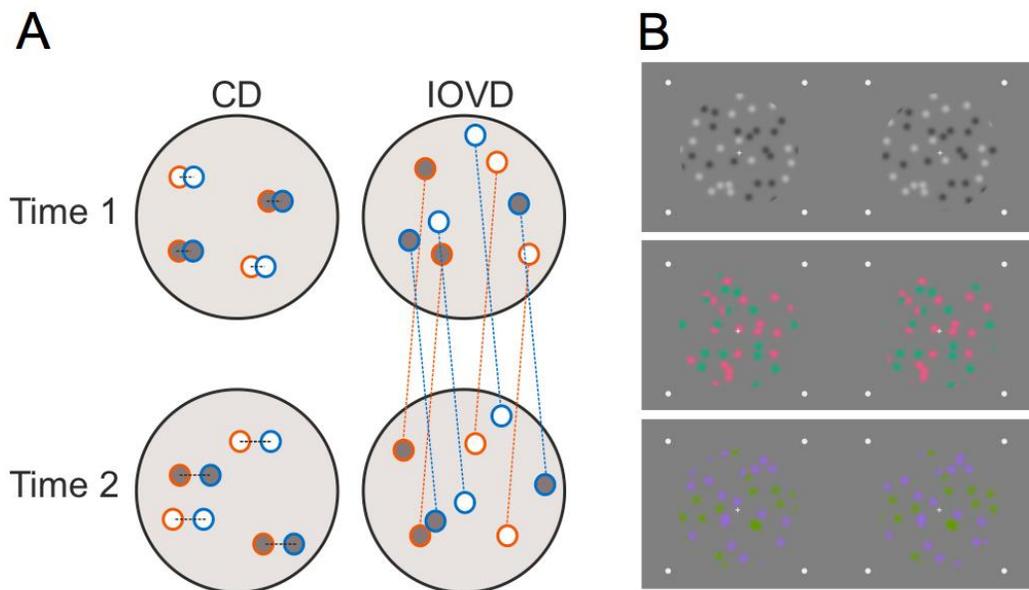
Figure 2.2 *Stimulus design (A) and examples of stimuli used (B). Panel A illustrates how CD and IOVD cues were isolated. For CD, left eye elements (orange) and right eye elements (blue) are paired across eyes and their binocular disparity changes between frames. Dots regenerate in new positions between frames to eliminate monocular motion. For IOVD, left and right eye elements are unpaired between the eyes, but individual elements persist across frames to generate monocular motion cues. Left and right eye elements move in opposite directions between frames. Panel B shows examples of CD L+M+S, L-M and S-(L+M) stimuli used in the experiments.*

2.3.4   CD stimuli

To isolate the CD cue, MID must be conveyed by changing binocular disparity in the absence of any monocular motion. We achieved this by drawing a new random dot pattern on each frame so that no systematic horizontal motion energy was present on average (**Figure 2.2**, panel A). Binocular disparity varied between +-11.94 arcmin with a temporal sine wave profile of one cycle per second. This is well below the temporal limit for stereoscopic motion perception (Norcia & Tyler, 1984). The initial direction of motion was randomized (either towards or away from the mid-point). A full cycle would have lasted 1s, but a single trial was presented for 500ms such that participants perceived half a sine-wave of MID. This was equivalent to motion towards-away, or away-towards. To ensure that mean disparity relative to fixation across each trail could not be used as a cue to MID direction, we added an additional starting disparity that varied between -2 and +2 arcmin. The binocular percept was one of a plane of dots moving smoothly towards and away from the observer. Closing one eye resulted in the percept of random flicker.

2.3.5   IOVD stimuli

For the IOVD cue, MID is conveyed by monocular motion vectors in horizontally opposite directions. Spatial correlations between element pairs must be degraded to avoid CD cue leakage. In our stimuli, we decorrelated the dot patterns between the two eyes so that different dot patterns were shown in each half-image (**Figure 2.2**, panel B) – a technique used by other groups to isolate IOVD motion (S Shioiri et al., 2000). In our stimuli, individual elements travelled at $\pm4°$/sec and persisted across 8 frames (10ms at 75Hz refresh rate).

Spurious binocular matches are difficult to quantify in this type of stimulus (Peng & Shi, 2014), however, the CD cue begins to break down with increasing velocity. Sensitivity to detecting MID from a CD cue is approaching zero at 2.7°/sec per eye (Czuba et al., 2010). Thus, any CD leakage in our rapid IOVD stimulus was not likely to contribute to the MID percept.

The monocular percept in an IOVD stimulus is coherent lateral motion. In theory, participants could correctly guess the direction of MID by identifying the monocular direction of motion. However, utrocular discrimination is generally poor, especially at low contrasts and high spatial frequencies, and does not improve with feedback (Blake & Cormack, 1979). Thus it is unlikely that participants used utrocular discrimination to perform our MID task.

The IOVD cue cannot provide a cue to absolute depth *position* (Rokers et al., 2008). Our stimuli therefore generated a sense of motion towards or away, but this did not look like a plane of motion in the same way the CD stimulus did. We matched the MID trajectory to the CD stimulus such that each trial lasted 500ms, following half a sine-wave of MID.

2.3.6   Participant training

Participants were shown several training stimuli to accustom them to the experimental tasks and equipment. These included stimuli to introduce the concept of fusion in the mirror stereoscope, and practice trials on high-contrast, 100% coherent versions of the experimental

stimulus. Participants were required to indicate the direction of MID. Auditory feedback was provided in these initial training trials, but was switched off for data collection.

Once performance stabilised at around 90% correct, participants progressed to the testing phase. Training did not exceed two sessions at one hour each. Some participants were able to reach criterion after only one session, whereas others required more time to accurately determine MID. To minimise order effects, experimental trials were fully counterbalanced between participants.

2.3.7   Procedure

Participants were tested over the course of a number of weeks in around 10 sessions. Each session lasted up to an hour. During testing, participants were required to indicate the *initial* direction of motion (towards or away) by pressing the correct button on a response pad.

Each trial lasted for 500ms and was preceded by an auditory tone to indicate stimulus onset. MID was perceived as half a sine-wave of motion, either towards-away, or away-towards. After each presentation, the screen went blank and participants had unlimited time to indicate the initial direction of MID perceived (towards or away). Immediately after the participant had responded, an auditory tone indicated the onset of the next trial.

Testing was split into three phases. In phase one, we used a Bayesian adaptive staircase (Kontsevich & Tyler, 1999) to vary the LMS cone contrast of a 100% coherent CD or IOVD cue. Signal elements were either achromatic, isoluminant red/green, or isoluminant blue/yellow. Participants completed three runs of each condition, and each run consisted of 10 training and 90 test trials.

In phase two, we used the same approach to measure coherence thresholds in achromatic, isoluminant red/green and isoluminant blue/yellow CD or IOVD stimuli. The staircase sampled between 0-100% coherence. Participants completed five runs of each condition, and each run consisted of 10 training and 90 test trials.

In phase three, we measured coherence thresholds in isoluminant red/green and isoluminant blue/yellow CD or IOVD stimuli. Here, signal elements were chromatic but noise elements were achromatic. Participants completed five runs of each condition, with the same number of training and test trials as in phase two.

To minimise order effects, experimental trials were fully counterbalanced between participants in phases two and three. Measurements from phase one were used to set the contrast of stimuli in phase two and three so we did not counterbalance here.

2.4 Results

2.4.1 *Contrast thresholding*

For each participant, we measured the percent cone contrast required to perceive the direction of MID for each cue type within each chromatic condition. Data from three repeats of each condition were collapsed together, and a logistic function was fit to the proportion of correct trials at each contrast level tested. The threshold value was taken at 90% correct, yielding contrast values for each participant for achromatic, L-M and S isolating CD and IOVD stimuli (see **Figure 2.3**, panel A). We used these values in coherence thresholding experiments to set the contrast of signal and noise elements.

Contrast thresholds for achromatic and L-M stimuli were around 2-3% for all participants irrespective of MID cue type (see **Figure 2.3**, panel B). For S-cone stimuli however, up to 10x this value was required. A repeated measures ANOVA using the Greenhouse-Geisser correction for sphericity found a significant main effect of chromaticity ($F(2,4.03) = 76.01$, $p = .001$), which was driven by the contrasts between the L+M+S and S cone ($p = .003$), and the L-M and S cone ($p = .003$), stimuli. This is similar in magnitude to the ten-fold increase in contrast required to match the 2D speed of an S cone stimulus to an achromatic standard (Dougherty et al., 1999).
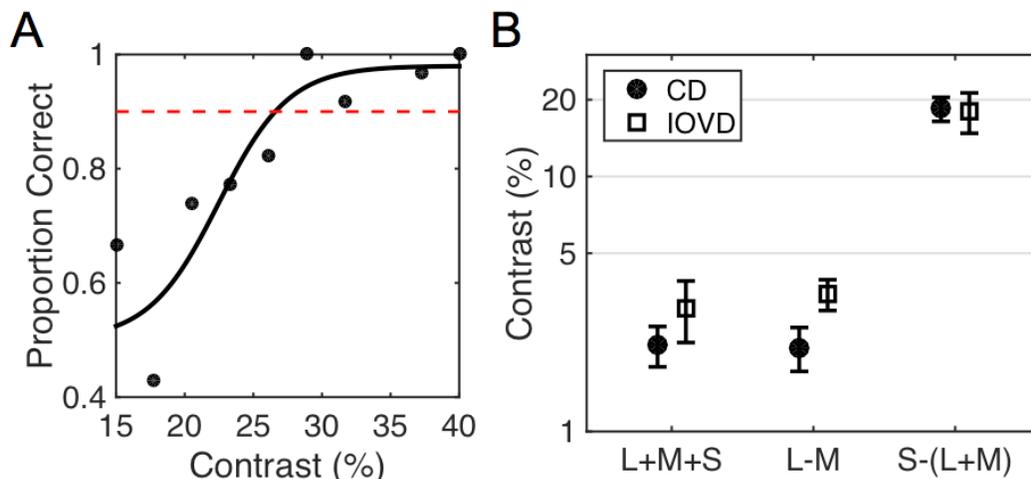
Figure 2.3 *Contrast thresholding data for an individual participant (A) and the group (B). Panel A illustrates how contrast threshold values are determined on an individual level for each cue type, at each chromatic condition. The proportion of correct responses was plotted at each contrast level tested, and a logistic function was fit (in this example, data from the S-isolating IOVD cue are shown). Proportion correct at 0.9 (indicated by the red line) is taken as the threshold contrast value – for this example, around 26% contrast. Panel B shows mean threshold values in % contrast required to perceive the direction of motion in depth from achromatic, L-M and S cone isolating CD and IOVD stimuli. Error bars indicate ±1 SEM and the Y-axis is log-scaled.*

The level of contrast required to perceive direction of MID did not differ between the CD and IOVD cues. A repeated measures ANOVA found no main effect of cue type ($F(1,4)$ = 1.52, p = .716). There appears to be a difference between CD and IOVD for L-M stimuli, but a paired-samples t-test showed this was not significant ($t(4)$ = -2.18, p = .095). This implies both CD and IOVD mechanisms require similar levels of input from the chromatic and achromatic pathways in order to signal the direction of MID.

2.4.2 Coherence thresholding

In this experiment, asked whether IOVD and CD mechanisms were equally sensitive to noise across the three chromatic conditions. To do this, we fixed the contrast of the stimuli based on the contrast thresholding results so that they generated identification accuracies of approximately 90% at a coherence level of 100%. We then varied the signal to noise ratio in our displays using a Bayesian adaptive staircase searching for a 75% correct threshold. We determined these coherence thresholds for each cue type (CD or IOVD) in each chromatic condition (L+M+S, L-M and S-(L+M)).

To combine data across multiple repeats of the same condition robustly, we fitted a logistic function to the entire set of trial values for each condition and each subject and used the value at 75% correct as the threshold criterion. Taking the average value of the five staircase endpoints produced almost identical results. Group data were computed by taking the mean threshold across participants for each stimulus condition (see **Figure 2.4**, panel A).
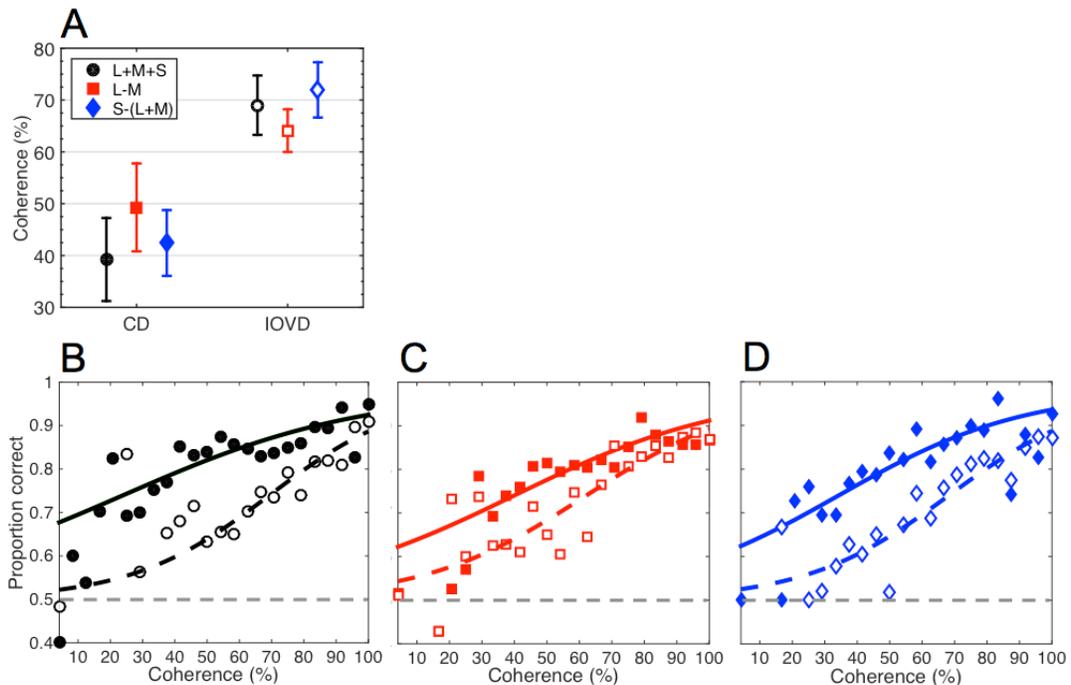
Figure 2.4 Results from the coherence thresholding experiments. Panel A shows mean coherence thresholds across participants for two different MID types (CD or IOVD) in three different chromatic conditions (achromatic, isoluminant L-M, and S-cone isolating). Error bars represent ±1 SEM. Panels B, C and D show psychometric functions fitted to all data collapsed across all participants for L+M+S (B), L-M (C) and S-(L+M) (D) CD (filled points) and IOVD (unfilled points) stimuli. The points represent proportion of correct responses at each coherence level, and the best-fitting logistic functions are plotted. The grey dashed line indicates chance level performance (50% correct).

A repeated measures ANOVA found a significant main effect of MID cue type ($F(1,4)$ = 11.29, p = .028). Participants required almost double the signal to noise ratio to perceive direction of MID from the IOVD cue than the CD cue. This implies that the IOVD mechanism is more sensitive to noise, and degrades at a faster rate than the CD mechanism. Because detection thresholds for the two MID types were identical at 100% coherence, we suggest that this difference is not due to differences in overall task difficulty between the conditions, but rather reflects intrinsic properties of the CD and IOVD mechanisms.

No significant effect of stimulus chromaticity was found ($F(2,8)$ = 0.71, p = .846). There was no interaction between chromaticity and cue type ($F(2,8)$ = 0.94, p = .430). Overall, this

result suggests that all three precortical pathways can contribute to CD and IOVD mechanisms, and the pattern of sensitivity to noise within each cue type is similar irrespective of which channel the signals are drawn from.

We fit psychometric functions to data collapsed across all trials and all participants to obtain threshold and slope estimates at the group level (**Figure 2.4**, panels B, C and D). Logistic functions were fit using the Palamedes toolbox (version 1.8.2; Prins & Kingdom, 2009) for Matlab. The logistic function fit was determined by searching for alpha (threshold) values between 0 and 100 in steps of 5, and beta (slope) values between 0.1 and 5 in steps of 0.1. A parametric bootstrap was used to find the optimal fit and determine the standard errors of the threshold and slope. Parameters from the function fits are listed in **Table 2.2**.

Table 2.2 Parameters for group logistic function fits. The alpha (threshold) at 75% correct and beta (slope) parameters for each cue type and each chromatic condition are given. Standard errors were determined using bootstrapping.

Chromatic condition	Cue type	Alpha (SE)	Beta (SE)
L+M+S	CD	24.08 (6.05)	0.03 (0.00)
	IOVD	69.39 (2.20)	0.05 (0.01)
L-M	CD	39.89 (3.29)	0.03 (0.00)
	IOVD	62.09 (2.63)	0.04 (0.00)
S-(L+M)	CD	34.32 (3.63)	0.03 (0.00)
	IOVD	68.26 (2.05)	0.05 (0.01)

The alpha (threshold at 75% correct performance) parameter for the IOVD cue is consistently higher than the alpha parameter for the CD cue. This implies that CD is more robust across a wider range of noise, as well as at lower levels of coherence.

All function fits are similar across different chromatic conditions. These data are in line with the group analysis above, indicating that the IOVD cue breaks down more rapidly with the introduction of noise irrespective of stimulus chromaticity.

2.4.3 Achromatic masking

Finally, we asked whether achromatic noise elements can interfere with the perception of MID in a chromatic stimulus. To do this, we masked chromatic signal elements with achromatic noise. We determined the percentage of L-M or S cone isolating signal elements required to perceive the direction of MID in a CD or an IOVD stimulus. **Figure 2.5** shows group mean threshold values where the direction of MID was detected at 75% accuracy. Data from the achromatic masking conditions were compared to data collected in the coherence thresholding experiment, where signal and noise elements were of identical chromaticity and contrast.
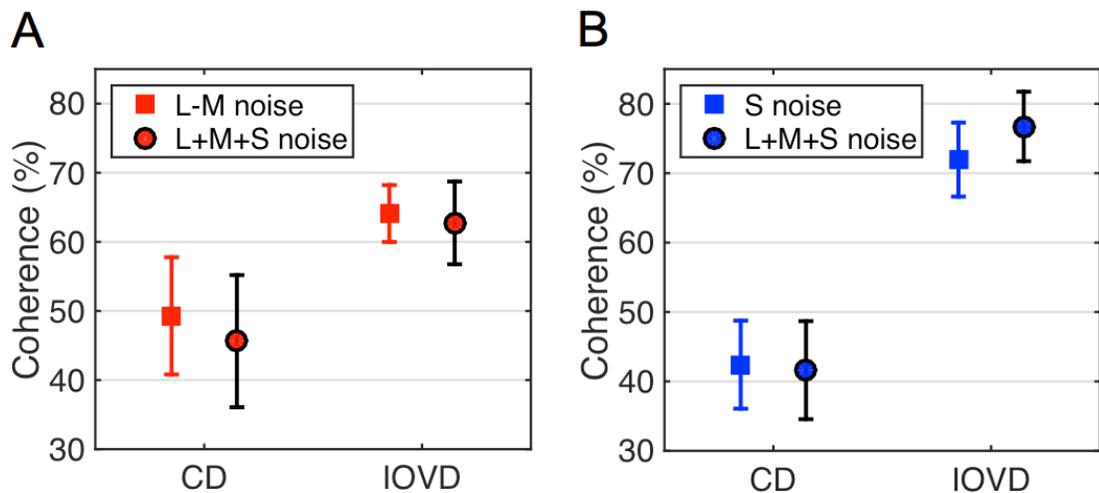
Figure 2.5 *Mean coherence thresholds across participants for two different motion in depth types (CD or IOVD) in four different chromatic conditions. Panel A shows data from stimuli with L-M isolating signal dots masked by either L-M or achromatic noise dots. Panel B shows data from stimuli with S signal dots and S or achromatic noise dots. Coherence is measured by the percentage of signal dots required to determine the direction of motion in depth with 75% accuracy. Error bars represent ±1 SEM.*

A repeated measures ANOVA found no significant effect of the chromaticity of the noise elements ($F(1,4) = 0.05$, $p = .839$). There was no difference in participants' ability to determine the direction of MID whether signal elements were masked with chromatic or achromatic noise.

We do not find evidence for an advantage to MID perception by segmenting the signal from the noise elements on the basis of chromatic identity. This supports the view that both CD and IOVD mechanisms are computed after signals from the chromatic pathways are pooled, in a manner analogous to 2D motion discrimination (Dougherty et al., 1999).

2.5   Discussion

2.5.1   Chromatic input to MID mechanisms

The relative sensitivities of two different MID cues to chromatic and achromatic signals were revealed by our contrast thresholding data. Within any single chromatic pathway, CD and IOVD mechanisms require similar levels of cone contrast to achieve the same performance level. Across pathways, sensitivity to S-cone stimuli was around one-tenth of that for L+M or L-M stimuli. This implies that while both CD and IOVD mechanisms are able to draw on chromatic information, their sensitivity to signals initiated in the S cones is relatively poor.

Our findings are similar to psychophysical and neuroimaging studies that have assessed the contribution of S cone signals to 2D motion processing. S cone stimuli have been found to convey weaker signals to 2D motion detection, and require a significant increase in contrast in order to match the motion of an L or M cone stimulus (Dougherty et al., 1999; Hawken, Gegenfurtner, & Tang, 1994; Lee & Stromeyer, 1989). This can be explained by the pattern of S cone input to motion selective areas in cortex. In order to generate a similar fMRI response amplitude in motion area hMT+, the contrast of an S cone stimulus must be increased tenfold in to match an L or an M cone stimulus (Dougherty et al., 1999) – an effect size similar to the tenfold increase required in chromatic contrast for S cone stimuli in this experiment. Thus 3D mechanisms may draw on S cone signals in a similar manner to 2D motion mechanisms.

We also asked whether achromatic noise can interfere with the perception of MID generated in a chromatic CD or IOVD stimulus. For 2D motion, it has been found that a luminance mask does not affect coherence thresholds in chromatic stimuli, although the same mask impairs sensitivity to perceiving the direction of luminance motion. This was taken as evidence for a colour-selective 2D motion detector (Cropper & Derrington, 1996).

In contrast, we found that the chromaticity of noise elements makes no difference to coherence thresholds in 3D motion stimuli. The same percentage of signal is required whether the CD or IOVD signal is masked by achromatic noise, or noise of the same chromaticity. This implies a 'colour-blind' MID mechanism that pools luminance and chromatic signals from all three precortical pathways before MID is computed.

Dougherty, Press and Wandell suggested a putative colour energy signal that passes signals to motion selective cortex. Their colour energy model draws on information across all chromatic channels, and is summarized as

$$E_C = \lambda\,(L+M)^2 + \rho\,(L-M)^2 + \beta\,(S-[L+M])^2$$

where λ, ρ and β are different weightings attributed to the L+M, L-M and S cone signals respectively (Dougherty et al., 1999). This model was originally suggested based on 2D motion judgements; we find evidence that 3D motion detectors draw on a similar colour energy signal, where the weighting of the S cone signal is around one log unit lower than that of the L+M and L-M signals. For 2D motion, the weighting of the three pathways is comparable, with the L+M signal providing the greatest input, followed by the L-M signal and a small S-cone weighting.

Our results cannot be accounted for by residual luminance leakage. Firstly, we corrected for individual variation in isoluminant points by measuring isoluminance on a subject-by-subject basis using a minimum motion technique (Anstis & Cavanagh, 1983).

Secondly, for our gamma-corrected 8-bit stimulus presentation system, the error ('splatter') onto the luminance axis when stimulus contrast is converted from LMS space to RGB values was 1 part in ~200. At a maximum contrast of 20% for our S cone stimuli, this implies less than 0.1% luminance contrast – a stimulus level that was far below the detection threshold. For L-M stimuli residual leakage was even lower (around 0.01%). As shown by our contrast thresholding data, the L+M+S MID signal in CD and IOVD stimuli becomes effective only between 2-3% contrast. Therefore, this residual luminance leakage was unlikely to contribute significantly to MID judgements.

2.5.2   Properties of the CD and IOVD mechanisms

Developing stimuli that establish a fair comparison between CD and IOVD cues is challenging, given that each cue operates optimally under different spatiotemporal characteristics (Czuba et al., 2010; Shioiri et al., 2000). A common approach is to balance stimuli on their low-level properties such as element size and stimulus field of view (Nefs et al., 2010). Our CD stimulus moved much more slowly than our IOVD stimulus, as the CD cue, in contrast to IOVD, is more sensitive at lower speeds (Czuba et al., 2010). Furthermore, we have taken the additional step of setting the contrast of our stimuli such that performance at 100% coherence is equal across cue type and chromatic condition. Thus, we were able to independently assess the effects of varying the signal to noise ratio on sensitivity to CD and IOVD cues.

Data from our coherence thresholding experiment showed, firstly, that the chromaticity of the stimulus had no effect on MID coherence thresholds. This is further evidence that MID mechanisms pool signals across all three precortical pathways, integrating chromatic and isoluminant signals.

Our participants did show a higher sensitivity to the CD cue overall, where the MID percept remained robust at relatively low levels of signal. Participants required between 40-50% signal to discriminate the direction of a CD stimulus, whereas for IOVD most participants required around 70% signal. This is consistent with other studies showing that the majority of participants are more sensitive to CD when coherence thresholds are measured (Nefs et al., 2010). These differences in sensitivity to noise further imply the independence of CD and IOVD mechanisms, suggesting that they are processed by a distinct neural circuitry.

2.5.3   Conclusions

Our results suggest that CD and IOVD mechanisms provide independent contributions to MID perception. MID mechanisms appear to draw from information encoded in all three precortical pathways, in a manner analogous to 2D motion detectors. We show that L+M+S, L-M and S-(L+M) cone signals are capable of driving both mechanisms, provided they are

scaled appropriately by contrast. Finally, we show that the IOVD mechanism is more susceptible to noise than the CD mechanism, irrespective of stimulus chromaticity.

Chapter 3.   Distinct response patterns for achromatic and S-cone inputs to cortical motion in depth mechanisms

3.1   Abstract

Motion in depth (MID) can be cued by both changes in binocular retinal disparity over time (CD), and inter-ocular velocity differences (IOVD). IOVD depends on differencing a coarse retinal motion signal, whereas CD requires the extraction of the first temporal derivative of a high resolution disparity signal. These differences suggest that CD and IOVD might be computed in very different visual pathways – with different spatial and temporal resolutions. Here, we used fMRI and cone-isolating stimuli to examine the contributions of two different precortical visual pathways to motion in depth.

Achromatic and S-cone-isolating stimuli that isolated the CD and IOVD mechanisms, as well as matched control stimuli that did not convey MID, were presented in an interleaved rapid event-related fMRI design. S-cone stimuli were set to each participant's subjective S-cone isoluminance point, determined using a minimum flicker photometry task performed in situ prior to scanning. A combination of retinotopic mapping and motion localisers were used to delineate visual areas including V3A/B, IPS-0, hMT and hMST. A whole brain analysis averaging across participants and a participant-level region of interest (ROI) analysis were carried out, using the general linear model (GLM) to model neural responses to different MID stimuli.

Overall, we found that both CD and IOVD are processed in an overlapping network of areas, which includes early visual areas, parts of the dorsal and ventral streams, and motion-selective hMT+. Crucially, however, we measured an interaction between MID type and chromaticity, where CD responses were mainly driven by achromatic inputs and IOVD responses were driven strongly by S-cone inputs.

These findings provide novel evidence for the role of S-cone signals in MID processing and suggest that CD and IOVD mechanisms can be dissociated on the basis of inputs from early precortical pathways. Specifically, they suggest that dichoptic S-cone motion signals may be combined in an opponent manner at an early stage of visual processing, and that these signals contribute to IOVD-based MID computations.

3.2   Introduction

Two binocular cues support our perception of motion towards and away from an observer (motion in depth, MID; Rushbass & Westheimer, 1961; Regan, 1993). These two cues depend on fundamentally different sources of information. The first, changing disparity (CD), is based on monitoring changes in binocular disparity over time. An object in space stimulates anatomically distinct parts of the left and right retinae, and the horizontal offsets between these two retinal images – the binocular disparity – provides a strong depth signal. Systematic increases or decreases in retinal disparity over time therefore allow us to estimate MID.

The second cue, the inter-ocular velocity difference (IOVD), is based on a comparison of binocular opponent motion vectors. As an object moves towards or away from the eyes it generates motion vectors pointing in opposing directions between the eyes. Comparing the sign and magnitude of these motion vectors provides an estimate of the speed and angle of MID.

Although both cues co-exist in the natural world, each are sufficient to generate a MID percept in isolation (Brooks, 2002; Cumming & Parker, 1994; Fernandez & Farell, 2005; Joo, Czuba, Cormack, & Huk, 2016; Nefs, O'Hare, & Harris, 2010; Rokers, Cormack, & Huk, 2008, 2009; Shioiri, Saisho, & Yaguchi, 2000; for a review of the psychophysical literature, see Harris, Nefs, & Grafton, 2008). Due to the constraints placed on the disparity and velocity computations they depend on, CD and IOVD operate optimally across reasonably distinct spatial and temporal ranges (Czuba, Rokers, Huk, & Cormack, 2010, Shioiri et al., 2008) and thus may also be subserved by dissociable neural mechanisms.

Recent neuroimaging studies emphasise the role of hMT+ in processing both CD and IOVD (Rokers et al., 2009), whilst corresponding neurophysiological evidence has identified cells tuned to 3D motion direction in this area (Czuba, Huk, Cormack, & Kohn, 2014; Sanada & DeAngelis, 2014). Although hMT+ integrates both motion and disparity cues (Movshon & Newsome, 1996; Ponce, Hunter, Pack, Lomber, & Born, 2011; Ponce, Lomber, & Born, 2008; Smolyanskaya, Haefner, Lomber, & Born, 2015), no evidence for cross-cue adaptation between CD and IOVD has been found (Joo et al., 2016). This implies that separate sub-populations of neurons are tuned to either CD or IOVD within a common network of areas (for a recent review, see Cormack, Czuba, Knöll, & Huk, 2017).

Other work suggests differences between CD and IOVD processing, both in an extended network of regions outside hMT+ as well as in the pathways that relay cues to hMT+. By comparing the fMRI response to a CD-type stimulus against the response to a static disparity plane, an area anterior to hMT+ – the putative 'cyclopean stereo-motion' (CSM) area – has been identified as the potential locus of stereo-defined MID processing (Likova & Tyler, 2007). V3A and regions in the parietal cortex, including the intraparietal sulcus (IPS), have been identified in an electroencephalography (EEG) study where responses to MID stimuli were mainly driven by disparity cues (Cottereau, McKee, & Norcia, 2014). Finally, it has been suggested that, whilst a direct motion pathway from V1 to hMT+ may subserve IOVD computations, an indirect, parallel pathway via V2 and V3 relays disparity cues from V1 to hMT+ (Ponce et al., 2008). Thus, the network of areas involved in CD motion processing may extend beyond areas involved in IOVD processing. Disparity and velocity signals may reach common MID areas, such as hMT+, via different, parallel pathways.

Cortical mechanisms underlying CD and IOVD can be dissected in greater detail by drawing on the chromatic specializations and response dynamics of early precortical pathways. Generally, motion processing is dominated by achromatic signals carried by the magnocellular (MC) pathway, which constitutes the majority of inputs to MT+ (Born & Bradley, 2005). Some achromatic inputs may also be conveyed by the parvocellular (PC) pathway (Benardete & Kaplan, 1999; Billock, 1995), whose inputs reach MT+ via V1 and V2 (Born & Bradley, 2005). However, MT+ also receives direct, subcortical inputs from the S-cone driven koniocellular (KC) layers of the LGN (Sincich, Park, Wohlgemuth, & Horton, 2004), and there is substantial evidence that S-cone isolating stimuli can convey an equivalent motion percept when differences in contrast sensitivity are accounted for (Cavanagh & Anstis, 1991;

Dougherty, Press, & Wandell, 1999; Gegenfurtner & Hawken, 1996; McKeefry & Burton, 2009; Mullen & Boulton, 1992).

The spatial resolution of S-cone signals is constrained from the front-end of the system, given the sparse tiling of the S-cones in the retina, and the lack of S-cones in the fovea (Curcio et al., 1991). In the LGN, cells in the KC layers have comparatively large receptive fields (Irvin, Norton, Sesma, & Casagrande, 1986; Norton & Casagrande, 1982; Norton, Casagrande, Irvin, Sesma, & Petry, 1988; Xu et al., 2004), though fMRI and electrophysiological measurements indicate that the relationship between spatial frequency tuning and receptive field size may break down in V1 (Movshon, Thompson, & Tolhurst, 1978; Welbourne, Morland, & Wade, 2018). Because of these properties, we hypothesised that an early, low-resolution S-cone signal may be particularly suited to conveying the coarse retinal motion vectors that are necessary for computing IOVD. Indeed, it has recently been shown that an S-cone isolating stimulus is able to induce a 3D motion after-effect generated by adapting to monocular 2D motion (Shioiri, Yoshizawa, Ogiya, Matsumiya, & Yaguchi, 2012).

Here, we used fMRI to investigate the neural correlates of binocular MID perception. We manipulated stimulus chromaticity to investigate whether achromatic and S-cone pathways differentially contribute to CD and IOVD mechanisms. Based on previous fMRI research, we expected both cue types to engage motion pathways including areas hMT and hMST (Joo et al., 2016; Rokers et al., 2009), with possible additional CD responses in parietal areas and a stereo-motion area anterior to hMST (Likova & Tyler, 2007). Additionally, we hypothesized that the S-cone pathway might be particularly suited to carrying the low-resolution motion signals required to compute IOVD.

3.3   Materials and methods

3.3.1   Participants

Participants (N=17, aged 21-45 years, seven male) with normal or corrected to normal vision were recruited. For the whole-brain analysis, data from all participants were used. For the region of interest (ROI) analysis, data from 6 participants were discarded due to poor fits in the general linear model (GLM < 5% variance explained across ROIs), leaving a final N of

11 for that analysis. Two participants were authors on this paper (MK and ARW), the rest were naïve. All participants had normal stereo-acuity (below 120 arcsec, measured using the TNO test, 19th edition, Laméris Ootech, Ede, The Netherlands) and normal colour vision (tested using Ishihara plates, 24 plate edition).

Before scanning, participants practised the isoluminance setting task and viewed high-visibility exemplars of the MID stimuli. These were 100% coherent CD or IOVD stimuli oscillating continually in depth, with identical parameters to those shown during the experiment (described in detail in a subsequent section). 'Coherence' here refers to the signal to noise ratio in the stimulus display, where in a 100% coherent stimulus all dots contribute to the motion in depth signal and there is zero noise. Participants were required to trace the MID trajectory with their fingers to indicate accurate perception of the MID. All participants were able to perceive MID from both stimulus types. Written informed consent was obtained in accordance with the Declaration of Helsinki, and the study was approved by the York NeuroImaging Centre Board of Ethics.

3.3.2   Apparatus

For pre-testing, stimuli were displayed on a VIEWpixx 3D LCD system with 1920 x 1080 pixel resolution, running at 120 Hz. Stereo presentation was achieved using wireless NVIDIA GeForce 3D vision LCD shutter goggles an infra-red emitter that synchronized the frame rate of the display with the goggles (VPixx technologies, Saint-Bruno, Canada).

During scanning, a PROpixx DLP LED projector (VPixx technologies, Saint-Bruno, Canada) at 1920 x 1080 pixel resolution and running at 120 Hz was used to back-project stimulus images on to a silver screen positioned behind the participant. Stereoscopic stimulus presentation was achieved using a circular polarizer (DepthQ Polarization Modulator, VPixx technologies, Saint-Bruno, Canada) placed in front of the long-throw lens and passive 3D glasses worn by the participant. Stimuli were viewed on a first-surface mirror mounted on the head coil (57cm viewing distance, including the optical pathway of the mirror), yielding a viewing angle of 41° × 23.5°. Maximum luminance, as measured through the polariser and glasses, was 390 cd/m^2.

Both display systems were photometrically calibrated using a fibre-optic photospectrometer (Ocean Optics, Dunedin, FL) which measured both the gamma and the spectral irradiance of each R, G and B channel as seen by each eye. To achieve this, the cable tip was positioned behind the goggles using a polystyrene mannequin head to match the viewing distance and position used in the experiments and behavioural testing. Left and right eye measurements were taken, and as there were no significant differences between the eyes, an average was taken for colour calibration.

Stimulus presentation was controlled from a Shuttle PC with Intel Core i7-4790K processor at 4.0 GHz and an NVIDIA GeForce GTX970 graphics card with 4 GB DDR5 memory. All stimuli were designed and run from Matlab 8.5.0 (2015a; The MathWorks Inc., Natick, MA, USA) in conjunction with the Psychtoolbox 3.0.12 routines (Brainard, 1997; Pelli, 1997). Behavioural responses and scanner trigger pulses to synchronise stimulus onset were transmitted using a four-button fibre-optic response pad (Current Designs, Philadelphia, PA).

3.3.3   Stimulus design

We designed stimuli to isolate CD and IOVD cues independently. We also generated appropriate null 'controls' for each stimulus type. Control stimuli matched the low-level properties of the MID stimuli but did not contain cues to generate a coherent percept of MID.

All stimuli shared some basic parameters. They were all variants of dynamic random dot stereograms (DRDS; Julesz, 1971; Julesz & Payne, 1968), where anti-aliased dots were 0.5° in diameter, within a cosine envelope that gradually smoothed the edges over 0.15°. Dots were pseudo-randomly positioned on a mean luminance grey background (390 cd/m^2). The dot centres were at least 0.5° apart in any direction, and dots were assigned with a .5 probability to be either positive or negative contrast polarity – for achromatic stimuli, this was along the L+M+S colour axis, and for S-cone stimuli, this was along the S-(L+M) colour axis. S-cone dots were displayed at the maximum possible contrast given the display gamut (90% on our system). To balance the extent to which this cone contrast drives the blood oxygenation level dependent (BOLD) signal in the early visual cortex (Wandell et al., 1999), the achromatic dot contrast was set to 10% of this value (9%). The DRDS was viewed through a circular aperture with edges smoothed by a Gaussian kernel (0.5° FWHM) with a 0.5° inner and a 5° outer radius. A fixation cross (0.2° wide/high) was placed at the centre of

the annulus to control eye position. Central (0.4° radius, centred around fixation) and peripheral (11.75° from fixation) achromatic fixation rings helped stabilise the MID percept. Stimuli were presented for 3s with a cosine ramp to avoid signal transients in the timecourse, and the stimulus was at peak contrast for 1.5s. The binocular view of the stimulus and a representation of the appearance of achromatic and S-cone isolating dots is illustrated in **Figure 3.1**, panel F.
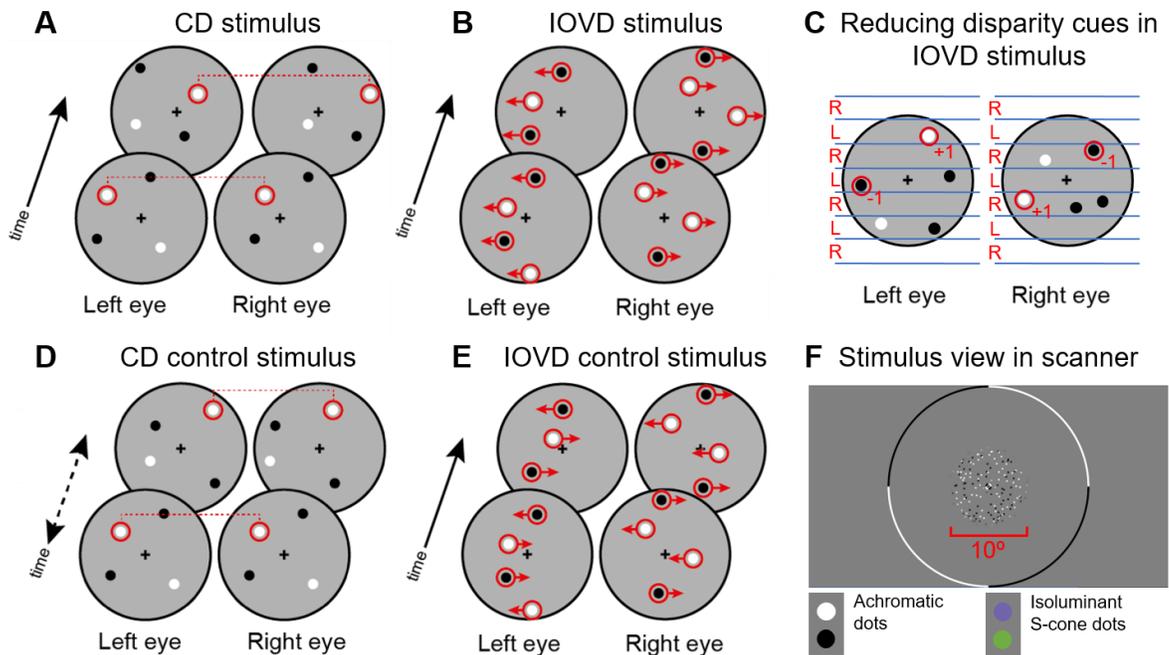
Figure 3.1 *Stimulus design for the MID fMRI experiment. MID stimuli isolated mechanisms based on changes in retinal disparity over time (CD; panel A) or interocular velocity differences (IOVD; panel B). Disparity cues were eliminated from the IOVD stimulus using a combination of decorrelation, anticorrelation, and spatial alternation of dot patterns in the left and right eyes (panel C). Control stimuli were matched to the MID stimuli on low-level properties but nulled the MID cue (panels B and E). All stimuli were presented with either achromatic or S-cone isolating dot patterns (panel F).*

3.3.4   CD stimulus

The CD stimulus generated a MID percept by systematically increasing and decreasing the retinal disparity between pairs of dots in the left and right eye (**Figure 3.1**, panel A). The stimulus oscillated sinusoidally in depth at a frequency of 1.4Hz, with a maximum of ±24 arcmin disparity (±12 arcmin shift per eye). This is well within the limits of the disparity mechanism for both achromatic and S-cone stimuli (around ±32 arcmin; (Wilson, Blake, & Pokorny, 1988). The location of each pair of dots was refreshed with each frame (refresh rate

120 Hz), and dots were positioned pseudorandomly with a monocular density of 1 dot/deg^2. Thus, the CD stimulus contained binocular correlations, but no monocular correlations over time, eliminating any coherent lateral motion (or IOVD) from the stimulus. This generated a percept of a plane of dots oscillating sinusoidally through depth.

