1. Introduction

A brown bear pads across a snowfield. It is watched (hopefully from a safe distance) by an observer. As the bear tramps through the snow, a bear-shaped patch of darkness is projected onto the back of the observer’s eye. The motion of this image across the observer’s otherwise brightly illuminated retina causes a series of changes in the activity of densely packed photoreceptors that are sensitive to changes in light intensity. The observer’s visual system can, as the bear progresses, perform the remarkable feat of computing its speed and direction of motion – and the speed and direction of each of the bear’s constituent parts - from many million changes in neural firing rate. This ability has clear evolutionary advantages, and as such it has been widely selected for in the animal kingdom.

Less common is the ability to detect motion that is not based on changes in luminance. To return to our wintery example, the force and direction of the wind or the presence of a smaller animal burrowing under the snow can be determined by detecting changes in the pattern of random flicker caused as flakes of snow on the ground are disturbed. Here, the changes in luminance do not, in themselves, signal any consistent speed or direction of motion, but the movement can clearly be seen.

In this chapter, we review key aspects of visual motion perception with a particular emphasis on the cortical areas thought to be involved. We begin with the integration of motion signals across extended regions of the visual field. This is central to the ability of the visual cortex to bind multiple features together into a coherent, stable visual percept. We then move on to the question of plasticity within the early cortical areas responsible for motion perception and review the brain regions thought to be involved in the processing of complex motion information such the motion signals that flow over the retina as an observer moves around in their environment. In the final two sections of the chapter we consider the mechanisms involved in the perception of motion in the absence of useful luminance information and the consequences of lesions and abnormal development affecting the cortical areas responsible for motion perception.
2. Cortical processing of visual information

Visual processing has often been thought of as being subdivided into two parallel processing streams known as the parvocellular (also known as ventral) and magnocellular (also known as dorsal) pathways (Ungerleider & Mishkin, 1982; Goodale & Milner, 1992). This delineation begins at the level of retinal ganglion cells. Small (midget) and larger (parasol) ganglion cells project, respectively, to the distinct parvocellular (“P”) and magnocellular (“M”) layers of the lateral geniculate nucleus of the thalamus (LGN) (Derrington & Lennie, 1984; Merigan et al., 1991). In turn, these M and P LGN cells project to distinct sub-regions of layer 4c within the primary visual cortex (Hubel & Wiesel, 1972). According to the “dual pathway” model, the parvocellular pathway, primarily carrying high spatial frequency (fine detail) and colour information, then projects to ventral areas of the extrastriate cortex such as V4. Conversely, the magnocellular pathway, primarily carrying low spatial (coarse detail) and high temporal frequency information, innervates dorsal extrastriate regions such as the middle temporal visual area (MT) also known as V5 and the middle superior temporal visual area (MST). These projections are thought to produce cortical pathways specialized for form processing and spatial position/motion perception respectively (Ungerleider & Haxby, 1994). It is now clear that there is considerable crosstalk between these two pathways and that other connections exist between the retina and the brain that include the koniocellular layers of the LGN and other thalamic regions such as the superior colliculus and pulvinar (see de Haan & Cowey, 2011 for a recent review). However the concept of parallel processing has provided a useful framework for the investigation of motion perception and has inspired a large number of psychophysical studies in this area.

2.1. Local motion analysis – V1

The first port of call for the majority of visual information in the cortex is the primary visual cortex, or V1. It is also often referred to as the “striate cortex” due to its stratified appearance under the microscope. Thanks to the Nobel prize winning experiments by Hubel and Wiesel, we know that single striate cortex neurons in the cat (Hubel & Wiesel, 1959) and monkey (Hubel & Weisel, 1968) respond best to oriented lines and that some of these neurons are also selective for the direction in which a luminance-defined stimulus is moved across their receptive field. The neural architecture necessary to achieve luminance-based motion detection and discrimination is, therefore, already in place at the relatively low level of V1. Various computational theories have successfully modelled the detection of this type of luminance-based motion (Adelson & Bergen, 1985; van Santen & Sperling, 1985; Watson & Ahumada, 1985). The method common to all of these models is to combine the outputs of two neurons, one whose receptive field has an “odd” (sine phase) spatial profile and an “even” (cosine phase) temporal profile and another neuron whose receptive field has an even spatial profile and an odd temporal profile. By appropriately combining the outputs of non-directional V1 neurons whose spatial and temporal responses are 90° out-of-phase (often referred to as “quadrature pairs”), motion energy can be recovered (Adelson &
Bergen, 1985). Alternative models have been suggested based on inhibitory interactions between adjacent regions within the receptive field (Barlow & Levick, 1965), spatiotemporal differencing (Marr & Ullman, 1981) or spatiotemporal gradients (Johnston et al., 1992), but the motion energy model is currently the dominant model of V1 motion selectivity.

2.2. Global motion analysis – V3A and the middle temporal visual area (MT/V5)

The year 1985 was a seminal year for the study of visual motion (Burr & Thompson, 2011) seeing, as it did, publication of several influential models of local motion processing (Adelson & Bergen, 1985; van Santen & Sperling, 1985; Watson & Ahumada, 1985). Although it had taken a great leap forward, the race to understand the processing of visual motion was, however, far from over. It is one thing to understand how local motion selectivity arises in single neurons, but quite another to understand how these local motion signals are combined across space to give the perception of moving edges, surfaces and objects. The major hurdle (and it is a significant one) is that the output of individual local motion detectors, such as those present in V1 and those modelled in the literature mentioned above, is often ambiguous.

