Visual motion and the human brain: what has neuroimaging told us?

Jody Culham a,*, Sheng He b,1, Sean Dukelow c,2, Frans A.J. Verstraten d,3

a Department of Psychology, University of Western Ontario, London, Ont., Canada N6A 5C2
b Department of Psychology, University of Minnesota, 75 E. River Rd., Minneapolis, MN 55455, USA
c Department of Physiology, University of Western Ontario, London, Ont., Canada N6A 5C1
d Department of Psychology/Psychonometrics division, Universiteit Utrecht, Heidelbergaan 2, NL-3584 CS
Utrecht, The Netherlands

Received 3 August 2000; received in revised form 11 December 2000; accepted 18 December 2000

Abstract

Recently, neuroimaging techniques have been applied to the study of human motion perception, complementing established techniques such as psychophysics, neurophysiology and neuropsychology. Because vision, particularly motion perception, has been studied relatively extensively, it provides an interesting case study to examine the contributions and limitations of neuroimaging to cognitive neuroscience. We suggest that in the domain of motion perception neuroimaging has: (1) revealed an extensive network of motion areas throughout the human brain, in addition to the well-studied motion complex (MT+); (2) verified and extended findings from other techniques; (3) suggested extensive top-down influences on motion perception; and (4) allowed experimenters to examine the neural correlates of awareness. We discuss these contributions, along with limitations and future directions for the neuroimaging of motion. © 2001 Elsevier Science B.V. All rights reserved.

PsycINFO classification: 2323; 2540
Keywords: Motion; Vision; Neuroimaging; fMRI; MT; Cognitive neuroscience

1. Introduction and historical overview

Interest in motion perception has a long history. Several great philosophers, like Euclid, Lucretius, Ptolemy, Ibn al Haytham, have commented on different types of motion perception (e.g., apparent motion, induced motion, aftereffects of motion, etc.; see Wade, 1998). This interest is not at all surprising given that it is an extremely important faculty of the human mind. Life becomes incredibly difficult without the ability to see motion, for example as a result of a brain lesion (Riddoch, 1917; Zihl, Von Cramon, & Mai, 1983).

Experimental research on motion perception, on the other hand, has a rather short history. Without minimizing the impact of earlier work, one could argue that the modern study of motion perception started with Exner (1888). He argued that motion is not an inference - something was there, now it is here, so it must have moved – but rather a perceptual quality in its own right. Another landmark discovery in motion perception came from the late Werner Reichardt (1961). He showed that two interacting ommatidia in the eye of a fly can code for space and time and therefore for motion. His model is now known as the Reichardt-detector, upon which most motion detection models are based.

Before the advent of neuroimaging techniques, unraveling the processes and structures underlying motion perception came from two main research areas, visual psychophysics and electrophysiology. The shortcoming with the former was that one could only speculate about the brain areas involved in motion perception. The problems for electrophysiology were also not trivial. Can the activity of a single cell or a small sample of the neuronal population tell us how motion is encoded within a network? It did tell us which areas have cells that are responsive to moving stimuli, but it focussed only on a few established motion areas, making it hard to obtain a global picture of the brain’s response to motion. The most difficult restriction with neurophysiology is probably the inherent differences between monkeys and humans. Although the monkey and human brain have many things in common, we were never sure that we could generalize from the results from the monkey to humans. Motion perception is a very active process. Although there are ingenious ways to train monkeys to perform complex tasks, questions related to attention and perceptual experience (visual awareness) are still best answered with human observers.

Modern neuroimaging techniques, functional magnetic resonance imaging (fMRI) in particular, provide some answers to questions that are not possible to answer with psychophysics and single-unit neurophysiology alone. It became possible to study human motion perception with all the psychophysical manipulations while simultaneously monitoring the observer’s cortical activity. In this paper we consider several motion phenomena to illustrate how neuroimaging has begun to both extend electrophysiological findings to the human brain and to ask new
questions about motion processing areas in the human brain that were either previously unexplored or are functionally different from those in the monkey.

2. What was known about motion perception prior to neuroimaging?

Over the last three decades, much effort went into detailed parametric studies of motion sensitivity. To study motion-sensitive mechanisms, psychophysicists focussed on performance indicators, such as motion sensitivity, direction discrimination, and speed discrimination. These studies primarily focussed on motion arising from changes in luminance over space and time, the stimuli which drive motion detectors (Reichardt, 1961). However, evidence has accumulated to suggest that observers can see various kinds of motions that cannot be explained by simple detectors. Examples include “second-order” motion (Chubb & Sperling, 1988; Derrington & Badcock, 1985; Lelkens & Koenderink, 1984), and attention-based motion perception (Cavanagh, 1992). Psychophysical experiments have demonstrated that additional mechanisms must exist and the relations among various motion mechanisms have been postulated (e.g., Lu & Sperling, 1995; Nishida & Ashida, 2000). However, behavioral techniques only allow speculation regarding the underlying neural substrates.