3.3.5   CD control stimulus

To generate the CD control stimulus, individual frames in the pre-generated CD stimuli were shuffled over time (**Figure 3.1**, panel D). This preserved the retinal disparity information in each frame but eliminated the smooth changes in disparity over time that generate the MID percept. Thus, on average, the CD control stimulus contained the same range of retinal disparities but did not convey MID.

3.3.6   IOVD stimulus

The IOVD stimulus consisted of dots that were moving in opposite directions between the left and the right eyes, creating motion signals in each eye that were equal in magnitude but opposite in direction (**Figure 3.1**, panel D). Dot patterns were unpaired ('decorrelated') between the eyes, with a monocular dot density of 1 dot/deg^2. The stimulus oscillated sinusoidally in depth at a frequency of 1.1Hz with a maximum lateral shift of ±200 arcmin between the eyes (±100 arcmin monocular horizontal displacement). Each dot had a maximum lifetime of 50ms, and visual transients were avoided by regenerating the same number of dots in new locations in each video frame. The perceptual quality of the IOVD stimulus was that of a cloud of dots oscillating towards and away from the observer, with no concrete sense of position in depth due to the lack of depth-from-disparity cues (Rokers et al., 2008).

A significant challenge in designing IOVD stimuli is to eliminate the possibility of binocular matches that could result in CD leakage (Peng & Shi, 2014). Previously, this has been achieved by anticorrelating the contrast polarity of binocular dot pairs (Czuba et al., 2010; Joo et al., 2016; Rokers et al., 2008, 2009), thereby degrading the disparity cue (Cogan, Kontsevich, Lomakin, Halpern, & Blake, 1995; Cumming, Shapiro, & Parker, 1998; Neri, Parker, & Blakemore, 1999). Alternatively, left and right eye displays can be divided into

'stripes', where dots are presented in alternating bands in the left and right eyes (Shioiri, Nakajima, Kakehi, & Yaguchi, 2008; Shioiri et al., 2000). Finally, dot patterns can be decorrelated between the left and right eyes (Nefs et al., 2010).

In this study, we combined all three of these approaches (**Figure 3.1**, panel C). Displays were divided into stripes, and decorrelated dot patterns were shown in alternate stripes between the two eyes. If two dots fell into close proximity at the borders of these stripes between the eyes, their contrast polarity was anticorrelated. In this manner, the CD cue was effectively eliminated in the IOVD stimulus.

3.3.7   IOVD control stimulus

The control for the IOVD stimulus contained the same lateral motion energy and low-level stimulus properties as its counterpart but did not convey any MID. Dots moved in both directions within a single eye, nulling the opponent motion signal between the eyes that generates the MID signal (**Figure 3.1**, panel E). Other aspects of the stimulus, such as the decorrelation of elements between the eyes, speed and lifetime of the dots, and the dividing of the stimulus into stripes, were identical to the IOVD stimulus.

3.3.8   Isoluminance setting

Stimuli were specified initially in LMS cone-excitation space. Matrices for the conversion from LMS to RGB values were given as a product of the Stockman and Sharpe (2000) 10° fundamentals for the L, M and S-sensitive cones and the spectral power distribution of the RGB phosphors for each eye. Because there are significant individual differences in macular pigment density, S-cone stimuli were adjusted to each participants' subjective point of isoluminance using heterochromatic flicker photometry (Walsh, 1953). This was performed *in situ* before scanning commenced. For this task, participants viewed a field of dots presented to either the left or the right eye. Dots alternated at 7.5 Hz between positive (violet) and negative (lime) contrast polarity along the S-(L+M) colour axis. Within each run, participants made small adjustments to the amount of L+M contamination until the minimum amount of flicker was perceived. Dots had a circular profile (0.5° diameter) and were positioned pseudorandomly with a density of 1 dot/deg^2, where each dot centre was separated by at

least 0.5°. The field of dots was viewed though a hard-edged annular window with a 1° inner radius around fixation and a 6° outer radius. The position of dots was refreshed with each left or right eye trial but stayed in the same position for each set of adjustments made by the participant. Participants completed three sets of adjustments for each eye separately.

The average isoluminance setting for each eye and each participant was used to specify the S-cone dots in the fMRI experiment.

3.3.9   MRI parameters

High-resolution anatomical T1-weighted scans (TR = 7.8ms; TE = 3.0ms; TI = 600ms; flip angle = 20°; FOV = 25.6 x 25.6 cm; matrix size = 256 x 256; voxel resolution = 1.0 x 1.0 x 1.0mm; 176 coronal slices to cover the whole head) were taken in a separate scanning session and were collected on a 3T SIGNA HDx Excite MRI scanner with an 8-channel whole-head phased-array coil (MRI Devices Corporation).

Functional data were collected on the same scanner with a 16-channel half-head phased-array coil (Novamed) to improve signal-to-noise in the occipital lobe. For the experimental session, the scan sequence was as follows: one high-resolution T1 anatomical reference scan for registration of the functional data to MNI space was collected using the same slice prescription as the functional data (TR = 2100ms; TE = 8.6ms; flip angle = 12°; FOV = 19.2 x 19.2cm; matrix size = 512 x 512; voxel resolution = 0.38 x 0.38 x 2.5mm; 39 quasi-axial, contiguous slices oriented parallel to the calcarine sulcus and covering the occipital lobe). Standard gradient-echo EPI scans for functional data collection included two runs with motion localiser stimuli (TR = 3000ms; TE = 30ms; flip angle = 90°; 124 TRs including 4 dummy volumes; FOV = 19.2 x 19.2cm; matrix size = 96 x 96; voxel resolution = 2.0 x 2.0 x 2.5mm) and seven runs with MID stimuli (TR = 3000ms; TE = 30ms; flip angle = 90°; 114 TRs including 4 dummy volumes; FOV = 19.2 x 19.2cm; matrix size = 96 x 96; voxel resolution = 2.0 x 2.0 x 2.5mm).

3.3.10   fMRI procedure and task

Before scanning, participants completed the isoluminance task detailed above. The first two functional scans were motion localiser scans designed to tease apart hMT and hMST from within the hMT+ complex (Fischer, Bülthoff, Logothetis, & Bartels, 2012; Huk, Dougherty, & Heeger, 2002). Moving and static stimuli were presented in a blocked design, where the four stimulus conditions (full-field coherent radial motion, coherent radial motion restricted to the left or right hemifield, and static dots) were presented for 12s each, followed by a 12s blank fixation-only block. Six such cycles were completed in each fMRI scan, with each run taking 6 minutes.

Following the motion localisers, participants completed seven fMRI runs where MID stimuli were presented. The 9 stimulus conditions (CD achromatic, CD achromatic control, CD S-cone, CD S-cone control, IOVD achromatic, IOVD achromatic control, IOVD S-cone, IOVD S-cone control, blank fixation-only) were presented in a rapid event-related design, with inter-stimulus intervals determined using Optseq2 (Dale, 1999). Each stimulus was presented for 3s with a cosine ramp to avoid transients in the timecourse, and the inter-stimulus interval varied between 3 and 12s. The fixation cross and two fixation rings were presented throughout the whole scan to maintain stable fixation. There were five repeats of each condition in each run, giving a total of thirty-five repeats of each stimulus condition across all seven fMRI runs. Each run took 5 minutes 42 seconds.

During all fMRI scans, participants completed a challenging attention task at fixation to control eye position and the allocation of spatial attention. The fixation cross alternated between two different shades of grey, given by the RGB values [0 0 0] and [0.7 0.7 0.7]. These changes occurred at intervals drawn randomly from a uniform distribution ranging between 1500 – 7500ms. Participant were required to track these subtle changes by pressing alternate buttons on a response pad.

A

Phase map   Eccentricity map

B

Full field > Static   Left hemifield > Static   Right hemifield > Static

-log(*p*)

-log(*p*)

-log(*p*)

-30   -15   0   15   30

C

Wedge   Ring

D

FFM   LHM   RHM   Static
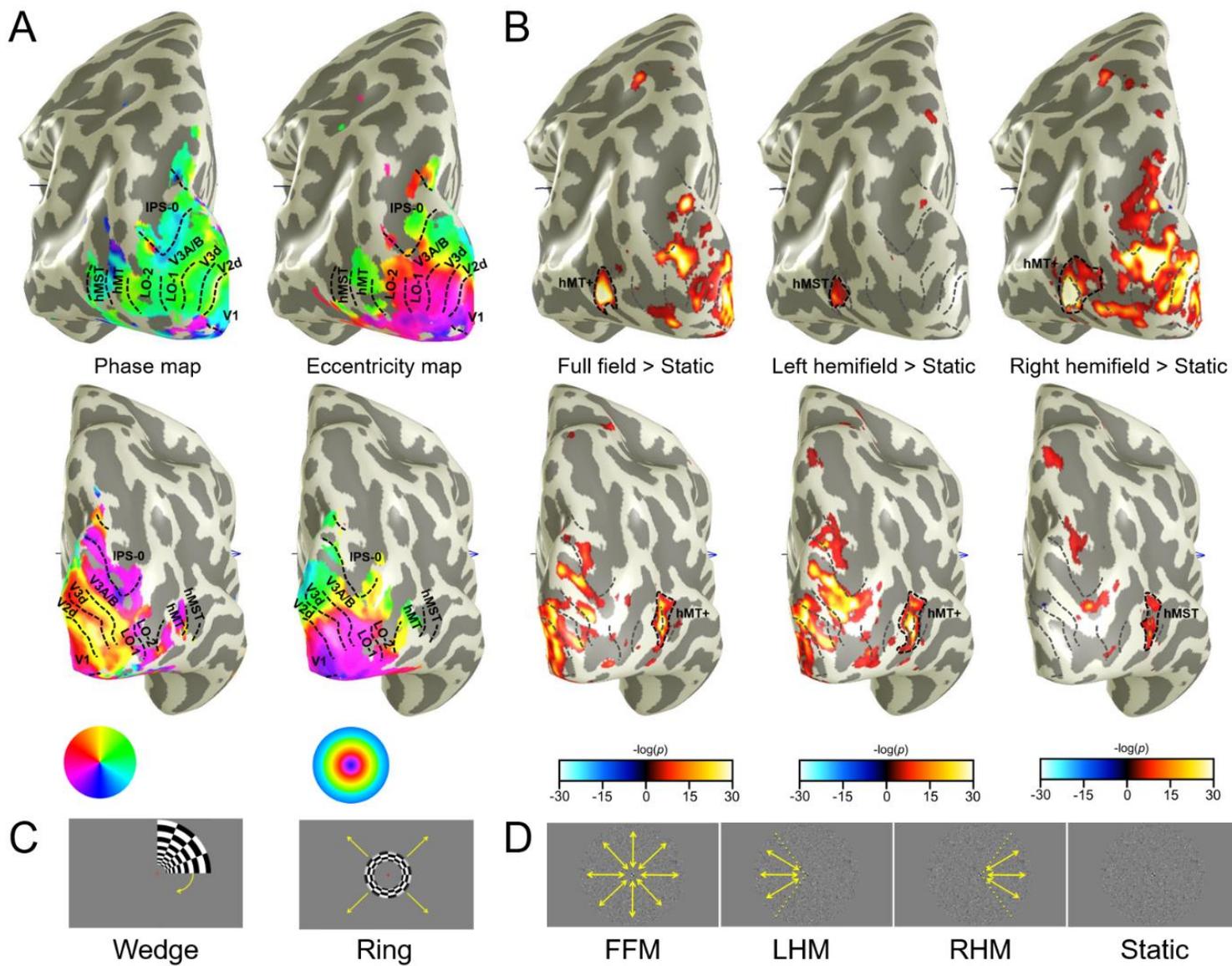
Figure 3.2 *ROI definition in an example participant, in the left (top row) and right (bottom row) hemispheres. V1, V2, V3, V4, V3A/B, IPS-0, LO-1, and LO-2 were defined based on the characteristic phase reversals in phase and eccentricity retinotopic maps (panel A, coherence threshold at 0.4), using periodically rotating wedge or expanding ring stimuli (panel C). hMT and hMST were identified using motion localisers (panel B, threshold set at p > .00001) with full-field motion (FFM), left hemifield motion (LHM), right hemifield motion (RHM) and static stimuli (panel D). Contrasting left hemifield motion, and right hemifield motion with static dots revealed subsets of voxels in hMT+ that were assigned to hMST in the ipsilateral hemisphere. Note that in the contralateral hemisphere, these contrasts often revealed a larger extent of activation than the full-field motion > static comparison in hMT+ (clearly seen in the top row of panel B).*

3.3.11 Mapping regions of interest

Regions of interest (ROIs – V1, V2, V3, V4, V3A/B, IPS-0, LO-1, LO-2, hMT and hMST) were mapped on an individual level using a combination of retinotopic mapping (e.g. Wandell & Winawer, 2011, and Wandell, Dumoulin & Brewer, 2007) and motion localisers (Fischer et al., 2012; Huk et al., 2002). Both techniques are illustrated in **Figure 3.2**. V1, V2, V3 (Dougherty et al., 2003; Schira, Tyler, Breakspear, & Spehar, 2009; Sereno et al., 1995), V4 (Brewer, Liu, Wade, & Wandell, 2005; Hansen, Kay, & Gallant, 2007; Wade, Brewer, Rieger, & Wandell, 2002; Winawer, Horiguchi, Sayres, Amano, & Wandell, 2010), LO-1, LO-2 (Larsson & Heeger, 2006), V3A/B and IPS-0 (Press, Brewer, Dougherty, Wade, & Wandell, 2001; Swisher, Halko, Merabet, McMains, & Somers, 2007; Tootell, Hadjikhani, Hall, et al., 1998) were determined based on characteristic phase reversals in response to standard retinotopic mapping stimuli (typically, each voxel's average response across 3 to 5 scans consisting of 8 cycles of a rotating checkerboard wedge or an expanding ring) collected in a separate scan session (**Figure 3.2**, panel A).

To avoid conflating the stimulus-driven response in V1, V2 and V3 with negative BOLD effects in the periphery (Shmuel et al., 2002; Wade & Rowland, 2010), we restricted these early ROIs to the retinotopic extent that corresponded to the size of the MID stimuli. This was done using a contrast map from the MID fMRI experiment, comparing the BOLD response to all stimuli (MID and control stimuli of both chromaticities) against fixation. We refined the restricted ROIs using the eccentricity maps from retinotopic data to ensure correspondence with the known stimulus size.

Motion sensitive ROIs were identified using a motion localiser designed to identify the hMT+ complex and segregate it into its hMT and hMST subcomponents (**Figure 3.2**, panel B). It was modelled on hMT/hMST localisers described previously (Fischer et al., 2012; Huk et al., 2002; Maloney et al., 2013). Briefly, moving black and white dots on a mean grey background (density 9.9 dots/deg^2, smoothed Gaussian profile σ = 0.04°, dot speed 5.3° s^{-1}) either filled an annulus extending from 0.5° - 11.75° eccentricity or were constrained to the left or right 120° of the display. Responses to these motion stimuli were contrasted against responses to a static dot stimulus consisting of randomly selected frames from the full-field motion stimulus, updating at 0.33 Hz. In hemifield motion conditions, the sections of moving dots were embedded in this static dot pattern. Stimuli were shown for 12 s in a blocked design, where each full cycle of stimuli (full-field motion, left hemifield motion, right hemifield motion, static dots) were interleaved with the blank fixation-only block. There were six stimulus cycles per fMRI run. The same central fixation task used during the MID scans was used here.

The BOLD response across visual areas was modelled using a General Linear Model (GLM). Contrasting the response between full-field motion against static conditions resulted in strong activations in V3A/B, IPS0 and hMT+.

As in earlier visual areas, neurons in hMT receive inputs primarily from the contralateral visual hemifield. However, the receptive fields of neurons in hMST extend into the ipsilateral hemifield. Therefore, these two areas can be dissociated based on their differential responses to ipsilateral motion (Huk et al., 2002). For example, contrasting left hemifield motion against static resulted in strong activations in hMT+ in the right hemisphere but only in a subset of voxels in the hMT+ complex in the left hemisphere. These left hemisphere voxels were assigned to hMST, whereas the remaining voxels where assigned to hMT. After these subdivisions were made, we iteratively refined the borders of motion-sensitive ROIs using each subject's retinotopic data.

Additionally, we defined a putative 'cyclopean stereo-motion' (CSM) area based on Talairach coordinates given by Likova and Tyler (2007). This area, anterior to hMT+, has been suggested as the main locus for CD processing. We used coordinates that were determined in their paper using a cyclopean stereo-motion localiser, where responses to a CD type stimulus were contrasted against responses to a static plane of dots at zero disparity (Likova & Tyler, 2007). Talairach co-ordinates were [-42.9 -65.9 1.1] in the left hemisphere,

and [44.4 -61.9 0.1] in the right hemisphere. To draw this ROI we grew a 5 mm sphere centred on these coordinates.

Combining retinotopic mapping with motion localisers allowed for the reliable definition of V1, V2, V3, V4, LO-1, LO-2, V3A/B, IPS-0, hMT and hMST in all of our participants, with the addition of a CSM area defined using Talairach co-ordinates. These ROIs were chosen on the basis of previous fMRI studies implicating the hMT+ complex and surrounding areas in MID perception (Joo et al., 2016; Likova & Tyler, 2007; Rokers et al., 2009).

3.3.12 Whole-brain analysis

For the whole-brain analysis, fMRI data were processed using a standard FEAT pipeline (version 6.00, part of the FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). The first four dummy volumes were deleted to account for initial changes in signal intensity before equilibrium was reached. Non-brain structures were removed from each functional scan using BET (Smith, 2002), and signal intensity was normalised across each 4D dataset by a multiplicative factor of the grand mean. Motion correction was applied using MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002). The time-series of each voxel was temporal high-pass filtered to remove slow signal drift (Gaussian-weighted least-squares straight line fitting, sigma = 50.0s) and smoothed using a Gaussian kernel at 3mm FWHM. To register fMRI data to a standard-space image, the structural reference scan taken with the same slice prescription as the functional runs was FAST-corrected (Zhang, Brady, & Smith, 2001) to improve signal drop-off at the front of the head, and non-brain tissue was removed. This image was then aligned to the MNI-152 2mm brain using FLIRT (Jenkinson et al., 2002; Jenkinson & Smith, 2001) and the transformation matrix was applied to the corresponding fMRI datasets.

A general linear model (GLM) with nine predictors for each stimulus type was applied to each 4D dataset using FILM (Woolrich, Ripley, Brady, & Smith, 2001) with local autocorrelation correction. Events were convolved with a standard Gamma function (3 s std, 6 s lag) to model the BOLD response, and the resulting beta weights gave estimates of each voxel's response to a particular stimulus. A mixed-effects analysis was carried out to combine data across scans and participants using FILM (FMRIB's Local Analysis of Mixed Effects; Beckmann, Jenkinson, & Smith, 2003; Woolrich, 2008; Woolrich, Behrens, Beckmann,

Jenkinson, & Smith, 2004) and single group averages were generated. The resulting Z statistic images for a predetermined set of contrasts were cluster corrected at a significance level of $p < .050$.

3.3.13 ROI analysis

For the individual-level ROI analysis, data were processed in mrVista (https://web.stanford.edu/group/vista/cgi-bin/wiki/index.php/Software; Vista Lab, Stanford University) and Matlab 8.5.0 (2015a; The MathWorks Inc., Natick, MA, USA). Four dummy volumes were discarded from the fMRI timecourse, and motion correction was carried out within and between scans. fMRI data were aligned to a high-resolution anatomical scan taken in a separate scan session, using the FAST-corrected and BET-extracted reference anatomical scan to transform the low-resolution data to the high-resolution T1 anatomy. Alignment between the reference anatomical scan and the high-resolution T1 was achieved using the Nestares algorithm (Nestares & Heeger, 2000). For volume- and surface-based reconstructions, grey and white matter segmentation of the high-resolution T1 scans were carried out using automated algorithms implemented in Freesurfer version 5.3. Using this segmentation, fMRI activation was restricted to the grey layers of individual brains.

A GLM analysis was carried out on grey-layer voxels by convolving event sequences for 9 different stimulus types with a 'difference of Gammas' (from the SPM 8 toolbox, http://www.fil.ion.ucl.ac.uk/spm/) hemodynamic response function (3s std, 6s lag) and fitting the modelled timecourse to the timecourse of each voxel. This yielded 9 beta weights corresponding to 9 stimulus types for each voxel. After ROI definition, the beta weights from each voxel were extracted. To generate estimates for the responses to specific stimulus types, the responses to control stimuli were subtracted from responses to MID stimuli, yielding estimates for responses to achromatic CD (achromatic CD – achromatic CD control), achromatic IOVD (achromatic IOVD – achromatic IOVD control), S-cone CD (S-cone CD – S-cone CD control) and S-cone IOVD (S-cone IOVD – S-cone IOVD control) in each participant and each ROI. The GLM variance explained for each voxel in each ROI was extracted in a similar way, and data from participants where the mean variance explained across ROIs was less than 5% were discarded (N = 6 of a total of 17).

3.4   Results

3.4.1   Whole-brain results

Group results from the whole-brain, mixed-effects analysis are shown in **Figure 3.4**. Z-statistic maps were generated by comparing the response to MID stimuli against their respective controls, for both achromatic (top row) and S-cone isolating (bottom row) conditions. For panels A and D, responses from both motion types (CD and IOVD) were taken together, whilst panels B, C, E and F show responses broken down by MID type and chromaticity.

A widespread network of areas involved in the computation of achromatic MID cues can be seen in panel A. This network includes visual areas V1, V2, V3, V3A/B, as well as ventral area V4 and motion-selective hMT and hMST. Activity extends dorsally to visually-driven areas in the intraparietal sulcus, including IPS-0.

In comparison, the network of areas involved in the S-cone MID response is restricted to earlier visual areas, and responses in dorsal areas such as IPS-0, hMT and hMST are weaker or absent.

The group maps hint at an interaction between the chromaticity of the input and the MID cue type. The achromatic MID response shown in A appears largely driven by the achromatic CD response shown in panel B, where the activation patterns are very similar. The achromatic IOVD map (C) is sparse; conversely, the S-cone IOVD response (F) is stronger and appears similar to the overall S-cone MID response (D). In this case, the S-cone CD response (E) is weaker than the achromatic CD response (B). To quantify these differences, an individual-level ROI analysis was carried out in which responses to each MID type and chromaticity were extracted and compared.
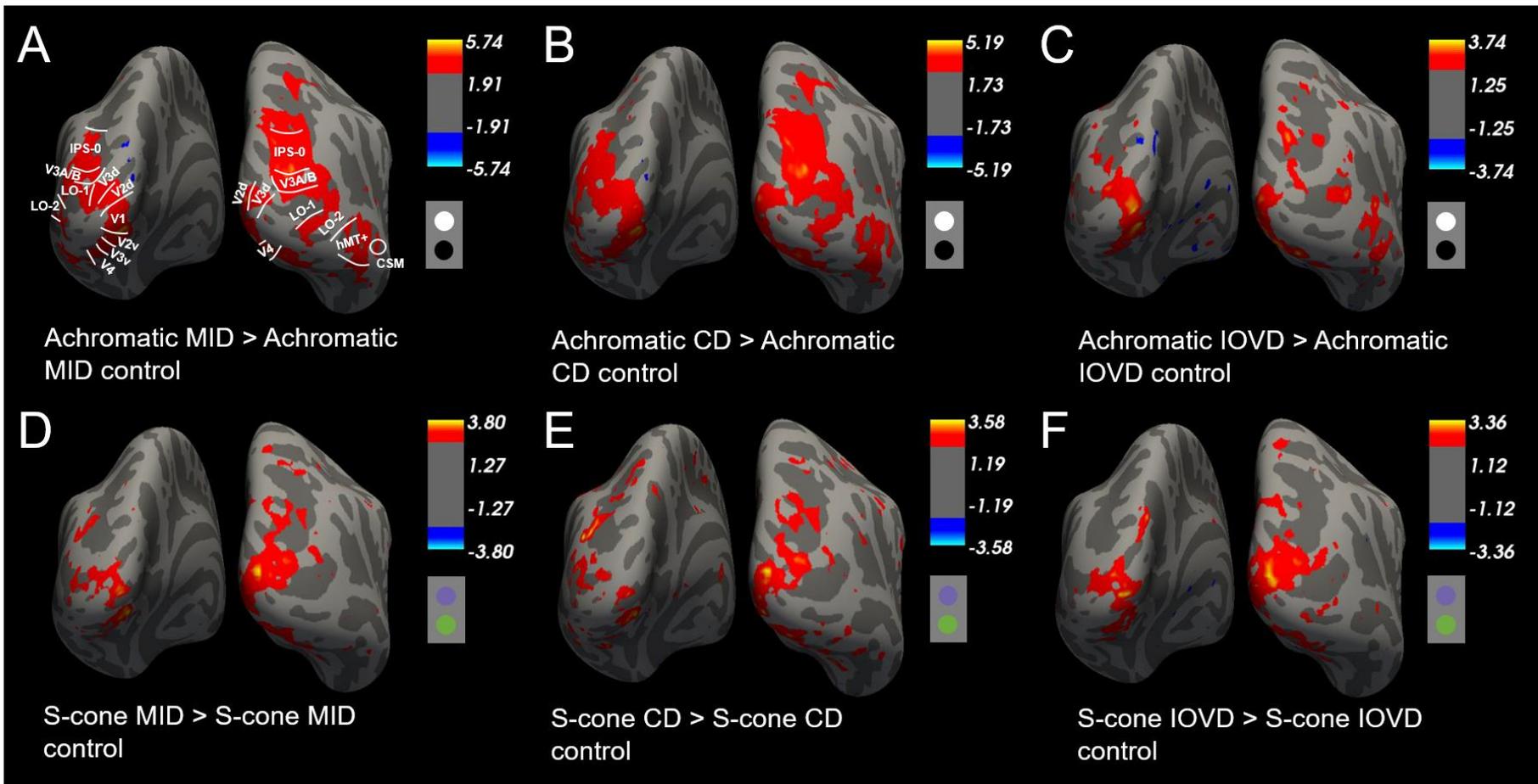
Figure 3.3 *Results from the whole-brain analysis showing group-level Z-statistic maps for responses to achromatic (A, B, C) and S-cone (D, E, F) MID stimuli. Dots below the scale bar illustrate the stimulus chromaticity. Panels A and D show the combined responses to both CD and IOVD stimuli, compared to the combined response to CD and IOVD control stimuli. Panels B and E show the CD response, and panels C and F show the IOVD response*

3.4.2   Region of interest results

Results from the ROI analysis are shown in **Figure 3.4**. Beta weights representing the response to each of the nine stimulus conditions were extracted for each participant, and mean responses to the MID stimuli were calculated by subtracting the modelled response to each control stimulus from the modelled response to each MID stimulus ($\Delta\beta$). These differences were entered into a 10 x 2 x 2 repeated measures ANOVA modelling the response in ten ROIs (V1, V2, V3, V3A/B, IPS-0, V4, LO-1, LO-2, hMT and hMST) for two chromaticities (achromatic and S-cone) and two MID types (CD and IOVD).

We asked whether MID mechanisms can be driven equally by CD and IOVD cues. If both cues are extracted within a similar network of areas, as suggested by the group maps, we would expect both CD and IOVD stimuli to elicit a similar BOLD response in each ROI. In line with this, the ANOVA found no main effect of MID type ($F(1, 10) = 1.46$, $p = .255$, ε^2 partial = .13, Greenhouse-Geisser correction for sphericity), implying that on average there was no difference between the effects of the two cues. In addition, we found no significant interaction between MID type and ROI ($F(3.00, 26.66) = 0.67$, $p = .576$, ε^2 partial = 0.06, Greenhouse-Geisser correction for sphericity). This indicates that, in line with our hypothesis, a similar network of areas is involved in computing MID from disparity and velocity mechanisms.

If these MID mechanisms were to be driven largely by achromatic inputs, we would expect to see a significant main effect of chromaticity or an interaction between chromaticity and ROI. If this were the case, the response to achromatic MID stimuli would be higher in some or all ROIs. However, the ANOVA found **no** main effect of chromaticity ($F(1, 10) = 0.06$, $p = .815$, ε^2 partial = 0.01, sphericity assumed) and no significant interaction between chromaticity and ROI ($F(2.66, 26.67) = 2.08$, $p = .132$, ε^2 partial = 0.17, Greenhouse-Geisser correction for sphericity). These results have two main implications. Firstly, that by contrast-scaling our stimuli, we succeeded in balancing the extent to which achromatic and S-cone inputs drive the BOLD response across ROIs, thereby avoiding bias by favouring either pathway. Secondly, that when this bias is avoided, MID mechanisms can be driven by both achromatic and S-cone inputs.
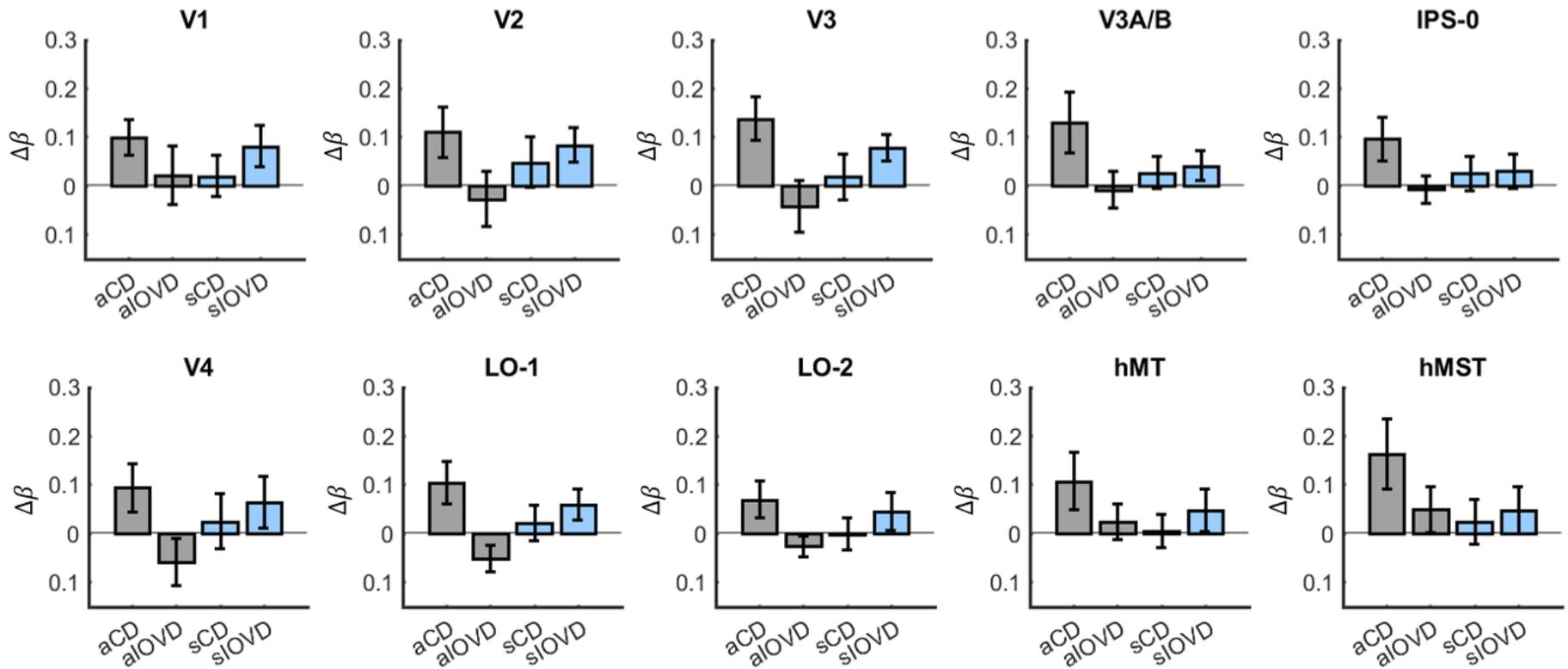
Figure 3.4 *Results from the ROI analysis showing relative responses to different MID stimuli in a network of areas. For each data point, beta weights were extracted from voxels in each ROI in each participant. The difference, $\Delta\beta$, in mean response across voxels to each MID control stimulus was subtracted from the mean response to the respective MID stimulus. This difference is taken as the mean response to achromatic CD (aCD), achromatic IOVD (aIOVD), S-cone CD (sCD), and S-cone IOVD (sIOVD) stimuli. Y-axis values greater than zero represent a stronger response to the MID stimulus, and values less than zero represent a stronger response to the MID control stimulus. Error bars are ±1 SEM.*

Thus, there appears to be no overall difference in the networks of areas involved in processing CD and IOVD and no overall difference in the extent to which achromatic and S-cone information can contribute to MID. However, clearly, the sources of information – disparity and velocity – for both MID types are vastly different. Is there then, a difference in the manner in which the early visual pathways convey these sources of information?

Crucially, the ANOVA revealed a significant interaction between MID type and chromaticity, across all ROIs ($F(1, 10) = 10.31$, $p = .009$, ε^2 partial = 0.51, sphericity assumed). This indicates a dissociation of the chromatic inputs into the MID mechanisms. The CD response was larger when it was driven by achromatic input, but the IOVD response was greater when it was driven by S-cone input. This pattern was consistent across ROIs, implying that whilst the general network of areas involved in processing CD and IOVD are similar, the two cues can be differentiated on the basis of early chromatic inputs.

To clarify this finding, results from different ROIs were averaged and grouped (**Figure 3.5**). The response to each control stimulus was subtracted from the response to its respective MID stimulus, and the resulting beta differences ($\Delta\beta$) were grouped into early visual areas (V1, V2 and V3), dorsal areas (V3A/B and IPS-0), ventral areas (V4, LO-1 and LO-2) and motion areas (hMT and hMST). Within each group of ROIs, the differences between chromatic and achromatic stimuli for the same MID type were compared using a paired-samples t-test.

Both CD and IOVD stimuli elicited reliable responses across all four grouped ROIs – but the consistency of this response was dependent on chromaticity. For the CD mechanism, this response was driven by achromatic input (see figure 5, top row). For the IOVD mechanism, it was the S-cone stimulus that resulted in reliable responses (Figure 5, bottom row). Thus, although both types of MID are processed in a similar network, they appear to be optimally conveyed by different chromatic mechanisms.

The S-cone contribution to the CD response was weak in early visual and dorsal areas and negligible in ventral and motion-selective ROIs. Paired *t*-tests comparing the S-cone and the achromatic CD responses showed stronger contributions from the achromatic pathway in

early visual areas ($t(32) = 3.10$, $p = .004$), dorsal areas ($t(21) = 2.17$, $p = .041$), ventral areas ($t(32) = 2.53$, $p = .017$) and motion areas ($t(21) = 2.78$, $p = .011$).
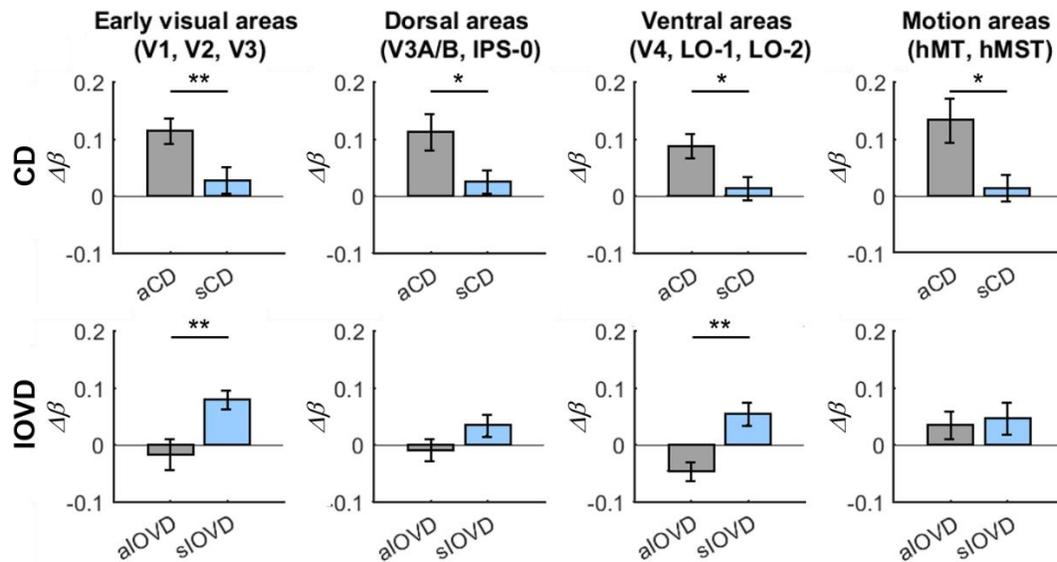
Figure 3.5 *Results from grouped ROIs showing the difference between MID and control stimuli for CD (top row) and IOVD (bottom row) mechanisms. Responses to achromatic stimuli are illustrated in grey, and responses to S-cone stimuli are illustrated in blue. Y-axis values (beta difference, $\Delta\beta$ = MID beta value – MID control beta value) greater than zero represent a stronger response to the MID stimulus, and values less than zero represent a stronger response to the MID control stimulus. Error bars are ±1 SEM. Asterisk annotation: * p < .050, ** p < .010*

In contrast to this, S-cone stimuli consistently elicited a stronger IOVD response than achromatic stimuli did – a pattern that was particularly striking in early visual and ventral areas. Here, paired t-tests revealed significantly larger S-cone responses than achromatic responses ($t(32) = -2.85$, $p = .008$ for early visual areas, and $t(32) = -3.44$, $p = .002$ for ventral areas).

In comparison to this dominant S-cone input, achromatic contributions to the IOVD mechanism were weak. In early visual areas and in dorsal areas, the achromatic IOVD response was at zero. In ventral areas, the achromatic IOVD responses was negative, implying that these areas respond more strongly to the control stimulus which contained lateral motion energy but no MID. In fact, contributions from the achromatic pathway to the

IOVD mechanism emerged only in the motion-selective ROIs, where stimulus-evoked responses were roughly equal regardless of chromaticity (t(21) = -0.36, p = .725).