Since V1 neurons (or a hypothetical local motion detector) only “see” a small portion of the world, they do not respond maximally to a single stimulus form, but to a whole family of motions. This problem, referred to as the aperture problem, is demonstrated in Figure 1.

By integrating the outputs of many motion detectors over larger and larger portions of the visual field, the visual system can disambiguate these signals and extract genuine motion from the inherently ambiguous local motion signals, a processing stage known as the extraction of global motion. Such a process is thought to occur in extrastriate cortical areas such as V3A and MT, both of which contain cells that respond to coherent global motion (Allman et al., 1985; Movshon et al., 1985; Newsome & Pare, 1988; Rodman & Albright, 1989; Salzman et al., 1992; Heeger et al., 1999; Braddick et al., 2001). Similar properties seem to be present in the human analogue of MT known as V5 or hMT+ (Beckers & Zeki, 1995; Tootell et al., 1995; Huk & Heeger, 2002; Cowey et al., 2006). For example, we have recently demonstrated (Figure 2) that inhibitory repetitive transcranial magnetic stimulation (rTMS) of human V5 can impair the combination of local motion signals into a global motion percept (Thompson et al., 2009). With such abundant evidence that this cortical region is crucial to the perception of global motion, a question still to be answered is what sort of combinatorial processes are actually occurring in V5?

Several models of local motion combination have been proposed. A widely cited early model is the intersection-of-constraints (Adelson & Movshon, 1982) in which local one-dimensional (1D) motions (see legend of Figure 1) are extracted from the two-dimensional (2D) visual stimulus, their respective constraint lines are computed and the motion of the pattern corresponds to the intersection of these constraint lines. The intersection-of-constraints is a geometrical solution to the aperture problem and therefore provides a useful point of comparison for psychophysical data and resulting models aiming to understand
how global motion is computed within the brain. A number of psychophysical results have suggested that perception does not always follow the intersection of constraints rule (Ferrera & Wilson, 1990; Yo & Wilson, 1992) and, as such, alternative models have been developed. One key alternative is the vector sum model (Wilson et al., 1992; Wilson & Kim, 1994) in which global motion is computed as the sum of the local motion vectors.

Figure 1. The aperture problem. A) An edge, represented by the black line, moves in a particular direction, at a particular speed, through the receptive field of a motion-sensitive neuron (blue circle). The direction in which the edge is moving is represented by the red arrow, and the speed at which it is moving is represented by the length of the arrow. Assume that the neuron is sensitive to stimuli of this particular orientation moving at this particular speed, and so fires. The aperture problem arises because exactly the same response can be achieved by moving another edge, oriented at the same angle, through the receptive field in a completely different direction, at a much faster speed. Since the edge is uniform and featureless, the neuron is unable to determine its direction and speed of motion. The global motion of the edge is therefore ambiguous. B) It turns out that exactly the same stimulus can be reproduced by a family of different combinations of speed and direction constrained by a constraint line (dashed line) perpendicular to the end point of the shortest (i.e. slowest) motion vector. As such, only motion perpendicular to a contour can be detected by a single detector. It is therefore referred to as one-dimensional motion.

Both of these models are, however, two-stage operations in which the 2D stimulus is decomposed into its local 1D motions and then reconstructed according to a mathematical rule. There are a number of psychophysical results that are not consistent with this decomposition-recombination approach. The ability to discriminate the direction of motion of a plaid appears to depend critically upon the speed of the 2D features or “blobs” in the
Figure 2. Evidence for the role of human V5 in the processing of global motion. The top panel depicts plaid stimuli that are constructed from two gratings moving in different directions within a circular aperture. If the spatial properties of the gratings are sufficiently dissimilar then the two gratings are seen to drift over one another in directions orthogonal to their orientation due to the aperture problem described above (top-left image). However, if the gratings are similar to one another they may be perceived as a coherent patterned surface moving globally in a different direction from either of the component gratings (top-right image). The lower panel shows average data from 11 participants who viewed ambiguous plaid patterns that could be perceived as either coherent or incoherent before and after inhibitory repetitive transcranial magnetic stimulation of either V1 or V5. The probability of perceiving the stimuli as moving coherently is shown for three different time points after stimulation. rTMS of V5 significantly reduced the probability that participants would perceive plaids as moving coherently directly after stimulation, suggesting that this area is centrally involved in global motion perception. Stimulation of V1 had the opposite effect. Data re-plotted from (Thompson et al., 2009).

stimulus, not the speed of its components (Derrington & Badcock, 1992; Wright & Gurney, 1992). In addition, the size and number of blobs within plaid stimuli influence the perceived motion direction and the direction of the associated motion after-effects (Alais et al., 1994;
Alais et al., 1996; Alais et al., 1997). Furthermore, physiological data show that motion-selective cells in V1, that have broad orientation-tuning, respond to the motion of 2D features in the stimulus (Tinsley et al., 2003) and presumably feed this information forward to MT/V5, where it is combined with 1D local motion signals to influence perception. The visual system has for some time been likened to a Fourier analyser, that breaks down the complicated visual stimulus into its simpler 1D components, then recombines it (Campbell & Robson, 1968). The work of Tinsley et al. and Alais and colleagues is by no means fatal to the two-stage conception of global motion analysis, but rather reminds us that the first-stage is not a straightforward global Fourier analysis. In addition, these data suggest that the visual system may shift between different strategies for combining local motion based on the task demands and available information (Nishida, 2011).

2.2.1. Spatial suppression

Centre-surround antagonism is a common feature of the visual system that begins at the level of the retina and assists segmentation of the visual image into objects and background. Briefly, stimulation in a neuron’s central receptive field causes excitation, whilst stimulation in the receptive field surround results in inhibition. These two signals may cancel each other out if stimulation of the centre and surround is uniform. This results in neurons that are selective for discontinuities in the visual image. At the level of the cortex, excitatory and inhibitory regions are organised in such a way that neurons show selectivity for a variety of properties such as orientation, spatial frequency or motion within their receptive field centres. These are sometimes referred to as the "classical" receptive field. Many other cortical neurons also show a substantial inhibitory surround whose stimulus selectivity is antagonistic to the receptive field centre. It has been argued that this fundamental property of the visual system may give rise to some curious psychophysical correlates, including one well-known motion-based perceptual effect termed “spatial suppression.”

As the retinal size of a stimulus increases, one would expect motion detection and discrimination to improve according to spatial summation of contrast resulting from the recruitment and integration of progressively larger numbers of motion sensors. Contrary to this idea, Tadin et al. (2003) noted that observers got worse at discriminating the direction of motion as the size of a high contrast stimulus was increased. Spatial summation occurred as expected only if the patch was low contrast, resulting in better performance as size increased. Tadin et al. proposed that their psychophysical results were a perceptual correlate of centre-surround antagonism in motion-selective neurons in cortical visual area V5, and have recently reported TMS findings to support this view (Tadin et al., 2011). The rationale is that large, high contrast stimuli activate both the excitatory centre and inhibitory surround of cells in V5, resulting in a less robust neural representation of motion. Further, this effect does not occur for low contrast stimuli, as the inhibitory surround requires high contrasts to become active.

It was subsequently reported that older observers (over 60 years of age) showed weaker spatial suppression, which paradoxically led to better performance than younger observers.
in the high contrast conditions (Betts et al., 2005). The authors proposed that this weaker spatial suppression in older observers is a perceptual correlate of age-related changes in GABA-mediated inhibitory processes in the brain (Leventhal et al., 2003). Similar claims regarding a weakening (or not) of centre-surround antagonism in cortical areas have since been made using the same psychophysical technique for observers with schizophrenia (Tadin et al., 2006), depression (Golomb et al., 2009) and migraine (Battista et al., 2010).

However, the proposed correlation of this psychophysical effect to centre-surround antagonism in cortical area V5, and the connection with age- or psychiatric-related changes in the GABA inhibitory system, has been criticised on the basis that the evidence to date is circumstantial (Aaen-Stockdale et al., 2009; Wallisch & Kumbhani, 2009).

Chen and colleagues propose that spatial suppression is indeed a consequence of centre-surround antagonism, but that this surround inhibition is not occurring in V5 (Chen, 2011). They argue that the simple gratings or 100% coherent dot stimuli used in previous studies of these effects would primarily activate earlier motion-sensitive areas such as V1, that also show centre-surround antagonism. Rather than use these types of stimuli, they instead used a variable-coherence random dot stimulus. This sort of stimulus requires the involvement of integration and segregation mechanisms found at higher levels of visual cortex, and would result in greater activation of V5. Using this global motion stimulus, Chen et al. tested a group of schizophrenic participants and found that, contrary to previous studies, the influence of the surround was stronger in schizophrenics (Chen et al., 2008) not weaker, suggesting that these patients have weaker inhibitory mechanisms in early visual cortex, not V5.

The data from migraineurs are also difficult to explain purely on the basis of centre surround interactions in V5. Migraineurs show stronger than normal spatial suppression, rather than the weaker suppression found in depressives and schizophrenics (Battista et al., 2010). This is unexpected considering that one of the dominant theories of migraine proposes that it results from cortical hyperexcitability caused by reduced cortical inhibition, a theory for which there is some physiological (Aurora et al., 2005; Chadaide et al., 2007; Brighina et al., 2009) and psychophysical (Palmer et al., 2000) support. The psychophysical results of Battista et al. using the spatial suppression paradigm (Tadin et al., 2003) would therefore seem to be at odds with the reduced inhibition model of migraine. This contradiction could be resolved if migraine resulted from primary neural hyperexcitability, but it is unclear whether this is the case (Aurora & Wilkinson, 2007; Coppola et al., 2007).