3.4.3   Analysis of the 'cyclopean stereo-motion' area

Finally, we analysed the activation pattern to achromatic and S-cone CD and IOVD stimuli in the CSM (cyclopean stereo-motion) ROI. This region was first described by Likova and Tyler (2007), who measured a strong CD motion response in an area anterior to hMT and hMST. The CSM can be localized using Talairach coordinates provided in their paper (Likova & Tyler, 2007). Using this approach, we extracted beta weights for responses in the CSM ROI using the same analysis pipeline described in the previous section, and found that overall response amplitudes were weaker than those measured in hMT and hMST. Notably, the response to achromatic CD stimuli was significantly lower in the CSM than in hMST (t(10) = 2.77, p = .020, paired samples t-test). Results are shown in **Figure 3.6**.
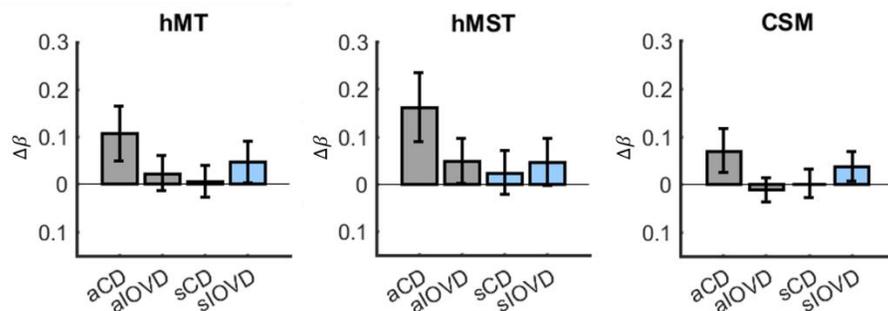
Figure 3.6 *Results from the hMST and CSM ROIs showing response patterns to achromatic CD, achromatic IOVD, S-cone CD, and S-cone IOVD stimuli. Values on the y-axis are the difference in beta amplitude between MID and MID control stimuli (Δβ). The overall response profiles between ROIs are similar, but responses in the CSM were weaker in general.*

We found that in over half of all cases (six of eleven participants used in the analysis), the CSM partially overlaps area hMST in at least one hemisphere. Overlaps with hMT occurred in three cases. Indeed, the amplitude of the achromatic CD response (taken as the difference between the CD stimulus and the CD control stimulus) was highly correlated between the CSM and hMT (R = .89, p < .001), and the CSM and hMST (R = .88, p < .001). Correlation results are shown in **Figure 3.7**. Overall, we therefore found no evidence to suggest that the CSM is uniquely involved in CD processing, although populations of cells here may contribute to CD processing more generally in a manner analogous to cells in hMT+.
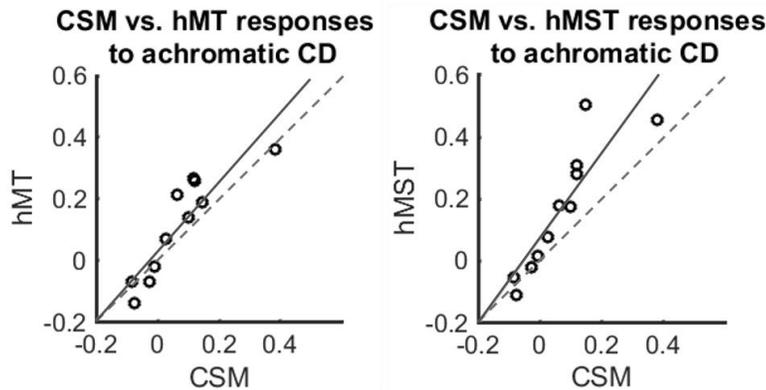
Figure 3.7 *Correlations between hMT and CSM, and hMST and CSM responses to achromatic CD stimuli (taken as the difference between the CD stimulus and CD control stimulus response). The dashed line indicates a perfect correlation (R = 1), and the data points are fitted with a least-squares line. Each data point represents the response for an individual subject. The CSM response was highly similar to the responses in hMT and hMST, though the amplitude of the response was weaker.*

3.5   Discussion

In this study, we used achromatic and S-cone isolating random-dot stimuli that engaged CD or IOVD mechanisms to probe the neural pathways involved in MID processing. Broadly, we found that both CD and IOVD stimuli elicit BOLD responses in a network of areas that includes early visual cortex, parts of the dorsal and ventral streams, and motion-selective areas hMT and hMST. Because we measured no significant differences between these two cues overall, and no interaction between ROI and MID type, our findings are consistent with previous studies suggesting that signals for both cues are multiplexed in a common network of areas with different neural sub-populations tuned to either CD or IOVD (Huk, 2012; Joo et al., 2016).

However, our novel finding is that within this network, achromatic and S-cone signals contribute to a different degree to IOVD and CD. The CD cue appears to depend primarily on achromatic inputs, and S-cone CD responses were weak. Conversely, the S-cone IOVD stimulus elicited a strong response in several ROIs, including early visual areas and ventral areas. Achromatic IOVD responses were weak and only began to emerge in later motion-selective ROIs. Thus, our findings suggest a critical dissociation in the way that early chromatic pathways contribute to CD and IOVD mechanisms.

3.5.1 Achromatic and chromatic inputs to motion in depth mechanisms

As mentioned in the section above, the CD and IOVD responses we measured across ROIs were dependent on the chromaticity of the stimulus. This interaction cannot be explained by overall differences in chromaticity – our stimuli were contrast-scaled such that on average there were no differences in the extent to which achromatic and S-cone signals drive activity across ROIs. This implies that achromatic signals conveyed by the magnocellular (MC) and parvocellular (PC) pathways, and S-cone signals conveyed by the koniocellular (KC) pathway, both contribute to MID processing, though signals carried in these pathways contribute differentially to the CD and IOVD mechanisms.

Across all ROIs, the CD response was primarily driven by achromatic stimuli, whilst S-cone contributions were weak in early areas and negligible in motion-selective areas. Due to the low spatial resolution of the KC pathway, neural populations driven by S-cone inputs are limited in their ability to perform the highly precise spatial matching required to resolve fine retinal disparity. S-cones are able to provide inputs to disparity mechanisms through low spatial frequency channels only (Wilson et al., 1988). Despite the weak cortical responses we measured here, our participants were able to perceive S-cone CD MID during the pre-testing phase of the experiment, perhaps due to coarse disparity processing (Wilcox & Hess, 1997). Nonetheless, our results suggest that the CD mechanism depends primarily on achromatic inputs.

In direct contrast to this, the IOVD signals we measured were highly biased towards S-cone inputs. We measured consistent BOLD modulations across ROIs to S-cone IOVD, whilst the achromatic stimulus appeared to contribute to IOVD mechanisms primarily in later, motion-sensitive areas. This may explain why previous fMRI research using achromatic IOVD stimuli has emphasised the role of hMT+ as the locus of IOVD processing (Rokers et al., 2009).

KC signals appear to be relayed particularly rapidly to extrastriate, motion-selective areas (Morand et al., 2000), suggesting an efficient mechanism through which the S-cones can contribute to motion processing. The precise source of S-cone signals in MT has been controversial. Direct anatomical projections from the KC layers of the LGN to MT have been used to explain sensitivity to moving isoluminant S-cone stimuli, measured perceptually

(Dougherty et al., 1999) as well as with fMRI (Wandell et al., 1999) and with electrophysiological recordings in MT (Seidemann, Poirson, Wandell, & Newsome, 1999). Alternatively, S-cone signals may 'piggyback' on the MC pathway, with some evidence suggesting S-cone inputs to around 10% of cells in the MC layers of the LGN (Chatterjee & Callaway, 2002). Thus, S-cone signals could in principle arrive at MT though the dominant MC pathway input (Calkins, 2001). In this study however, we found an early dissociation in the extent to which achromatic and S-cone signals contribute to IOVD MID, suggesting at least two different underlying mechanisms. This dissociation was particularly evident in early visual (V1, V2, V3 and V4) and ventral (LO-1 and LO-2) areas. The differences were smaller in hMT+, which may imply a convergence of MC, PC and KC signals in later, motion-selective ROIs.

We suggest that the large receptive field sizes in the KC layers of the LGN provide an early mechanism for computing IOVD-based MID. Because the IOVD cue depends on motion vectors generated at the level of the retina and does not require spatial matching between the eyes, it may integrate across larger portions of the visual field to generate reliable estimates of MID. Indeed, very sparse IOVD stimuli can convey MID percepts (Rokers, Czuba, Cormack, & Huk, 2011). Our findings suggest that early, low-resolution S-cone signals are combined in an opponent manner, and these signals contribute towards IOVD mechanisms through a network of ROIs.

3.5.2   *Motion in depth signals in primary visual cortex*

Both CD and IOVD stimuli elicited responses as early as V1 that were comparable in amplitude to those measured in later, motion-selective ROIs. Previous neuroimaging studies have reported only small responses here, with the strongest MID-driven responses recorded in hMT+ (Rokers et al., 2009). There may be several reasons for this. First, our CD stimulus contained little depth context compared to the stimulus used by Rokers et al., who divided their stimulus into quadrants moving in opposing directions. Our stimulus contained only the fixation point and the fixation ring, meaning that it engaged more neurons tuned to absolute rather than relative disparity. This profile is consistent with the tuning properties of binocular disparity neurons in V1 but not of those further upstream (Cumming & Parker, 1999), and may explain why we measured larger CD responses in V1 than previous studies.

Secondly, the CD control stimulus, which consisted of temporally scrambled frames from the CD stimulus, could lead to a more stochastic response from disparity-tuned neurons. In theory, two consecutive frames in the control stimulus can excite, then inhibit, a single neuron. This may not provide sufficient integration times for neurons to fire an action potential, leading to weak local field potentials and a weak BOLD response.

Finally, our CD stimulus contained a large number of contrast transients due to the limited lifetime of the dots – though it is important to note that this was balanced in the control stimulus and thus is unlikely to account for differences between CD and CD control stimuli.

Classical motion energy models (e.g. Adelson and Bergen, 1985) would not predict a strong V1 response to IOVD motion, given that V1 cells have small receptive fields that are primarily tuned to component motion (Hubel & Wiesel, 1968; Movshon & Newsome, 1996). In early visual cortex, the motion direction of individual components can be represented by populations of cells tuned to a given orientation and direction of motion (Rust, Mante, Simoncelli & Movshon, 2006). Furthermore, V1 cells do not exhibit strong motion opponency (Heeger, Boynton, Demb, Seidemann, & Newsome, 1999). Based on these properties, neural populations in V1 could provide early velocity estimates that are combined at a later stage to generate estimates of MID.

However, recent models of binocular motion perception in MT suggest that V1 inputs should exhibit motion opponent suppression and that these signals arise before binocular integration in V1 (Baker & Bair, 2016). Monocular motion opponency has also been proposed to drive pattern motion cells in MT, which are tuned to the combined outputs of several component motion cells in V1 (Rust, Mante, Simoncelli, & Movshon, 2006; Tailby, Majaj, & Movshon, 2010). Finally, there has also been some electrophysiological evidence for motion opponent suppression in V1, although these signals were weak, and it is unclear whether their source was monocular or binocular (Qian & Andersen, 1995).

The IOVD responses we measured in V1 in this study also suggest an early motion opponent signal. These signals could arise from joint motion and eye selective fields in V1, or early motion opponent inputs into binocular V1 cells. Crucially, the IOVD signals we measured in V1 were driven primarily by the S-cone stimulus; we did not measure strong achromatic IOVD responses. Thus, these signals may be characteristic of KC inputs.

93

These results suggest that dichoptic S-cone signals are combined in an opponent manner prior to V1. Such signals may arise from directionally-selective cells in the KC layers of the LGN (Casagrande, 1994; Tailby, Dobbie, et al., 2010), some of which receive binocular inputs (Zeater, Cheong, Solomon, Dreher, & Martin, 2015). It has recently been suggested that the direction selectivity measured in blue-on cells in the KC layers of the LGN are generated by latencies between the 'on' and 'off' subfields of small bistratified ganglion cells in the retina (Tailby, Dobbie, et al., 2010). These mechanisms could provide a very early basis for extracting the binocular motion-opponent signals in V1 that support IOVD.

3.5.3   Other areas involved in the extraction of 3D motion

The role of hMT+ in CD and IOVD processing has been documented previously (Czuba et al., 2014; Rokers et al., 2009; Sanada & DeAngelis, 2014), with emphasis on 2D and 3D motion being processed by the same cortical pathways (Huk, 2012). In addition to MID responses in classical motion pathways, from V1 to hMT+, we measured strong CD-driven responses in area IPS-0. The human intraparietal sulcus (IPS) is involved in a variety of cognitive functions, including the top-down control of visual attention and eye movements, which modulates activity in earlier visual areas (Corbetta et al., 1998; Corbetta, Kincade, Ollinger, McAvoy, & Shulman, 2000; Wojciulik & Kanwisher, 1999). In addition, the IPS also contains distinct populations of neurons that are sensitive to motion (Orban et al., 2006) and 3D structure from motion (Orban, Sunaert, Todd, Van Hecke, & Marchal, 1999; Vanduffel et al., 2002). This may explain why activation in IPS-0 was more pronounced for CD stimuli than for IOVD stimuli. Because IOVD stimuli lack the concrete depth information provided by the binocular disparity cues in the CD stimulus (Rokers et al., 2008), they are much less likely to convey shape or form information and are thus less likely to engage form-from-motion mechanisms. IPS-0 activation observed here may constitute a part of the MID pathway that is involved in extracting 3D shape from disparity and the allocation of visual attention, rather than in extracting 3D motion *per se.*

Previously, an area anterior to hMT+, the 'cyclopean stereo-motion' (CSM) area, has been suggested as the main locus for stereo-defined MID processing (Likova & Tyler, 2007). However, we found no compelling evidence that this area is uniquely and crucially involved in the extraction of CD MID signals. Although we measured activation to the CD stimulus

anterior to hMT+, this was not restricted to this area alone, and we also measured strong modulations in hMT+ itself. Additionally, an ROI analysis of the CSM area, defined using Talairach co-ordinates given in the Likova & Tyler (2007) paper, showed a similar response profile to hMT and hMST across all stimulus types, with weaker responses in the CSM.

Unlike Likova and Tyler, we found substantial overlaps between the CSM ROI and our own hMST definitions. This is likely to drive similarities in responses between these two regions. However, even in hMT, which was clearly distinct from the CSM in almost all participants, response profiles were very similar and the achromatic CD response was highly correlated between the CSM and hMT. Thus, we suggest that CD is extracted primarily in hMT and hMST.

3.6   Conclusions

We measured responses to CD and IOVD stimuli in a network of areas that included early visual areas, parts of the dorsal and ventral streams, and motion selective hMT and hMST. Overall, both achromatic and S-cone stimuli provided inputs to these areas – suggesting that signals carried in the MC, PC and KC pathways contribute to MID processing. However, we found that achromatic and S-cone signals contribute in a different manner to CD and IOVD mechanisms. CD responses were most strongly driven by achromatic inputs, whereas the S-cone stimuli elicited only weak responses. Thus, we suggest that the CD mechanism depends more heavily on achromatic inputs. Conversely, IOVD mechanisms across a hierarchy of areas were driven most strongly by S-cone inputs, with achromatic inputs generating a comparable response only in later, motion-selective ROIs. The IOVD S-cone signals were robust even in V1, suggesting that KC signals are combined in an opponent manner at a very early stage in visual processing. Overall, we have shown that achromatic and S-cone signals contribute differentially to CD and IOVD mechanisms, within a broader, shared architecture of MID pathways.

Chapter 4.   Decoding eye of origin in and beyond primary visual cortex

4.1   Abstract

The inter-ocular velocity difference (IOVD) mechanism for conveying motion in depth (MID) depends on the simultaneous availability of eye of origin (EOO) and motion direction signals. Because IOVD computations likely occur at an extrastriate locus, we sought to identify the presence of EOO and motion direction information in and beyond primary visual cortex. Participants (N=8) viewed moving dot patterns during fMRI scanning, where dot patterns were presented to the left eye, the right eye or to both eyes, and motion could be either up/down or left/right. We estimated voxel-wise beta amplitudes for the stimulation condition (left, right, binocular, fixation) and motion direction (up/down or left/right), and used a combination of multivariate pattern analysis (MVPA) and classification techniques to decode EOO and motion direction across a hierarchy of visual areas. We were unable to identify EOO information in V1 or in any other areas. However, motion direction tuning was revealed in hMT. Therefore, we did not identify regions of interest (ROIs) where both types of information were preserved. This implies that IOVD mechanisms inherit EOO information from V1. Though EOO information may be represented in a distributed manner in extrastriate areas, we suggest that our analysis techniques were not sensitive enough to identify its availability.

4.2   Introduction

Motion towards and away from an observer generates two dissociable cues that each independently contribute to the perception of motion in depth (MID). One of these cues depends on incremental changes over time in the retinal disparity of the moving object. The second cue depends on a comparison of motion vectors originating from each eye. Both cues are mathematically equivalent, but differ in the order of computation and can be expressed as

$$\frac{d(r-l)}{dt} = \frac{dr}{dt} - \frac{dl}{dt}$$

where dt is the temporal derivative of the left (l) and right (r) retinal images (Rushbass & Westheimer, 1961). The left side of the equation relates to the disparity mechanism (changing disparity, CD), where ($r - l$) is the horizontal disparity. The right side of the equation relates to the inter-ocular velocity difference mechanism (IOVD), where the displacement of the right retinal image (dr/dt) is compared against the displacement of the left retinal image (dl/dt).

Although both cues depend on a comparison of monocular signals, the point at which these monocular signals are combined to form the binocular cue to MID differs between CD and IOVD. For CD, signals from the left and the right eye are combined first to generate the retinal disparity cue, before the temporal derivative is taken. For the IOVD cue, the temporal derivative is taken first, to generate velocity vectors in each eye. These velocity vectors are then compared between the eyes at a later computational stage. The two processes are illustrated in **Figure 4.1**.
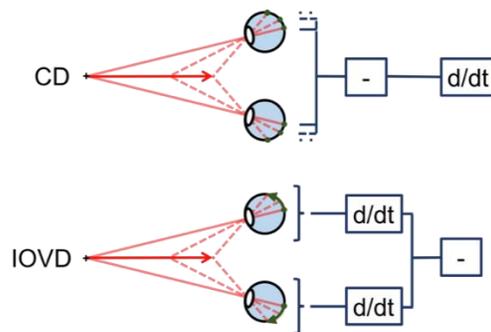
Figure 4.1 *Schematic of the CD and IOVD mechanisms, showing differences in computational stages. For the CD mechanism, information from the left and the right eyes is combined first before the temporal derivative is taken. For the IOVD mechanism, the integration of left and right eye signals occurs at a later stage.*

These computational stages imply the comparatively late integration of monocular signals to form the IOVD cue. Thus, the site at which the final, binocular MID signal is extracted must maintain access to eye of origin (EOO) information.

In primates, EOO information is represented in ocular dominance columns in primary visual cortex (V1), where alternating columns of cells are predominantly driven by left or right eye inputs. This characteristic functional organization has been revealed in humans using post-mortem histological staining (Adams, Sincich, & Horton, 2007; Horton & Hedley-Whyte, 1984) as well as with functional magnetic resonance imaging (fMRI) of the cortical sheet (Cheng, Waggoner, & Tanaka, 2001; Dechent & Frahm, 2000; Goodyear & Menon, 2001; Menon, Ogawa, Strupp, & Uğurbil, 1997; Yacoub, Shmuel, Logothetis, & Uğurbil, 2007). Although ocular dominance columns extend somewhat into V2 (Adams et al., 2007), they do not appear to exist further upstream (Hubel & Livingstone, 1987; Nasr, Polimeni, & Tootell, 2016; Tootell & Hamilton, 1989; Ts'o, Frostig, Lieke, & Grinvald, 1990), and beyond V1 monocular information is thought to merge into a single binocular representation. Indeed, disparity selectivity emerges as early as V1 (Gian F. Poggio, Motter, Squatrito, & Trotter, 1985), and motion aftereffects transfer readily from one eye to the other (Blakemore & Campbell, 1969), implying early integration of those eye-specific signals that CD and IOVD mechanisms depend upon.

This gives rise to two possibilities. First, that IOVDs are extracted at an early stage of visual processing, drawing directly on information encoded in ocular dominance columns. Alternatively, EOO information may be maintained to later stages in the visual hierarchy than previously assumed, represented in a distributed fashion in the absence of strict columnar organisation. The majority of evidence to date suggests an extrastriate locus for IOVD computation; IOVDs can be computed over spatial scales larger than V1 receptive fields (Rokers, Czuba, Cormack, & Huk, 2011), and fMRI studies reliably measure IOVD-driven signals in hMT and hMST (Joo, Czuba, Cormack, & Huk, 2016; Rokers, Cormack, & Huk, 2009; see also Chapter 3 of this thesis).

Recent neurophysiological findings in macaque MT indicate that neurons selective for 3D motion also show a monocular tuning bias (Czuba, Huk, Cormack, & Kohn, 2014). For each recorded neuron, Czuba et al. calculated a 'monocularity index', given by normalising the absolute difference in magnitude between left and right eye responses by their sum. In comparison to neurons with 2D motion tuning, or those neurons with only a weak bias for 3D motion, the response profiles of neurons with a strong 3D motion selectivity also resulted in the largest monocularity indices. This implies that EOO information may be preserved specifically in those neural populations that extract the IOVD cue.

In this study, we attempted to unify the availability of EOO information with possible sites for IOVD computation, using fMRI and multivariate pattern analysis (MVPA) techniques. Extracting EOO information using fMRI is technically challenging, as the average width of ocular dominance columns in V1 is approximately 1mm (Adams et al., 2007). At 3 Tesla, and using standard gradient-echo blood oxygen level-dependent (GE BOLD) fMRI sequences, voxel resolution is typically between 2-3mm and the point-spread function of the fMRI signal is roughly 3.5mm (Engel, Glover, & Wandell, 1997; Parkes et al., 2005; Shmuel, Yacoub, Chaimow, Logothetis, & Ugurbil, 2007). Clearly, this results in substantial blurring in the neurovascular signals that sample the underlying columnar physiology, meaning that standard mass univariate analyses, where responses are averaged across voxels, cannot dissociate between left and right eye signals. Even within a single voxel, responses can be blurred across multiple ocular dominance columns.

MVPA and classification techniques have become widely popular in fMRI research, and can be used to decode information represented at the sub-millimetre scale. MVPA methods are highly sensitive to small changes in voxel activity patterns, which can be detected even when there are no mean changes in activity level across voxels (Haxby et al., 2001). Thus, these analyses can be employed when less sensitive univariate analyses fail to extract meaningful information, and have been used previously to decode EOO information in early visual areas (Larsson, Harrison, Jackson, Oh, & Zeringyte, 2016; Schwarzkopf, Schindler, & Rees, 2010; Shmuel, Chaimow, Raddatz, Ugurbil, & Yacoub, 2010).

The ability to extract information represented at a much smaller scale than the typical fMRI sampling resolution has been hypothesised to depend on hyperacuities in individual voxels (Haynes & Rees, 2006; Kamitani & Tong, 2005, 2006; Kriegeskorte, Goebel, & Bandettini, 2006). The signal from a voxel that overlaps several ocular dominance columns may be biased in favour of input from one eye, depending on the proportion of left vs. right eye cell populations the voxel (indirectly) samples from. MVPA may exploit these voxel biases to reveal information represented at a fine scale.

A contrasting viewpoint is that the ability to decode EOO in V1 need not depend on an underlying columnar organisation, and may instead reflect biased sampling of elongated blood vessels along non-isotropic ocular dominance columns (Shmuel et al., 2010). Indeed, spatial smoothing, which blurs responses across voxels, and thus results in a loss of small

biases in the responses of individual voxels, has little effect on decoding accuracy (Op de Beeck, 2010). Additionally, a large-scale naso-temporal bias reflecting a preference for the contralateral eye in each hemisphere, has been found to drive EOO decoding in V1 to V3 (Larsson et al., 2016). These findings imply that MVPA techniques depend on biased draining regions in the vasculature, rather than on hyperacuities in individual voxels. Accordingly, MVPA should be able to detect EOO information in areas that lack ocular dominance columns, where this information is inherited from earlier visual areas and is encoded in a distributed manner.

Here, we used MVPA techniques to investigate how eye-specific signals are maintained throughout the visual hierarchy to support the extraction of the IOVD cue. To do this, we presented moving dot patterns to the left eye, right eye and binocularly, and attempted to decode both EOO and motion direction from voxel response patterns across multiple visual areas. Decoding EOO should be possible in V1, whilst we expected to decode motion direction in motion-selective areas hMT and hMST. As both EOO and motion direction signals are necessary for the extraction of IOVD, we hypothesised that there should be a convergence of these signals in extrastriate areas involved in the computation of IOVD. Thus, areas where both EOO and motion direction information can be decoded reflect candidate regions for the binocular integration stage of the IOVD mechanism.

4.3   Methods

4.3.1   Participants

Ten participants (age 21-45 years, 3 male) with normal or corrected-to-normal vision were recruited. Two participants (MK and ARW) were authors, the rest were naïve. Of these, 2 participants were excluded from analysis due to poor fits in the modelled fMRI timecourse (general linear model [GLM] mean variance explained across regions of interest [ROIs] < 5%) leaving 8 participants in total. This study was approved by the York Neuroimaging Centre ethical review panel, and written informed consent was obtained from all participants according to guidelines set out in the Declaration of Helsinki.

4.3.2   Apparatus

A PROpixx DLP LED projector (VPixx Technologies, Saint-Bruno, Canada) at 1920 x 1080 pixel resolution and running at 120 Hz was used to back-project stimuli on to a silver screen positioned behind the participants' heads during scanning. Stimuli were viewed on a first-surface mirror mounted on the head coil (57cm viewing distance, including the optical pathway of the mirror), giving a viewing angle of 41º x 23.5 º. Stereoscopic presentation was achieved using a circular polarizer (DepthQ Polarization Modulator, VPixx technologies, Saint-Bruno, Canada) placed in front of the long-throw lens, and passive 3D glasses worn by the participant. This system ensures efficient isolation of images targeted at each eye, with measured contrast and luminance crosstalk levels well below 1% (Baker, Kaestner, & Gouws, 2016).

Stimulus presentation was controlled from a Shuttle PC with Intel Core i7-4790K processor at 4.0 GHz and an NVIDIA GeForce GTX970 graphics card with 4 GB DDR5 memory. All stimuli were designed and run from Matlab 8.5.0 (2015a; The MathWorks Inc., Natick, MA, USA) running the Psychtoolbox 3.0.12 routines (Brainard, 1997; Pelli, 1997). Behavioural responses and scanner trigger pulses to synchronise stimulus onset were transmitted using a four-button fibre-optic response pad (Current Designs, Philadeliphia, PA).

4.3.3   Stimuli and design

For the EOO experiment, stimuli were designed to excite the receptive fields of neurons across a range of regions of interest (ROIs), including early visual areas V1, V2 and V3, and motion-sensitive areas hMT and hMST. They were spatially and temporally broadband, to accommodate the differences in preferred spatial frequency of monocular and binocular neurons (Schwarzkopf et al., 2010), variations in velocity tuning (Rodman & Albright, 1987), and different receptive field sizes. Stimuli consisted of randomly positioned isotropic Laplacian-of-Gaussian (LoG; Dakin & Mareschal, 2000) elements (dot density = 1.5 dots/deg^2, dot sigma = 0.05º), presented at 50% contrast on a mean grey background. Elements were generated by convolving the x and y point positions with a LoG function given by

$$\nabla^2 G(x, y, \sigma) = \frac{1}{\sigma^2}\left(1 - \frac{x^2+y^2}{\sigma^2}\right)\exp\left(-\frac{(x^2+y^2)}{2\sigma^2}\right),$$

where $\sigma = 0.05$ (example shown in **Figure 4.2**). Dots moved in two cardinal directions (up/down or left/right) during different instances of stimulus presentation, and at four different speeds within each display (0.20, 0.68, 2.34 and 8.00º/s). Because of the balance of contrast polarity in each individual dot, LoG profiles avoid artefactual "motion streaks" (Apthorp, Cass, & Alais, 2011; Geisler, 1999) that can confound the decoding of motion direction (Clifford, Mannion, & McDonald, 2009; Maloney, Watson, & Clifford, 2014).
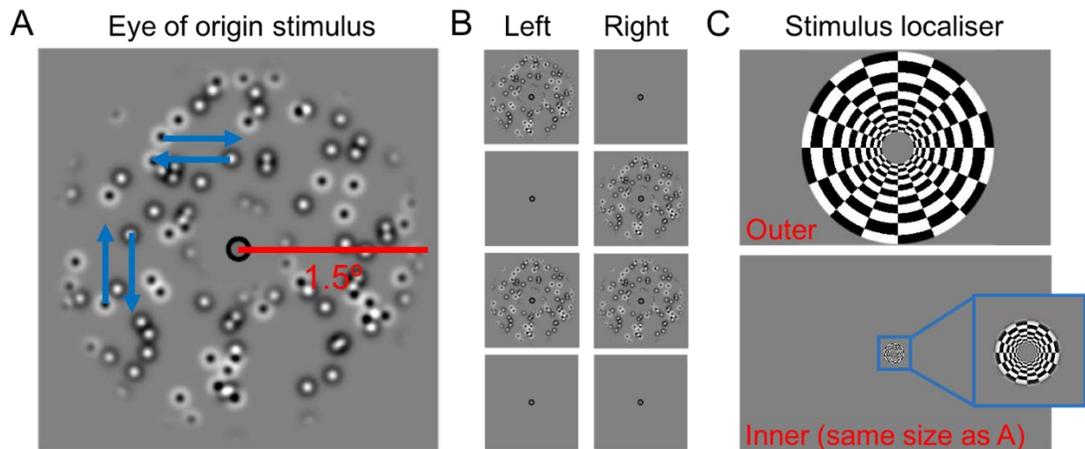
Figure 4.2 *Stimulus conditions for the EOO (panels A and B) and stimulus localiser (panel C) fMRI scans. A zoomed-in binocular view of the EOO stimulus is shown in panel A. Elements moved at 4 different speeds, either leftward and rightward, or upwards and downwards (indicated by blue arrows). Each eyes' views during different stimulation conditions are shown in panel B. From top to bottom: left eye stimulated, right eye stimulated, binocular stimulation, baseline fixation. Panel C shows the stimuli used to map the retinotopic extent of the EOO stimulus (bottom – same size as the stimulus in panel A), and the unstimulated periphery (top). The whole field of view is shown, giving an estimate of the size of the EOO stimulus in context.*

Dot fields were viewed through a circular aperture (0.25º inner and 1.5º outer radius, edges ramped with a cosine envelope of 0.4º). This scale ensured visual stimulation well away from the blind spot (at an eccentricity of roughly 15º from the fovea, with a monocular representation in V1; Tong & Engel, 2001; Tootell, Hadjikhani, Vanduffel, et al., 1998). Stimuli were presented to the left eye, right eye, or both eyes simultaneously. In the monocular conditions, the unstimulated eye viewed the mean grey background with the same fixation point that was also present during stimulus conditions. This fixation point (diameter 0.22º) was presented at the centre of the display to control the eye position of participants. To ensure participants were fixating, and to control the allocation of spatial attention, participants

pressed a button when the centre of the fixation point changed subtly in luminance. Luminance changes occurred randomly over the course of a scan, where the time between changes varied between 1.5s and 9s.

Stimuli were presented for 3s – at onset and offset, stimulus contrast was ramped over 300ms with a raised cosine envelope, and stimuli were at maximum contrast for 2.4s. They were presented in a dense, rapid event-related design, where the inter-stimulus interval (ISI) varied between 3s (the length of 1TR) and 12s. Stimulus onsets were determined separately for each scan using Optseq2 (https://surfer.nmr.mgh.harvard.edu/optseq). Stimulation conditions were: left eye stimulation, right eye stimulation, binocular stimulation, and a fixation-only baseline condition. Each condition occurred 12 times in each scan – of these, motion in the stimulus could be up/down or left/right with a 50/50 split. The EOO stimulus and presentation conditions are depicted in **Figure 4.2**, panels A and B.

We also designed localiser stimuli to map the retinotopic extent of the EOO stimulus, and the unstimulated periphery. fMRI responses to the localisers allowed early visual areas (V1, V2 and V3) to be split along the eccentricity dimension, into stimulated 'inner' and unstimulated 'outer' ROIs. Localiser stimuli were contrast-reversing log-scaled radial checkerboard rings updating at 1Hz, where the 'inner' ring was the same size as the EOO stimulus (extending 0.25º to 1.5º from fixation) and the 'outer' stimulus mapped the surrounding area (extending 2º to 11.75º from fixation, leaving a 0.5º gap between the 'inner' and 'outer' rings). Stimuli were presented at 50% contrast on a mean grey background, to the left eye, to the right eye, or to both eyes. A fixation point at the centre of the display stabilised eye gaze, and the centre of the fixation point changed in contrast in the same manner as for the EOO stimulus.

In the stimulus localiser fMRI run, each stimulus condition (inner ring left eye, inner ring right eye, inner ring binocular, outer ring right eye, outer ring left eye, outer ring binocular, and a blank baseline fixation condition) was presented in 9s blocks with 3 repeats of each condition. The localiser stimuli are illustrated in **Figure 4.2**, panel C.

4.3.4   MRI parameters

High-resolution T1-weighted anatomical scans (TR = 7.8ms; TE = 3.0ms; TI = 600ms; flip angle = 20°; FOV = 25.6 x 25.6 cm; matrix size = 256 x 256; voxel resolution = 1.0 x 1.0 x 1.0mm; 176 coronal slices to cover the whole head) were acquired on a 3T GE SIGNA HDx Excite MRI scanner for each participant in a separate scanning session, using an 8-channel whole-head phased-array coil (MRI Devices Corporation).

During the experimental session, data were collected using a 16-channel phased-array half-head coil (Novamed) to improve the signal-to-noise ratio in the occipital cortex. To co-register functional data with the high-resolution anatomical scan, a proton-density (PD) weighted reference scan with the same slice prescription as the EPI scans was collected for each participant (TR = 2700ms; TE = 38ms; flip angle = 90°; FOV = 19.2 x 19.2cm; matrix size = 512 x 512; voxel resolution = 0.38 x 0.38 x 2.5mm; 39 quasi-coronal, contiguous slices oriented along the calcarine sulcus and covering the occipital lobe). Standard gradient-echo EPI sequences (TR = 3000ms; TE = 30ms; flip angle = 90°; FOV = 19.2 x 19.2cm; matrix size = 96 x 96; voxel resolution = 2 x 2 x 2.5mm; 116 volumes including 4 dummy volumes, total scan time = 5 minutes 48 seconds) were used for one stimulus localiser scan and seven EOO scans.

Finally, each participant completed two motion localiser scans to identify motion-sensitive areas V3A/B, hMT and hMST. In addition, standard retinotopic mapping scans (typically five wedge and two ring scans, with eight stimulus cycles each) were carried out to delineate early visual areas. These data were collected in separate scan sessions, with fMRI parameters similar to those described above, and are described in more detail in Chapter 3.

4.3.5   Mapping regions of interest

ROIs (V1, V2, V3, V4, V3A/B, IPS-0, LO-1, LO-2, hMT and hMST) were defined in each subject, as described Chapter 3, using a combination of retinotopic mapping and motion localisers. Briefly, areas V1, V2, V3 (Dougherty et al., 2003; Schira, Tyler, Breakspear, & Spehar, 2009; Sereno et al., 1995), V4 (Brewer, Liu, Wade, & Wandell, 2005; Hansen, Kay, & Gallant, 2007; Wade, Brewer, Rieger, & Wandell, 2002; Winawer, Horiguchi, Sayres, Amano,

& Wandell, 2010), LO-1, LO-2 (Larsson & Heeger, 2006), and IPS-0 (Press, Brewer, Dougherty, Wade, & Wandell, 2001; Swisher, Halko, Merabet, McMains, & Somers, 2007; Tootell, Hadjikhani, Hall, et al., 1998), were delineated using characteristic reversals in the polar angle phase map acquired using conventional retinotopic mapping methods. Motion-sensitive ROIs V3A/B, hMT and hMST (Amano, Wandell, & Dumoulin, 2009; Fischer, Bülthoff, Logothetis, & Bartels, 2012; Huk, Dougherty, & Heeger, 2002) were defined iteratively using both motion localisers and retinotopic mapping.

Early visual areas V1, V2 and V3 were restricted in the eccentricity dimension into six ROIs, using the stimulus localiser described above. 'Inner' ROIs corresponded to the cortical surface that was directly stimulated by the retinotopic extent of the EOO stimulus (from 0.25º to 1.5º from fixation). The 'outer' ROIs mapped the periphery (from 2º to 11.75º from fixation). Voxels in the 'outer' ROIs were not driven by the experimental stimulus, and some of the underlying receptive fields may be actively suppressed, resulting in a negative BOLD response (Shmuel et al., 2002). This dissociation allowed us to investigate whether the suppressive response also contains some monocular tuning.

Finally, the fusiform face area (FFA) was defined as a control ROI. This area is strongly driven by face stimuli. The FFA was chosen as it is visually responsive, but its high degree of category selectivity implies that it is unlikely to maintain any EOO tuning, given that this information is redundant for its functional specialism. The FFA was defined in each subject by centring a 5mm sphere on Talairach co-ordinates given in the original fMRI paper by Kanwisher et al., where the amplitude of the BOLD response was compared between faces, objects and houses (Kanwisher, McDermott, & Chun, 1997). These co-ordinates were set at [-35 -63 -10] in the left hemisphere, and [40 -55 -10] in the right hemisphere. The size of the sphere was chosen because the mean size of the FFA is $5mm^3$ in the left hemisphere, and $10mm^3$ in the right hemisphere (Kanwisher et al., 1997).

4.3.6   MRI and fMRI pre-processing

High-resolution T1-weighted anatomical scans were automatically segmented into grey and white matter using Freesurfer (https://surfer.nmr.mgh.harvard.edu/). To align functional data to anatomical data, the PD-weighted reference image was skull-stripped and the signal drop-off due to distance from the coil was corrected using tools available in the FSL toolbox

(Brain Extraction Tool, BET, Smith, 2002; FMRIB's Automatic Segmentation Tool, FAST, Zhang, Brady, & Smith, 2001 – available from the FMRIB Software Library, www.fmrib.ox.ac.uk/fsl). The corrected PD image was then aligned to the segmented, subject-specific high-resolution T1-weighted anatomical scan using the Nestares algorithm (Nestares & Heeger, 2000). The resultant transformation matrix was applied to the fMRI datasets to align these to the high-resolution anatomical scan.

Individual subject analyses were carried out in mrVista (https://web.stanford.edu/group/vista/cgi-bin/wiki/index.php/Software; VistaLab, Stanford University) and Matlab 8.5.0 (2015a; The MathWorks Inc., Natick, MA, USA). Four dummy volumes were clipped from the start of each 4D dataset to account for initial net magnetization instabilities, and motion correction was carried out between and within each fMRI scan. Using the segmentation of the high-resolution anatomical image, functional voxels were restricted to the grey layers for volume and surface-based analyses. The timecourse of each voxel was high-pass filtered to remove slow drifts in the scanner signal, and normalisation was carried out by dividing each timecourse by the mean response across all voxels.

4.3.7   *Eye of origin univariate analysis*

The response in each grey-layer voxel was modelled using a GLM, with predictors for left eye stimulation, right eye stimulation, binocular stimulation, and the baseline fixation condition. Additional predictors modelled the mean response across each scan, allowing the scans to be concatenated. The design matrix was convolved with a standard 'difference of Gammas' hemodynamic response function from the SPM 8 toolbox (http://www.fil.ion.ucl.ac.uk/spm). The resulting beta weights modelled the BOLD amplitude for each voxel during each stimulation condition (the mean response across all events of the same type).

Voxel beta weights were extracted for each participant, in each bilateral ROI (V1 inner, V1 outer, V2 inner, V2 outer, V3 inner, V3 outer, V4, V3A/B, IPS-0, LO-1, LO-2, hMT, hMST, FFA). First, we analysed the univariate responses by averaging all voxel beta weights within an ROI, within and across participants with GLM variance explained greater than 5% (N=8). These results gave the mean ROI responses to each stimulus condition, and were

entered in a 14 x 4 repeated measures analysis of variance to determine main and interaction effects (ANOVA; 14 ROIs by 4 stimulus conditions).

4.3.8   4.3.8   Eye of origin voxel analysis

Because EOO information is represented at a sub-millimetre scale in primary visual cortex (Adams et al., 2007), averaging responses across voxels in each ROI is unlikely to discriminate between left and right eye stimulation. In primary visual cortex, each 2mm voxel spans multiple ocular dominance columns, pooling across populations of neurons whose receptive fields are tuned to either eye. These overlaps imply that voxels may, at best, be *biased* towards left- or right-eye inputs. Such biases can be revealed by cross-correlating patterns of voxel responses in each ROI and during each stimulus condition. A voxel that largely overlaps a left-eye ocular dominance column will be driven strongly by left-eye and binocular stimulation, but only slightly by right-eye stimulation. Correlating across voxels with a monocular bias should therefore yield a small correlation coefficient between left and right eye stimulation, and a larger correlation coefficient between monocular and binocular stimulation. In the ideal case, voxel response patterns could be *anti*-correlated, where the voxel is driven when its preferred eye is stimulated, but is suppressed when its non-preferred eye is stimulated. ROIs that contain populations of neurons with monocular tuning should show these types of response patterns. The logic of this approach is illustrated in **Figure 4.3**.
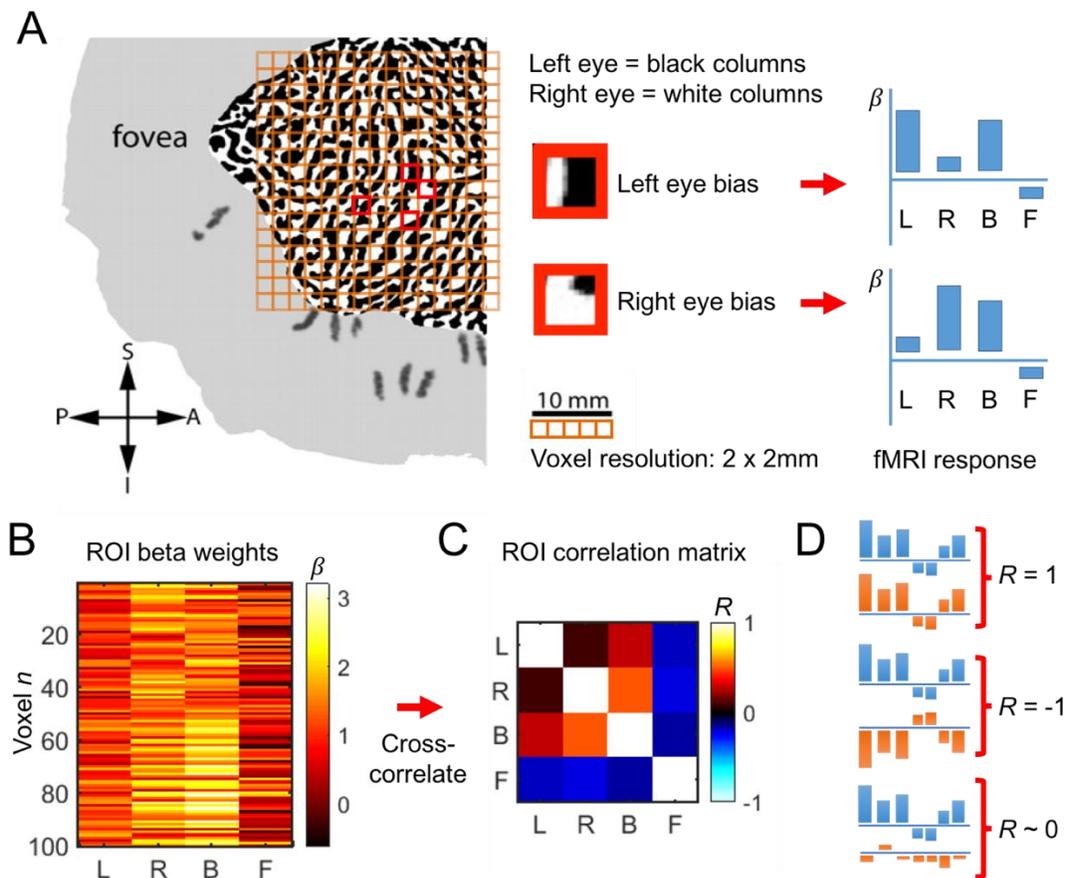
Figure 4.3 *EOO analysis. Panel A shows a flattened patch of human primary visual cortex with stained ocular dominance columns (adapted from Adams et al., 2007). A simplified representation of the fMRI sampling resolution in this is study is overlaid, to scale, as an orange grid on the ocular dominance pattern, illustrating bias towards left and right eye inputs in individual voxels. Expected beta weight responses for each of these voxels is shown – a large response when its preferred eye is stimulated, a smaller response when its non-preferred eye is stimulated, and a large binocular response. The response during fixation is negative or near zero. Panel B shows a simulated pattern of beta weights in an ROI, with beta amplitudes representing responses to left eye stimulation (L), right eye stimulation (R), binocular stimulation (B), and to fixation (F). These patterns are cross-correlated, and the correlation matrix is shown in panel C. Correlations between 0 and 1 imply that voxels are stimulated in the same direction, for two types of event. Correlations between 0 and -1 imply that voxels are stimulated in the opposite direction for two types of event. Correlations around 0 imply that there is no relationship between response patterns across voxels to two types of event (panel D). In this example, voxel responses to L and R events result in a lower correlation than monocular and binocular conditions, implying EOO tuning –voxels are responding differently to left and right eye inputs. If there were no monocular bias, and no tuning, the correlations between left and right eye events would be the same as the correlations between left and binocular and right and binocular events.*

ROI response patterns were generated in participants where the GLM variance explained was greater than 5% (N = 8). Voxel beta weights for each stimulus condition were extracted, and patterns were normalised by subtracting the mean response across voxels from each voxel within each ROI. The bootstrapped cross-correlation (1000 samples) was computed across voxel response patterns, resulting in 4 x 4 correlation matrices representing the similarity between voxel responses to left eye stimulation, right eye stimulation, binocular stimulation, and during baseline fixation conditions. Cross-correlation was carried out in each subject individually, and the bootstrapped coefficients were averaged across participants.

To quantify monocular tuning, we computed an EOO index during each iteration of the cross-correlation, given by the formula

$$Index = \left(\frac{R_{LvB} + R_{RvB}}{2} \right) - R_{LvR}$$

where the correlation coefficient in each ROI to left vs. right eye stimulation (R_{LvR}) was subtracted from the mean left vs. binocular (R_{LvB}) and right vs. binocular (R_{RvB}) correlation coefficient. Positive values indicate a smaller left vs. right correlation than a monocular vs. binocular correlation, indicating EOO tuning. A value of zero implies no difference between monocular and binocular stimulation protocols, where correlation amplitudes are similar between these three conditions. Negative values would imply larger LvR correlations than monocular to binocular correlations – a scenario that is unlikely due to binocular suppression (Tong, Meng, & Blake, 2006). Small negative values may occur by chance or due to the presence of noise. This bootstrapped index was averaged across participants to reveal the presence of EOO information in different visual areas.

4.3.9   Decoding motion direction

In a complementary analysis, we used a support vector machine (SVM implemented in LIBSVM; Chang & Lin, 2011) to decode the direction of motion in the EOO stimulus (up/down or left/right). Analysis was carried out in Matlab 8.5.0 (2015a; The MathWorks Inc., Natick, MA, USA), using custom routines for pre-processing and response modelling of voxel data. For each participant and in each ROI we selected the top 100 most informative voxels, based

on the voxel mean GLM variance explained across all 7 scans (for the selected voxels, mean R^2 across ROIs and across participants = 0.23). For each voxel, we estimated one beta weight per stimulus event (across all 7 scans, this was 126 up/down and 126 left/right events). Each raw timecourse was zero-centred and normalised by fitting and subtracting a polynomial function, and dividing by the standard deviation across voxels. Voxel response amplitudes were estimated using a least-squares fitting approach. The design matrix, containing one predictor per event, was convolved with a standard difference-of-gammas HRF (6s lag), and its pseudo-inverse was multiplied by each voxel's timecourse. Across all scans, this resulted in 252 event beta estimates per voxel.

To increase the reliability of these beta weights, we calculated bootstrapped mean estimates by averaging across one-third of all events of the same type for each voxel, resulting in a total of 42 estimates per voxel, per up/down or left/right event. These estimates were fed into the SVM with a radial basis function kernel. Classification accuracy was determined using cross-validation where the data were split into 10 equal parts, and one subset was tested using the classifier trained on the remaining 9. Voxel event beta estimates were resampled and the direction of motion was decoded across 1000 iterations. Cross-validation accuracy was averaged across these iterations for each ROI and each participant, resulting in the final, group-level decoding accuracy.

4.4   Results

4.4.1   Eye of origin univariate analysis

The BOLD response was modelled using predictors for each stimulus presentation type: left eye stimulated, right eye stimulated, binocular stimulation, and baseline fixation. Beta weights for each voxel in each ROI were extracted on a subject by subject basis, and were averaged within each ROI, before being entered into a 14 x 4 repeated measures ANOVA (14 bilateral ROIs – V1 inner, V1 outer, V2 inner, V2 outer, V3 inner, V3 outer, V3A/B, IPS-0, V4, LO-1, LO-2, hMT, hMST and FFA, by 4 stimulation conditions – left, right, binocular, and fixation). Beta weights were averaged across subjects to generate group-level results. **Figure 4.4** illustrates these results in the visually-driven ROIs, and **Figure 4.5** shows results in the non-visually-driven and control ROIs.
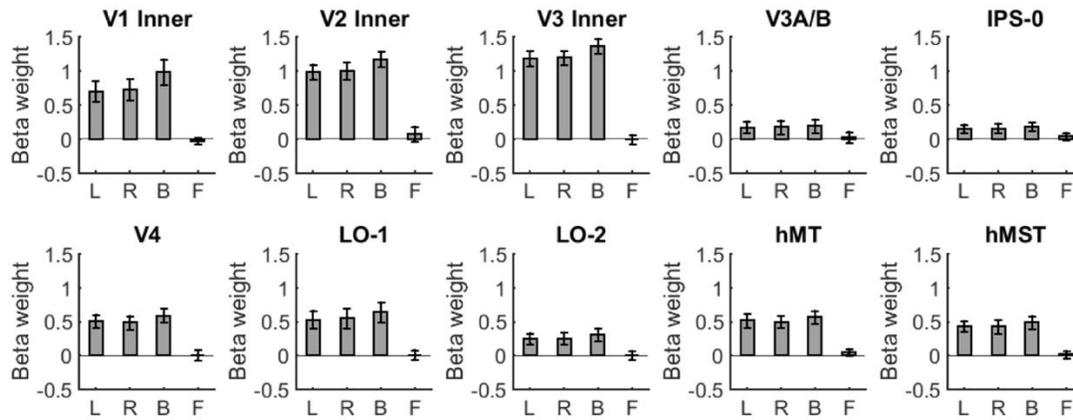
Figure 4.4 *Univariate responses across visually-driven ROIs, group average taken over participants with >5% variance explained in the GLM (N=8). Left (L) and right (R) eye stimulation conditions are well-balanced; in most ROIs binocular (B) stimulation results in a higher BOLD amplitude. Baseline fixation (F) response is not subtracted, and is around zero in all ROIs. Error bars represent ±1 SEM.*

In contrast to the 'Inner' ROIs, all 'Outer' ROIs showed near-zero responses to all types of stimulation, as well as to the fixation events. This is because, by definition, 'outer' ROIs map the peripheral aspects of V1, V2 and V3, including those parts of the cortex that are not driven by the retinotopic extent of the stimulus. These ROIs were defined to identify any tuning in the negative BOLD response, which arises when voxels are actively suppressed when they lie outside the area of active visual stimulation (Gouws et al., 2014). The negative BOLD response is lost in the univariate average, but some voxels did indeed show this tuning profile – for a comparison of beta weight distributions in 'inner' and 'outer' V1, V2 and V3, see section 4.7: Supplementary information, **Figure 4.11**.
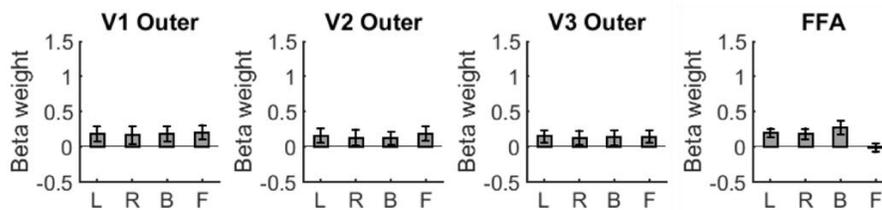
Figure 4.5 *Univariate responses across 'Outer' (non-visually-stimulated) ROIs, and the FFA control ROI which is visually driven but presumably has no access to EOO information. Average taken across the same participants as previously (N=8). Error bars represent ±1 SEM.*

The ANOVA revealed a significant main effect of ROI ($F(13, 91) = 13.43$, $p < .001$, partial $\varepsilon^2 = 0.66$), where the mean BOLD modulation was largest in the visually stimulated

'inner' V1, V2 and V3 ROIs. Intermediate responses were measured in motion-selective hMT and hMST, as well as areas in the ventral stream (V4, LO-1 and LO2). The smallest responses were measured in V3A/B, IPS-0 and the control ROI (FFA), as well as in non-visually-driven ROIs (V1 outer, V2 outer and V3 outer).

There was also a significant main effect of stimulation type ($F(3, 21) = 70.16$, $p < .001$, partial $\varepsilon^2 = 0.91$), driven by mean differences between left (L), right (R) and binocular (B) stimulation versus fixation (F); all significance values between these three contrasts were below $p = .001$. Crucially, the comparisons between left and right eye stimulation, and monocular to binocular stimulation conditions, were non-significant.

Finally, there was a significant interaction between ROI and stimulation type ($F(39, 273) = 22.78$, $p < .001$, partial $\varepsilon^2 = 0.77$). As can be seen in **Figure 4.4**, the main visually-driven ROIs showed a similar response pattern, where activity was strongly driven by left, right and binocular stimulation, but fixation responses were near zero. In some ROIs, especially in V1 inner, the response during binocular stimulation was marginally higher than the response during monocular stimulation. In contrast, non-visually-driven ROIs (V1 outer, V2 outer and V3 outer; **Figure 4.5**) showed near-zero responses to all conditions, and unsurprisingly there was no difference between monocular and binocular stimulation relative to fixation. It is important to note that the control ROI, the FFA, showed a response pattern that was similar to many of the visually-driven ROIs, though responses were weak overall. This shows that our stimuli do indeed drive neural activity here.

The univariate analysis demonstrates that our stimulation protocol drives visual responses in key ROIs, including V1, V2, V3, hMT and hMST, as well as the control FFA ROI. As expected, it cannot tease apart differences between left and right eye stimulation.

4.4.2   *Eye of origin voxel analysis*

Beta weights for each voxel for four stimulation types (left eye stimulated, right eye stimulated, binocular stimulation, and fixation only) were extracted for each ROI in each participant. ROI beta weights were randomly sampled with replacement and cross-correlated between all stimulation types for each participant, and the bootstrapped cross-correlations

were averaged at the group level yielding group-level 4 x 4 cross-correlation matrices for each ROI. Results for the visually-driven ROIs are shown in **Figure 4.6** below.
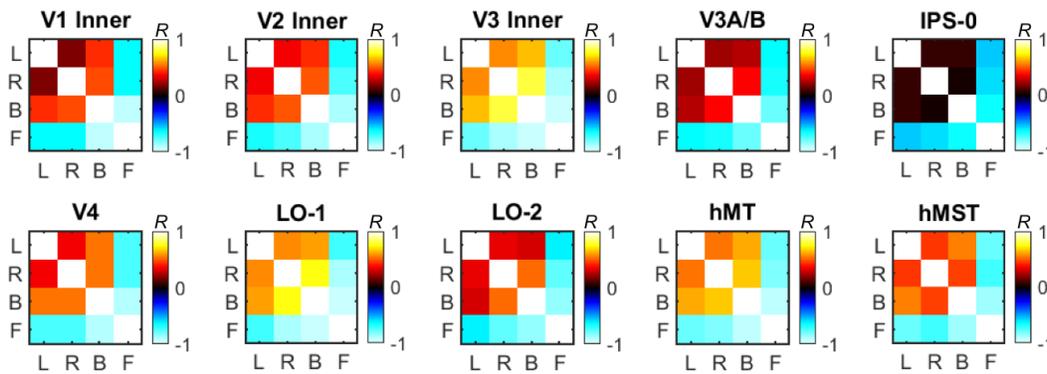
Figure 4.6 Bootstrapped cross-correlation matrices in visually-driven ROIs, for left eye stimulated (L), right eye stimulated (R), binocular stimulation (B), and fixation only EOO stimulus presentation conditions. The colour scale represents Pearson's R. Hot colours indicate correlations greater than zero (voxels are responding similarly between conditions); cold colours indicate correlations smaller than zero (anticorrelation; voxels are responding differently between conditions). Black represents zero correlation (no relationship between voxel responses in different conditions. Low positive or negative correlations between left and right eye conditions imply EOO tuning, especially when those correlations are significantly smaller than the monocular vs. binocular correlations.

Across ROIs shown in **Figure 4.6** (V1 inner, V2 inner, V3 inner, V3A/B, IPS-0, V4, LO-1, LO-2, hMT and hMST), correlations between visual stimulation conditions (left eye, right eye or binocular stimulation) were generally positive, with variations in the magnitude of the correlation coefficient. In contrast to this, correlating visual stimulation conditions against the fixation condition resulted in negative correlation, with the largest negative correlations between binocular stimulation and fixation. During left, right and binocular stimulus presentation conditions, most voxels were stimulated and therefore generated a pattern of mostly positive beta weights in each ROI. During fixation conditions however, many voxels are likely to be unstimulated or even suppressed, resulting in near-zero or negative beta weight patterns. This explains why cross-correlating visual stimulation conditions results in positive correlation coefficients, whereas correlating against fixation results in anticorrelation.

The overall pattern of cross-correlations is distinctive across ROIs. For example, in IPS-0, the correlation coefficients between left, right and binocular stimulation conditions were very small and just above zero. By comparison, in V3A/B the correlation coefficients were much higher. This is despite similar univariate responses across voxels (see **Figure 4.5**), and

thus is unlikely to be caused by differences in the overall response magnitude of the voxels. Rather, this implies that voxels in V3A/B are responding in a more consistent manner to all types of visual stimulation, than voxels in IPS-0.

As might be expected, the magnitude of correlations between left, right and binocular stimulation increases from V1 to V3, as voxels lose EOO tuning and respond more similarly between different stimulation conditions. EOO tuning is indicated by lower correlation coefficients between left and right stimulation, in comparison to the correlation coefficients representing the similarity of voxel responses between monocular and binocular stimulation conditions.

To quantify this, we computed an EOO index where the left vs. right (LvR) correlation coefficient was subtracted from the mean monocular vs. binocular (LvB and RvB) correlation coefficient. Values above zero indicate smaller LvR correlations relative to LvB and RvB correlations. Results are shown in **Figure 4.7**. EOO indices for each participant and each ROI were entered into a repeated measures ANOVA (eleven ROIs – V1 inner, V2 inner, V3 inner, V3A/B, IPS-0, V4, LO-1, LO-2, hMT, hMST, FFA). Post-hoc pairwise comparisons with Bonferroni adjustment across ROIs were extracted to compare each ROI against the control FFA ROI.
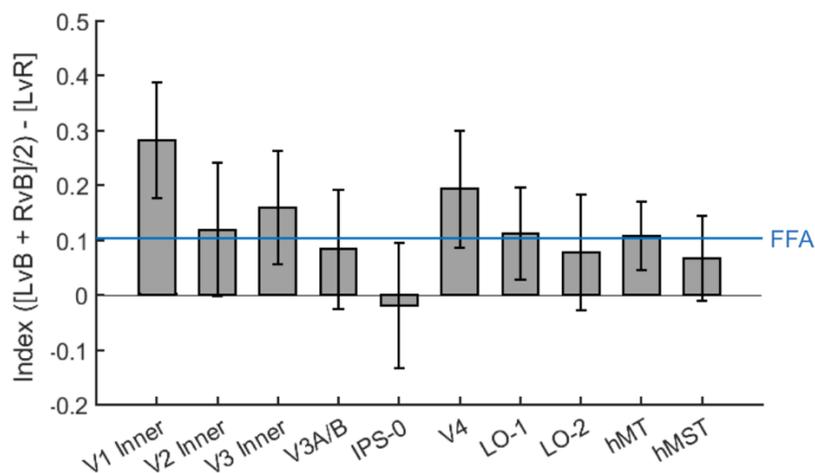
Figure 4.7 *EOO index across ROIs. The index was calculated by subtracting the left vs. right (LvR) correlation coefficient from the mean correlation coefficient for left vs. binocular (LvB) and right vs. binocular (RvB) stimulation. Values above zero indicate a smaller LvR correlation relative to monocular vs. binocular stimulation. The index in the control ROI, the FFA, is shown as a blue reference line. Indices significantly higher than the FFA index imply EOO tuning. Error bars represent ±1 SEM.*

The largest EOO index was measured in V1 Inner, followed by V4 and V3. In motion-sensitive areas V3A/B, hMT and hMST the index was between zero and that of the FFA. However the ANOVA found no significant main effect of ROI ($F(10, 70) = 1.12$, $p = .3111$, partial $\varepsilon^2 = 0.15$). Furthermore, none of the measured ROI indices were significantly different from the control FFA ROI, suggesting that we could not reliably detect EOO tuning in any of the ROIs.

Cross-correlations were analysed in the same manner in non-visually-driven 'outer' ROIs, as well as in the control ROI. Cross-correlation matrices for these ROIs are shown in **Figure 4.8**. Generally, correlation coefficients for LvR, LvB and RvB conditions were lower in the 'outer' ROIs than in their 'inner' counterparts, most notably in V1 where responses to left vs. right stimulation were anticorrelated. This implies that the responses to left eye stimulation were systematically different than responses to right eye stimulation, which could occur if a voxel is driven by one condition but is suppressed by the other.
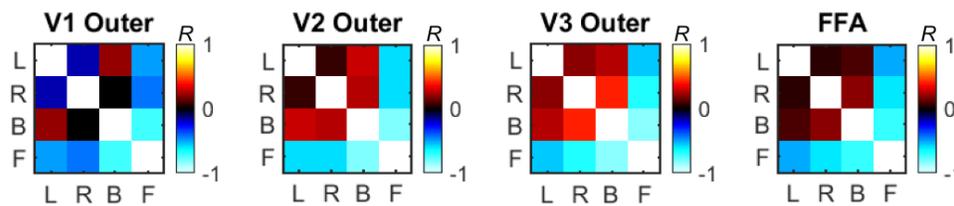
Figure 4.8 *Bootstrapped cross-correlation matrices in non-visually-driven and control ROIs, for left eye stimulated (L), right eye stimulated (R), binocular stimulation (B), and fixation only EOO stimulus presentation conditions. The colour scale represents Pearson's R. Weak positive or even negative correlations between left and right eye conditions imply EOO tuning, especially when those correlations are significantly smaller than the monocular vs. binocular correlations.*

Reassuringly, the correlations between L, R and B stimulation in the FFA were weak, but positive, and were similar between all stimulation conditions. Correlations were negative when compared to fixation. This is to be expected in an ROI that is visually driven, but not selective for any particular stimulation condition – in other words, for an ROI that contains no monocular tuning. Uniform low correlations may result from higher variability in voxel responses – in other words, from noisy response patterns. Thus, the FFA provides a good baseline for the amplitude of correlations that could be expected when voxels are driven by the stimulus but are not responding systematically.

Because EOO tuning is indicated by smaller LvR correlations compared to the monocular vs. binocular stimulation conditions, we again computed the EOO index in 'outer' ROIs and analysed these using a repeated measures ANOVA (four ROIs – V1 outer, V2 outer, V3 outer, FFA), followed by post-hoc pairwise comparisons. Results are shown in **Figure 4.9**.

The largest EOO index was measured in V1 Outer, but there was a high degree of variability across subjects. Again, there was no significant main effect of ROI ($F(3, 21) = 1.14$, $p = .360$, partial $\varepsilon^2 = 0.14$), and none of the indices were significantly different from the FFA.
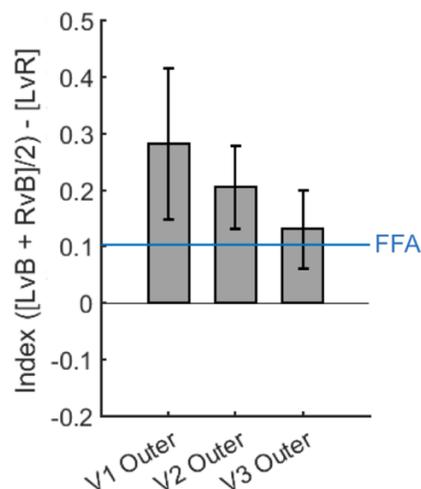
Figure 4.9 *EOO index across 'outer', unstimulated ROIs. The index was calculated by subtracting the left vs. right (LvR) correlation coefficient from the mean correlation coefficient for left vs. binocular (LvB) and right vs. binocular (RvB) stimulation. Values above zero indicate a smaller LvR correlation relative to monocular vs. binocular stimulation. The index in the control ROI, the FFA, in shown as a blue reference line. Indices significantly higher than the FFA index imply EOO tuning. Error bars represent ±1 SEM.*

4.4.3   Motion direction tuning

To complement the EOO voxel analysis, we also decoded tuning for motion direction in each voxel by training a support vector machine to classify event betas for up/down or left/right motion present in the EOO stimulus. Classification accuracy was determined using cross-validation where data were split into 10 even parts. This was carried out over 1000 iterations and the mean classification accuracy across voxels, across participants and in each ROI was calculated. Results are shown in **Figure 4.10**

One-sample *t*-tests were used to compare decoding performance against chance (50% accuracy). The support vector machine was unable to correctly label up/down and left/right voxel responses in all ROIs except in hMT, where performance was above chance (*t*(7) = 3.00, *p* = .020 – before Bonferroni correction). In other motion sensitive ROIs (V3A/B and hMST), performance was at chance.
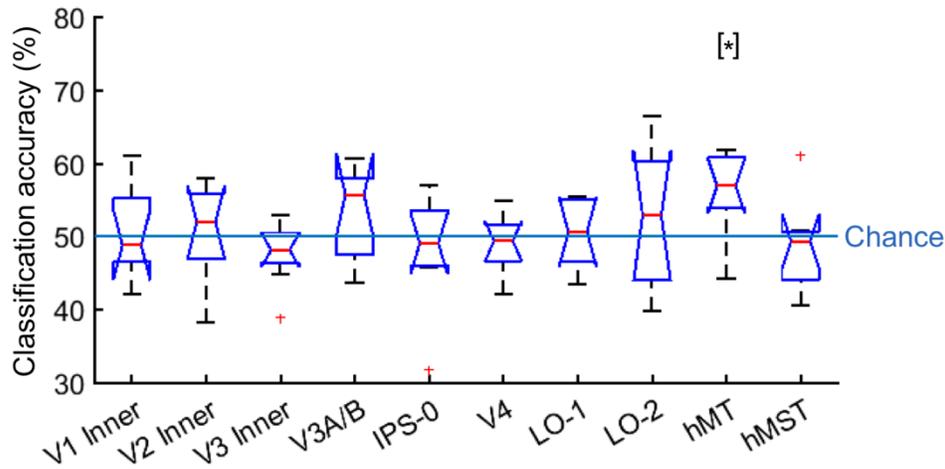
Figure 4.10 *Box plots showing classification accuracy for decoding motion direction tuning in voxels across 1000 folds of cross-validation. Motion could be up/down or left/right, therefore chance performance was at 50% and is indicated with a blue reference line. Decoding performance was above chance in hMT (indicated by [*]), but not in hMST or V3A/B.*

4.5   Discussion

In this study, we attempted to unify the availability of eye-specific (EOO) information with possible sites for IOVD computation. Both motion direction and eye-specific information are necessary during the binocular integration stage of the IOVD mechanism, where velocity vectors are combined between the eyes to generate an estimate of MID. To reveal possible ROIs where these signals can be extracted, participants viewed moving dot patterns that were shown to the left eye, right eye and both eyes during fMRI scanning, and we used MVPA and classification techniques to extract EOO and motion direction tuning from voxel response patterns.

An initial univariate analysis confirmed that our stimulation protocol evoked strong BOLD responses in key ROIs, including V1 inner, V2 inner, V3 inner, hMT, and hMST. We

measured weaker, but systematic, responses in the the control FFA region. In non-visually-driven ROIs (V1 outer, V2 outer, and V3 outer), mean responses during all stimulus conditions were near zero, with a subset of voxels showing negative BOLD responses – this allowed us to investigate EOO tuning in the suppressive surround of extraclassical receptive fields (Webb, Dhruv, Solomon, Tailby, & Lennie, 2005). Finally, the univariate analysis confirmed that the mean response across voxels was equal between left and right eye stimulation, implying that stimulus conditions were well-balanced. In averaging responses across voxels, information encoded at the small spatial scales of cortical columns is lost.

For this reason, we used MVPA and classification techniques to identify EOO and motion direction tuning, detectable in small variations in voxel response patterns within an ROI. For EOO, we extracted one beta weight per voxel per stimulation condition (left eye stimulated, right eye stimulated, binocular stimulation, and baseline fixation-only). Within each ROI, we cross-correlated voxel response patterns between these conditions.

Cross-correlation patterns were distinctive across ROIs. Correlating between any type of visual stimulation and the baseline fixation condition resulted in negative correlation coefficients, where the beta weight patterns were generally positive during left eye, right eye or binocular stimulation, whereas voxel responses were negative or suppressed during fixation conditions.

EOO tuning can be revealed by smaller correlation coefficients between left and right eye stimulation, than between left eye and binocular and right eye and binocular stimulation. This is because voxels that are biased towards either the left or the right eye should be driven strongly by stimulating the preferred eye, but the response should be weak or suppressed when its non-preferred eye is stimulated. Even in areas where there is a strict columnar organisation for eye-specific inputs (such as in V1), a large proportion of neurons are tuned to binocular inputs (Barlow, Blakemore, & Pettigrew, 1967; Hubel & Wiesel, 1970; Poggio & Fischer, 1977; Poggio et al., 1985) and therefore we would expect some positive correlations between left and right eye stimulation. Correlations between monocular and binocular conditions should, therefore, be larger, where stimulating the left eye only drives responses of left eye tuned and binocular voxels, and stimulating binocularly drives responses in all voxels irrespective of tuning. Thus there should be a greater similarity between voxel response patterns in these cases.

To quantify this, we computed an EEO index by subtracting the left vs. right correlation coefficient from the mean monocular vs. binocular correlation coefficient. Values greater than zero are expected because a larger left vs. right correlation than monocular vs. binocular correlations could only result by chance from noise. The larger the value, the greater the difference between monocular and binocular stimulation protocols and therefore the more eye-specific tuning is present.

The largest EOO index was measured in V1, which contains ocular dominance columns and therefore contains a systematic (and relatively coarse) representation of EOO information. However, the index measured here did not differ from the index measured in other ROIs. In addition, there were no differences between indices in ROIs relative to the FFA control ROI. Therefore, contrary to previous reports (Larsson et al., 2016), our stimulation protocol did not allow us to reliably decode EOO information. We found no evidence that IOVD mechanisms inherit EOO information from areas outside V1, though this null result does not exclude the possibility that these signals may be represented in a distributed manner in extrastriate areas (Czuba et al., 2014).

We also investigated possible monocular tuning in the negative BOLD response. The negative BOLD response is thought to originate from the suppressive mechanism in the extraclassical receptive field (Pasley, Inglis, & Freeman, 2007). Extraclassical receptive fields show some monocular tuning, especially in the input layers of primary visual cortex (Webb et al., 2005). Although we measured the largest EOO index in the 'outer' V1 ROI, this was not significantly different from the FFA index. More careful extraction of voxels that do indeed show the negative BOLD response (rather than selecting all those in the periphery – see section 4.7: Supplementary information, **Figure 4.11**) should reduce variation and may show more conclusive monocular tuning indices.

Indeed, it is important to note that the restriction of early ROIs to the stimulus-driven area of cortex, as well as the small stimulus size, is likely to have a large effect on the results presented in this chapter. This is true for the correlation analyses for detecting eye of origin information, as well as for the SVM decoding for detecting motion direction tuning. Because voxels in V1, V2 and V3 were selected depending on whether or not they were directly driven by the stimulus, these ROIs contain lower levels of noise, thereby increasing decoding

reliability. This may also explain why the BOLD response in these areas was larger than the BOLD response measured elsewhere (see **Figure 4.4**).

We used SVM decoding, rather than correlations, to reveal motion direction tuning in individual voxels. SVMs require a large number of observations relative to features, so we estimated one beta weight per stimulus event for each voxel. To increase the stability of these estimates we generated bootstrapped averages of several event betas. These mean event betas were split equally into six partitions and fed into the SVM, which was trained on five partitions and tested on the sixth. Decoding accuracy was estimated across 1000 iterations.

The SVM was able to decode motion direction (up/down or left/right) significantly above chance in hMT, an area that is characterised by its motion selectivity and contains many directionally tuned neurons (Born & Bradley, 2005). However, decoding accuracy in other ROIs that also contain directionally-tuned neurons, such as V1, V3A/B and hMST, was at chance.

Motion direction tuning has previously been decoded in early visual areas as well as in hMT+, albeit at higher field strengths and finer voxel resolutions (Beckett, Peirce, Sanchez-Panchuelo, Francis, & Schluppeck, 2012; Kamitani & Tong, 2006). Because large-scale biases have been found to drive decoding accuracy for both motion direction and EOO (Beckett et al., 2012; Larsson et al., 2016), the voxel resolution is unlikely to be the limiting factor in our own analysis.

Instead, optimising the stimulation protocol is expected to improve decoding accuracy for both motion direction and EOO. Here, we used a dense event-related design – previous research has employed longer stimulus durations and blocked designs to facilitate estimation of response amplitudes (Beckett et al., 2012; Kamitani & Tong, 2005; Larsson et al., 2016; Schwarzkopf et al., 2010). Because our events were close together, and could in principle occur at the onset of successive TRs, the fMRI signal often contained interactions between different event types. Increasing the ISI could reduce these interactions and improve the stability of amplitude estimates – although the linearity of the BOLD response holds for jittered ISIs as short at 2s (Boynton, Engel, Glover, & Heeger, 1996; Burock, Buckner, Woldorff, Rosen, & Dale, 1998; Dale, 1999; Dale & Buckner, 1997), and ISIs that are too long

are undesirable for a number of reasons including a reduction in statistical power (Serences, 2004).

Alternatively, a more nuanced approach to data analysis could improve the decoding of EOO and motion direction in our study. For example, instead of using a pre-defined, standard HRF, deconvolving the mean timecourse in each ROI and each individual with respect to event timings could yield an estimated HRF based on, and tailored to, the functional data. Because the HRF can vary substantially across individual subjects and across brain regions (Aguirre, Zarahn, & D'Esposito, 1998; Handwerker, Ollinger, & D'Esposito, 2004; Miezin, Maccotta, Ollinger, Petersen, & Buckner, 2000), this approach has been found to significantly improve estimates of response amplitudes in dense event-related designs (Hinrichs et al., 2000).

In addition to this, different modelling techniques extract amplitude estimates with varying efficiency, depending on the fitting procedure, and the choice of this fitting procedure can have a substantial impact on decoding performance (Pedregosa, Eickenberg, Ciuciu, Thirion, & Gramfort, 2015). Pedregosa et al. propose a method that combines both HRF estimation and the estimation of voxel-wise event amplitudes, using least-squares error minimization based on rank-one matrices. This effectively constrains the HRF to be the same across events (avoiding over-fitting), but allows it to differ between voxels. This approach was found to outperform decoding based on amplitude estimates generated by a standard GLM with more free parameters. Because our data show some promise in the ability to decode EOO in V1, and motion direction in hMT, we suggest that optimising our analysis techniques using these techniques may improve decoding accuracy in other ROIs.

4.6   Conclusions

In this study, we investigated whether EOO information is maintained beyond V1 to motion-selective areas, and which cortical regions may be involved in the extraction of the IOVD cue to MID. Because we were unable to reliably decode EOO, we did not identify areas where motion direction tuning and EOO information converge. Such an area could reflect the binocular integration stage of the IOVD mechanism. Our result implies that extrastriate areas inherit EOO information from V1, but we do not exclude the possibility that EOO information

is represented in a distributed fashion in extrastriate areas. We suggest that optimising the estimation of voxel-wise event response amplitudes will improve decoding accuracy, allowing the use of a SVM to decode the presence of EOO information in extrastriate areas.

4.7　　Supplementary information
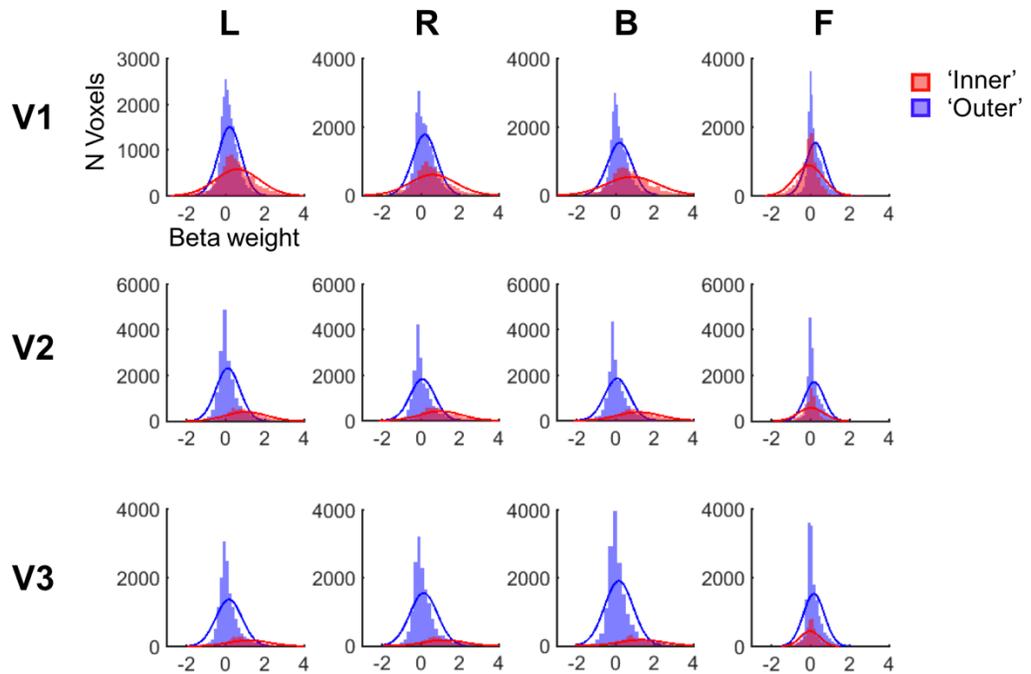
Figure 4.11 *Beta weight distributions across participants in inner (in red) and outer (in blue) V1, V2 and V3 ROIs, showing a leftward shift (into more negative values, indicating negative BOLD responses) for outer ROIs in all conditions (right eye stimulated, R; left eye stimulated, L; and binocular stimulation, B) except fixation (F). Note distributions for inner ROIs are flatter than for outer ROIs because the inner ROIs have fewer voxels in total, and there is a larger spread in responses as most 'outer' voxels are unstimulated and clump around 0.*

Chapter 5.   Attentional modulation of human cortical areas involved in perceiving motion in depth

5.1   Abstract

In humans, discrete areas on the lateral occipital cortex are associated with both two- and three-dimensional motion processing. These areas can be driven in a bottom-up manner by two, potentially independent cues to motion in depth (MID): changing disparity (CD) and inter-ocular velocity differences (IOVD). However, the response profiles of areas involved in the active perception of MID remain unclear. Here, we used fMRI and an attentional manipulation to highlight neuronal populations involved in attention *to* – and perception *of* – these two MID cues.

A random-dot stimulus isolating either CD or IOVD cues was constantly presented throughout a set of fMRI scans. During each scan, brief events (subtle changes in motion in depth, local contrast or fixation letter contrast) were continuously interleaved at random. Subjects were cued to detect a single class of stimulus change in blocks of 15 seconds at a time. Thus, attentional state changed systematically throughout the scan, but the stimulus remained constant on average.

Attention to either contrast or MID resulted in greater modulations in early visual areas (V1, V2, V3 and V4) than the fixation task. The largest attention-generated fMRI responses were found in areas in and around the human MT complex (hMT+) and were driven by attention to MID. However, no differences between CD and IOVD attention conditions were found at a univariate level in these areas.

We also measured the blood oxygen level-dependent (BOLD) response as a function of probe detection: areas that had stronger responses in 'hit' compared to 'miss' conditions may

have a causal role in MID perception. We found strong evidence of hit/miss dependency in hMT+ for CD but not for IOVD stimuli suggesting that while conventional motion responsive areas may be driven by IOVD stimuli, they may not be the site at which the final, perceptually-relevant 3D motion signal is extracted.

5.2   Background

There are two distinct binocular sources of motion in depth (MID) information – changes in retinal disparity over time (changing disparity; CD), and differences in the sign and amplitude of motion in the left and right eyes (inter-ocular velocity difference; IOVD). These cues can be dissociated mathematically (Regan, 1993; Rushbass & Westheimer, 1961), and recent psychophysical evidence suggest that they can independently drive a percept of MID (Brooks, 2002; Czuba, Rokers, Huk, & Cormack, 2010; Fernandez & Farell, 2005; Nefs, O'Hare, & Harris, 2010; Rokers, Cormack, & Huk, 2008; Shioiri, Saisho, & Yaguchi, 2000). Because CD and IOVD depend on two different early sources of information – retinal disparities in the case of CD, and monocular velocity estimates in the case of IOVD – they may be dissociated at the level of the cortex. Here, we used an attentional manipulation to reveal differences in the processing of the two cues.

Neuroimaging studies investigating the cortical pathways involved in CD and IOVD processing typically employ a bottom-up, stimulus driven paradigm. In these studies, the low-level cues that convey CD and IOVD are carefully isolated to reveal populations of neurons that are involved in some stage of MID computation. For example, one fMRI study compared the blood oxygen level-dependent (BOLD) response to CD and IOVD stimuli against a set of control stimuli that contained matched disparity or velocity signals, but did not convey MID (Rokers, Cormack, & Huk, 2009). These contrasts subtract the constituents of the stimuli, leaving behind only those signals that convey coherent MID and revealing cortical areas involved in CD and IOVD processing. Such areas include hMT and hMST for both CD and IOVD, leading to the suggestion that binocular MID signals share pathways with mechanisms involved in the extraction of 2D motion and depth (Cormack, Czuba, Knöll, & Huk, 2017; Huk, 2012).

Another approach has been to use fMRI adaptation protocols. In areas containing neural populations tuned to MID direction, such as hMT+, adapting to a particular direction of MID resulted in a decreased BOLD amplitude when the test stimulus was moving in the same direction as the adaptation stimulus (Huk, 2012; Ponce, Lomber, & Born, 2008). Testing the CD and IOVD cues separately resulted in similar adaptation effects, revealing those areas that selectively code for CD- or IOVD-defined direction of MID (Joo, Czuba, Cormack, & Huk, 2016). However, this study found no evidence for cross-cue adaptation, where the BOLD response in hMT+ was not suppressed when the test stimulus contained a different MID cue than the adaptation stimulus. This suggests that whilst CD and IOVD engage similar cortical regions, different sub-populations of neurons within these areas are tuned to either CD or IOVD motion.

These bottom-up approaches have been fruitful in revealing a network of areas that are driven by CD and IOVD stimuli; however, these methods cannot dissociate those areas that are merely responding to the stimulus, from those that are critical for driving the MID percept. Furthermore, these studies have not directly related neural activation patterns to participant behaviour during the scan. It is unclear where neural populations crucial for solving a MID-related task sit within the identified network of areas involved in the computation of CD and IOVD.