With regard to the weaker spatial suppression reported in older observers (Betts et al., 2005), subsequent studies have failed to replicate this effect (Karas & McKendrick, 2011) and other studies, again using stimuli designed to selectively target V5, have concluded that any motion deficits in older observers are primarily a result of contrast sensitivity loss (Allen et al., 2010). Intrigued by the counterintuitive idea that older observers were performing better than their younger counterparts, we carried out a series of experiments in which we reproduced a “suppressive” effect in younger observers very similar to that of Tadin et al., and we also showed that this effect was absent in older observers, akin to the study of Betts
et al. (Aaen-Stockdale et al., 2009). However, we also obtained contrast thresholds for all observers at all stimulus sizes and calculated the suprathreshold contrast for each stimulus. In this analysis, we found that the “suppressive” effect (and its absence in older observers) was entirely predictable from the observer’s contrast threshold. This explanation of psychophysical spatial suppression based on low-level visual mechanisms has however been disputed (Glasser & Tadin, 2010).

The prevailing interpretation of psychophysical spatial suppression rests upon the idea that surround-inhibition is weaker at low contrasts. However, induced motion, an illusion in which the motion of a central patch is biased by the motion of its surround – a phenomenon which is almost certainly mediated by centre-surround antagonism – is stronger at low contrasts rather than weaker (Hanada, 2010). On the balance of it, whether centre-surround antagonism in V5 is directly to blame for paradoxical psychophysical results or whether other factors are responsible or contribute to the effect remains unresolved.

2.2.2. Perceptual learning

A number of psychophysical studies have demonstrated that specific aspects of motion perception can improve with repeated exposure, a phenomena known as perceptual learning (Fine & Jacobs, 2002). For example, repeated practise of a task that involves discrimination of the motion direction of a field of moving dots, results in significant improvements in task performance (Ball & Sekuler, 1982; Ball & Sekuler, 1987). The fact that such improvements are often highly specific for particular aspects of the trained stimulus such as motion direction and location within the visual field led to the suggestion that learning, and the associated neural plasticity, takes place at a relatively early stage of visual motion processing such as MT. There is additional evidence supporting the idea that MT plays a causal role in perceptual learning of motion tasks. Lesions of MT in monkeys result in an inability to demonstrate perceptual learning for tasks involving the detection of a coherent motion within a random dot kinematogram (Rudolph & Pasternak, 1999). This particular stimulus consists of two populations of moving dots, one moving in a coherent (signal) direction and the other moving in random (noise) directions. The task is to identify the signal direction and task difficulty is manipulated by varying the signal to noise ratio within the stimulus (Newsome & Pare, 1988). In addition, it has been demonstrated psychophysically that perceptual learning of a challenging motion orientation discrimination task is impaired or absent when the ability of MT neurons to encode the motion signal is compromised (Lu et al., 2004). This was achieved by constraining the local motion of pairs of dots within the training stimulus to be equal and opposite. The theory was that this would activate suppressive motion opponent mechanisms within MT (Qian & Andersen, 1994), which would interfere with the global processing of the stimulus and disrupt perceptual learning. In support of this concept, it was subsequently found that while the motion opponent stimulus impaired learning, simply removing the motion opponency from the stimulus by altering the phase of local dot motions resulted in pronounced perceptual learning (Thompson & Liu, 2006). It has also been reported that patients who have hemianopia due to V1 lesions that do not extend to MT are able to learn a motion
coherence task in their blind hemifield (Huxlin et al., 2009). This raises the intriguing possibility that V1 may not be necessary for perceptual learning of specific types of motion stimuli.

It would appear, therefore, that MT is centrally involved in perceptual learning of motion-based tasks. However, the available neurophysiological data, collected using random dot kinematogram stimuli, do not directly support the hypothesis that perceptual learning leads to long-term plastic changes within MT. For example, Zohary et al (1994), found that neurons within MT and MST narrowed their directional tuning and increased their firing rate as monkeys improved their performance on a motion coherence task within a single training session. However these changes did not persist across multiple training sessions. More recent neurophysiological work, also in monkeys, has implicated the lateral intraparietal area in perceptual learning of coherent motion perception (Law & Gold, 2008). This suggests that perceptual learning of specific types of motion stimuli may rely on changes in the way that the responses of cells within MT are ‘read out’ by higher level extrastriate areas. Whether this is the case for the human brain and for other types of motion tasks is yet to be established.

2.3. Complex motion analysis – the middle superior temporal area (MST) and V6

Still higher cortical areas respond to complex motion signals such as global expansion, contraction and rotation. These types of motion are particularly interesting, because they are generated by the interaction between an observer and the environment. For example, radial patterns of motion such as expansion and contraction occur on the retina when objects approach or recede from an observer, respectively. These patterns of motion could be caused by motion of the object, the head and body or both. Similarly, rotational patterns of motion can be caused either by tilting of the head or physical rotation of an object. In other words, neurons selective for these motion patterns could encode “optic flow” and allow us to navigate in the world (Gibson, 1950).