Here, we isolated top-down attentional mechanisms to address some of these outstanding questions. We asked participants to attend to different features within a CD or an IOVD stimulus, where attended features were related to the MID percept itself or a subtle, local change in stimulus contrast. Attending to a specific stimulus feature has been shown to enhance the response gain of neurons tuned to that feature (Martinez-Trujillo & Treue, 2004; Treue & Martinez-Trujillo, 1999; Treue & Maunsell, 1996). This 'feature-similarity gain' mechanism has been observed across multiple regions in visual cortex, including hMT+ when subjects attended to the direction or speed of a moving stimulus (Beauchamp, Cox, & Deyoe, 1997; M. Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1990; Treue & Martinez-Trujillo, 1999). More recently, fMRI pattern classification approaches have demonstrated that the direction of stimulus motion can be decoded from attentional modulations in motion-selective areas, even when the stimulus was outside the spatial receptive fields of the neurons (Serences & Boynton, 2007). Whilst attention-driven gain modulations are tightly related to task, the *spread* of this mechanism can extend beyond the spatially attended area – even occurring in the absence of visual stimulation – and is not restricted to neural populations that

would be stimulus-driven (Chawla, Rees, & Friston, 1999; Kastner, Pinsk, De Weerd, Desimone, & Ungerleider, 1999; Reynolds & Heeger, 2009; Saenz, Buracas, & Boynton, 2002; Serences & Boynton, 2007). Thus, featural attention is a mechanism that can be decoupled from bottom-up neural activity, isolating neural populations that are maximally informative for solving a specific task (Verghese, Kim, & Wade, 2012).

We used fMRI to investigate how the BOLD signal changes as subjects attend to either CD or IOVD stimuli. Participants viewed a continually oscillating CD or IOVD stimulus, in which we embedded occasional, independent probes defined by subtle changes in MID, contrast, or a change in fixation point colour. The timings of all three probe types were random over the course of a scan, but attentional state was manipulated systematically by instructing participants to attend to only one of the three probe types at any time. We hypothesised that attending to contrast or attending to MID would engage different cortical networks, with the MID task engaging motion-selective areas hMT and hMST. Furthermore, we asked whether we could identify areas whose activity was correlated with the *perception* of MID (as opposed to the presence of the bottom-up stimulus change). By analysing responses in different cortical areas when participants correctly identified attended targets ('hit'), versus when those attended targets were not detected ('miss'), we expected to identify neural populations whose response profile tracks the perceptual experience of the participants, and are thus crucial for solving a MID task.

5.3   Methods

5.3.1   Participants

Twelve participants (age 22 to 45 years, 5 male) with normal or corrected-to-normal vision and normal stereo-acuity (below 120 arcsec, measured using the TNO test, 19th edition, Laméris Ootech, Ede, The Netherlands) were recruited for this study. Two subjects were authors (MK and ARW) and were experienced psychophysical observers; the rest were naïve. This study was approved by the ethics committee of the York NeuroImaging Centre, and conformed to the standards set by the Declaration of Helsinki. Informed written consent was obtained from all participants.

Prior to scanning, participants were invited to the lab for an initial screening session. Participants were asked to indicate the motion trajectory of a CD or IOVD-isolating stimulus that was continually oscillating forwards and backwards in depth. The screening stimulus was based on fMRI stimulus, with identical parameters but no embedded events. All participants were able to distinguish the two directions of 3D motion in these stimuli. To familiarise participants with the fMRI task (detailed below, section 5.3.4), each participant completed a practise fMRI run outside the scanner for CD and IOVD motion in depth.

5.3.2   Apparatus

The experiment was run from a Shuttle XPC SZ87R6 high-end graphics system fitted with an Intel Core i7-4790K processor at 4.0 GHz, and an NVIDIA GeForce GTX970 graphics card with 4 GB DDR5 memory. Stimuli were generated and controlled using Matlab 8.5.0 (2015a) in conjunction with the Psychtoolbox 3.0.12 routines (Brainard, 1997; Pelli, 1997).

During scanning, stimuli were presented using a PROPixx DLP LED projector (VPixx technologies, Saint-Bruno, Canada) running at 240 Hz with a pixel resolution of 1920 x 1080. A long-throw lens back-projected the stimulus images through a waveguide located behind the scanner bore, on to a silver screen positioned behind the participant. Images were viewed through a mirror mounted on the head coil (57 cm viewing distance, including the optical pathway of the mirror), with a viewing angle of 41.0° x 23.5°.

Binocular stimulus presentation was achieved using a circular polarization modulator (DepthQ, Lightspeed Designs, Bellevue, USA) placed in front of the projector lens. Left and right eye images were temporally interleaved, where the switch between polarizer states was synchronized with the video frame refresh rate, yielding a 120 Hz refresh rate per eye. Stimuli were viewed through MR-safe glasses fitted with polarizing filters that separated the left and right eye images.

5.3.3   3D motion stimuli

Stimuli were variants on dynamic random dot stereograms (DRDS; Julesz, 1971) and isolated the CD and IOVD cues. They were identical to those described in Chapter 3, but they were viewed through a larger Gaussian aperture (0.5° inner and 6° outer radius, edges smoothed using a kernel at 0.5° FWHM). A schematic of MID stimuli as seen by participants is shown in **Figure 5.1**, panel D. For all MID stimuli, an inner fixation ring and a fixation ring in the periphery stabilised binocular fusion. A red letter positioned in the centre of the inner fixation ring (M, C, F, or R) was used to cue the attentional state of the participant. All stimuli were shown at 100% contrast on a mean grey background.

In brief, the CD stimulus generated a percept of sinusoidal oscillation in depth by gradually incrementing and decrementing the retinal disparity between pairs of left and right eye dots, to a maximum of ±12 arcmin monocular disparity (**Figure 5.1**, panel A). The 3D oscillation rate was 1.4Hz. Dots were refreshed at the onset of each video frame, eliminating any local motion or IOVD cues.

In the IOVD stimulus, dot patterns were decorrelated between the left and right eyes, and all dots moved in a coherent direction within each eye (**Figure 5.1**, panel B).). Motion between the eyes was in opposite directions. The difference in sign between the motion vectors generated in the left and right eyes gave the cue to MID. Dots moved to a maximum lateral shift of ±100 arcmin in each eye and had a maximum lifetime of 50ms. The rate of the MID oscillation was 1.1Hz.

The probability of spurious disparity cues from chance matches of elements between the eyes in the IOVD stimulus was reduced by dividing the stimulus into stripes, and presenting dots in alternating stripes between the eyes. This method has previously been used to isolate the IOVD cue (Shioiri, Nakajima, Kakehi, & Yaguchi, 2008; Shioiri et al., 2000). Furthermore, when dots fell close to the borders of these stripes and in close proximity between the eyes, their contrast polarity was anticorrelated to degrade the disparity cue (Czuba et al., 2010; Rokers et al., 2008, 2009). These methods are illustrated in **Figure 5.1**, panel C.
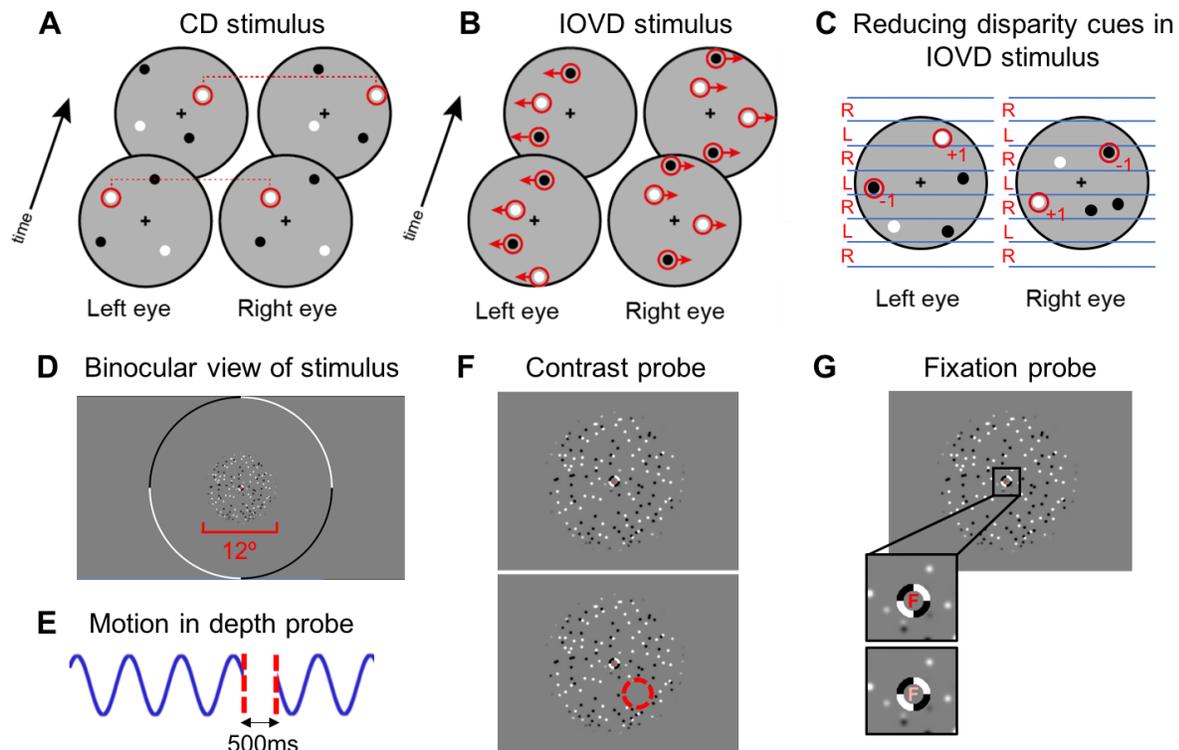
Figure 5.1 *Stimuli and probe types for the attention fMRI study. For the CD stimulus, the retinal disparity of binocular dot pairs increased or decreased systematically over time, but the dot pattern was refreshed at each frame (panel A). For the IOVD stimulus, dots are unpaired between the left and right eye views, but are moving in an opponent direction between the eyes (panel B). Binocular matches are reduced in the IOVD stimulus by decorrelating dots in the left and right eyes, spatial alternation of strips of dots shown to the left and right eyes, and anticorrelation of dot contrast in nearby dots that fall close to the edges of left and right eye strips (panel C). The base stimulus shown to participants was a continually oscillating 3D motion stimulus, given by either the CD or the IOVD cue (panel D). Two fixation rings in the centre and the periphery helped stabilise the percept. The attentional state was cued using the letters M (detect motion in depth probes), C (detect contrast probes), F (detect fixation probes) or R (rest) at fixation. During M probes, the 3D motion oscillation was interrupted (panel E). During C probes, dots inside a Gaussian window presented at any angle within the field of dots decreased 100% in contrast (panel F). During F probes, the contrast of the fixation letter reduced. All probes lasted for 500ms and were continually presented in three interleaved timing sequences throughout the fMRI run. Participants responded by button press when the attended probe type was detected.*

5.3.4   Attention task

This study aimed to isolate top-down signals relating to the attentional state of the participant. Holding all stimulus-driven signals constant throughout the scan was crucial, with the exception of the brief probe events which occurred at random. To ensure this, we presented the CD or the IOVD stimulus continually throughout an entire fMRI run (6 minutes each). We embedded three independent timing sequences for three attentional probe types

(MID, contrast, or fixation probes – detailed below) within this 3D motion stimulus. Timings for the onset of probes was based on an exponential distribution with a mean of 7.5s, and lower and upper bounds of of 1.5s and 14.5s. To avoid overlaps probes never began less than 500ms before the end of an 'attentional state' block. Because the timing sequences were fully independent, probes of different types could co-occur. The key manipulation was to direct the participant to attend to only one of these three probe types, ignoring the other two, and alternating between probe types in 15s 'attentional state' blocks. Participants were cued with a red letter M, C, F or R at fixation. On average, each probe type was likely to occur twice in each block. When participants detected a probe that corresponded to the attentional cue, they were asked to press a button. Thus, whilst the attentional state of the participant varied systematically across each fMRI run, the low-level cues did not. A schematic timecourse of a single fMRI run is shown in **Figure 5.2**.
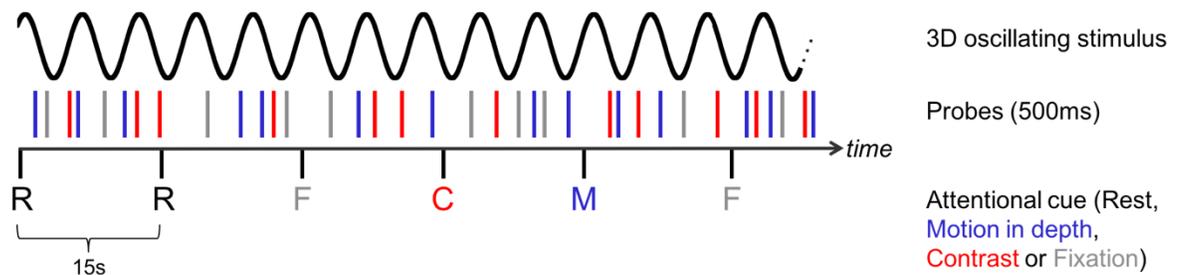
Figure 5.2 *Design of each fMRI run. A 3D oscillating stimulus (CD or IOVD) is continually presented. Three independent timing sequences for motion in depth, contrast or fixation probes are embedded within the main stimulus. Participants are cued with an M, C, F or an R to detect the motion in depth (M) probes, detect the contrast (C) probes, detect the fixation (F) probes, or to rest (R), respectively. All probes can occur at any time during the run. On average, the stimulus is constant whereas attentional state varies systematically.*

Probe types are illustrated in **Figure 5.1**, panels E-G. During the motion in depth probe, the oscillatory motion of the 3D stimulus was briefly interrupted. For the CD stimulus this was achieved by holding the retinal disparity constant at the disparity prior to the onset of the probe, and for the duration of the probe, before continuing on its increment or decrement. Effectively, the plane of oscillating dots appeared to freeze at a given depth (although dot update continued) before continuing on its previous trajectory.

For the IOVD stimulus, the 3D motion was interrupted by switching to a 'nulled IOVD' stimulus that contained the same lateral motion energy, without the 3D motion cue – this was identical to the 'IOVD control stimulus' described in Chapter 3. In this stimulus, dots moved

laterally in both directions and in both eyes – a mixture of the two eyes' views of the IOVD stimulus. Because the average velocity signal on the two retinae was equal, no 3D motion vector could be extracted. Other than this disruption, the low level properties of the stimulus were identical to the IOVD stimulus. The percept of the switch between the IOVD stimulus and this 'nulled IOVD' stimulus was similar to a break in binocular fusion.

During the contrast probe, the dots within a Gaussian window (2.12º at full-width half-maximum) decreased 100% in contrast. This Gaussian window could be placed at any angle, but always at a fixed radius (half-way between the inner and the outer edge of the stimulus, such that the peak of the Gaussian was centred at 3.5 º from fixation), within the field of dots. Perceptually this was described as a contrast 'blip', where a small portion of the dots became invisible.

To avoid the effects of spatial attention, both MID and contrast probes required that the participant spread their attention across the whole field of dots – either to generate a reliable estimate of the 3D motion, or because the location of the contrast decrement was unpredictable.

During the fixation probe, the fixation letter (always an M, C, F or R to cue the attentional state of the participant) dropped in colour saturation from red to pink, where the RGB values changed from [255 0 0] to [255 128 128]. This probe was included as a baseline to the MID and contrast probes, and is similar to fixation tasks often used in stimulus-driven fMRI to exclude attentional effects.

We also included a 'rest' condition where the participant was not required to respond to any cues, but passively viewed the stimulus. All fMRI runs began with two rest blocks, followed by one fixation block. After that the attentional blocks were alternated randomly to avoid predictive effects.

Participant performance (given by the proportion of correctly identified probes; **Figure 5.3**) was below ceiling for MID and contrast probes, but well above chance level. The chance was calculated assuming that responses made by the participant were unrelated to the incidence of a probe by shuffling button response data 1000 times with respect to probe

onset, for each separate subject and each separate fMRI scan. This preserved the frequency at which participants made a response but removed any relationship to the probe timings. The proportion of those responses that fell, by chance, within 1.5s window from the onset of a probe was calculated. Assuming this estimated guess rate, the proportion of probes detected by chance was around 0.05.
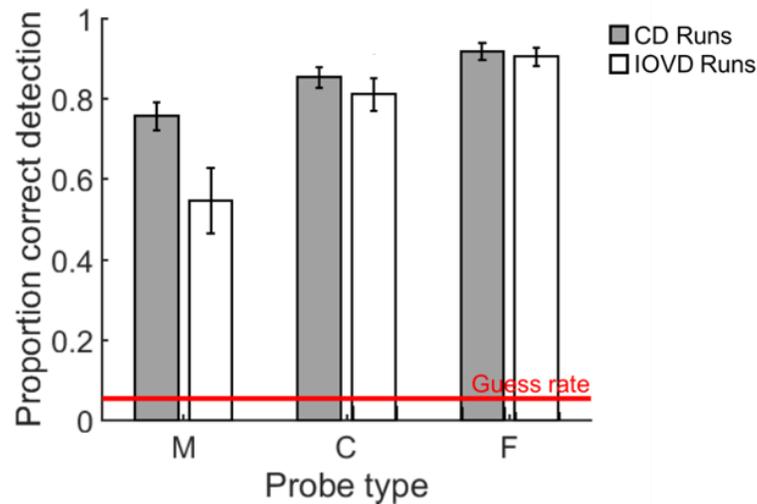
Figure 5.3 *Subject performance, given by the proportion of correctly identified MID (M), Contrast (C), and Fixation (F) targets. Data from CD runs are in grey, and data from IOVD runs are in white. Error bars are ±1 SEM. The guess rate (red reference line) was calculated by shuffling participants' responses relative to the onset of the probe times.*

Because performance was unequal between conditions, the proportion of correctly detected probes for each participant were entered into a repeated measures ANOVA with two factors – one for motion type (CD or IOVD) and one for probe type (MID hit, contrast hit, and fixation hit). There were significant main effects for both motion type ($F(1,11) = 7.08$, $p = .022$, partial $\varepsilon^2 = 0.39$) and probe type ($F(1.26, 13.84) = 21.96$, $p < .001$, partial $\varepsilon^2 = 0.67$, Greenhouse-Geisser correction for sphericity), with a significant interaction between these two factors ($F(1.31,14.35) = 5.61$, $p = .025$, partial $\varepsilon^2 = 0.34$, Greenhouse-Geisser correction for sphericity). Overall, performance was higher during CD runs than IOVD runs (CD > IOVD, $p = .022$). In addition, performance was lowest on the MID probe detection task, (MID < C, $p = .013$, and MID < F, $p < .001$), and highest on the fixation probe detection task (F > C, $p = .007$, and F > MID, $p < .001$). Finally, the interaction term was driven by greater performance on the MID probe detection task for CD runs than for IOVD runs ($M_{CD} > M_{IOVD}$, $p = .020$). All reported p values for post-hoc comparisons are Bonferroni corrected.

5.3.5 MRI parameters

During the main scan session, one anatomical reference scan, one stimulus localiser scan, and six attention fMRI scans were collected on a 3.0T GE Signa Excite HDX MRI scanner, using a 16-channel phased-array receive coil (Novamed) covering the back of the head to maximise the signal-to-noise ratio. A T1-weighted anatomical reference scan (TR = 2100ms; TE = 8.6ms; flip angle = 12°; FOV = 19.2 x 19.2cm; matrix size = 512 x 512; voxel resolution = 0.38 x 0.38 x 2.5mm; 39 quasi-coronal, contiguous slices oriented along the calcarine sulcus and covering the occipital lobe) was acquired for registration of the functional data to a high-resolution anatomical scan.

Gradient-echo sequences were used for the functional scans. The stimulus localiser scan (TR = 3000ms; TE = 30ms; 116 volumes including 3 dummy volumes; flip angle = 90°; FOV = 19.2 x 19.2cm; matrix size = 96 x 96; voxel resolution = 2 x 2 x 2.5mm) used the same slice prescription as the reference anatomical scan. Data from this scan were used to restrict visual areas V1, V2 and V3 to the retinotopic extent of cortex directly stimulated by the 3D motion stimulus (see below, section 5.3.6: Identifying ROIs). Six attention fMRI scans (3 x CD and 3 x IOVD) were identical to the stimulus localiser scan, but 125 volumes (including 5 dummy volumes) were collected. Stimulus presentation was synchronized with data acquisition on the scanner, accounting for five dummy volumes.

High-resolution, whole brain T1-weighted structural scans (TR = 7.8ms; TE = 3.0ms; TI = 600ms; flip angle = 20°; FOV = 25.6 x 25.6 cm; matrix size = 256 x 256; voxel resolution = 1.0 x 1.0 x 1.0mm; 176 coronal slices to cover the whole head) were acquired on the same scanner for each participant in a separate scanning session, using an 8-channel whole-head phased-array coil (MRI Devices Corporation).

Finally, we acquired two motion localizer scans per participant to identify areas V3AB, hMT and hMST. The motion localiser scans are described in Chapter 3. Standard retinotopic mapping scans (typically five wedge and two ring scans, with eight stimulus cycles each) were carried out to delineate early visual areas. These data were collected in separate scan sessions, using fMRI parameters similar to those detailed above.

5.3.6   Identifying ROIs

Regions of interest (ROIs; V1, V2, V3, V4, V3A/B, IPS-0, LO-1, LO-2, hMT and hMST) were defined in a similar manner to the methods described in Chapter 3, with the addition of a stimulus localiser to restrict early visual areas in the eccentricity dimension.

Briefly, visual areas V1, V2, V3 (Dougherty et al., 2003; Schira, Tyler, Breakspear, & Spehar, 2009; Sereno et al., 1995), V4 (Brewer, Liu, Wade, & Wandell, 2005; Hansen, Kay, & Gallant, 2007; Wade, Brewer, Rieger, & Wandell, 2002; Winawer, Horiguchi, Sayres, Amano, & Wandell, 2010), LO-1, LO-2 (Larsson & Heeger, 2006) and IPS-0 (Press, Brewer, Dougherty, Wade, & Wandell, 2001; Swisher, Halko, Merabet, McMains, & Somers, 2007; Tootell et al., 1998) were delineated using phase reversals in the polar angle maps acquired using conventional retinotopic mapping methods (Engel, Glover, & Wandell, 1997; Wandell, Dumoulin, & Brewer, 2007; Wandell & Winawer, 2011). V1, V2 and V3 were restricted to match the extent of cortex that was directly stimulated by the MID stimuli. To do this, participants completed a stimulus localizer fMRI scan, where a contrast reversing checkerboard (50% contrast, phase updated at 1 Hz) alternated in 9s blocks between an 'inner' annulus and an 'outer' annulus. The size of the 'inner' annulus matched the size of our experimental stimuli (extending from $0.5º – 6º$ from fixation), whereas the 'outer' annulus mapped the periphery (extending from $7º – 11.75 º$ from fixation). ROIs were restricted based on a general linear model (GLM) contrast comparing the BOLD response during presentation of the 'inner' vs. the 'outer' stimulus. We checked these ROIs against their retinotopic eccentricity maps to ensure correspondence to the known stimulus size.

Motion-selective areas V3AB, hMT and hMST were identified using motion localisers and retinotopic mapping (Amano, Wandell, & Dumoulin, 2009; Fischer, Bülthoff, Logothetis, & Bartels, 2012; Huk, Dougherty, & Heeger, 2002). A detailed description of this process is given in Chapter 3.

5.3.7   Analysis

For the whole-brain analysis of the attention fMRI runs, data from each subject were analysed using a standard FEAT pipeline (fMRI Expert Analysis Tool, version 6.0, part of

FMRIB's Software Library; www.fmrib.ox.ac.uk/fsl). The first five dummy volumes were deleted to allow scanner magnetization to stabilise, and signal intensity was normalised in each 4D dataset by a multiplicative factor of the grand mean. The time-series of each voxel was temporal high-pass filtered to remove slow signal drift (Gaussian-weighted least-squares straight line fitting, sigma = 50.0s). Non-brain structured were removed using BET (Smith, 2002) and motion correction was performed on each dataset using MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002). Timeseries were smoothed using a Gaussian kernel at 3mm FWHM.

To register each functional dataset to the 2mm MNI-152 mean brain template, the structural reference scan was FAST-corrected (Zhang, Brady, & Smith, 2001) and BET extracted (Smith, 2002) to improve the signal drop-off at the front of the head, and remove non-brain structures. This was aligned to the mean brain using FLIRT (Jenkinson et al., 2002; Jenkinson & Smith, 2001), and the resultant transformation matrix was applied to the corresponding 4D datasets.

Attentional state blocks (attend motion, attend contrast, attend fixation and rest) were modelled as separate predictors in a general linear model (GLM) using FILM (Woolrich, Ripley, Brady, & Smith, 2001), where each predictor was convolved with a standard hemodynamic response function (3s std, 6s lag). Planned contrasts were carried out to compare attentional states to fixation and rest, as well as to compare attentional states to one another. Z-statistic images were cluster corrected at a significance level of $p \le .050$. Results for the CD runs, IOVD runs, and All runs were averaged at the individual subject level, before averaging at the group level in a mixed-effects analysis (Beckmann, Jenkinson, & Smith, 2003; Woolrich, 2008; Woolrich, Behrens, Beckmann, Jenkinson, & Smith, 2004).

For the ROI analyses, data were processed on an individual level in mrVista (https://web.stanford.edu/group/vista/cgi-bin/wiki/index.php/Sortware; Vista Lab, Stanford University) and Matlab 8.5.0 (2015a; TheMathWorks Inc., Natick, MA, USA). Five dummy volumes were discarded and motion correction was carried out between and within the scans. The reference T1-weighted image, and the functional data, were aligned to a high-resolution T1 anatomical scan for each participant using the Nestares algorithm (Nestares & Heeger, 2000); to improve this alignment we used a FAST-corrected and BET-extracted reference scan. For volume and surface-based reconstruction, the high-resolution anatomical scan was

segmented into grey and white matter using automated algorithms implemented in Freesurfer v5.3. Using this segmentation, activation in the fMRI datasets was restricted to the grey layers only. Subsequent analyses were carries out on the grey-layer voxels only.

We ran two different GLM analyses on concatenated CD scan timecourses, and on concatenated IOVD scan timecourses. First, we modelled the BOLD response to all probe types by the attentional state the participant was in, giving a total of twelve predictors (eg. motion in depth probe during attend motion in depth block, motion in depth probe during attend contrast block, etc.). Predictors were convolved with a 'difference of Gammas' HRF (from the SPM 8 toolbox, http://www.fil.ion.ucl.ac.uk/spm/). This yielded an estimate of the BOLD amplitude, given by beta weight, for all probe types under all attentional states.

Second, we modelled the BOLD response to 'hit' and 'missed' attentional probes. In this analysis, there were 6 predictors in total – 'hit' and 'miss' responses within all three attentional conditions (motion in depth probe hit, motion in depth probe miss, contrast probe hit, contrast probe miss, fixation probe hit and fixation probe miss). We considered a probe to be 'hit' when the participant pressed the response button within 1.5s of the onset of the probe, when that probe corresponded to the cued attentional state.

Beta weights for both of these GLMs were extracted for each voxel in each ROI (V1, V2, V3, V4, V3AB, LO-1, LO-2, IPS-0, hMT and hMST), and were averaged across voxels to give one beta estimate per predictor, per ROI and per participant. Finally, these ROI beta weights were averaged across participant for a group-level analysis.

5.4   Results

5.4.1   *Whole-brain analysis of attentional states*

The BOLD response during an attentional block (15 seconds each of Attend Motion, Attend Contrast, or Attend Fixation, with additional Rest blocks) was modelled with a GLM. We modelled the response across all 3 CD scans, all 3 IOVD scans, and all 6 scans together for each participant, before using a mixed effects analysis to average across all 12 subjects. Cluster-corrected group maps set to a significance threshold of $p < .050$ were plotted on the

MNI-152mm brain, showing changes in the BOLD response during different attentional states (**Figure 5.4**).
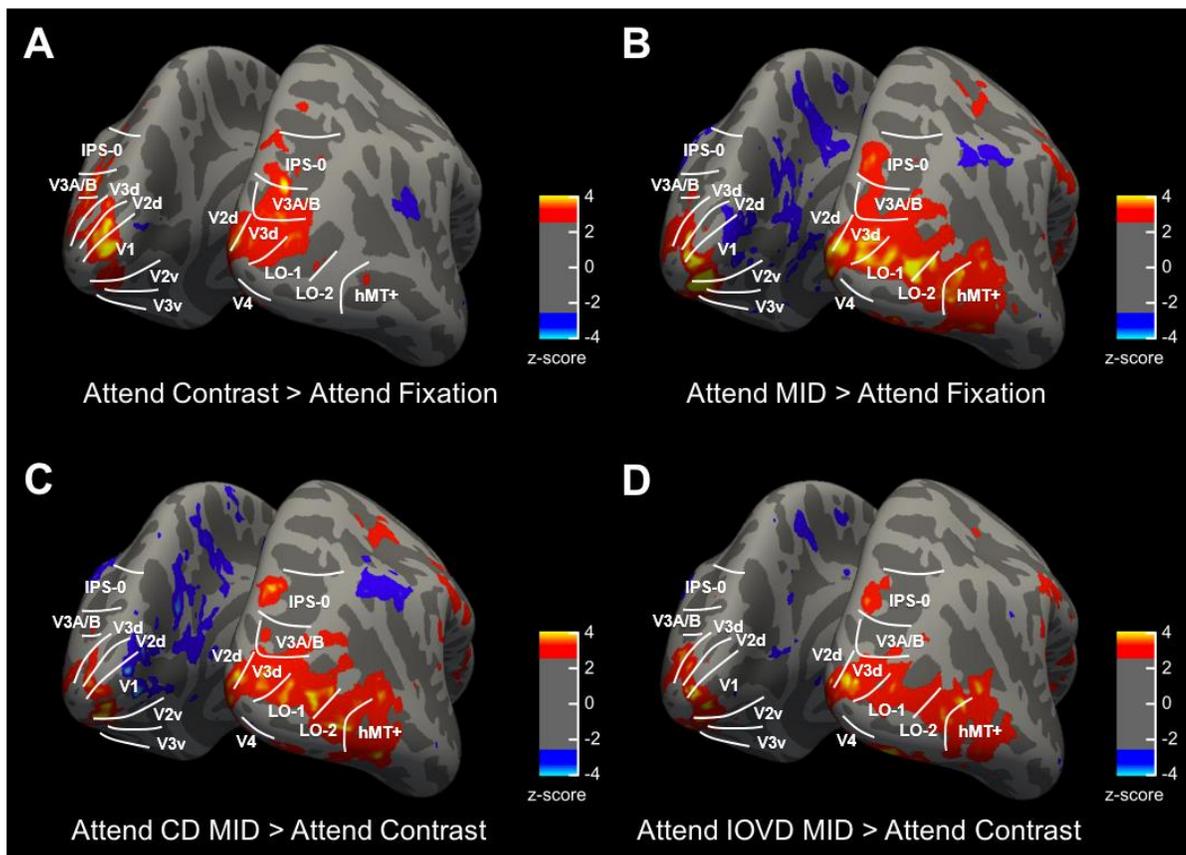
Figure 5.4 *Group-level (N=12) changes in the BOLD response as a function of attentional state. Comparing the contrast map in Panel A (Attend Contrast > Attend Fixation) against Panel B (Attend MID > Attend Fixation), reveals how the BOLD response shifts as participants switch their attention from the contrast changes in the stimulus to the MID changes in the stimulus. Both attention states result in modulation in early visual areas, but only the MID attentional blocks result in activation in motion-selective hMT+. Panels C and D show similar modulations irrespective of whether the MID stimulus was CD- or IOVD-driven. CD MID and IOVD MID attentional blocks are compared against Contrast attentional blocks to avoid spatial attention effects. Data were smoothed with a 3mm Gaussian kernel at FWHM, Z-statistic maps are cluster-corrected and set to a significance threshold of p < .050 (z-score of 2.5 and higher).*

All low-level properties of the stimulus were constant throughout all scans, where the 3D motion stimulus was presented continually throughout the scan and all three attention probes (events consisting of changes in MID, contrast and at fixation) could occur at any time. Thus, each attentional block contained, on average, the same low-level image properties. Contrasting the BOLD response between blocks thus isolated the feedback mechanisms resulting from changes in attention, rather than in feed-forward signals

pertaining to the stimulus features. As illustrated in **Figure 5.4**, panels A and B, the whole-brain analysis revealed a large-scale shift in the BOLD response as a function of a shift in the attentional state of the participants. These comparisons conflate spatial and featural attention because detecting changes in contrast and MID require attending to the whole stimulus field, whilst detecting changes at fixation require attending to the central part of the visual field only. Attending to both contrast changes and MID changes in the stimulus resulted in large-scale BOLD modulations in early visual cortex (V1, V2, V3 and V4) and extending into the intraparietal sulcus (IPS). As predicted, for MID attentional states only, the BOLD response shifted to include hMT+.

Similar BOLD modulations were observed irrespective of whether the MID stimulus isolated CD or IOVD cues (**Figure 5.4**, panels C and D). We contrasted the BOLD responses against contrast attention blocks, to avoid the effects of spatial attention – both MID and contrast blocks required that the participant spread their attention across the whole stimulus field.

There are several noteworthy results from this analysis. Firstly, BOLD modulations for the MID attentional blocks were larger than for the contrast attentional blocks in early visual areas. Secondly, attending to both CD and IOVD MID resulted in modulations that included V1, V2, V3, and V4, and extended into the IPS and hMT+. Finally, attending to CD MID resulted in a tighter band of activation in more central parts of V1, V2 and V3, compared to attending to IOVD MID. The activation in peripheral, non-stimulated parts of these ROIs, seen in blue in **Figure 5.4**, panel C, is probably caused by suppression and negative BOLD from the MID attention task (Gouws et al., 2014). If suppression here were a result of attending to contrast, we would expect to see it for both CD and IOVD stimuli. Attending to MID resulted in similar BOLD amplitudes in CD and IOVD runs in early visual areas (see **Figure 5.15** for an ROI breakdown), indicating that the suppressive response seen in in CD runs only is not driven by differences in overall response amplitude, but is more likely to reflect the strategy participants are using in the CD MID attention blocks. These findings imply that the CD MID task may be solved with a more foveal attention bias, compared to the IOVD MID task.

Finally, we also noted an asymmetry in the extent of BOLD modulation across the hemispheres. In general, activity was modulated more strongly in the right rather than the left hemisphere. This was true for all attentional conditions and all contrasts shown in **Figure 5.4**.

5.4.2 ROI analysis of attention probes: CD runs

An ROI analysis was carried out in each subject in visual areas V1, V2, V3, V3A/B, IPS-0, V4, LO-1, LO-2, hMT and hMST. For clarity, analysis in the restricted ROIs (V1, V2 and V3) was performed separately from analysis in the non-restricted ROIs (V3A/B, IPS-0, V4, LO-1, LO-2, hMT and hMST).

A GLM was used to model the response during each attentional probe, as a function of the attentional block the probe occurred in, during all CD scans. This resulted in twelve predictors within each ROI: three probe types (MID probe, contrast probe, fixation probes) by four blocks (attend motion, attend contrast, attend fixation, rest). Beta weights for each predictor were averaged across all voxels in one ROI, to give a mean set of beta weights per ROI, per participant. Mean beta weights in each ROI were averaged across participants to generate group-level plots.

Beta weights were entered into two separate analyses of variance (ANOVA), for restricted and non-restricted ROIs respectively. For restricted ROIs, the ANOVA was a 3 x 4 x 3 repeated measures design (3 ROIs [V1, V2 and V3] by 4 blocks [attend MID, attend contrast, attend fixation, rest] by 3 probe types [MID change, contrast change, fixation change]). For non-restricted ROIs, the ANOVA was a 7 x 4 x 3 repeated measures design (7 ROIs [V3A/B, IPS-0, V4, LO-1, LO-2, hMT and hMST] by 4 blocks [attend MID, attend contrast, attend fixation, rest] by 3 probe types [MID change, contrast change, fixation change]).
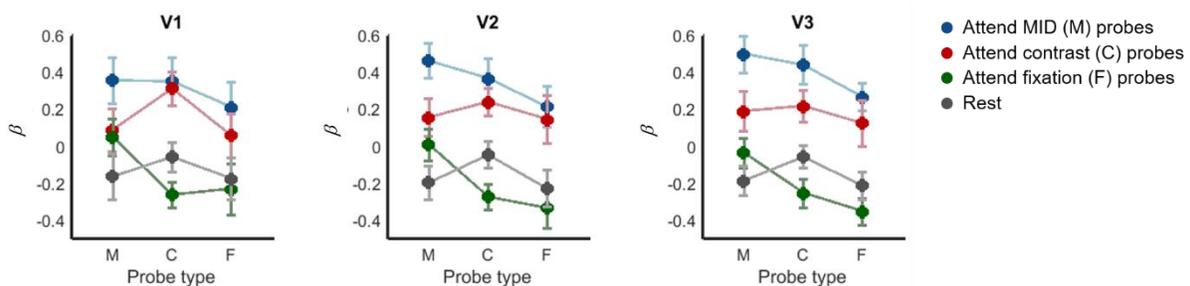
Figure 5.5 *Results from the ROI analysis of restricted ROIs, for all attention probes, from the CD MID stimulus runs. Beta weights are illustrated as a function of ROI, probe type (M = MID probe, C = contrast probe, F = fixation probe) and attention block (attend MID in blue, attend contrast in red, attend fixation in green, rest in grey). Error bars are ±1 SEM.*

Results from the restricted ROIs are shown in **Figure 5.5**. Overall, the BOLD response to specific attentional probes during 'attend MID' blocks was higher than the BOLD response to specific probes during all other blocks, including the 'attend contrast' blocks, and in accordance with this the ANOVA revealed a significant main effect of attention block ($F(3, 33)$ = 29.73, p < .001, partial ε^2 = 0.73). During 'attend MID' and 'attend contrast' blocks, the mean amplitude independent of probe types was well above zero, indicating an elevated BOLD response during these attentional conditions.

In contrast, during the 'attend fixation' blocks, responses to MID, contrast and fixation probes were generally below zero. This is because when participants are attending at fixation, the attentional probe is occurring in a part of the visual field that is *outside* the field of attention, and outside the extent of the restricted ROIs, which excluded the fovea. Thus, the response within the ROI may be actively suppressed (Gouws et al., 2014). In the 'rest' blocks, which is the equivalent of free viewing, responses were equally weak (or suppressed) and were generally similar to the 'attend fixation' blocks.

The main effects of ROI ($F(2, 22)$ = 0.27, p = .766, partial ε^2 = 0.02) and probe type ($F(2, 22)$ = 3.43, p = .051, partial ε^2 = 0. 24) were nonsignificant. However, there was a significant interaction between ROI and attention block ($F(6, 66)$ = 3.53, p < .004, partial ε^2 = 0.24), indicating that ROIs can be identified by the amplitude of the BOLD response during different attentional states, and that these amplitudes vary depending on the ROI in question.

All other interaction terms were nonsignificant; however some interesting trends can be observed for probe types and attentional state – most notably for the contrast and motion probes (see the red and blue data points in **Figure 5.5**). For both of these conditions, we see a larger BOLD amplitude when the attentional state of the participants corresponded to the presented probe. When attention wass directed towards another aspect of the stimulus, the response was weaker. This could hint at the compound effect of bottom-up stimulus features and attention, as illustrated by the Reynolds and Heeger normalization model of attentional gain (Reynolds & Heeger, 2009).
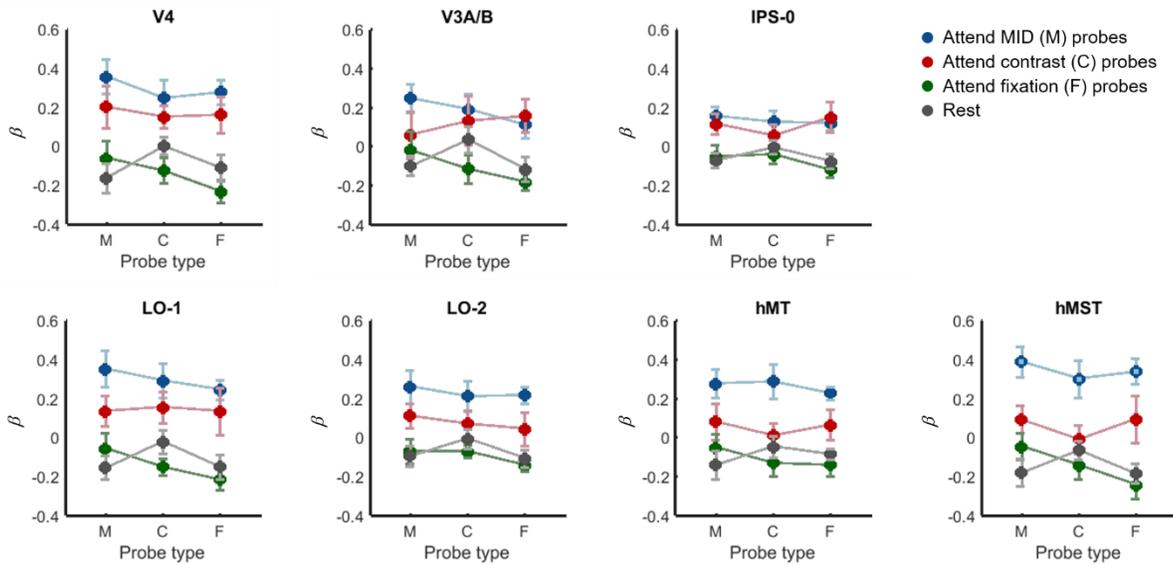
Figure 5.6 Results from the ROI analysis of non-restricted ROIs, for all attention probes, from the CD MID stimulus runs. Beta weights are illustrated as a function of ROI, probe type (M = MID probe, C = contrast probe, F = fixation probe) and attention block (attend MID in blue, attend contrast in red, attend fixation in green, rest in grey). Error bars are ±1 SEM.