Physiological work has shown that the medial superior temporal area, MST (Saito et al., 1986; Tanaka et al., 1986; Tanaka et al., 1989; Tanaka & Saito, 1989; Duffy & Wurtz, 1991b; Duffy & Wurtz, 1991a; Britten & van Wezel, 1998) is selective for such translational, radial and rotational motion patterns over wide areas of the visual field. Neuroimaging has demonstrated the presence of a human homologue of MST in what has become known as the hMT+ complex or V5, which also includes area MT (Morrone et al., 2000; Dukelow et al., 2001; Huk et al., 2002). The extraction of patterns of complex motion such as radial and rotational motion, has been modelled from the combined outputs of MT neurons (Perrone, 1992; Grossberg et al., 1999). Recently, a second visual area responsive to wide-angle flow fields has been identified (Pitzalis et al., 2010) and is thought to be a homologue of macaque area V6 (Galletti et al., 1996). Although most work on optic flow has concentrated on area MST, neurons in this second region have many similar characteristics and the two areas are strongly connected. Pitzalis et al. propose that MST and V6 work in concert, the former analysing the motions of objects in the world and the latter extracting self-motion.
Whether MST neurons are responsive to only the cardinal motion directions (radial, rotational and translational) as suggested by some psychophysical work (Morrone et al., 1999; Burr et al., 2001), or whether other intermediate forms of motion such as spiral motion are encoded directly, is still a matter of some debate. In support of the direct detection of spiral motions, it has been suggested that summation of mechanisms tuned to purely cardinal motion directions is insufficient to explain the psychophysical data (Snowden & Milne, 1996; Meese & Harris, 2001; Meese & Anderson, 2002) and some physiological work seems to have identified neurons tuned to spiral motions (Graziano et al., 1994; Geesaman & Andersen, 1996). A particularly interesting study found that continuously-changing flow stimuli, obtained by morphing one flow field into the next, lead to a continuum of responses in MST (Paolini et al., 2000). This supports the “generalised spiral” model of MST, in which the tuning of MST neurons is a continuum from pure contraction, though clockwise-contraction, to clockwise-rotation, to clockwise-expansion, to pure expansion, and so on.

2.4. Biological motion perception – the superior temporal sulcus (STS)

Human observers are acutely sensitive to the complex pattern of motion trajectories generated by other people and animals known as biological motion (Johansson, 1973; Mather & West, 1993). Investigations of biological motion often use stimuli constructed from dots or “point lights” that represent the joints of an actor (Troje, 2002). When stationary, these displays appear as an elongated group of dots, however when set in motion, a vivid percept of a person moving is generated. Sufficient information can be extracted from dynamic point light stimuli to allow for identification of a wide range of complex attributes such as gender (Mather & Murdoch, 1994) and mood (Dittrich et al., 1996), and observers are sensitive to these stimuli in both the central and peripheral visual field (Thompson et al., 2007). Although point light stimuli contain both motion information and configural, form-based cues (the relative position of the points), motion cues do appear to be centrally involved in the processing of biological motion. For example, observers can still perform above chance on a walking direction discriminating task when the relative positions of each of the dots representing the joints of an actor are scrambled (Troje & Westhoff, 2006). Interestingly, simply inverting the dots representing the feet of walking humans or animals disrupts discrimination of walking direction suggesting that the characteristic local motion cues specific to biological motion, may be processed by dedicated “life detector” mechanisms (Troje & Westhoff, 2006).

Initial insights into the regions of the visual cortex involved in biological motion perception were provided by the neurophysiological investigations of Oram and Perrett (1994) in the monkey. Cells were found within the superior temporal polysensory area, a region anterior to MT and MST within the superior temporal sulcus, that were sensitive to biological motion stimuli. Subsequently, a large number of neuroimaging studies have been conducted in humans with the aim of identifying the cortical areas involved in biological motion perception. It is now apparent that biological motion perception recruits a distributed neural circuit in humans which includes the posterior region of the superior temporal sulcus (Grossman et al., 2000; Grezes et al., 2001; Servos et al., 2002; Grossman et al., 2005; Pelphrey
et al., 2005; Peelen et al., 2006) and may also involve “mirror neurons” in the ventral pre-motor cortex (Saygin et al., 2004) along with a range of additional visual areas including the posterior middle temporal gyrus and regions known as the extrastriate and fusiform body areas (Jastorff & Orban, 2009). A recent meta-analysis of neuroimaging data in humans has emphasised the importance of the pSTS in processing motion cues from biological motion stimuli and also identified a region within the hMT+ complex that may play a role in the perception of human body movement (Grosbras et al., 2012).

2.5. Structure-from-motion – the lateral occipital sulcus (LOS) and the intraparietal sulcus (IPS)

As well as being able to extract biologically relevant information from motion patterns, the visual system is also able to extract three-dimensional (3D) structure from moving two-dimensional stimuli, often referred to as the “kinetic depth effect” or “structure-from-motion” (SFM) (Wallach & O’Connell, 1953). Several models have been developed to explain how the visual system achieves this remarkable feat, some based on the tracking of local positional cues (Ullman, 1984; Grzywacz & Hildreth, 1987; Shariat & Price, 1990; Snowden et al., 1991) and others based on the use of local motion information (Clocksin, 1980; Longuet-Higgins & Prazdny, 1980; Koenderink & van Doorn, 1986; Husain et al., 1989; Hildreth et al., 1995). The weight of evidence currently seems to suggest that motion rather than positional information is crucial to extracting SFM (Andersen & Bradley, 1998; Farivar, 2009; Farivar et al., 2009).