Results from the non-restricted ROIs are shown in **Figure 5.6**. These ROIs were not tightly defined to the extent of the stimulus, and thus may also include voxels that fall outside the area of cortex that corresponds to the attentional window. In reflection of this, the BOLD response during the 'attend fixation' conditions (in green) is not as strongly suppressed as for the restricted ROIs. Nonetheless, the amplitudes for 'attend motion; and 'attend contrast' are still generally higher across ROIs, and the ANOVA again found a main effect of attentional block ($F(3, 33) = 7.35$, $p < .001$, partial $\varepsilon^2 = 0.74$).

Again, there was a significant interaction between ROI and attentional state ($F(18, 198) = 7.05$, $p < .001$, partial $\varepsilon^2 = 0.39$). In this case, this appears to be driven by the relatively large BOLD amplitude during the 'attend MID' blocks in the motion ROIs (in particular in hMST). It is also interesting to note that the *difference* between the attend MID and attend contrast blocks increases in motion-selective ROIs, where the attend MID amplitudes are elevated but the attend contrast amplitudes collapse. This is particularly clear when comparing the responses in V4 to the responses in hMT and hMST, reflecting the known functional properties of these ROIs.

Similar to in the restricted ROIs, all other main effects and interaction terms were nonsignificant. Overall, the pattern of results for the CD runs suggest that the largest effect comes from manipulating the attentional state of the participants, with some interesting interactions with ROIs and probe types that reflect the functional specializations of the underlying neurons.

5.4.3   ROI analysis of attention probes: IOVD

For IOVD runs, the pattern of responses to all probe types during different attentional conditions was measured in the same way as in the previous section. Restricted ROIs were analysed separately from non-restricted ROIs. The BOLD response was modelled using a GLM and the resultant beta weights were entered into a repeated measures ANOVA.
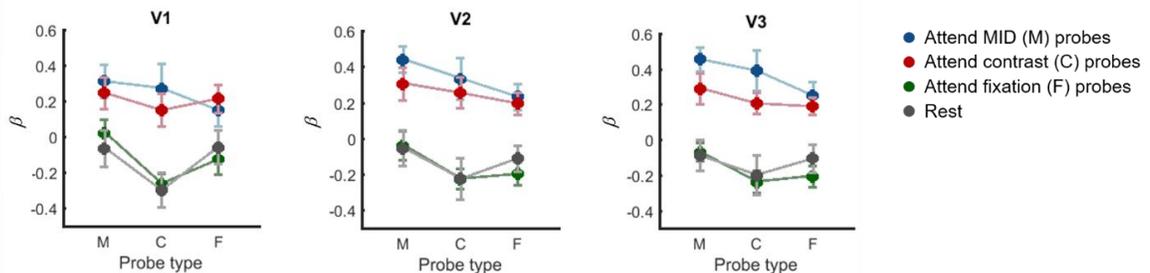
Figure 5.7 *Results from the ROI analysis of restricted ROIs, for all attention probes, from the IOVD MID stimulus runs. Beta weights are illustrated as a function of ROI, probe type (M = MID probe, C = contrast probe, F = fixation probe) and attention block (attend MID in blue, attend contrast in red, attend fixation in green, rest in grey). Error bars are ±1 SEM.*

Results for the restricted ROIs are shown in **Figure 5.7**. The pattern of results was similar to the pattern observed in the same ROIs during the CD runs, with a significant main effect of attentional block ($F(3, 33) = 6.26$, $p < .001$, partial $\varepsilon^2 = 0.77$). Here, 'attend motion' and 'attend contrast' conditions resulted in similarly elevated BOLD amplitudes across all probe types. 'Attend fixation' and 'rest' conditions were also similar to one another, and overall the BOLD amplitudes for these were low.

As in CD runs, the interaction between ROI and attention block was significant ($F(6, 66) = 3.67$, $p =.003$, partial $\varepsilon^2 = 0.25$) indicating that each ROI is identifiable by its pattern of

attentional modulation. Additionally, there was an interaction between ROI and probe type ($F(4, 44) = 2.98$, $p = .029$, partial $\varepsilon^2 = 0.21$), where the BOLD modulation for the motion and contrast probes increased overall throughout V1, V2 and V3. The modulation for the contrast probes remained low but stable, thus driving the interaction term. Again, this is possibly due to the restriction of the ROIs, where the fixation probe occurs outside the actual stimulus extent and therefore does not drive any voxels inside the ROIs. All other main effects and interactions were non-significant.
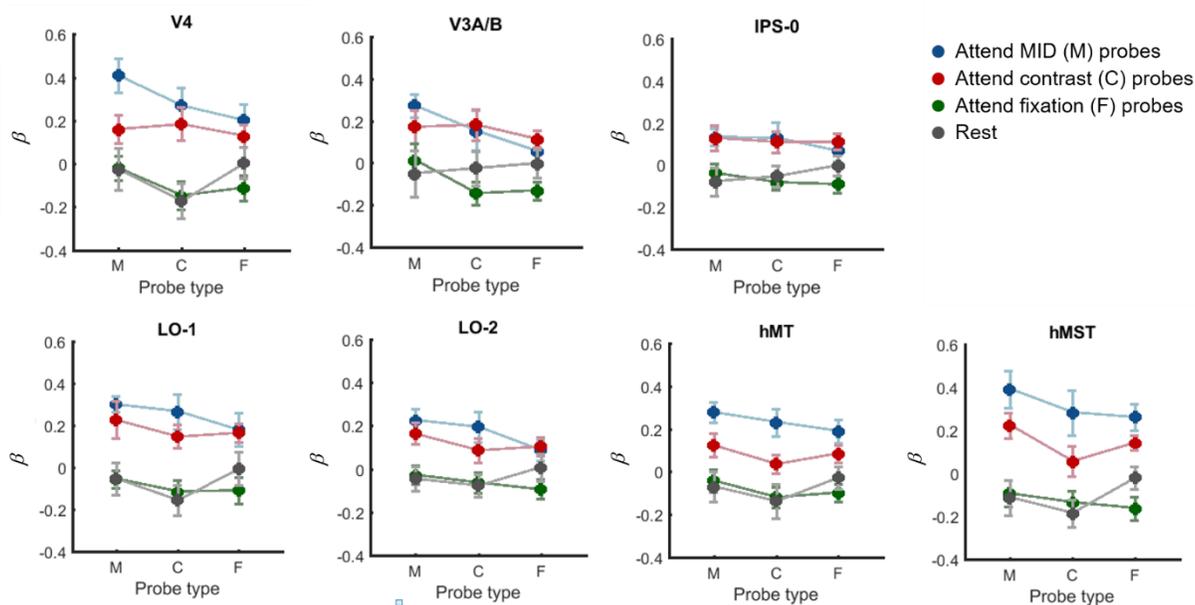
Figure 5.8 *Results from the ROI analysis of non-restricted ROIs, for all attention probes, from the IOVD MID stimulus runs. Beta weights are illustrated as a function of ROI, probe type (M = MID probe, C = contrast probe, F = fixation probe) and attention block (attend MID in blue, attend contrast in red, attend fixation in green, rest in grey). Error bars are ±1 SEM.*

Results for the non-restricted ROIs are shown in **Figure 5.8**. For these ROIs, there was a significant main effect of ROI ($F(6, 66) = 2.74$, $p = .020$, partial $\varepsilon^2 = 0.20$) and attention block ($F(3, 33) = 23.07$, $p < .001$ partial $\varepsilon^2 = 0.68$), whilst the interaction between ROI and attention block was also significant ($F(18, 198) = 5.83$, $p < .001$, partial $\varepsilon^2 = 0.35$). All other interaction terms and main effects were non-significant. This is a similar result to the CD runs, where the profile for each ROI was distinctive and the largest effects were driven by the attentional state of the participant, irrespective of the underlying probe type.

5.4.4 Comparison of CD and IOVD ROI results

To summarise changes in the extent to which attending to MID differed in relation to attending to contrast across ROIs, and to compare CD and IOVD runs more directly, we calculated the difference between the mean BOLD amplitude to all 3 probe types in 'attend MID' and 'attend contrast' blocks. Results for the CD and IOVD runs are shown in **Figure 5.9**.
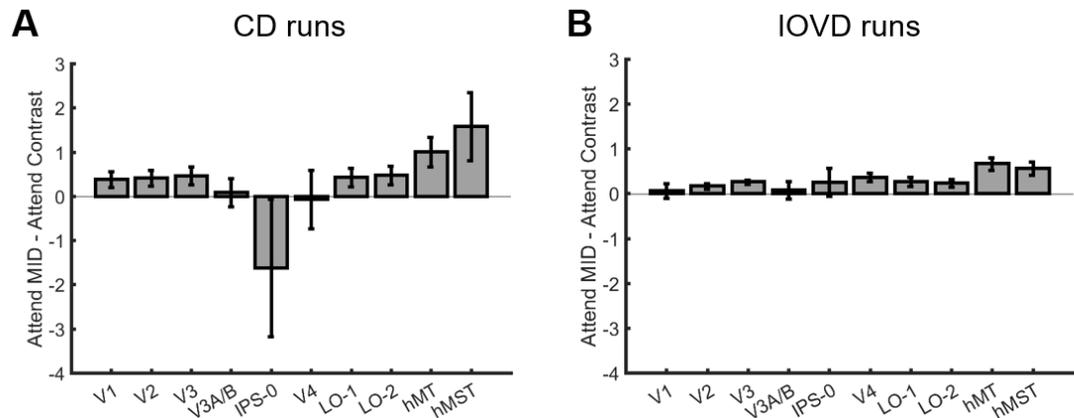
Figure 5.9 *Differences in the BOLD response during the attend MID and attend contrast blocks. Beta amplitudes were averaged in each ROI across all probe types to give a mean response during attend MID and attend contrast blocks, before subtracting the attend contrast value from the attend MID value. In the CD runs (panel A), there was a steady increase in this difference towards hMT and hMST, with the exception of IPS-0 where the difference was more varied across participants. For IOVD runs (panel B), the same trends were seen although overall the magnitude of the difference were weaker – note the difference in scale on the y-axis. Error bars are ±1 SEM.*

There was a similar trend in CD and IOVD scans, where the magnitude of the BOLD response to probes during 'attend MID' blocks was greater than the magnitude of the BOLD response to probes during 'attend contrast' blocks, and this magnitude increased systematically towards motion-selective areas hMT and hMST. This is generally true although the overall magnitude of the differences was much weaker in IOVD runs than CD runs. This relationship was less varied during CD scans. These trends dovetail with the observation that switching attention from contrast to MID changes in the stimulus shifts the amplitude of the BOLD response to extend into hMT+, as discussed in the whole brain analysis.

5.4.5　ROI analysis of hits and misses: CD runs

Finally, we investigated whether the BOLD amplitude across ROIs is different depending on whether an attentional probe was detected or missed. For this, the analysis was restricted to those probes that corresponded to the attentional state the participant was in, and whether these were classed as a 'hit' (participants responded within 1.5 seconds to a probe) or a 'miss' (participants did not respond within 1.5 seconds). We modelled the BOLD response using a GLM where predictors were MID probe hit, MID probe miss, contrast probe hit, contrast probe miss, fixation probe hit, and fixation probe miss. The resulting set of beta weights were averaged across voxels in each ROI and each participant. These beta weights were averaged across subjects to yield group-level results, and we used paired-samples t-tests to compare the amplitude between 'hits' and 'misses' within probe types.

Attentional mechanisms should pick out those neural populations that are maximally informative for solving a specific task. Therefore, in ROIs that are selective for, or directly involved in, solving a specific task, we would expect to see a larger BOLD response when attentional probes are correctly identified, versus when they are not identified.

To identify these ROIs, beta weights were entered into a repeated measures ANOVA where we were particularly interested in extracting pairwise comparisons for hits and misses within a given ROI, for a given probe type. Restricted ROIs were analysed separately from non-restricted ROIs as before. For the CD runs, data from V1, V2 and V3 were entered into a 3 x 3 x 2 repeated measures design ANOVA (3 ROIs [V1, V2 and V3] by 3 probe types [MID, contrast and fixation] by 2 response types [hit or miss]). Results are shown in **Figure 5.10**.

For these early ROIs, the ANOVA found a significant main effect of probe type ($F(2, 22)$ = 48.71, $p < .001$, partial $\varepsilon^2 = 0.82$), where the BOLD response to MID probes was generally higher than the response to contrast or the fixation probes. There was also a significant main effect of participant response ($F(1, 11) = 21.53$, $p = .001$, partial $\varepsilon^2 = 0.66$), where 'hit' probes that were correctly identified corresponded with a higher BOLD amplitude than unidentified probe types did.
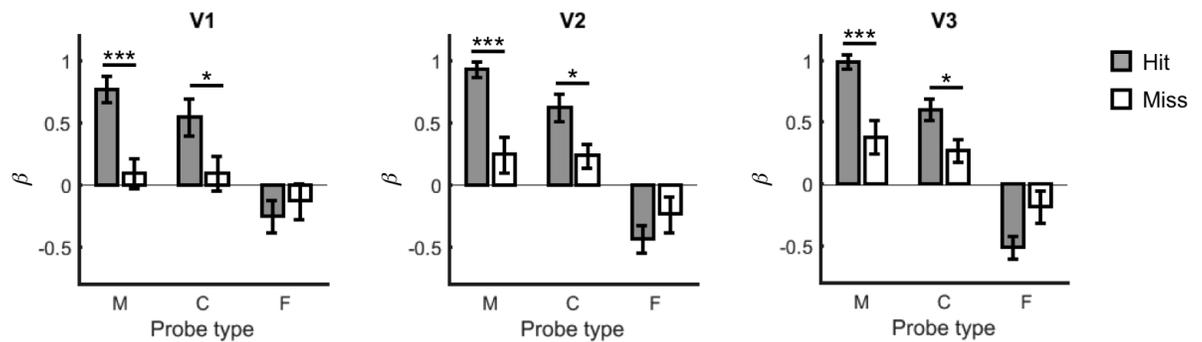
Figure 5.10 *ROI analysis of hits and misses in CD runs, restricted ROIs. Beta weights model the amplitude of the BOLD response when attended probe types (M = MID probe, C = contrast probe, F = fixation probe) are correctly detected ('Hit', in grey) or not detected ('Miss', in white). *** = p < .001, ** = p < .010, * = p < .050 . All p-values are Bonferroni corrected. Error bars are ±1 SEM.*

However, the most telling results stem from the pairwise comparisons carried out on the three-way interaction term. Overall this effect was non-significant ($F(4, 44) = 0.45$, $p = .771$, partial $\varepsilon^2 = 0.04$), but post-hoc tests revealed that the BOLD amplitude for correctly identified ('hit') vs. unidentified ('miss') MID probes was significantly higher in all three restricted ROIs (V1, V2 and V3). This was also true for the contrast probe types, implying that activity in these areas is modulated depending on whether or not participants were able to detect changes in both CD-defined MID and contrast. Note that the response to detected and undetected fixation probes was negative across all ROIs; this reflects the restriction of the ROI to exclude the fixation point and also explains why there were no significant differences here. The *p* values for these pairwise comparisons are given in **Table 5.1**. Results from the CD runs for the 'Hit' / 'Miss' pairwise comparisons for MID, contrast and fixation probes. Data from restricted ROIs. All *p*-values were extracted post-hoc and are Bonferroni corrected.

Table 5.1 *Results from the CD runs for the 'Hit' / 'Miss' pairwise comparisons for MID, contrast and fixation probes. Data from restricted ROIs. All p-values were extracted post-hoc and are Bonferroni corrected.*

	MID probe	**Contrast probe**	**Fixation probe**
V1	< .001	.040	.568
V2	.001	.029	.319
V3	.001	.038	.061

Responses from non-restricted ROIs were analysed in the same manner, and results are shown in **Figure 5.11**.
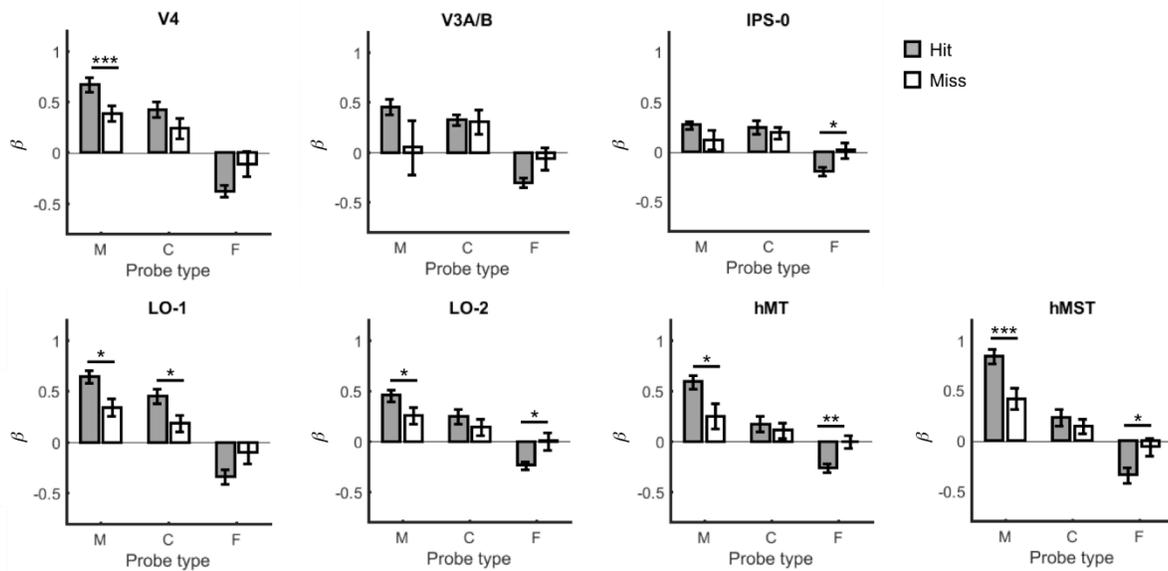
Figure 5.11 ROI analysis of hits and misses in CD runs, non-restricted ROIs. Beta weights model the amplitude of the BOLD response when attended probe types (M = MID probe, C = contrast probe, F = fixation probe) are correctly detected ('Hit', in grey) or not detected ('Miss', in white). *** = p < .001, ** = p < .010, * = p < .050. All p-values are Bonferroni corrected. Error bars are ±1 SEM.

For non-restricted ROIs, the ANOVA found main effects of ROI ($F(6, 66)$ = 4.22, p = .001, partial ε^2 = 0.28) and probe type ($F(2, 22)$ = 40.10, p =< .001, partial ε^2 = 0.79). In comparison to the restricted ROIs, this illustrates the greater variability in mean BOLD amplitude across ROIs. The main effect of participant response ('hit' or 'miss') was non-significant ($F(1, 11)$ = 1.94, p = .191, partial ε^2 = 0.15), also implying that the relationship between BOLD amplitude and correctly identified target was more varied.

Again, the main analysis of interest concerned the pairwise comparisons from the non-significant three-way interaction term ($F(12, 132)$ = 1.18, p = .303, partial ε^2 = 0.10). Here, a significant difference in BOLD amplitude for 'hit' and 'missed' MID probes was seen in ventral areas V4, LO-1 and LO-2, as well as in motion selective hMT and hMST. Crucially, these motion-selective ROIs did not show an effect for the contrast probes, where the BOLD amplitude was weak generally and did not distinguish between the identification of a contrast probe. This reflects the functional specialization of these ROIs, as well as their role in CD-

defined MID processing. Furthermore this suggests that these areas are crucial for the top-down *perception* of MID, in addition to their involvement in the bottom-up processing of stimulus driven activity.

Area LO-1 did track the identification of the contrast probe, though other areas of the ventral stream did not, implying that the majority of contrast perception occurs in early areas V1, V2 and V3. All Bonferroni-corrected *p*-values are given in **Table 5.2**.

Table 5.2 *Results from the CD runs for the 'Hit' / 'Miss' pairwise comparisons for MID, contrast and fixation probes. Data from non-restricted ROIs. All p-values were extracted post-hoc and are Bonferroni corrected.*

	MID probe	Contrast probe	Fixation probe
V4	.001	.204	.070
V3A/B	.244	.865	.065
IPS-0	.213	.466	.022
LO-1	.018	.028	.063
LO-2	.041	.332	.022
hMT	.032	.574	.004
hMST	< .001	.389	.044

Note also that some ROIs tracked the perception of the fixation probe, now that these ROIs were not restricted and included the foveal representation. This included IPS-0, LO-2, hMT and hMST. In other areas variability was high between subjects and no significant difference was measured. It is suggested that, because performance on the contrast detection task was near ceiling, this may simply reflect rare occasions when participants blinked or, through some other attentional lapse, lost sight of the fixation point. Here, the BOLD response was still negative, perhaps because of the suppressive effect of the surrounding stimulus.

5.4.6   ROI analysis of hits and misses: IOVD runs

Beta weights for the IOVD runs were analysed in the same manner as for the CD runs, with separate analyses carried out for the restricted and non-restricted ROIs. Results for V1, V2 and V3 are shown in **Figure 5.12**.
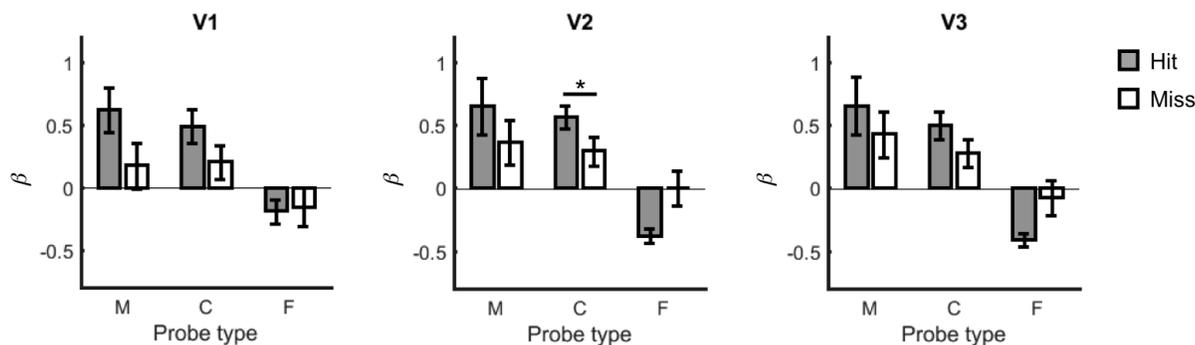
Figure 5.12 ROI analysis of hits and misses in IOVD runs, restricted ROIs. Beta weights model the amplitude of the BOLD response when attended probe types (M = MID probe, C = contrast probe, F = fixation probe) are correctly detected ('Hit', in grey) or not detected ('Miss', in white). *** = $p < .001$, ** = $p < .010$, * = $p < .050$. All p-values are Bonferroni corrected. Error bars are ±1 SEM.

For the IOVD runs, there was a significant main effect of probe type in early visual areas ($F(2, 22) = 16.15$, $p < .001$, partial $\varepsilon^2 = 0.60$), where, similar to in the CD runs, the BOLD response was highest for MID probes. The amplitude for the contrast probes was similar, whilst responses during the fixation probes were negative. However, the main effects for ROI ($F(2, 22) = 1.72$, $p = .203$, partial $\varepsilon^2 = 0.14$) and participant response ($F(1, 11) = 0.84$, $p = .379$, partial $\varepsilon^2 = 0.07$) were non-significant, indicating weaker 'hit' / 'miss' dependencies in these ROIs than during the CD runs.

Table 5.3 Results from the IOVD runs for the 'Hit' / 'Miss' pairwise comparisons for MID, contrast and fixation probes. Data from restricted ROIs. All p-values were extracted post-hoc and are Bonferroni corrected.

	MID probe	Contrast probe	Fixation probe
V1	.106	.055	.871
V2	.371	.041	.053
V3	.514	.117	.060

The pattern of results from pairwise comparisons in the three-way interaction term was less clear in the IOVD runs than for the CD runs, whilst also being non-significant ($F(4, 44) = 1.72$, $p = .163$, partial $\varepsilon^2 = 0.14$). In early ROIs, only V2 showed a significant difference between correctly identified and non-identified contrast probes. All other paired comparisons were non-significant. Exact *p*-values are reported in **Table 5.3**. However, this is likely due to a high variability across participants – trends can still be observed in the MID and contrast 'hit' and 'miss' amplitudes, hinting at a similar result to that of the CD runs.

Analyses for the non-restricted ROIs were carried out in the same manner and are illustrated in **Figure 5.13**.
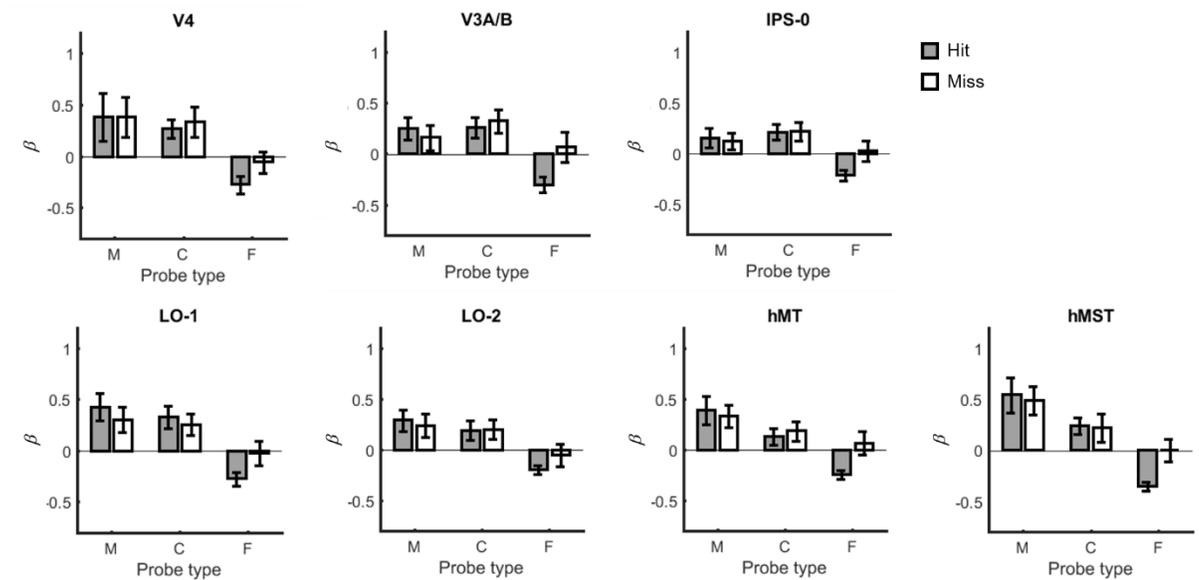
Figure 5.13 *ROI analysis of hits and misses in IOVD runs, non-restricted ROIs. Beta weights model the amplitude of the BOLD response when attended probe types (M = MID probe, C = contrast probe, F = fixation probe) are correctly detected ('Hit', in grey) or not detected ('Miss', in white). *** = p < .001, ** = p < .010, * = p < .050. All p-values are Bonferroni corrected. Error bars are ±1 SEM.*

Significant main effects from the ANOVA were, similar to in the CD runs, those of ROI ($F(6, 66) = 2.75$, $p = .019$, partial $\varepsilon^2 = 0.20$) and probe type ($F(2, 22) = 14.42$, $p < .001$, partial $\varepsilon^2 = 0.14$). The main effect of participant response was non-significant ($F(1, 11) = 0.55$, $p = .473$, partial $\varepsilon^2 = 0.05$), indicating that the amplitude was similar irrespective of whether participants were able to correctly identify target.

Again, pairwise comparisons from the non-significant three-way interaction term ($F(12, 132) = 0.76$, $p = .690$, partial $\varepsilon^2 = 0.07$) revealed a less clear picture than for the CD runs. The only significant differences between correctly identified and non-identified targets were in hMT and hMST for fixation probes. No other ROIs appeared to show any BOLD modulation as a function of whether participants were able to correctly detect a probe type. Exact *p*-values are given in **Table 5.4**.

Table 5.4 *Results from the IOVD runs for the 'Hit' / 'Miss' pairwise comparisons for MID, contrast and fixation probes. Data from non-restricted ROIs. All p-values were extracted post-hoc and are Bonferroni corrected.*

	MID probe	Contrast probe	Fixation probe
V4	.998	.683	.098
V3A/B	.636	.667	.091
IPS-0	.815	.965	.054
LO-1	.564	.599	.063
LO-2	.733	.915	.177
hMT	.769	.634	.041
hMST	.797	.872	.013

Could the differences seen in hMT and hMST between CD and IOVD cues be related to variability in subject performance? Between-subject variability in task performance (see **Figure 5.3**) and variability in beta amplitudes were greater in IOVD than in CD scans. Therefore, we investigated the relationship between subject performance and the magnitude of the difference in hit/miss amplitudes (β MID hit – β MID miss) for IOVD MID targets in hMT and hMST using a linear regression model.

Scatterplots are shown in **Figure 5.14**, and a linear regression model was fit to the data points to generate R^2 values and significance levels. In hMT and hMST, there was a moderate relationship between performance and the difference in amplitude between hit and missed events, where the magnitude of this difference was positively correlated with participant performance (hMT: $R^2 = 0.39$, $F(1, 10) = 6.27$, $p = .031$; hMST: $R^2 = 0.44$, $F(1, 10) = 7.87$, $p = .019$). Thus, the extent to which activity in hMT and hMST tracks the perception of IOVD targets depends on the performance of the participant and how able they are to interpret the IOVD-defined MID.

A hMT
$R^2 = 0.39$
$p = .031$

B hMST
$R^2 = 0.44$
$p = .019$
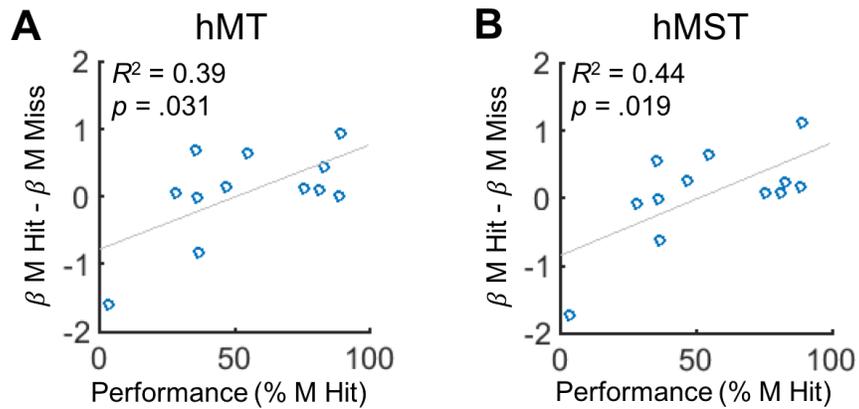
Figure 5.14 *The relationship between performance and the amplitude of the difference between 'Hit' and 'Miss' IOVD MID probes, in hMT (panel A) and hMST (panel B). Each data point represents the performance on the IOVD task (percentage of correctly identified MID targets) and difference in response amplitude (beta amplitude 'hit' MID target – beta amplitude 'missed' MID target) for a single participant, and data are fit with a regression line. There is a significant, positive relationship between performance and activation patterns in hMT and hMST.*

In summary, the ROI 'hits' and 'misses' analysis reveals that early visual areas are involved in the perception of MID and contrast probes. In motion-selective hMT and hMST, there was also a stronger BOLD response to detected MID probes – but this was only true for the CD probes. For IOVD probes, there was a significant relationship between the magnitude of the BOLD difference between MID hits and misses, and subject performance, in hMT and hMST.

5.5   Discussion

Previous research has used bottom-up, stimulus-driven approaches to identify a network of areas containing neural populations that can be driven by binocular MID cues (Joo et al., 2016; Likova & Tyler, 2007; Rokers et al., 2009; see also Chapter 3). This study investigated whether activity in this 3D motion network can be modulated by top-down, attentional mechanisms. This approach revealed populations of neurons that are critically involved in the perception of MID, and whose activity is correlated with the detection of MID probes. Participants viewed a continually oscillating CD or IOVD stimulus during fMRI scanning, whilst attending to subtle changes in MID, contrast, or at fixation. Attentional state varied systematically over time, revealing shifts in the BOLD response as participants shifted their attention between these three aspects of the stimulus. Participants were also asked to

identify probes relating to changes in MID, contrast, or fixation, highlighting those cortical areas whose activity is predictive of task performance.

Because participants varied the object of their attention between changes in MID, a local (but unpredictable) contrast change, and a change at fixation, resultant BOLD changes could be driven by a combination of featural and spatial attention. The switch between MID and contrast engaged mainly featural attention, because participants were required to detect a change that occurred across the whole stimulus field (MID) or at an unpredictable location (contrast change). It is perhaps counter-intuitive that a local contrast change can only be solved by spreading attention across the whole stimulus; however, theoretically a whole-field change in contrast could be solved by narrowing the attentional window to only a small portion of the stimulus. Although the contrast change occurred locally, and always at a fixed radius, we varied the angular location of this change so that participants were required to attend globally. We think it is unlikely that participants used a narrowed attentional window to solve the contrast task, however, because of the fixed radius we cannot wholly exclude this possibility. Finally, because the fixation task occurred in the centre of the stimulus (outside the annulus of dots), this task required a switch in both featural and spatial attention and thus can be used as a baseline. The most informative comparisons for featural attention are those between contrast and MID tasks.

The whole-brain analysis revealed large-scale shifts in the BOLD response as participants switched their attention between MID and contrast changes in the stimulus. When participants were attending to contrast, large BOLD modulations were observed in early visual areas including V1, V2 and V3 (Ress, Backus, & Heeger, 2000; Ress & Heeger, 2003). When participants switched their attention to the MID aspect of the stimulus, for both CD and the IOVD stimuli, these modulations extended into motion-sensitive areas such as V3A/B and hMT+, and included those areas that have been identified previously in a bottom-up manner (Joo et al., 2016; Likova & Tyler, 2007; Rokers et al., 2009; see also Chapter 3). BOLD modulations also extended into the inter-parietal sulcus (IPS), which has been implicated in a variety of cognitive functions including the allocation of visual attention (Maurizio Corbetta et al., 1998; Maurizio Corbetta, Kincade, Ollinger, McAvoy, & Shulman, 2000; Wojciulik & Kanwisher, 1999), and has been associated with stereo-defined MID in particular (Cottereau, McKee, & Norcia, 2014; see also Chapter 3).

There were some small differences in the BOLD modulations for attending to CD- or IOVD-defined MID. Most notably, we measured suppressive responses in peripheral aspects of early visual areas (V1-V3) that were outside the retinotopic area that would be stimulus-driven. Negative BOLD responses have been observed in spatial attention tasks, where the unattended portion of the visual field is suppressed (Gouws et al., 2014). In our study, such responses were only measured during CD runs. They are unlikely to originate from attending to contrast, because this condition was the identical in CD and IOVD runs. Suppressive modulations are more likely related to the MID attention task, and may reflect differences in attentional strategy that participants are using when they attend to CD vs. IOVD. A possible difference is that the CD task could be solved with a more foveal attention bias. In theory, CD MID can be perceived from a small portion of the visual field, given that the cue is based on retinal disparities. Because IOVD depends on a global estimate of velocity differences between the two eyes, participants may be spreading their attention across a larger part of the visual field, integrating across many individual spatial locations in order to generate a more reliable MID percept. To test this hypothesis, future work could investigate the area of integration for CD and IOVD mechanisms, where the prediction would be that neurons tuned to CD MID have smaller receptive fields than neurons tuned to IOVD MID.

BOLD modulations observed in the whole-brain analysis were not driven by changes in the stimulus itself – the moving dot stimulus was continually presented over the course of the scan, and all MID and contrast changes occurred randomly and in an interleaved manner, irrespective of the probe type the participant was instructed to attend to. Thus, observed BOLD changes were induced by fluctuations in the attentional state of the participant, suggesting that those areas that can be driven by MID cues are also modulated by top-town attentional demands relating to the perception of MID.

We observed an interesting hemispheric asymmetry on the whole brain maps, where the BOLD response was more strongly modulated in the right than the left hemisphere when participants attended to both CD and IOVD MID. This is not surprising; the control of visual spatial attention is known to originate in a right-lateralized network of fronto-parietal areas (Corbetta, Miezin, Shulman, & Petersen, 1993; Maurizio Corbetta & Shulman, 2002; Georgieva, Peeters, Kolster, Todd, & Orban, 2009) which would more strongly modulate visual areas in the right hemisphere and the right IPS through top-down mechanisms (Nobre et al., 1997).

In addition, there is evidence that disparity processing occurs in a right-dominated network of areas. Patients with lesions in the right hemisphere present with deficiencies in disparity processing (Cowey & Porter, 1979; Szczepanski, Konen, & Kastner, 2010), and patient and monkey studies show that total removal of the right hemispheres abolishes all stereo processing (Carmon & Bechtoldt, 1969; Ptito, Zatorre, Larson, & Tosoni, 1991). Finally, fMRI studies in healthy populations similarly show right hemispheric dominance in extracting shape and structure from stereo-defined motion (Sunaert, Van Hecke, Marchal, & Orban, 1999; Vaina, 1989) and disparity stimuli (Baecke et al., 2009; Ip, Minini, Dow, Parker, & Bridge, 2014; Kwee, Fujii, Matsuzawa, & Nakada, 1999; Orban, Sunaert, Todd, Van Hecke, & Marchal, 1999). Therefore, the right hemispheric dominance seen in our study could result from a combination of the allocation of visual spatial attention, and the extraction of motion from disparity.

The nature of these whole-brain BOLD changes were investigated in more detail in an ROI analysis. First, responses to all probe types (MID, contrast and fixation probes) across all attentional states (attend motion, attend contrast, attend fixation, and a baseline 'rest' condition) were investigated. Results from CD and IOVD runs were similar, where the largest responses during the presentation of probes occurred when participants were attending to MID – irrespective of which probe type was present. Attending to contrast also increased the BOLD response across all three probe types, relative to attending to fixation or during rest periods – though to a smaller degree than when participants were attending to MID.

The magnitude of the *difference* between attending to MID and attending to contrast varied across ROIs. In early visual areas (V1, V2, V3, V3A/B and V4), the difference in response magnitude was less pronounced than in later, motion selective ROIs (hMT and Hmst, see **Figure 5.9**). Here, there was a large difference between attention to MID and attention to contrast. Indeed, there was a significant interaction between attentional block and ROI, where ROIs could be distinguished based on their net response across probe types during different attentional conditions. This is similar to results from the whole-brain analysis, and reflects the functional specialization of hMT and hMST.

There was no interaction between attentional state and probe type in this analysis, suggesting that the magnitude of the BOLD response when participants were attending towards different aspects of the stimulus was independent of the underlying changes in the

stimulus. Such an interaction would be reflective of a multiplicative 'feature similarity response gain' mechanism that acts to increase the response of neurons with tuning profiles that are congruent with the attentional task, independent of any low-level changes in visual input. This was surprising, as it might be expected that attention towards a certain feature should also results in larger BOLD amplitudes when a probe pertaining to that feature occurred. Recent models of attention propose that, whilst attentional state results in a general increase in response, this is compounded by underlying changes in stimulus (the 'stimulus drive'; Reynolds & Heeger, 2009).

It may be that the univariate analysis carried out in this study was not sensitive enough to distinguish such patterns. There were some indications that attentional state does compound with underlying changes in stimulus response, but these did not reach significance. An example of this is the pattern of responses during CD runs for the contrast probe during the 'attend contrast' condition (see **Figure 5.5**). In V1, the BOLD response is enhanced during the occurrence of a contrast probe, specifically during the 'attend contrast' blocks and relative to the MID and fixation probes during the same attentional state. This enhancement dies off progressively through V2 and V3, and the pattern becomes inverted in later ROIs. For MID probes, a similar trend is seen in V4 and hMST, and perhaps V2, but nowhere else. However, these relationships are weak – especially so during the IOVD scans. A multivariate approach, such as multivoxel pattern analysis (MVPA; Haynes & Rees, 2005; Kamitani & Tong, 2005, 2006; Ress et al., 2000; for a review of the method see Haxby, 2012), would be more suited to tease out these relationships.

Finally, we investigated the relationship between BOLD activity across ROIs and the participant's performance on the MID task. Some areas showed a stronger BOLD response when participants correctly identified an attended MID target, compared to when that target was present in the stimulus, attended to, but not perceived. For the CD stimulus, this included early visual areas (V1, V2, V3 and V4), hMT and hMST. These ROIs contain populations of neurons that are maximally informative for solving the MID task, and that are crucial for the perception of CD defined MID. Weak activity in these areas was predictive of poor task performance, whereas larger BOLD modulations were associated with a perceived change in the MID of the CD stimulus. Although other ROIs were strongly driven by attending to MID, responses here did not reflect the participant's performance.