SFM is usually investigated by using random dot stimuli in which the dots are randomly distributed across the image, but are projected onto an “invisible” object, which is then rotated with the direction and speed of the dots being dictated by their position on the object. Provided that the dot density is high enough, SFM can be perceived with very short, two-frame displays in which dots simply jump from one position to another (Lappin et al., 1980). SFM is not perceived if a very small number of dots are used. However, periodically re-positioning a similarly small number of dots across the stimulus, allows the surface of the object to be reconstructed via interpolation (Husain et al., 1989; Treue et al., 1995). Functional magnetic resonance imaging (fMRI) suggests that SFM is carried out in a network of cortical regions: V5, lateral occipital sulcus (LOS) and several sites along the intraparietal sulcus (IPS) (Orban et al., 1999; Peuskens et al., 2004).

3. Second-order motion

The preceding discussion has dealt mainly with the processing of luminance-based motion, often called first-order motion. However, at the input stage, motion can also be defined by characteristics other than luminance, such as flicker, texture and contrast. To return to our snowfield example, the force and direction of the wind or the presence of a small animal burrowing under the snow can be detected visually by detecting changes in the pattern of random flicker caused as flakes of snow on the ground are disturbed. Motion that is defined by modulation of a property other than luminance is referred to as second-order motion.
(Cavanagh & Mather, 1989) and second-order motion is invisible to first-order, i.e. luminance-based, motion sensors (Chubb & Sperling, 1988).

### 3.1. Local second-order motion

Currently, it is a mystery how the visual system detects second-order motion, as the primary input to the visual system represents changes in retinal illumination. The visual system has a small compressive non-linearity, probably at the level of the photoreceptors (Scott-Samuel & Georgeson, 1999) that could transform second-order information into a weak luminance signal. This weak internal artefact could mean that second-order motion is actually detected by first-order mechanisms and could explain the (usually) weaker performance for purely second-order motion stimuli. However, the distortion product measured by Scott-Samuel & Georgeson is only detectable in high speed, high contrast modulation stimuli. Since second-order motion is still visible in slow moving, or low-modulation, stimuli, the early non-linearity is unlikely to be able to fully explain many of the experimental results. Figure 3 shows a schematic wiring diagram of the motion processing hierarchy and the dotted orange arrow shows the presence of these “pseudo-second-order” motion signals.

**Figure 3.** A schematic view of the orthodox model of motion processing in the cortex. FO = first order, SO = second order, MT = middle temporal visual area, MST middle superior temporal visual area, STS = superior temporal sulcus, IPS = intraparietal sulcus, LOS = lateral occipital sulcus. See text for further details.
If second-order motion was ultimately detected by first-order mechanisms, we might expect the two types of stimuli to interact. This does not seem to be the case for local motion. Temporally interleaving first- and second-order stimuli in alternate frames of a motion stimulus fails to generate the percept of smooth motion, suggesting that the two systems do not interact at this level (Ledgeway & Smith, 1994). Adaptation to one type of motion does not impair detection of the other type (Nishida et al., 1997), and the second-order system does not seem to be able to discriminate the direction of motion at detection threshold, unlike the first-order system (Smith & Ledgeway, 1997), which can discriminate motion direction as soon as motion is detected. It is therefore likely that first- and second-order motion are initially analysed in parallel by separate processing streams (shown in figure 3 in red and blue, respectively).

Several neuroimaging studies have investigated the possibility that some cortical areas may be selective for second order motion, as predicted by the dual-pathway hypothesis. To date, however, the results from these studies have been mixed. Second-order specific responses have been reported in areas such as V3 (Smith et al., 1998), but other studies have found either substantial overlap of first and second-order motion responses throughout the visual cortex (Dumoulin et al., 2003) or no anatomical segregation of areas responsive to first- and second-order motion (Nishida et al., 2003; Seiffert et al., 2003; Ashida et al., 2007). The idea of an anatomically distinct second-order pathway may be rescued if different neurons in the same anatomical area responded selectively to different types of motion, or if the same neurons responded to both first- and second-order motion, but had different spatial/temporal tuning for first-order motion than for second-order motion. This latter contention is supported by some neurophysiological investigations of MT in the primate (O’Keefe & Movshon, 1998) and areas 17 and 18 in the cat (Mareschal & Baker, 1998). However, the idea that different neuronal populations for first- and second-order motion exist, but share common anatomical locations has also found support in a human neuroimaging study using fMRI adaptation (Ashida et al., 2007). In this technique, repeated presentation of similar stimuli causes a reduction of the blood oxygen level–dependent (BOLD) response in cortical regions containing neurons that cannot differentiate between the stimuli, whereas little or no reduction occurs if the different stimuli activate distinct populations of neurons. Ashida et al. found direction-selective fMRI adaptation for stimuli of the same type, but no cross-adaptation between first- and second-order motion. These fMRI results provide persuasive evidence that neural populations differentially selective for first- and second-order motion co-exist in motion sensitive regions of the human brain such as V3a, MT and MST.

### 3.2. Higher level second-order motion

Assuming segregation of first- and second-order motion at early stages of visual motion processing, at what point in the visual motion hierarchy are the two types of motion combined? Models of the mammalian visual motion processing hierarchy (Wilson et al., 1992; Lu & Sperling, 1995; Lu & Sperling, 2001) usually integrate first- and second-order streams at, or before, the level of global motion analysis (see Figure 3) and insensitivity to
such low-level stimulus characteristics, termed “cue-invariance”, has been found in neurons at the level of MT (Albright, 1992; O’Keefe & Movshon, 1998) and MST (Geesaman & Andersen, 1996). The argument that extrastriate areas are cue-invariant has also been supported by a TMS study in humans (Cowey et al., 2006).

There is, however, plenty of counter-evidence to the presence of cue-invariance at MT and MST in the physiological literature (Olavarri et al., 1992; Churan & Ilg, 2001) and some human psychophysical studies have suggested that the two processing streams could remain segregated at even higher levels than MT. For example, Badcock and colleagues have shown that the addition of second-order noise dots to either a first-order global motion stimulus or complex motion stimulus does not impair detection and discrimination of the first-order signal (Edwards & Badcock, 1995; Badcock & Khuu, 2001; Cassanello et al., 2011). Several studies have suggested that the contribution of the second-order pathway to mechanisms responsible for extraction of structure-from-motion is weak or non-existent (Dosher et al., 1989; Mather, 1989; Landy et al., 1991; Hess & Ziegler, 2000), while research on whether second-order motion can support biological motion has produced differing results (Mather et al., 1992; Ahlstrom et al., 1997; Bellefeuille & Faubert, 1998).

Although these findings are certainly of interest, it seems counterintuitive that the visual system would maintain two separate pathways for the analysis of different types of high-level motion stimuli that differ only in terms of their local characteristics. Some recent studies by ourselves and our collaborators support the conventional concept of a functional integration of first- and second-order motion at higher levels of the motion hierarchy. Ledgeway, et al. (2002) and Aaen-Stockdale et al. (forthcoming) have argued that the relative visibility of the first- and second-order dots in the stimuli used by Badcock and colleagues may not have been equalised (Edwards & Badcock, 1995; Badcock & Khuu, 2001). Although the static first- and second-order dots were highly visible, their relative visibility to first- or second-order motion sensors (which have different spatio-temporal characteristics) (Derrington et al., 1993; Ledgeway & Hess, 2002) may not have been equal. Ledgeway, Hess & McGraw (2002) demonstrated that reducing the visibility of the luminance-modulated (first-order) dots resulted in interactions between first- and second-order dots in global motion stimuli consistent with a combination of the two motion cues within extrastriate visual areas. Aaen-Stockdale et al. (forthcoming) showed that these visibility-dependent interactions also occurred with complex (radial and rotational) motion stimuli and that weakening the first-order signal by increasing the size of dot displacements between frames resulted in similar interactions. This latter study also pitted opposing first- and second-order motion signals against each other to demonstrate that impairments of first-order motion discrimination caused by the inclusion of second-order dots within the stimulus was genuinely a result of a cue-invariant motion system attempting to integrate these separate signals, and not simply the result of increased noise. Subsequent work by us has used these same techniques to show that other types of higher-order motion perception are similarly cue-invariant. For example we (Aaen-Stockdale, et al., 2008) found that when relative stimulus visibility was varied, first- and second-order elements interacted to mask biological motion of the opposite class suggesting that biological motion perception is cue
invariant. Similarly, it had been proposed that the contribution of the second-order pathway to structure-from-motion mechanisms was weak or non-existent (Dosher et al., 1989; Mather, 1989; Landy et al., 1991; Hess & Ziegler, 2000), but reducing the relative visibility of the first-order elements in structure-from-motion displays results in almost linear summation between first- and second-order mechanisms, suggesting that this modality may also be cue-invariant (Aaen-Stockdale et al., 2010). These findings have highlighted the importance of ensuring that local first-order and second-order motion signals are of equal strength when comparing the two systems.

4. Abnormalities of motion processing

4.1. Akinetopsia

The first widely accepted report of akinetopsia or “motion blindness” due to bilateral lesions affecting the hMT+ complex but sparing the primary visual cortex, was reported by Zihl and colleagues (Zihl et al., 1983; Zeki, 1991). This patient, LM, did not have a scotoma (region of blindness) as would be expected from damage to the primary visual cortex; however LM exhibited a severe and selective impairment of motion perception. Subsequent work with this patient indicated that LM could perceive motion direction and structure from motion under certain circumstances as long as the stimuli did not contain noise elements such as static or randomly moving dots. However, as soon as noise was introduced into the stimulus, task performance was dramatically reduced (Baker et al., 1991; Rizzo et al., 1995). This pattern of deficits is similar to that reported by Rudolph and Pasternak (Rudolph & Pasternak, 1999) who studied the effects of MT lesions in macaque monkeys. Immediately after the lesions the animals exhibited a pronounced and general impairment in motion perception. However, over time, performance on motion tasks that did not include noise elements recovered, whereas performance for tasks requiring signal/noise segregation, such as those involving RDKs, did not recover. As a whole, these data from both humans and primates suggest that MT and MST may have a particular specialization for the detection of motion in a noisy environment.