The hit/miss dependency in V4 is an interesting finding because it was only measured for CD stimuli, and not for IOVD stimuli. As V4 is part of the ventral visual pathway, and contains neurons tuned to shape representations (Desimone & Schein, 1987; Gallant, Braun, & Van Essen, 1993; Kobatake & Tanaka, 1994; Pasupathy & Connor, 1999, 2001, 2002), we suggest that activity here may be related to extracting form from motion, or motion from form. The retinal disparity cues in the CD stimulus convey strong sensations of depth and shape. Because these cues are absent in the IOVD stimulus, IOVD is less likely to engage form-from-motion mechanisms. This may reflect differences in the functional utility of CD and IOVD cues, a suggestion that has also been made based on differences in spatial and temporal properties of both cues (Czuba et al., 2010).

Areas hMT and hMST have been proposed as the critical site for IOVD computation. Single-cell recordings in macaques have identified neurons here that are tuned to the direction of IOVD-defined MID (Czuba, Huk, Cormack, & Kohn, 2014; Sanada & DeAngelis, 2014), and fMRI research has consistently revealed stimulus-driven BOLD activity in these ROIs (Joo et al., 2016; Rokers et al., 2009). Surprisingly therefore, the univariate analysis here measured no strong hit/miss dependency in either hMT or hMST for IOVD MID probe detection. This implies that, although these areas can be driven by an IOVD stimulus, and although activity here can also be modulated by attentional mechanism, the relationship between neural activity and *perception* of the IOVD cue is weaker.

However, these results may be due to individual differences in IOVD performance. On average, participants performed significantly worse on the IOVD probe detection task than on the CD probe detection task, and variability between subjects was higher. An investigation into the individual differences in MID perception across the general population has revealed that most participants prefer the CD cue over the IOVD cue (Nefs et al., 2010); however, most fMRI studies to date have tested only small samples of highly-trained individuals (Joo et al., 2016; Rokers et al., 2009). Indeed, the IOVD stimulus is greatly impoverished and ambiguous to interpret (Cormack et al., 2017). A natural stimulus moving in depth would not exert identical amplitude or direction of velocity vectors across different parts of the retina, as the IOVD stimulus does, nor would it be divorced from stereo cues that provide strong depth sensations to disambiguate the direction of MID. Because our probe detection task relies on interpretation of the MID in the stimulus, this may explain why performance was poorer on the IOVD task.

These individual differences on task performance allowed us to correlate IOVD probe detection against the magnitude of the difference between hit/miss BOLD responses in hMT and hMST. When this more detailed view was taken into account, we indeed measured a strong relationship between task performance and BOLD modulation in these ROIs. This suggests that the extent to which activity in hMT and hMST tracks perception is dependent on how successfully participants are able to interpret the MID stimulus. Therefore, our results are not inconsistent with previous research suggesting a critical role of hMT and hMST in the perception of IOVD. In those participants who perform well on the IOVD probe detection task, activity in these ROIs indeed reflects whether probes were correctly identified or not.

5.6   Conclusions

In this study, we have shown that attention alone modulates neuronal responses to moving dot stimuli, and that patterns of responses are different depending on the attentional state of the participant (when detecting motion or contrast). Attending to contrast modulates responses in early visual areas. Attending to MID additionally drives lateral and dorsal activity, extending into the IPS and hMT+. These response patterns were similar for CD and IOVD-defined MID, although some differences were observed – attending to CD resulted in more suppression in peripheral aspects of V1, V2 and V3, suggesting a greater foveal attention bias than for IOVD.

Across ROIs, we found that modulations were largest when participants were attending to MID – but these modulations did not interact directly with bottom-up stimulus cues. Such interactions may be revealed with more sensitive analyses such as MVPA – in our study, interactions between attentional gain and stimulus drive are likely to be lost in the univariate average.

Finally, an analysis of hit/miss responses revealed which areas contain populations of neurons whose response profile tracks the perception of MID. Early areas showed a strong hit/miss dependency for MID probe detection. CD probe detection was also correlated with BOLD modulation in hMT and hMST. For IOVD, these patterns were revealed only when subject performance was taken into account. For those participants who were able to reliably

interpret the IOVD cue, we measured a relationship between hit/miss amplitudes and detection of IOVD probes.

Together, these data show that attentional mechanisms modulate activity through a hierarchy of areas involved in the perception of MID, and emphasises the roles of hMT and hMST in the perception of CD and IOVD.

5.7   Supplementary information
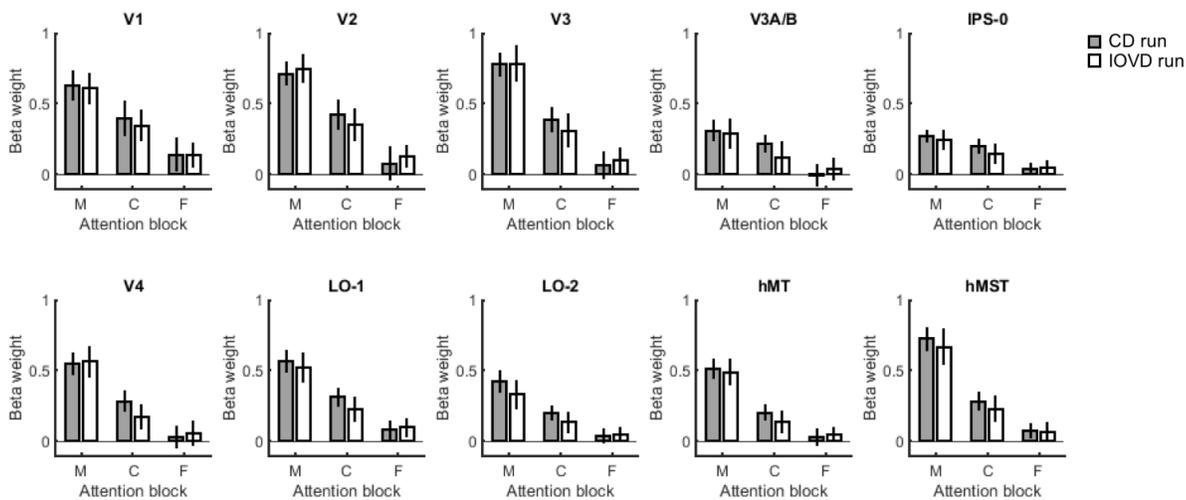
Figure 5.15 *ROI analysis of changes in BOLD amplitude as a function of attentional state. In this analysis we modelled the BOLD response in each ROI when participants were attending to MID (M), Contrast (C), and Fixation (F), as well as during rest blocks. The response during 'rest' blocks has been subtracted from M, C and F amplitudes as proxy for a baseline condition. Responses across CD (in grey) and IOVD (in white) runs were calculated separately. Overall there is no difference in amplitudes within each attention block, between CD and IOVD runs. The BOLD response during 'M' blocks was generally higher than the BOLD response during 'C' blocks, and the lowest responses were measured for the 'F' blocks.*

Chapter 6.   General discussion

6.1   Summary of key findings

This thesis examined the neural pathways that underpin our ability to perceive MID, focusing on two binocular cues – CD and IOVD. Previous research, reviewed in Chapter 1, has revealed cortical sites that can be driven in a feed-forward manner by both cue types. The work presented here adds to this body of research by probing these neural mechanisms in more detail, asking how information necessary for the computation of CD and IOVD is represented in a hierarchy of visual areas. Key questions included the neural locus of CD processing, how information in early precortical pathways contribute to MID processing, and how eye-specific information necessary for computing IOVD is maintained in extrastriate visual areas. Finally, attentional feedback mechanisms were isolated to investigate how activity in areas associated with CD and IOVD processing is affected by task demands.

In the first experimental chapter, it was demonstrated that signals carried in all three precortical pathways can contribute to MID mechanisms, in a manner analogous to 2D motion mechanisms. Contrast-scaled achromatic, isoluminant L-M, and S-cone isolating stimuli were able to drive MID percepts in stimuli that isolated the CD and IOVD mechanisms. Adding noise to the stimulus degraded the ability of participants to discriminate the direction of the MID, and the IOVD signal was more susceptible to noise than the CD stimulus. However, masking the chromatic signal with achromatic noise elements did not affect performance. These findings provided the first demonstration that both CD and IOVD mechanisms draw on signals carried in chromatic pathways.

The second experimental chapter described fMRI responses to CD and IOVD stimuli, where responses were driven in a feed-forward manner. The effects of disparity and 2D motion, as well as other low-level properties, were removed by subtracting responses to control stimuli. First, this study replicated previous findings that implicated hMT+ in CD and IOVD computation. Responses to both stimulus types were measured throughout a hierarchy of visual areas, with no overall differences between the two cue types. Secondly, this study found no evidence that the putative CSM area is particularly critical for computing CD.

Although responses to CD stimuli were correlated with those measured in hMST, they were generally weaker and as such contributed only to a lesser extent to the CD mechanism.

This study also discriminated between achromatic and S-cone inputs to CD and IOVD mechanisms. Although no overall differences in chromaticity were measured, there was a significant interaction between MID cue type and chromaticity. S-cone IOVD cues elicited particularly strong responses, even in primary visual cortex, where responses were larger even than to the achromatic IOVD stimulus. This suggests an early, binocular motion-opponent signal that is most effectively carried in the S-cone pathway.

In contrast, the S-cone contribution to the CD signal was weak. This is somewhat surprising given that S-cone and achromatic stimuli were found to contribute equally to CD MID perception in the first experimental chapter. Taken together, these two findings would imply that whilst the S-cone pathway can contribute to disparity mechanisms, and whilst S-cone CD MID can be perceived, these signals do not translate to large BOLD modulations in motion-sensitive areas.

However, it is more likely that the experimental paradigm used in the psychophysical study was not sensitive enough to detect such differences in chromatic pathway input. Indeed, the measured thresholds were generally high and the slopes of the psychometric functions were shallow, indicating a high degree of uncertainty. It is suggested that the dissociation between S-cone inputs to psychophysical judgements of CD or IOVD-defined MID might be more carefully teased out by simplifying the experimental task (for example, to a two-alternative forced-choice task where participants choose the interval that contains MID). This may be more suited than the challenging task of identifying the initial direction of MID in a stimulus that moves in both directions within a single trial. I would predict that, under this type of paradigm, the results would show a strong koniocellular contribution to the IOVD mechanism, and a weak contribution to the CD mechanism, and thus would dovetail with the fMRI findings presented in this thesis.

The third experimental chapter investigated whether eye-specific signals remain available in cortical areas that are involved in the computation of IOVD, both inside and outside primary visual cortex. To increase the sensitivity of the analysis to detect information encoded on a fine scale, multivariate pattern analysis and classification techniques were

applied to fMRI data. Nonetheless, the analysis was limited by the difficulty in estimating reliable event amplitudes in a dense event-related design. EOO information was decoded in V1, and motion direction was detected in hMT, but no regions where both sources converged were identified. This suggests that IOVD mechanisms inherits EOO information from ocular dominance columns in V1. Increasing the stability of generating beta estimates may reveal the presence of this information in extrastriate areas.

Finally, the fourth experimental chapter investigated how top-down, attentional demands modulate activity in those areas that can be driven by bottom-up CD and IOVD signals. Detecting changes in CD and IOVD-defined MID resulted in net changes in BOLD amplitude, with strong attentional effects measured as early as V1. Task demands therefore influenced neural activity even at very early stages. These changes were not driven by bottom-up changes in the stimuli, and varied instead depending on which aspect of the stimulus the participant was attending to. Furthermore, responses in hMT and hMST were correlated with the correct detection of CD and IOVD targets, demonstrating that neural populations in these areas are critically involved in detecting CD and IOVD. This study provided the first demonstration that neural populations in these motion-sensitive areas are crucially involved in detecting changes in MID, and are recruited by attentional mechanisms.

The work presented here supports the prevailing view that both CD and IOVD share common cortical loci, although they draw on different sources of visual information, and that signals for both cue types are multiplexed in areas known to be involved in the computation of binocular disparity and 2D motion. The overarching task of computing MID draws on many different sources of information. These include estimates of direction, speed and depth. Because motion signals and depth signals are inherently linked in the natural world, the combination of multiple sources of information computed though parallel pathways is likely to improve the quality of MID estimates extracted across a dynamic range of features.

6.2   Future directions

A question arising from the work presented here concerns the receptive field or integrative field sizes of the CD and IOVD mechanisms, which have not been explicitly compared. Differences in susceptibility to noise suggest that the CD mechanism may have a

smaller integrative field size than the IOVD mechanism, and the CD-specific foveal attention bias seen in the last data chapter may also reflect this. In addition, because the IOVD cue depends on a coarse motion vector generated at the retina, it may integrate across a larger portion of the visual field to improve the stability of the MID estimate. The initial encoding of the CD cue is necessarily tied to the limits of stereopsis. For IOVD, integrative field size could be measured by manipulating the starting position between binocular pairs of dots moving in the opposite direction. With increasing distance, the binocular percept of MID should degrade until the IOVD mechanism is no longer able to resolve the MID. This limit can be taken as the maximum integrative field size of the mechanism. Note that this may be different to the integrative field size of a general MID mechanism, which receives inputs from disparity, velocity, and optic flow.

The role of the KC pathway in the IOVD mechanism could also be explored further. The large S-cone driven response to IOVD MID measured in V1 suggests that this computation involves primary visual cortex, but activity in the LGN was not tested and inputs from the direct koniocellular projection from LGN to MT cannot be ruled out. Some interesting characteristics of the KC cells in the LGN – such as binocularity (Zeater, Cheong, Solomon, Dreher, & Martin, 2015) – suggest a potentially unique contribution from the KC pathway to the IOVD mechanism.

This S-cone motion-opponent signal measured in V1 is surprising because there is little evidence for binocular motion opponency this early on in visual cortex – perhaps because inputs from S-cone signals have not yet been considered explicitly. If this is true, this leads to a further hypothesis – that some motion processing occurs between the retina and V1. It has been suggested that the direction selectivity measured in blue-on cells in the KC layers of the LGN arise as a function of latencies between the 'on' and 'off' subfields of small bistratified ganglion cells in the retina (Tailby et al., 2010). This is rather like a Reichardt detector, but contained within a single receptive field. These mechanisms could provide a very early basis for extracting the binocular motion-opponent signals in V1 that support IOVD. To test this explicitly in humans, one might probe the inter-ocular transfer of motion aftereffects for S-cone stimuli. In achromatic stimuli, motion after-effects readily transfer from one eye to the other, providing evidence that motion is extracted beyond the point of binocular integration. This may not be true for S-cone isolating stimuli, which could suggest that an S-cone motion signal is extracted before information from the left and right eyes merge into a cyclopean stream.

A more general question to address concerns the manner in which the CD and IOVD cues are combined. Some psychophysical research has been carried out to this end, by testing the sensitivity profiles of CD and IOVD across the spatiotemporal parameter space (Czuba, Rokers, Huk, & Cormack, 2010). This suggests that a general MID mechanism weights the CD and IOVD cues differently depending on stimulus speed and eccentricity. Currently, our lab is using a combination of EEG and classification analysis to identify the point at which CD and IOVD merge into a general MID signal. The logic of this is that the classifier should be able to distinguish the CD and IOVD cues up to the point at which they are combined. This point could be localised across different electrodes (for example, integration at a later stage in visual processing), or at a distinct point in time.

In the real world, the two sources of MID information exist simultaneously. Experimentally, the both cue types can only be isolated by generating a highly reduced, artificial stimulus. For this IOVD cue this is particularly problematic, because the pattern of motion vectors on the retina once the optic flow cue is removed is unnatural, and results in cue conflict. This may explain why participants tend to find the IOVD cue more difficult to interpret – some participants systematically report perceiving MID in the opposite direction than the direction that is cued (Fulvio, Rosen, & Rokers, 2015). The 'binoptic flow field' stimulus proposed by Cormack et al. (2017) is attractive, because it combines CD, IOVD and optic flow information and thus will be instructive in understanding MID perception under more naturalistic viewing conditions.

6.3   Final conclusions

This thesis aimed to investigate the neural pathways and mechanisms involved in the perception of MID from CD and IOVD cues. Its principal findings, relating to the questions posed in the Introduction, are:

1. CD and IOVD share a common network of areas that can be driven in a feed-forward manner by stimuli that isolate these cues.
2. There is no evidence for a dedicated CSM area that is critical in the computation of CD MID. Instead, activity in hMT and hMST is strongly driven by CD stimuli.

3. Similar to 2D motion mechanisms, both CD and IOVD mechanisms can draw on information conveyed by the magnocellular, parvocellular and koniocellular pathways to support a MID percept.

4. The koniocellular pathway may be particularly suited to carrying an early, binocular motion-opponent signal than contributes to IOVD processing, and S-cone IOVD stimuli elicit robust fMRI responses as early as V1.

5. IOVD mechanisms likely draw on eye-specific signals encoded in ocular dominance columns in V1. No areas were identified that contained both motion direction and EOO information.

6. Activity in a network of areas that can be driven in a feed-forward manner, is also strongly modulated by attentional demands relating to the detection of CD and IOVD MID.

7. Areas hMT hMST contains populations of neurons that are maximally informative for solving MID tasks, where activity is correlated with the detection of both CD and IOVD stimuli.

Together, these findings detail a thorough investigation of the neural mechanisms that underpin our ability to perceive CD and IOVD MID. The findings add to our current understanding of CD and IOVD by considering both feedforward and feedback mechanisms, and by investigating how information carried in precortical pathways is sampled. Future research investigating the emergence of a generalised binocular MID signal is justified, given the co-occurrence of CD and IOVD in the natural world. This general MID signal is likely to weight the inputs from CD and IOVD mechanisms depending on the available information, contextual features such as speed and eccentricity, and task demands.

Abbreviations

ANOVA	Analysis of variance
BFF	Binoptic flow feld
BET	Brain extraction tool
BOLD	Blood oxygen level-dependent
CD	Changing disparity
CSM	Cyclopean stereo-motion area
CRT	Cathode ray tube
DLP	Digital light processing
DRDS	Dynamic random dot stereogram
EOO	Eye of origin
EPI	Echo-planar imaging
FAST	FMRIB's automatic segmentation tool
FFA	Fusiform face area
fMRI	Functional magnetic resonance imaging
FOV	Field of view
GLM	General linear model
GE	Gradient-echo
hMT	Human medial temporal visual area
hMT+	Human medial temporal and human medial superior temporal visual area, forming the human medial temporal complex (also referred to as V5)
hMST	Human middle superior temporal visual area
HRF	Hemodynamic response function
IOVD	Inter-ocular velocity difference

IP	Inferior pulvinar
IPS-0	Intraparietal sulcus area zero (also referred to as V7)
ISI	Inter-stimulus interval
KC	Koniocellular
LCD	Liquid crystal display
L *or* L-cone	Cone cells sensitive to long wavelengths (~564nm)
LGN	Lateral geniculate nucleus
	L+M Sum of long and medium wavelength cone responses, characteristic of cells in the magnoceulluar pathway
L-M	Difference between long and medium wavelength cone responses, characteristic of cells in the parvocellular pathway
LMS	Long, medium and short wavelength colour space
LO-1	Lateral occipital cortex visual area 1
LO-2	Lateral occipital cortex visual area 2
LoG	Laplacian-of-Gaussian
M *or* M-cone	Cone cells sensitive to medium wavelengths (~534nm)
MC	Magnocellular
MID	Motion in depth, 3D motion
MRI	Magnetic resonance imaging
MT	Medial temporal area in primates
MT+	Medial temporal and medial superior temporal visual area in primates, forming the medial temporal complex
MST	Medial superior temporal area in primates
MVPA	Multivariate pattern analysis
PC	Parvocellular
PD	Proton-density
RDS	Random dot stereogram

RDK   Random dot kinematogram

RGB   Red, green and blue phosphors

RGC   Retinal ganglion cell

ROI   Region of interest

S *or* S-cone   Cone cells sensitive to short wavelengths (~420nm)

SC   Superior colliculus

SEM   Standard error of the mean

S-(L+M)   Difference between the short wavelength cone responses, and the
                    sum of the long and medium wavelength cone responses,
                    characteristic of cells in the koniocellular pathway

SVM   Support vector machine

TE   Echo time

TI   Inversion time

TR   Repetition time

V1   Primary visual cortex

V2   Second visual area

V3   Third visual area

V3A/B   Visual area 'V3A/B'

V4   Fourth visual area

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. https://doi.org/10.1523/JNEUROSCI.2923-07.2007

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, *2*(2), 284–299. https://doi.org/10.1364/JOSAA.2.000284

Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). The variability of human, BOLD hemodynamic responses. *NeuroImage*, *8*(4), 360–369. https://doi.org/10.1006/nimg.1998.0369

Ajina, S., & Bridge, H. (2018). Subcortical pathways to extrastriate visual cortex underlie residual vision following bilateral damage to V1. *Neuropsychologia, in press*. https://doi.org/10.1016/j.neuropsychologia.2018.01.007

Ajina, S., Pestilli, F., Rokem, A., Kennard, C., & Bridge, H. (2015). Human blindsight is mediated by an intact geniculo-extrastriate pathway. *ELife*, *4*, 1-23. https://doi.org/10.7554/eLife.08935

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. https://doi.org/10.1152/jn.1984.51.1.16

Allen, B., Haun, A. M., Hanley, T., Green, C. S., & Rokers, B. (2015). Optimal combination of the binocular cues to 3D motion. *Investigative Ophthalmology & Visual Science*, *56*(12), 7589–7596. https://doi.org/10.1167/iovs.15-17696

Amano, K., Wandell, B. A., & Dumoulin, S. O. (2009). Visual field maps, population receptive field sizes, and visual field coverage in the human MT+ complex. *Journal of Neurophysiology*, *102*(5), 2704–2718. https://doi.org/10.1152/jn.00102.2009

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research*, *10*(12), 1411-1430. https://doi.org/10.1016/0042-6989(70)90092-1

Anstis, S. M., & Cavanagh, P. (1983). A minimum motion technique for judging equiluminance. *Color Vision: Physiology and Psychophysics*, 155–166.

Apthorp, D., Cass, J., & Alais, D. (2011). The spatial tuning of "motion streak" mechanisms revealed by masking and adaptation. *Journal of Vision*, *11*(7), 17–17. https://doi.org/10.1167/11.7.17

Baecke, S., Lützkendorf, R., Tempelmann, C., Müller, C., Adolf, D., Scholz, M., & Bernarding, J. (2009). Event-related functional magnetic resonance imaging (efMRI) of depth-by-disparity perception: Additional evidence for right-hemispheric lateralization. *Experimental Brain Research*, *196*(3), 453–458. https://doi.org/10.1007/s00221-009-1844-z

Baker, D. H., Kaestner, M., & Gouws, A. D. (2016). Measurement of crosstalk in stereoscopic display systems used for vision research. *Journal of Vision*, *16*(15), 14–14. https://doi.org/10.1167/16.15.14

Baker, J. F., Petersen, S. E., Newsome, W. T., & Allman, J. M. (1981). Visual response properties of neurons in four extrastriate visual areas of the owl monkey (Aotus trivirgatus): a quantitative comparison of medial, dorsomedial, dorsolateral, and middle temporal areas. *Journal of Neurophysiology*, *45*(3), 397–416. https://doi.org/10.1152/jn.1981.45.3.397

Baker, P. M., & Bair, W. (2016). A model of binocular motion integration in MT neurons. *Journal of Neuroscience*, *36*(24), 6563–6582. https://doi.org/10.1523/JNEUROSCI.3213-15.2016

Barendregt, M., Dumoulin, S. O., & Rokers, B. (2016). Impaired velocity processing reveals an agnosia for motion in depth. *Psychological Science*, *27*(11), 1474–1485. https://doi.org/10.1177/0956797616663488

Barlow H. B., Blakemore C., & Pettigrew J. D. (1967). The neural mechanism of binocular depth discrimination. *The Journal of Physiology*, *193*(2), 327–342. https://doi.org/10.1113/jphysiol.1967.sp008360

Beauchamp, M. S., Cox, R. W., & Deyoe, E. A. (1997). Graded effects of spatial and featural attention on human area MT and associated motion processing areas. *Journal of Neurophysiology*, *78*(1), 516–520. https://doi.org/10.1152/jn.1997.78.1.516

Beckett, A., Peirce, J. W., Sanchez-Panchuelo, R.-M., Francis, S., & Schluppeck, D. (2012). Contribution of large scale biases in decoding of direction-of-motion from high-resolution fMRI data in human early visual cortex. *NeuroImage*, *63*(3), 1623–1632. https://doi.org/10.1016/j.neuroimage.2012.07.066

Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2003). General multilevel linear modeling for group analysis in fMRI. *NeuroImage*, *20*(2), 1052–1063. https://doi.org/10.1016/S1053-8119(03)00435-X

Benardete, E. A., & Kaplan, E. (1999). Dynamics of primate P retinal ganglion cells: responses to chromatic and achromatic stimuli. *Journal of Physiology*, *519*(3), 775–790. https://doi.org/10.1111/j.1469-7793.1999.0775n.x

Beverley, K. I., & Regan, D. (1973). Evidence for the existence of neural mechanisms selectively sensitive to the direction of movement in space. *The Journal of Physiology*, *235*(1), 17–29. https://doi.org/10.1113/jphysiol.1973.sp010376

Billock, V. A. (1995). Cortical simple cells can extract achromatic information from the multiplexed chromatic and achromatic signals in the parvocellular pathway. *Vision Research*, *35*(16), 2359–2369. https://doi.org/10.1016/0042-6989(95)00002-H

Blake, R., & Cormack, R. H. (1979). On utrocular discrimination. *Perception & Psychophysics*, *26*(1), 53–68. https://doi.org/10.3758/BF03199861

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of Physiology*, *203*(1), 237–260. https://doi.org/10.1113/jphysiol.1969.sp008862

Born, R. T., & Bradley, D. C. (2005). Structure and function of visual area MT. *Annual Review of Neuroscience*, *28*(1), 157–189. https://doi.org/10.1146/annurev.neuro.26.041002.131052

Boynton, G. M., Engel, S. A., Glover, G. H., & Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *Journal of Neuroscience*, *16*(13), 4207–4221. https://doi.org/10.1523/JNEUROSCI.16-13-04207.1996

Braddick, O. (1974). A short-range process in apparent motion. *Vision Research*, *14*(7), 519–527.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.

Brewer, A. A., Liu, J., Wade, A. R., & Wandell, B. A. (2005). Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nature Neuroscience*, *8*(8), 1102–1109. https://doi.org/10.1038/nn1507

Bridge, H., & Parker, A. J. (2007). Topographical representation of binocular depth in the human visual cortex using fMRI. *Journal of Vision*, *7*(14), 15–15. https://doi.org/10.1167/7.14.15

Bridge, H., Hicks, S. L., Xie, J., Okell, T. W., Mannan, S., Alexander, I., … Kennard, C. (2010). Visual activation of extra-striate cortex in the absence of V1 activation. *Neuropsychologia*, *48*(14), 4148–4154. https://doi.org/10.1016/j.neuropsychologia.2010.10.022

Briggman, K. L., Helmstaedter, M., & Denk, W. (2011). Wiring specificity in the direction-selectivity circuit of the retina. *Nature*, *471*(7337), 183–188. https://doi.org/10.1038/nature09818

Britten, K. H., & Wezel, R. J. A. van. (1998). Electrical microstimulation of cortical area MST biases heading perception in monkeys. *Nature Neuroscience*, *1*(1), 59–63. https://doi.org/10.1038/259

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *Journal of Neuroscience*, *12*(12), 4745–4765.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1993). Responses of neurons in macaque MT to stochastic motion signals. *Visual Neuroscience*, *10*(06), 1157–1169.

Brooks, K. R. (2002). Interocular velocity difference contributes to stereomotion speed perception. *Journal of Vision*, *2*(3), 218–231. https://doi.org/10:1167/2.3.2

Brooks, K. R. (2002). Monocular motion adaptation affects the perceived trajectory of stereomotion. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(6), 1470–1482. https://doi.org/10.1037//0096-1523.28.6.1470

Brooks, K. R., & Stone, L. S. (2004). Stereomotion speed perception: Contributions from both changing disparity and interocular velocity difference over a range of relative disparities. *Journal of Vision*, *4*(12), 1061–1079. https://doi.org/10.1167/4.12.6

Brooks, K. R., & Stone, L. S. (2006). Spatial scale of stereomotion speed processing. *Journal of Vision*, *6*(11), 1257–66. https://doi.org/10.1167/6.11.9

Burock, M. A., Buckner, R. L., Woldorff, M. G., Rosen, B. R., & Dale, A. M. (1998). Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. *NeuroReport*, *9*(16), 3735.

Calkins, D. J. (2001). Seeing with S cones. *Progress in Retinal and Eye Research*, *20*(3), 255–287. https://doi.org/10.1016/S1350-9462(00)00026-4

Carmon, A., & Bechtoldt, H. P. (1969). Dominance of the right cerebral hemisphere for stereopsis. *Neuropsychologia*, *7*(1), 29–39. https://doi.org/10.1016/0028-3932(69)90042-6

Casagrande, V. A. (1994). A third parallel visual pathway to primate area V1. *Trends in Neurosciences*, *17*(7), 305–310. https://doi.org/10.1016/0166-2236(94)90065-5

Cavanagh, P., & Anstis, S. (1986). Do opponent-color channels contribute to motion? *Investigative Ophthalmology & Visual Science (Suppl.)*, *27*, 291.

Cavanagh, P., & Anstis, S. (1991). The contribution of color to motion in normal and color-deficient observers. *Vision Research*, *31*(12), 2109–2148. https://doi.org/10.1016/0042-6989(91)90169-6

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for Support Vector Machines. *ACM Transactions on Intelligent Systems Technology, 2*(3), 27:1–27:27. https://doi.org/10.1145/1961189.1961199

Chatterjee, S., & Callaway, E. M. (2002). S cone contributions to the magnocellular visual pathway in macaque monkey. *Neuron*, *35*(6), 1135–1146. https://doi.org/10.1016/S0896-6273(02)00874-7

Chawla, D., Rees, G., & Friston, K. J. (1999). The physiological basis of attentional modulation in extrastriate visual areas. *Nature Neuroscience*, *2*(7), 671–676. https://doi.org/10.1038/10230

Cheng, K., Waggoner, R. A., & Tanaka, K. (2001). Human ocular dominance columns as revealed by high-field functional magnetic resonance imaging. *Neuron*, *32*(2), 359–374. https://doi.org/10.1016/S0896-6273(01)00477-9

Cheong, S. K., Tailby, C., Solomon, S. G., & Martin, P. R. (2013). Cortical-like receptive fields in the lateral geniculate nucleus of marmoset monkeys. *Journal of Neuroscience*, *33*(16), 6864–6876. https://doi.org/10.1523/JNEUROSCI.5208-12.2013

Chichilnisky, E.-J., Heeger, D., & Wandell, B. A. (1993). Functional segregation of color and motion perception examined in motion nulling. *Vision Research*, *33*(15), 2113–2125. https://doi.org/10.1016/0042-6989(93)90010-T

Clifford, C. W. G., Mannion, D. J., & McDonald, J. S. (2009). Radial biases in the processing of motion and motion-defined contours by human visual cortex. *Journal of Neurophysiology*, *102*(5), 2974–2981. https://doi.org/10.1152/jn.00411.2009

Cogan, A. I., Kontsevich, L. L., Lomakin, A. J., Halpern, D. L., & Blake, R. (1995). Binocular disparity processing with opposite-contrast stimuli. *Perception*, *24*(1), 33–47. https://doi.org/10.1068/p240033

Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., … Shulman, G. L. (1998). A common network of functional areas for attention and eye movements. *Neuron*, *21*(4), 761–773. https://doi.org/10.1016/S0896-6273(00)80593-0

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, *3*(3), 292–297. https://doi.org/10.1038/73009

Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, G. L., & Petersen, S. E. (1990). Attentional modulation of neural processing of shape, color, and velocity in humans. *Science*, *248*(4962), 1556–1559. https://doi.org/10.1126/science.2360050

Corbetta, M., Miezin, F. M., Shulman, G. L., & Petersen, S. E. (1993). A PET study of visuospatial attention. *Journal of Neuroscience*, *13*(3), 1202–1226. https://doi.org/10.1523/JNEUROSCI.13-03-01202.1993

Corbetta, Maurizio, & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755

Corbetta, Maurizio, Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., … Shulman, G. L. (1998). A common network of functional areas for attention and eye movements. *Neuron*, *21*(4), 761–773. https://doi.org/10.1016/S0896-6273(00)80593-0

Corbetta, Maurizio, Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, *3*(3), 292–297. https://doi.org/10.1038/73009

Cormack, L. K., Czuba, T. B., Knöll, J., & Huk, A. C. (2017). Binocular mechanisms of 3D motion processing. *Annual Review of Vision Science*, *3*(1), 297–318. https://doi.org/10.1146/annurev-vision-102016-061259

Cottereau, B. R., McKee, S. P., & Norcia, A. M. (2014). Dynamics and cortical distribution of neural responses to 2D and 3D motion in human. *Journal of Neurophysiology*, *111*(3), 533–543. https://doi.org/10.1152/jn.00549.2013

Cowey, A., & Porter, J. (1979). Brain damage and global stereopsis. *Proceedings of the Royal Society B*, *204*(1157), 399–407. https://doi.org/10.1098/rspb.1979.0035

Cropper, S. J., & Derrington, A. M. (1996). Rapid colour-specific detection of motion in human vision. *Nature*, *379*(6560), 72–74. https://doi.org/10.1038/379072a0

Cumming, B. G. (1995). The relationship between stereoacuity and stereomotion thresholds. *Perception*, *24*(1), 105–114. https://doi.org/10.1068/p240105

Cumming, B. G., & Parker, A. J. (1994). Binocular mechanisms for detecting motion-in-depth. *Vision Research*, *34*(4), 483–495. https://doi.org/10.1016/0042-6989(94)90162-7

Cumming, B. G., & Parker, A. J. (1999). Binocular neurons in V1 of awake monkeys are selective for absolute, not relative, disparity. *Journal of Neuroscience*, *19*(13), 5602–5618.

Cumming, B. G., Shapiro, S. E., & Parker, A. J. (1998). Disparity detection in anticorrelated stereograms. *Perception*, *27*(11), 1367–1377. https://doi.org/10.1068/p271367

Curcio, C. A., Allen, K. A., Sloan, K. R., Lerea, C. L., Hurley, J. B., Klock, I. B., & Milam, A. H. (1991). Distribution and morphology of human cone photoreceptors stained with anti-blue opsin. *Journal of Comparative Neurology*, *312*(4), 610–624. https://doi.org/10.1002/cne.903120411

Czuba, T. B., Huk, A. C., Cormack, L. K., & Kohn, A. (2014). Area MT encodes three-dimensional motion. *The Journal of Neuroscience*, *34*(47), 15522–15533.

Czuba, T. B., Rokers, B., Guillet, K., Huk, A. C., & Cormack, L. K. (2011). Three-dimensional motion aftereffects reveal distinct direction-selective mechanisms for binocular processing of motion through depth. *Journal of Vision*, *11*(10), 18–18. https://doi.org/10.1167/11.10.18

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. *Journal of Neurophysiology*, *104*(5), 2886–2899. https://doi.org/10.1152/jn.00585.2009

Czuba, T. B., Rokers, B., Huk, A. C., & Cormack, L. K. (2012). To CD or not to CD: Is there a 3D motion aftereffect based on changing disparities? *Journal of Vision*, *12*(4), 1–3. https://doi.org/10.1167/12.4.7.Introduction

Dakin, S. C., & Mareschal, I. (2000). Sensitivity to contrast modulation depends on carrier spatial frequency and orientation. *Vision Research*, *40*(3), 311–329. https://doi.org/10.1016/S0042-6989(99)00179-0

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping*, *8*(2–3), 109–114.

Dale, A. M., & Buckner, R. L. (1997). Selective averaging of rapidly presented individual trials using fMRI. *Human Brain Mapping.*, *5*(5), 329–340. https://doi.org/10.1002/(SICI)1097-0193(1997)5:5%3C329::AID-HBM1%3E3.0.CO%3B2-5

De Monasterio F M, & Gouras P. (1975). Functional properties of ganglion cells of the rhesus monkey retina. *The Journal of Physiology*, *251*(1), 167–195. https://doi.org/10.1113/jphysiol.1975.sp011086

De Monasterio, F. M. (1978). Properties of ganglion cells with atypical receptive-field organization in retina of macaques. *Journal of Neurophysiology*, *41*(6), 1435–1449. https://doi.org/10.1152/jn.1978.41.6.1435

DeAngelis, G. C., & Newsome, W. T. (1999). Organization of disparity-selective neurons in macaque area MT. *The Journal of Neuroscience*, *19*(4), 1398–1415.

DeAngelis, G. C., & Newsome, W. T. (2004). Perceptual "read-out" of conjoined direction and disparity maps in extrastriate area MT. *PLOS Biology*, *2*(3), E77. https://doi.org/10.1371/journal.pbio.0020077

DeAngelis, G. C., & Uka, T. (2003). Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *Journal of Neurophysiology*, *89*(2), 1094–1111. https://doi.org/10.1152/jn.00717.2002

DeAngelis, G. C., Cumming, B. G., & Newsome, W. T. (1998). Cortical area MT and the perception of stereoscopic depth. *Nature*, *394*(6694), 677–680. https://doi.org/10.1038/29299

Dechent, P., & Frahm, J. (2000). Direct mapping of ocular dominance columns in human primary visual cortex. *NeuroReport*, *11*(14), 3247.

Derrington A. M., & Lennie P. (1984). Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *The Journal of Physiology*, *357*(1), 219–240. https://doi.org/10.1113/jphysiol.1984.sp015498

Desimone, R., & Schein, S. J. (1987). Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *Journal of Neurophysiology*, *57*(3), 835–868. https://doi.org/10.1152/jn.1987.57.3.835

Dougherty, R. F., Koch, V. M., Brewer, A. A., Fischer, B., Modersitzki, J., & Wandell, B. A. (2003). Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *Journal of Vision*, *3*(10), 586–598. https://doi.org/10.1167/3.10.1

Dougherty, R. F., Press, W. A., & Wandell, B. A. (1999). Perceived speed of colored stimuli. *Neuron*, *24*(4), 893–899. https://doi.org/10.1016/S0896-6273(00)81036-3

Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex*, *7*(2), 181–192. https://doi.org/10.1093/cercor/7.2.181

Felleman, D. J., & Kaas, J. H. (1984). Receptive-field properties of neurons in middle temporal visual area (MT) of owl monkeys. *Journal of Neurophysiology*, *52*(3), 488–513. https://doi.org/10.1152/jn.1984.52.3.488

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*(1), 1–47. https://doi.org/10.1093/cercor/1.1.1

Fernandez, J. M., & Farell, B. (2005). Seeing motion in depth using inter-ocular velocity differences. *Vision Research*, *45*(21), 2786–2798. https://doi.org/10.1016/j.visres.2005.05.021

Fernandez, J. M., & Farell, B. (2006). Motion in depth from interocular velocity differences revealed by differential motion aftereffect. *Vision Research*, *46*(8), 1307–1317. https://doi.org/10.1016/j.visres.2005.10.025

Field, G. D., & Chichilnisky, E. J. (2007). Information processing in the primate retina: Circuitry and coding. *Annual Review of Neuroscience*, *30*(1), 1–30. https://doi.org/10.1146/annurev.neuro.30.051606.094252

Fischer, E., Bülthoff, H. H., Logothetis, N. K., & Bartels, A. (2012). Visual motion responses in the posterior cingulate sulcus: A comparison to V5/MT and MST. *Cerebral Cortex*, *22*(4), 865–876. https://doi.org/10.1093/cercor/bhr154

Gallant, J. L., Braun, J., & Van Essen, D. C. (1993). Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science*, *259*(5091), 100–103. https://doi.org/10.1126/science.8418487

Geesaman, B. J., Born, R. T., Andersen, R. A., & Tootell, R. B. (1997). Maps of complex motion selectivity in the superior temporal cortex of the alert macaque monkey: a double-label 2-deoxyglucose study. *Cerebral Cortex*, *7*(8), 749–757. https://doi.org/10.1093/cercor/7.8.749

Gegenfurtner, K. R., & Hawken, M. J. (1996). Perceived velocity of luminance, chromatic and non-fourier stimuli: Influence of contrast and temporal frequency. *Vision Research*, *36*(9), 1281–1290. https://doi.org/10.1016/0042-6989(95)00198-0

Gegenfurtner, K. R., Kiper, D. C., Beusmans, J. M., Carandini, M., Zaidi, Q., & Movshon, J. A. (1994). Chromatic properties of neurons in macaque MT. *Visual Neuroscience*, *11*(3), 455–466.