4.2. A motion deficit in amblyopia?

Deficits in motion perception have also been reported for patients with a developmental disorder of the visual cortex known as amblyopia (or “lazy eye”). Unilateral amblyopia occurs when the images seen by each eye are poorly correlated during early visual development due to a chronically blurred image in one eye (anisometropia), a turned eye (strabismus) or less commonly a congenital cataract (Holmes & Clarke, 2006). This can result in abnormal development of the visual cortex and a visual impairment in the affected eye that is not due to a problem with eye itself, but is the result of abnormal processing of inputs from the amblyopic eye within the visual cortex (Hubel & Wiesel, 1965; Barnes et al., 2001) and possibly the lateral geniculate nucleus (Hess et al., 2009; Li et al., 2011).
Although amblyopia is typically regarded as a disorder of spatial vision, a number of studies have identified deficits in motion perception that appear to be independent of impairments of form perception (see Thompson et al., 2011 for a recent overview). Many of the motion deficits that have been reported are consistent with abnormalities at the level of hMT+ complex. For example, patients with amblyopia exhibit elevated motion coherence thresholds when viewing random dot kinematograms, even when abnormal contrast sensitivity is taken into account (Simmers et al., 2003; Constantinescu et al., 2005; Simmers et al., 2006). These deficits are present across spatial scale (Aaen-Stockdale & Hess, 2008), include both first- and second-order motion stimuli (Aaen-Stockdale et al., 2007), may involve deficits in spatial summation (Thompson et al., 2011) and appear to be related to the segregation of signal from noise (Mansouri & Hess, 2006; Thompson et al., 2008b). In agreement with these psychophysical findings, it has recently been reported that cells within MT of monkeys made experimentally strabismic are less direction selective and less tolerant of noise (El-Shamayleh et al., 2011). It would appear, therefore, that motion sensitive areas such as MT are susceptible to abnormal sensory experience during development and that this can result in specific deficits in motion perception.

In this context we were surprised to find that patients with amblyopia did not show pronounced impairments in the perception of coherent plaid stimuli (Thompson et al., 2008a). As described above, plaid stimuli have been used extensively to investigate the integration of local motion signals within the visual cortex and cells have been found within MT that respond selectively to the integrated “pattern motion” direction of coherent plaids. We therefore expected to find reduced levels of coherent motion perception in patients with amblyopia as would be predicted by deficient processing within MT. In contrast to this we found that within the small region of the parameter space where amblyopes did exhibit differences from controls, these differences were characterised by an increased probability of coherent motion perception (Thompson et al., 2008a). Since a similar change in plaid perception can be induced in the normal visual system when inhibitory rTMS is delivered over V1 (Thompson et al., 2009) (figure 2), the differences between amblyopic observers and controls in plaid perception may be due to abnormal processing within V1 rather than MT.

This seemingly anomalous result has recently been further explored using fMRI (Thompson et al., 2012). Coherent and incoherent plaid stimuli that were perceived in exactly the same way by control observers and amblyopes, activated distinct networks of brain areas when the plaids were viewed by non-amblyopic eyes vs. amblyopic eyes. For controls and patients viewing through their non-amblyopic eye, the hMT+ complex was differentially activated by coherent and incoherent plaids consistent with previous fMRI studies (Castelo-Branco et al., 2002) and the idea that this area is centrally involved in motion integration. However this was not the case for amblyopic eye viewing for which there appeared to be a selective loss of response within hMT+ to incoherent plaid motion. In patients for whom the hMT+ complex could be subdivided into MT and MST, this loss was apparent in both areas. It would appear, therefore, that areas other than hMT+ were involved in the normal perception of plaid patterns for amblyopic eye viewing. The fMRI data provided preliminary evidence for a preserved response to incoherent motion within the pulvinar complex of the patients with
amblyopia. This area has previously been shown to be involved in the processing of pattern motion (Merabet et al., 1998; Villeneuve et al., 2005) and has extensive reciprocal connections with extrastriate brain areas including MT (Casanova, 2004). These results in humans with amblyopia raise the intriguing possibility that alternate brain areas may play a role in motion perception if the function of hMT+ is sub-optimal or compromised. If this were to be the case, the current psychophysical data from patients with amblyopia and both humans and monkeys with MT lesions (described above) imply that this compensatory processing is highly sensitive to noise such as the random dots present in RDK stimuli.

5. Conclusion

The cortical mechanisms involved in the perception of visual motion represent one of the most widely studied and well-understood processes in neuroscience, however there are still many questions left to answer. Over the last few decades a picture has emerged of a motion processing system that is rigidly hierarchical, yet possesses considerable redundancy, and plasticity. In general terms, progressively higher-level areas of the brain integrate the outputs of areas below them in order to detect and discriminate increasingly complicated stimuli. However, with apologies to the Reverend William Paley, the visual brain is not an organ that has been designed, but one that has evolved over the millennia, and which demonstrates all the adaptations and redundancies that implies. A very much open question relates to the evolutionary origins of specialised forms of motion perception, such as second-order motion, structure-from-motion or biological motion; assuming that there actually are dedicated mechanisms for these type of motion, a question which is by no means settled.

Author details

Craig Aaen-Stockdale
Department of Optometry and Visual Science, Buskerud University College, Norway

Benjamin Thompson
Department of Optometry and Vision Science, University of Auckland, New Zealand

Acknowledgement

Craig Aaen-Stockdale is supported by an Yggdrasil international mobility grant from the Research Council of Norway (Forskningsrådet). Benjamin Thompson is supported by grants from the Health Research Council of New Zealand, the Marsden Fund, the University of Auckland and the Neurological Foundation of New Zealand.

6. References