Geisler, W. S. (1999). Motion streaks provide a spatial code for motion direction. *Nature*, *400*(6739), 65–69. https://doi.org/10.1038/21886

Georgieva, S., Peeters, R., Kolster, H., Todd, J. T., & Orban, G. A. (2009). The processing of three-dimensional shape from disparity in the human brain. *Journal of Neuroscience*, *29*(3), 727–742. https://doi.org/10.1523/JNEUROSCI.4753-08.2009

Gheorghiu, E., & Erkelens, C. J. (2005). Differences in perceived depth for temporally correlated and uncorrelated dynamic random-dot stereograms. *Vision Research*, *45*(12), 1603–1614. https://doi.org/10.1016/j.visres.2004.12.005

Girard, P., Salin, P. A., & Bullier, J. (1992). Response selectivity of neurons in area MT of the macaque monkey during reversible inactivation of area V1. *Journal of Neurophysiology*, *67*(6), 1437–1446. https://doi.org/10.1152/jn.1992.67.6.1437

Goodyear Bradley G., & Menon Ravi S. (2001). Brief visual stimulation allows mapping of ocular dominance in visual cortex using fMRI. *Human Brain Mapping*, *14*(4), 210–217. https://doi.org/10.1002/hbm.1053

Gouws, A. D., Alvarez, I., Watson, D. M., Uesaki, M., Rogers, J., & Morland, A. B. (2014). On the role of suppression in spatial attention: Evidence from negative BOLD in human subcortical and cortical structures. *Journal of Neuroscience*, *34*(31), 10347–10360. https://doi.org/10.1523/JNEUROSCI.0164-14.2014

Gray, R., & Regan, D. (1996). Cyclopean motion perception produced by oscillations of size, disparity and location. *Vision Research*, *36*(5), 655–665. https://doi.org/10.1016/0042-6989(95)00145-X

Greer, D. A., Bonnen, K., Huk, A. C., & Cormack, L. K. (2016). Speed discrimination in the far monocular periphery: A relative advantage for interocular comparisons consistent with self-motion. *Journal of Vision*, *16*(10), 7–7. https://doi.org/10.1167/16.10.7

Handwerker, D. A., Ollinger, J. M., & D'Esposito, M. (2004). Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *NeuroImage*, *21*(4), 1639–1651. https://doi.org/10.1016/j.neuroimage.2003.11.029

Hansen, K. A., Kay, K. N., & Gallant, J. L. (2007). Topographic organization in and near human visual area V4. *Journal of Neuroscience*, *27*(44), 11896–11911. https://doi.org/10.1523/JNEUROSCI.2991-07.2007

Harris, J. M., & Rushton, S. K. (2003). Poor visibility of motion in depth is due to early motion averaging. *Vision Research*, *43*(4), 385–392.

Harris, J. M., & Watamaniuk, S. N. J. (1995). Speed discrimination of motion-in-depth using binocular cues. *Vision Research*, *35*(7), 885–896. https://doi.org/10.1016/0042-6989(94)00194-Q

Harris, J. M., Nefs, H. T., & Grafton, C. E. (2008). Binocular vision and motion-in-depth. *Spatial Vision*, *21*(6), 531–547. https://doi.org/10.1163/156856808786451462

Hawken, M. J., Gegenfurtner, K. R., & Tang, C. (1994). Contrast dependence of colour and luminance motion mechanisms in human vision. *Nature*, *367*(6460), 268–270. https://doi.org/10.1038/367268a0

Haxby, J. V. (2012). Multivariate pattern analysis of fMRI: The early beginnings. *NeuroImage*, *62*(2), 852–855. https://doi.org/10.1016/j.neuroimage.2012.03.016

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(5539), 2425–2430. https://doi.org/10.1126/science.1063736

Haynes, J.-D., & Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience*, *8*(5), 686–691. https://doi.org/10.1038/nn1445

Haynes, J.-D., & Rees, G. (2006). Neuroimaging: Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, *7*(7), 523–534. https://doi.org/10.1038/nrn1931

Heeger, D. J., Boynton, G. M., Demb, J. B., Seidemann, E., & Newsome, W. T. (1999). Motion opponency in visual cortex. *Journal of Neuroscience*, *19*(16), 7162–7174.

Hendry, S. H. C., & Reid, R. C. (2000). The koniocellular pathway in primate vision. *Annual Review of Neuroscience*, *23*(1), 127–153. https://doi.org/10.1146/annurev.neuro.23.1.127

Hinrichs, H., Scholz, M., Tempelmann, C., Woldorff, M. G., Dale, A. M., & Heinze, H.-J. (2000). Deconvolution of event-related fMRI responses in fast-rate experimental designs: Tracking amplitude variations. *Journal of Cognitive Neuroscience*, *12*(supplement 2), 76–89. https://doi.org/10.1162/089892900564082

Hong, X., & Regan, D. (1989). Visual field defects for unidirectional and oscillatory motion in depth. *Vision Research*, *29*(7), 809–819. https://doi.org/10.1016/0042-6989(89)90093-X

Horton, J. C., & Hedley-Whyte, E. T. (1984). Mapping of cytochrome oxidase patches and ocular dominance columns in human visual cortex. *Philosophical Transactions of the Royal Society of London B*, *304*(1119), 255–272. https://doi.org/10.1098/rstb.1984.0022

Hubel, D. H., & Livingstone, M. S. (1987). Segregation of form, color, and stereopsis in primate area 18. *Journal of Neuroscience*, *7*(11), 3378–3415. https://doi.org/10.1523/JNEUROSCI.07-11-03378.1987

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, *195*(1), 215–243. https://doi.org/10.1113/jphysiol.1968.sp008455

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*(5227), 41–42. https://doi.org/10.1038/225041a0

Huk, A. C. (2012). Multiplexing in the primate motion pathway. *Vision Research*, *62*(4), 173–180. https://doi.org/10.1016/j.visres.2012.04.007

Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience*, *22*(16), 7195–7205.

Ip, I. B., Minini, L., Dow, J., Parker, A. J., & Bridge, H. (2014). Responses to interocular disparity correlation in the human cerebral cortex. *Ophthalmic and Physiological Optics*, *34*(2), 186–198. https://doi.org/10.1111/opo.12121

Irvin, G. E., Norton, T. T., Sesma, M. A., & Casagrande, V. A. (1986). W-like response properties of interlaminar zone cells in the lateral geniculate nucleus of a primate (Galago crassicaudatus). *Brain Research*, *362*(2), 254–270. https://doi.org/10.1016/0006-8993(86)90450-6

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*(2), 143–156. https://doi.org/10.1016/S1361-8415(01)00036-6

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*(2), 825–841. https://doi.org/10.1006/nimg.2002.1132

Joo, S. J., Czuba, T. B., Cormack, L. K., & Huk, A. C. (2016). Separate perceptual and neural processing of velocity- and disparity-based 3D motion signals. *Journal of Neuroscience*, *36*(42), 10791–10802. https://doi.org/10.1523/JNEUROSCI.1298-16.2016

Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.

Julesz, B., & Payne, R. A. (1968). Differences between monocular and binocular stroboscopic movement perception. *Vision Research*, *8*(4), 433–444. https://doi.org/10.1016/0042-6989(68)90111-9

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, *8*(5), 679–685. https://doi.org/10.1038/nn1444

Kamitani, Y., & Tong, F. (2006). Decoding seen and attended motion directions from activity in the human visual cortex. *Current Biology, 16*(11), 1096–1102. https://doi.org/10.1016/j.cub.2006.04.003

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*(11), 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

Kaplan, E., & Shapley, R. M. (1986). The primate retina contains two types of ganglion cells, with high and low contrast sensitivity. *Proceedings of the National Academy of Sciences, 83*(8), 2755–2757. https://doi.org/10.1073/pnas.83.8.2755

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron, 22*(4), 751–761. https://doi.org/10.1016/S0896-6273(00)80734-5

Kobatake, E., & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology, 71*(3), 856–867. https://doi.org/10.1152/jn.1994.71.3.856

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research, 39*(16), 2729–2737. https://doi.org/10.1016/S0042-6989(98)00285-5

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences, 103*(10), 3863–3868. https://doi.org/10.1073/pnas.0600244103

Kwee, I. L., Fujii, Y., Matsuzawa, H., & Nakada, T. (1999). Perceptual processing of stereopsis in humans: High-field (3.0-tesla) functional MRI study. *Neurology, 53*(7), 1599–1599. https://doi.org/10.1212/WNL.53.7.1599

Lages, M., Mamassian, P., & Graf, E. W. (2003). Spatial and temporal tuning of motion in depth. *Vision Research, 43*(27), 2861–2873. https://doi.org/10.1016/j.visres.2003.08.006

Larsson, J., & Heeger, D. J. (2006). Two retinotopic visual areas in human lateral occipital cortex. *Journal of Neuroscience, 26*(51), 13128–13142. https://doi.org/10.1523/JNEUROSCI.1657-06.2006

Larsson, J., Harrison, C., Jackson, J., Oh, S. M., & Zeringyte, V. (2016). Spatial scale and distribution of neurovascular signals underlying decoding of orientation and eye of origin from fMRI data. *Journal of Neurophysiology, 117*(2), 818–835. https://doi.org/10.1152/jn.00590.2016

Lee, B. B. (2010). Visual pathways and psychophysical channels in the primate. *The Journal of Physiology, 589*(1), 41–47. https://doi.org/10.1113/jphysiol.2010.192658

Lee, J., & Stromeyer, C. F. (1989). Contribution of human short-wave cones to luminance and motion detection. *J. Physiol.*, *413*(1), 563–593.

Leventhal, A. G., Rodieck, R. W., & Dreher, B. (1981). Retinal ganglion cell classes in the Old World monkey: morphology and central projections. *Science*, *213*(4512), 1139–1142. https://doi.org/10.1126/science.7268423

Likova, L. T., & Tyler, C. W. (2007). Stereomotion processing in the human occipital cortex. *NeuroImage*, *38*(2), 293–305. https://doi.org/10.1016/j.neuroimage.2007.06.039

Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, *240*(4853), 740–749. https://doi.org/10.1126/science.3283936

MacLeod, D. I. A., & Boynton, R. M. (1979). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *Journal of the Optical Society of America*, *69*(8), 1183. https://doi.org/10.1364/JOSA.69.001183

Maloney, R. T., Watson, T. L., & Clifford, C. W. G. (2013). Human cortical and behavioral sensitivity to patterns of complex motion at eccentricity. *Journal of Neurophysiology*, *110*(11), 2545–2556. https://doi.org/10.1152/jn.00445.2013

Maloney, R. T., Watson, T. L., & Clifford, C. W. G. (2014). Determinants of motion response anisotropies in human early visual cortex: The role of configuration and eccentricity. *NeuroImage*, *100*, 564–579. https://doi.org/10.1016/j.neuroimage.2014.06.057

Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Current Biology*, *14*(9), 744–751. https://doi.org/10.1016/j.cub.2004.04.028

Masland, R. H. (2012). The neuronal organization of the retina. *Neuron*, *76*(2), 266–280. https://doi.org/10.1016/j.neuron.2012.10.002

Maunsell, J. H. R., & Newsome, and W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, *10*(1), 363–401. https://doi.org/10.1146/annurev.ne.10.030187.002051

Maunsell, J. H., & Van Essen, D. C. (1983a). Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology*, *49*(5), 1148–1167. https://doi.org/10.1152/jn.1983.49.5.1148

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), 2563–2586. https://doi.org/10.1523/JNEUROSCI.03-12-02563.1983

Maunsell, J. H., Nealey, T. A., & DePriest, D. D. (1990). Magnocellular and parvocellular contributions to responses in the middle temporal visual area (MT) of the macaque monkey. *Journal of Neuroscience*, *10*(10), 3323–3334. https://doi.org/10.1523/JNEUROSCI.10-10-03323.1990

McKeefry, D. J., & Burton, M. P. (2009). The perception of speed based on L-M and S-(L+M) cone opponent processing. *Vision Research*, *49*(8), 870–876. https://doi.org/10.1016/j.visres.2009.03.004

Menon, R. S., Ogawa, S., Strupp, J. P., & Uğurbil, K. (1997). Ocular dominance in human V1 demonstrated by functional magnetic resonance imaging. *Journal of Neurophysiology*, *77*(5), 2780–2787. https://doi.org/10.1152/jn.1997.77.5.2780

Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E., & Buckner, R. L. (2000). Characterizing the hemodynamic response: Effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *NeuroImage*, *11*(6), 735–759. https://doi.org/10.1006/nimg.2000.0568

Morand, S., Thut, G., Peralta, D., Grave, R., Clarke, S., Khateb, A., … Michel, C. M. (2000). Electrophysiological evidence for fast visual processing through the human koniocellular pathway when stimuli move. *Cerebral Cortex*, *10*(8), 817–825. https://doi.org/10.1093/cercor/10.8.817

Morgan, M. J., & Ward, R. (1980). Conditions for motion flow in dynamic visual noise. *Vision Research*, *20*(5), 431–435. https://doi.org/10.1016/0042-6989(80)90033-4

Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1987). Receptive field organization of complex cells in the cat's striate cortex. *Journal of Physiology*, *283*(1), 79–99. https://doi.org/10.1113/jphysiol.1978.sp012489

Movshon, J. Anthony, & Newsome, W. T. (1996). Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *Journal of Neuroscience*, *16*(23), 7733–7741.

Movshon, J., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattass, & C. Gross (Eds.), *Pattern Recognition Mechanisms* (pp. 117–151). Rome: Vatican Press. Retrieved from https://nyuscholars.nyu.edu/en/publications/the-analysis-of-moving-visual-patterns-3

Mullen, K. T., & Boulton, J. C. (1992). Interactions between colour and luminance contrast in the perception of motion. *Ophthalmic and Physiological Optics*, *12*(2), 201–205. https://doi.org/10.1111/j.1475-1313.1992.tb00290.x

Nasr, S., Polimeni, J. R., & Tootell, R. B. H. (2016). Interdigitated color- and disparity-selective columns within human visual cortical areas V2 and V3. *Journal of Neuroscience*, *36*(6), 1841–1857. https://doi.org/10.1523/JNEUROSCI.3518-15.2016

Nefs, H. T., O'Hare, L., & Harris, J. M. (2010). Two independent mechanisms for motion-in-depth perception: evidence from individual differences. *Frontiers in Psychology*, *1*(10), 1–8. https://doi.org/10.3389/fpsyg.2010.00155

Neri, P., Bridge, H., & Heeger, D. J. (2004). Stereoscopic processing of absolute and relative disparity in human visual cortex. *Journal of Neurophysiology*, *92*(3), 1880–1891. https://doi.org/10.1152/jn.01042.2003

Neri, P., Parker, A. J., & Blakemore, C. (1999). Probing the human stereoscopic system with reverse correlation. *Nature*, *401*(6754), 695–698. https://doi.org/10.1038/44409

Nestares, O., & Heeger, D. J. (2000). Robust multiresolution alignment of MRI brain volumes. *Magnetic Resonance in Medicine*, *43*(5), 705–715.

Newsome, W. T., & Pare, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *The Journal of Neuroscience*, *8*(6), 2201–2211.

Nienborg, H., Bridge, H., Parker, A. J., & Cumming, B. G. (2004). Receptive field size in V1 neurons limits acuity for perceiving disparity modulation. *Journal of Neuroscience*, *24*(9), 2065–2076. https://doi.org/10.1523/JNEUROSCI.3887-03.2004

Nienborg, H., Bridge, H., Parker, A. J., & Cumming, B. G. (2005). Neuronal computation of disparity in V1 limits temporal resolution for detecting disparity modulation. *Journal of Neuroscience*, *25*(44), 10207–10219. https://doi.org/10.1523/JNEUROSCI.2342-05.2005

Nishihara, H. K. (1984). Practical real-time imaging stereo matcher. *Optical Engineering*, *23*(5), 536–545. https://doi.org/10.1117/12.7973334

Nobre, A. C., Sebestyen, G. N., Gitelman, D. R., Mesulam, M. M., Frackowiak, R. S., & Frith, C. D. (1997). Functional localization of the system for visuospatial attention using positron emission tomography. *Brain*, *120*(3), 515–533. https://doi.org/10.1093/brain/120.3.515

Norcia, A. M., & Tyler, C. W. (1984). Temporal frequency limits for stereoscopic apparent motion processes. *Vision Research*, *24*(5), 395–401. https://doi.org/10.1016/0042-6989(84)90037-3

Norton, T. T., & Casagrande, V. A. (1982). Laminar organization of receptive-field properties in lateral geniculate nucleus of bush baby (Galago crassicaudatus). *Journal of Neurophysiology*, *47*(4), 715–741. https://doi.org/10.1152/jn.1982.47.4.715

Norton, T. T., Casagrande, V. A., Irvin, G. E., Sesma, M. A., & Petry, H. M. (1988). Contrast-sensitivity functions of W-, X-, and Y-like relay cells in the lateral geniculate nucleus of bush baby, Galago crassicaudatus. *Journal of Neurophysiology*, *59*(6), 1639–1656. https://doi.org/10.1152/jn.1988.59.6.1639

Nover, H., Anderson, C. H., & DeAngelis, G. C. (2005). A logarithmic, scale-invariant representation of speed in macaque middle temporal area accounts for speed discrimination performance. *Journal of Neuroscience*, *25*(43), 10049–10060. https://doi.org/10.1523/JNEUROSCI.1661-05.2005

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. https://doi.org/10.1126/science.2396096

Op de Beeck, H. P. (2010). Against hyperacuity in brain reading: Spatial smoothing does not hurt multivariate fMRI analyses? *NeuroImage*, *49*(3), 1943–1948. https://doi.org/10.1016/j.neuroimage.2009.02.047

Orban, G. A., Claeys, K., Nelissen, K., Smans, R., Sunaert, S., Todd, J. T., … Vanduffel, W. (2006). Mapping the parietal cortex of human and non-human primates. *Neuropsychologia*, *44*(13), 2647–2667. https://doi.org/10.1016/j.neuropsychologia.2005.11.001

Orban, G. A., Sunaert, S., Todd, J. T., Van Hecke, P., & Marchal, G. (1999). Human cortical regions involved in extracting depth from motion. *Neuron*, *24*(4), 929–940. https://doi.org/10.1016/S0896-6273(00)81040-5

Parkes Laura M., Schwarzbach Jens V., Bouts Annemieke A., Deckers Roel h R., Pullens Pim, Kerskens Christian M., & Norris David G. (2005). Quantifying the spatial resolution of the gradient echo and spin echo BOLD response at 3 Tesla. *Magnetic Resonance in Medicine*, *54*(6), 1465–1472. https://doi.org/10.1002/mrm.20712

Pasley, B. N., Inglis, B. A., & Freeman, R. D. (2007). Analysis of oxygen metabolism implies a neural origin for the negative BOLD response in human visual cortex. *NeuroImage*, *36*(2), 269–276. https://doi.org/10.1016/j.neuroimage.2006.09.015

Pasupathy, A., & Connor, C. E. (1999). Responses to contour features in macaque area V4. *Journal of Neurophysiology*, *82*(5), 2490–2502. https://doi.org/10.1152/jn.1999.82.5.2490

Pasupathy, A., & Connor, C. E. (2001). Shape representation in area V4: Position-specific tuning for boundary conformation. *Journal of Neurophysiology*, *86*(5), 2505–2519. https://doi.org/10.1152/jn.2001.86.5.2505

Pasupathy, A., & Connor, C. E. (2002). Population coding of shape in area V4. *Nature Neuroscience*, *5*(12), 1332–1338. https://doi.org/10.1038/972

Pedregosa, F., Eickenberg, M., Ciuciu, P., Thirion, B., & Gramfort, A. (2015). Data-driven HRF estimation for encoding and decoding models. *NeuroImage*, *104*, 209–220. https://doi.org/10.1016/j.neuroimage.2014.09.060

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Peng, Q., & Shi, B. E. (2010). The changing disparity energy model. *Vision Research*, *50*(2), 181–192. https://doi.org/10.1016/j.visres.2009.11.012

Peng, Q., & Shi, B. E. (2014). Neural population models for perception of motion in depth. *Vision Research*, *101*(7), 11–31. https://doi.org/10.1016/j.visres.2014.04.014

Perry, V. 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. https://doi.org/10.1016/0306-4522(84)90006-X

Poggio, G. F., & Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *Journal of Neurophysiology*, *40*(6), 1392–1405. https://doi.org/10.1152/jn.1977.40.6.1392

Poggio, Gian F., Motter, B. C., Squatrito, S., & Trotter, Y. (1985). Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. *Vision Research*, *25*(3), 397–406. https://doi.org/10.1016/0042-6989(85)90065-3

Ponce, C. R., Hunter, J. N., Pack, C. C., Lomber, S. G., & Born, R. T. (2011). Contributions of indirect pathways to visual response properties in macaque middle temporal area MT. *Journal of Neuroscience*, *31*(10), 3894–3903. https://doi.org/10.1523/JNEUROSCI.5362-10.2011

Ponce, C. R., Lomber, S. G., & Born, R. T. (2008). Integrating motion and depth via parallel pathways. *Nature Neuroscience*, *11*(2), 216–223. https://doi.org/10.1038/nn2039

Portfors-Yeomans, C. V., & Regan, D. (1996). Cyclopean discrimination thresholds for the direction and speed of motion in depth. *Vision Research*, *36*(20), 3265–3279. https://doi.org/10.1016/0042-6989(96)00065-X

Press, W. A., Brewer, A. A., Dougherty, R. F., Wade, A. R., & Wandell, B. A. (2001). Visual areas and spatial summation in human visual cortex. *Vision Research*, *41*(10), 1321–1332. https://doi.org/10.1016/S0042-6989(01)00074-8

Prins, N., & Kingdom, F. A. A. (2009). Palamedes: Matlab routines for analyzing psychophysical data. Retrieved from http://www.palamedestoolbox.org/

Ptito, A., Zatorre, R. J., Larson, W. L., & Tosoni, C. (1991). Stereopsis after unilateral anterior temporal lobectomy - dissociation between local and global measures. *Brain*, *114*(3), 1323–1333. https://doi.org/10.1093/brain/114.3.1323

Qian, N., & Andersen, R. A. (1995). V1 responses to transparent and nontransparent motions. *Experimental Brain Research*, *103*(1), 41–50. https://doi.org/10.1007/BF00241963

Regan, D. (1993). Binocular correlates of the direction of motion in depth. *Vision Research*, *33*(16), 2359–2360. https://doi.org/10.1016/0042-6989(93)90114-C

Regan, D., & Beverley, K. I. (1973). Some dynamic features of depth perception. *Vision Research*, *13*(12), 2369–2379. https://doi.org/10.1016/0042-6989(73)90236-8

Regan, D., Erkelens, C. J., & Collewijn, H. (1986). Visual field defects for vergence eye movements and for stereomotion perception. *Investigative Ophthalmology & Visual Science*, *27*(5), 806–819.

Ress, D., & Heeger, D. J. (2003). Neuronal correlates of perception in early visual cortex. *Nature Neuroscience*, *6*(4), 414–420. https://doi.org/10.1038/nn1024

Ress, D., Backus, B. T., & Heeger, D. J. (2000). Activity in primary visual cortex predicts performance in a visual detection task. *Nature Neuroscience*, *3*(9), 940–945. https://doi.org/10.1038/78856

Reynolds, J. H., & Heeger, D. J. (2009). The normalization model of attention. *Neuron*, *61*(2), 168–185. https://doi.org/10.1016/j.neuron.2009.01.002

Richards, W., & Regan, D. (1973). A stereo field map with implications for disparity processing. *Investigative Ophthalmology & Visual Science*, *12*(12), 904–909.

Rodman, H. R., & Albright, T. D. (1987). Coding of visual stimulus velocity in area MT of the macaque. *Vision Research*, *27*(12), 2035–2048. https://doi.org/10.1016/0042-6989(87)90118-0

Rodman, H., Gross, C., & Albright, T. (1989). Afferent basis of visual response properties in area MT of the macaque. I. Effects of striate cortex removal. *Journal of Neuroscience*, *9*(6), 2033–2050.

Rokers, B., Cormack, L. K., & Huk, A. C. (2008). Strong percepts of motion through depth without strong percepts of position in depth. *Journal of Vision*, *8*(4), 1-10. https://doi.org/10.1167/8.4.6

Rokers, B., Cormack, L. K., & Huk, A. C. (2009). Disparity- and velocity-based signals for three-dimensional motion perception in human MT+. *Nature Neuroscience*, *12*(8), 1050–1055. https://doi.org/10.1038/nn.2343

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), 10–10. https://doi.org/10.1167/11.2.10

Rushbass C., & Westheimer G. (1961). Disjunctive eye movements. *Journal of Physiology*, *159*(2), 339–360. https://doi.org/10.1113/jphysiol.1961.sp006812

Rust, N. C., Mante, V., Simoncelli, E. P., & Movshon, J. A. (2006). How MT cells analyze the motion of visual patterns. *Nature Neuroscience*, *9*(11), 1421–1431. https://doi.org/10.1038/nn1786

Saenz, M., Buracas, G. T., & Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nature Neuroscience*, *5*(7), 631–632. https://doi.org/10.1038/nn876

Saito, H., Tanaka, K., Isono, H., Yasuda, M., & Mikami, A. (1989). Directionally selective response of cells in the middle temporal area (MT) of the macaque monkey to the movement of equiluminous opponent color stimuli. *Experimental Brain Research*, *75*(1). https://doi.org/10.1007/BF00248524

Sakano, Y., & Allison, R. S. (2014). Aftereffect of motion-in-depth based on binocular cues: Effects of adaptation duration, interocular correlation, and temporal correlation. *Journal of Vision*, *14*(8), 1–14. https://doi.org/10.1167/14.8.21

Sakano, Y., Allison, R. S., & Howard, I. P. (2012). Motion aftereffect in depth based on binocular information. *Journal of Vision*, *12*(1), 1–15. https://doi.org/10.1167/12.1.11

Sanada, T. M., & DeAngelis, G. C. (2014). Neural representation of motion-in-depth in area MT. *Journal of Neuroscience*, *34*(47), 15508–21. https://doi.org/10.1523/JNEUROSCI.1072-14.2014

Schira, M. M., Tyler, C. W., Breakspear, M., & Spehar, B. (2009). The foveal confluence in human visual cortex. *Journal of Neuroscience, 29*(28), 9050–9058. https://doi.org/10.1523/JNEUROSCI.1760-09.2009

Schwarzkopf, D. S., Schindler, A., & Rees, G. (2010). Knowing with which eye we see: Utrocular discrimination and eye-specific signals in human visual cortex. *PLOS ONE*, *5*(10), e13775. https://doi.org/10.1371/journal.pone.0013775

Seidemann, E., Poirson, A. B., Wandell, B. A., & Newsome, W. T. (1999). Color signals in area MT of the macaque monkey. *Neuron*, *24*(4), 911–917. https://doi.org/10.1016/S0896-6273(00)81038-7

Serences, J. T. (2004). A comparison of methods for characterizing the event-related BOLD timeseries in rapid fMRI. *NeuroImage*, *21*(4), 1690–1700. https://doi.org/10.1016/j.neuroimage.2003.12.021

Serences, J. T., & Boynton, G. M. (2007). Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron*, *55*(2), 301–312. https://doi.org/10.1016/j.neuron.2007.06.015

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., … Tootell, R. B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*(5212), 889–893.

Shioiri, S, Saisho, H., & Yaguchi, H. (2000). Motion in depth based on inter-ocular velocity differences. *Vision Research*, *40*(19), 2565–2572.

Shioiri, S., Kakehi, D., Tashiro, T., & Yaguchi, H. (2009). Integration of monocular motion signals and the analysis of interocular velocity differences for the perception of motion-in-depth. *Journal of Vision*, *9*(13), 1–17. https://doi.org/10.1167/9.13.10.Introduction

Shioiri, S., Nakajima, T., Kakehi, D., & Yaguchi, H. (2008). Differences in temporal frequency tuning between the two binocular mechanisms for seeing motion in depth. *Journal of the Optical Society of America A*, *25*(7), 1574–1585. https://doi.org/10.1364/JOSAA.25.001574

Shioiri, S., Saisho, H., & Yaguchi, H. (2000). Motion in depth based on inter-ocular velocity differences. *Vision Research*, *40*(19), 2565–2572. https://doi.org/10.1016/S0042-6989(00)00130-9

Shioiri, S., Yoshizawa, M., Ogiya, M., Matsumiya, K., & Yaguchi, H. (2012). Low-level motion analysis of color and luminance for perception of 2D and 3D motion. *Journal of Vision*, *12*(6), 33–33. https://doi.org/10.1167/12.6.33

Shioiri, Satoshi, Nakajima, T., Kakehi, D., & Yaguchi, H. (2008). Differences in temporal frequency tuning between the two binocular mechanisms for seeing motion in depth. *Journal of the Optical Society of America, A: Optics, Image Science, and Vision*, *25*(7), 1574–1585.

Shmuel, A., Chaimow, D., Raddatz, G., Ugurbil, K., & Yacoub, E. (2010). Mechanisms underlying decoding at 7 T: Ocular dominance columns, broad structures, and macroscopic blood vessels in V1 convey information on the stimulated eye. *NeuroImage*, *49*(3), 1957–1964. https://doi.org/10.1016/j.neuroimage.2009.08.040

Shmuel, A., Yacoub, E., Chaimow, D., Logothetis, N. K., & Ugurbil, K. (2007). Spatio-temporal point-spread function of fMRI signal in human gray matter at 7 Tesla. *NeuroImage*, *35*(2), 539–552. https://doi.org/10.1016/j.neuroimage.2006.12.030

Shmuel, A., Yacoub, E., Pfeuffer, J., Van de Moortele, P.-F., Adriany, G., Hu, X., & Ugurbil, K. (2002). Sustained negative BOLD, blood flow and oxygen consumption response and its coupling to the positive response in the human brain. *Neuron*, *36*(6), 1195–1210. https://doi.org/10.1016/S0896-6273(02)01061-9

Silvanto, J. (2015). Why is "blindsight" blind? A new perspective on primary visual cortex, recurrent activity and visual awareness. *Consciousness and Cognition*, *32*(3), 15–32. https://doi.org/10.1016/j.concog.2014.08.001

Sincich, L. C., & Horton, J. C. (2005). The circuitry of V1 and V2: Integration of color, form, and motion. *Annual Review of Neuroscience*, *28*(1), 303–326. https://doi.org/10.1146/annurev.neuro.28.061604.135731

Sincich, L. C., Park, K. F., Wohlgemuth, M. J., & Horton, J. C. (2004). Bypassing V1: a direct geniculate input to area MT. *Nature Neuroscience*, *7*(10), 1123–1128. https://doi.org/10.1038/nn1318

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*(3), 143–155. https://doi.org/10.1002/hbm.10062

Smolyanskaya, A., Haefner, R. M., Lomber, S. G., & Born, R. T. (2015). A modality-specific feedforward component of choice-related activity in MT. *Neuron*, *87*(1), 208–219. https://doi.org/10.1016/j.neuron.2015.06.018

Stockman, A., & Sharpe, L. T. (2000). The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Research*, *40*(13), 1711–1737. https://doi.org/10.1016/S0042-6989(00)00021-3

Stone, L. S., & Thompson, P. (1992). Human speed perception is contrast dependent. *Vision Research*, *32*(8), 1535–1549. https://doi.org/10.1016/0042-6989(92)90209-2

Sunaert, S., Van Hecke, P., Marchal, G., & Orban, G. A. (1999). Motion-responsive regions of the human brain. *Experimental Brain Research*, *127*(4), 355–370. https://doi.org/10.1007/s002210050804

Swisher, J. D., Halko, M. A., Merabet, L. B., McMains, S. A., & Somers, D. C. (2007). Visual topography of human intraparietal sulcus. *Journal of Neuroscience*, *27*(20), 5326–5337. https://doi.org/10.1523/JNEUROSCI.0991-07.2007

Szczepanski, S. M., Konen, C. S., & Kastner, S. (2010). Mechanisms of spatial attention control in frontal and parietal cortex. *Journal of Neuroscience*, *30*(1), 148–160. https://doi.org/10.1523/JNEUROSCI.3862-09.2010

Tailby C., Szmajda B. A., Buzás P., Lee B. B., & Martin P. R. (2008). Transmission of blue (S) cone signals through the primate lateral geniculate nucleus. *The Journal of Physiology*, *586*(24), 5947–5967. https://doi.org/10.1113/jphysiol.2008.161893

Tailby, C., Dobbie, W. J., Solomon, S. G., Szmajda, B. A., Hashemi-Nezhad, M., Forte, J. D., & Martin, P. R. (2010). Receptive field asymmetries produce color-dependent direction selectivity in primate lateral geniculate nucleus. *Journal of Vision*, *10*(8), 1–18. https://doi.org/10.1167/10.8.1

Tailby, C., Majaj, N. J., & Movshon, J. A. (2010). Binocular integration of pattern motion signals by MT neurons and by human observers. *Journal of Neuroscience*, *30*(21), 7344–7349. https://doi.org/10.1523/JNEUROSCI.4552-09.2010

Thiele, A., Dobkins, K. R., & Albright, T. D. (1999). The contribution of color to motion processing in macaque middle temporal area. *J. Neurosci.*, *19*(15), 6571–6587.

Thiele, A., Dobkins, K. R., & Albright, T. D. (2001). Neural correlates of chromatic motion perception. *Neuron*, *32*(2), 351–358. https://doi.org/10.1016/S0896-6273(01)00463-9

Tong, F., & Engel, S. A. (2001). Interocular rivalry revealed in the human cortical blind-spot representation. *Nature*, *411*(6834), 195–199. https://doi.org/10.1038/35075583

Tong, F., Meng, M., & Blake, R. (2006). Neural bases of binocular rivalry. *Trends in Cognitive Sciences*, *10*(11), 502–511. https://doi.org/10.1016/j.tics.2006.09.003

Tootell, R. B. H., Hadjikhani, N. K., Vanduffel, W., Liu, A. K., Mendola, J. D., Sereno, M. I., & Dale, A. M. (1998). Functional analysis of primary visual cortex (V1) in humans. *Proceedings of the National Academy of Sciences*, *95*(3), 811–817. https://doi.org/10.1073/pnas.95.3.811

Tootell, R. B. H., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., & Dale, A. M. (1998). The retinotopy of visual spatial attention. *Neuron*, *21*(6), 1409–1422. https://doi.org/10.1016/S0896-6273(00)80659-5

Tootell, R. B., & Hamilton, S. L. (1989). Functional anatomy of the second visual area (V2) in the macaque. *Journal of Neuroscience*, *9*(8), 2620–2644. https://doi.org/10.1523/JNEUROSCI.09-08-02620.1989

Treue, S., & Martinez-Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, *399*(6736), 575–579. https://doi.org/10.1038/21176

Treue, S., & Maunsell, J. H. R. (1996). Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature*, *382*(6591), 539–541. https://doi.org/10.1038/382539a0

Ts'o, D. Y., Frostig, R. D., Lieke, E. E., & Grinvald, A. (1990). Functional organization of primate visual cortex revealed by high resolution optical imaging. *Science*, *249*(4967), 417–420. https://doi.org/10.1126/science.2165630

Tyler, C. W. (1971). Stereoscopic depth movement: Two eyes less sensitive than one. *Science*, *174*(4012), 958–961. https://doi.org/10.1126/science.174.4012.958

Tyler, C. W., & Cavanagh, P. (1991). Purely chromatic perception of motion in depth: two eyes as sensitive as one. *Perception & Psychophysics*, *49*(1), 53–61.

Vaina, L. M. (1989). Selective impairment of visual motion interpretation following lesions of the right occipito-parietal area in humans. *Biological Cybernetics*, *61*(5), 347–359. https://doi.org/10.1007/BF00200800

Van Essen, D. C., Maunsell, J. H. R., & Bixby, J. L. (1981). The middle temporal visual area in the macaque: Myeloarchitecture, connections, functional properties and topographic organization. *Journal of Comparative Neurology*, *199*(3), 293–326. https://doi.org/10.1002/cne.901990302

Vanduffel, W., Fize, D., Peuskens, H., Denys, K., Sunaert, S., Todd, J. T., & Orban, G. A. (2002). Extracting 3D from motion: Differences in human and monkey intraparietal cortex. *Science*, *298*(5592), 413–415. https://doi.org/10.1126/science.1073574

Verghese, P., Kim, Y.-J., & Wade, A. R. (2012). Attention selects informative neural populations in human V1. *Journal of Neuroscience*, *32*(46), 16379–16390. https://doi.org/10.1523/JNEUROSCI.1174-12.2012

Wade, A. R., & Rowland, J. (2010). Early suppressive mechanisms and the negative blood oxygenation level-dependent response in human visual cortex. *Journal of Neuroscience*, *30*(14), 5008–5019. https://doi.org/10.1523/JNEUROSCI.6260-09.2010

Wade, A. R., Brewer, A. A., Rieger, J. W., & Wandell, B. A. (2002). Functional measurements of human ventral occipital cortex: retinotopy and colour. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *357*(1424), 963–973.

Walsh, J. W. T. (1953). *Photometry*. London, Great Britain: Constable & Company LTD. Retrieved from http://archive.org/details/Photometry

Wandell, B. A., & Winawer, J. (2011). Imaging retinotopic maps in the human brain. *Vision Research*, *51*(7), 718–737. https://doi.org/10.1016/j.visres.2010.08.004

Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron*, *56*(2), 366–383. https://doi.org/10.1016/j.neuron.2007.10.012

Wandell, B. A., Poirson, A. B., Newsome, W. T., Baseler, H. A., Boynton, G. M., Huk, A., … Sharpe, L. T. (1999). Color signals in human motion-selective cortex. *Neuron*, *24*(4), 901–909.

Wardle, S. G., & Alais, D. (2013). Evidence for speed sensitivity to motion in depth from binocular cues. *Journal of Vision*, *13*(1), 1–16. https://doi.org/10.1167/13.1.17

Watson, A. B., & Ahumada, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, *2*(2), 322–342. https://doi.org/10.1364/JOSAA.2.000322

Webb, B. S., Dhruv, N. T., Solomon, S. G., Tailby, C., & Lennie, P. (2005). Early and late mechanisms of surround suppression in striate cortex of macaque. *Journal of Neuroscience*, *25*(50), 11666–11675. https://doi.org/10.1523/JNEUROSCI.3414-05.2005

Welbourne, L. E., Morland, A. B., & Wade, A. R. (2018). Population receptive field (pRF) measurements of chromatic responses in human visual cortex using fMRI. *NeuroImage*, *167*, 84–94. https://doi.org/10.1016/j.neuroimage.2017.11.022

White, A. J. R., Solomon, S. G., & Martin, P. R. (2001). Spatial properties of koniocellular cells in the lateral geniculate nucleus of the marmoset Callithrix jacchus. *The Journal of Physiology*, *533*(2), 519–535. https://doi.org/10.1111/j.1469-7793.2001.0519a.x

Wilcox, L. M., & Hess, R. F. (1997). Scale selection for second-order (non-linear) stereopsis. *Vision Research*, *37*(21), 2981–2992. https://doi.org/10.1016/S0042-6989(97)00061-8

Williams, D. W., & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions (abstract only). *SIGGRAPH Computer Graphics*, *18*(1), 24–24. https://doi.org/10.1145/988525.988533

Wilson, H. R., Blake, R., & Pokorny, J. (1988). Limits of binocular fusion in the short wave sensitive ("blue") cones. *Vision Research*, *28*(4), 555–562. https://doi.org/10.1016/0042-6989(88)90176-9

Winawer, J., Horiguchi, H., Sayres, R. A., Amano, K., & Wandell, B. A. (2010). Mapping hV4 and ventral occipital cortex: the venous eclipse. *Journal of Vision*, *10*(5), 1.

Wojciulik, E., & Kanwisher, N. (1999). The generality of parietal involvement in visual attention. *Neuron*, *23*(4), 747–764. https://doi.org/10.1016/S0896-6273(01)80033-7

Woolrich, M. W. (2008). Robust group analysis using outlier inference. *NeuroImage*, *41*(2), 286–301. https://doi.org/10.1016/j.neuroimage.2008.02.042

Woolrich, M. W., Behrens, T. E. J., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage*, *21*(4), 1732–1747. https://doi.org/10.1016/j.neuroimage.2003.12.023

Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of fMRI data. *NeuroImage*, *14*(6), 1370–1386. https://doi.org/10.1006/nimg.2001.0931

Xu, X., Ichida, J. M., Allison, J. D., Boyd, J. D., Bonds, A. B., & Casagrande, V. A. (2004). A comparison of koniocellular, magnocellular and parvocellular receptive field properties in the lateral geniculate nucleus of the owl monkey (Aotus trivirgatus). *The Journal of Physiology*, *531*(1), 203–218. https://doi.org/10.1111/j.1469-7793.2001.0203j.x

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. https://doi.org/10.1126/science.1057916

Yacoub, E., Shmuel, A., Logothetis, N., & Uğurbil, K. (2007). Robust detection of ocular dominance columns in humans using Hahn Spin Echo BOLD functional MRI at 7 Tesla. *NeuroImage*, *37*(4), 1161–1177. https://doi.org/10.1016/j.neuroimage.2007.05.020

Yoshida, K., & Benevento, L. A. (1981). The projection from the dorsal lateral geniculate nucleus of the thalamus to extrastriate visual association cortex in the macaque monkey. *Neuroscience Letters*, *22*(2), 103–108. https://doi.org/10.1016/0304-3940(81)90071-9

Zeater, N., Cheong, S. K., Solomon, S. G., Dreher, B., & Martin, P. R. (2015). Binocular visual responses in the primate lateral geniculate nucleus. *Current Biology*, *25*(24), 3190–3195. https://doi.org/10.1016/j.cub.2015.10.033

Zeki, S. M. (1980). The response properties of cells in the middle temporal area (area MT) of owl monkey visual cortex. *Proceedings from the Royal Society London B*, *207*(1167), 239–248. https://doi.org/10.1098/rspb.1980.0022

Zeki, S. M. (1993). *A vision of the brain.* Oxford ; Boston: Blackwell Scientific Publications.

Zeki, S. M. (1974). Functional organization of a visual area in the posterior bank of the sperior temporal sulcus of the rhesus monkey. *Journal of Physiology, 236*(3), 549–573.

Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging, 20*(1), 45–57. https://doi.org/10.1109/42.906424