Physiological techniques have revealed a number of brain areas which are responsive to motion. Although motion-selective cells have been identified as early as the retina in species such as the rabbit (Barlow & Hill, 1963), in primates, direction-selective cells are observed only in primary visual cortex, V1 (Hubel & Wiesel, 1977), and later stages. Beginning in the lateral geniculate nucleus of the thalamus, neurons in the magnocellular pathway have faster and more transient responses (compared to those in the parvocellular pathway), allowing them to detect and encode movement in the later stages of the pathway. Within magno-recipient visual areas, there are progressively more direction-selective cells from area V1 to V3 (Felleman & Van Essen, 1987; Gegenfurtner, Kiper, & Fensstemaker, 1996; Gegenfurtner, Kiper, & Levitt, 1997; Hubel, 1988). The most well-studied areas in the neurophysiology of motion are MT/V5 (middle temporal area) and adjacent area MST (medial superior temporal), both in the superior temporal sulcus of the macaque. Almost all MT neurons exhibit directionally selective responses (Lagae, Maes, Raiguel, Xiao, & Orban, 1994) and MT lesions impair motion processing (Newsome & Paré, 1988; Newsome, Wurtz, Dursteler, & Mikami, 1985). Recordings from MT show a close correlation between the performance of a monkey on a direction discrimination task and the activity of motion-selective cells in MT (Newsome, Britten, & Movshon, 1989). Furthermore, microstimulation of a directionally selective column can bias a monkey’s choice toward this direction (Salzman, Britten, & Newsome, 1990). Area MT projects directly to MST, which also contains a high proportion of directionally selective cells (Lagae et al., 1994) with more complex properties than those in MT, responding for example to rotation and expansion/contraction (Saito et al., 1986). Microstimulation in area MST can bias a monkey’s perceived direction of heading (Britten & van Wezel, 1998). MT also sends projections to other “motion” areas:
FST (in the fundus of the superior temporal sulcus), VIP (in the fundus of the intraparietal sulcus), and STP in the superior temporal sulcus (Boussaoud, Ungerleider, & Desimone, 1990; Ungerleider & Desimone, 1986). Physiology has identified motion-selective cells in numerous areas, but has largely focussed on the best understood areas (MT and MST) and stimuli (luminance-based motion energy).

It is clear that psychophysics has characterized many aspects of human performance related to motion perception and neurophysiology has provided a large amount of data on the neural responses to motion in non-human primate brain. Neuroimaging has the potential to investigate both neural activity and behavior simultaneously in humans and, in conjunction with the established techniques, should enhance our understanding of motion.

3. What has neuroimaging added to our understanding of motion perception?

Neuroimaging techniques with good spatial resolution became available to a few groups in the 1980s and more widespread in the 1990s (for a review of the history of neuroimaging, see Savoy, 2001). Positron emission tomography (PET) allowed scientists to get a first glimpse of the human brain in action, and vision was one of the first subjects tackled (Fox et al., 1986; Roland, 1993; Zeki et al., 1991). In 1992, it became possible to use fMRI to image the intrinsic blood oxygen level-dependent (BOLD) signal reflecting neural activity (Kwong et al., 1992; Ogawa et al., 1992). Again, vision was one of the first subjects addressed (for reviews see Courtney & Ungerleider, 1997; Op de Beeck, Wagemans, & Vogels, 2001). At the beginning, imaging studies merely confirmed activation in visual cortex during visual stimulation. More recently, however, researchers have begun to ask theoretically motivated questions about the underlying mechanisms, in addition to simply “mapping” cortex. Below we review some of the findings from neuroimaging research that we believe have added to our understanding of human motion perception.

3.1. Neuroimaging has identified numerous motion-sensitive areas throughout the brain

To understand cortical motion mechanisms, one first needs to know what areas are involved. Perhaps one of the most salient findings of the neuroimaging of visual motion has been that many regions are more active when subjects view moving compared to stationary stimuli (see especially, Dupont, Orban, De Bruyn, Verbreggen, & Mortelmans, 1994; Sunaert, Van Hecke, Marchal, & Orban, 1999; Van Essen & Drury, 1997). For example, Fig. 1 shows a number of motion-responsive areas throughout the brain. The sheer number of regions is surprising, since much of electrophysiology has emphasized only one or two “motion areas” (namely MT and MST). One can get the sense in reviewing the neuroimaging literature that the entire visual brain is a “motion area”. However, given the range of functions that motion processing subserves (Nakayama, 1985; Smith & Snowden, 1994), it is not surprising that many different areas are involved in motion processing, and it is likely that these
areas participate in a diverse range of motion-related cognitive functions. The various areas, however, respond to different tasks and motion stimuli.

3.1.1. MT+ / V5

The first human motion area identified by neuroimaging was area V5, discovered by Zeki and colleagues in 1991 (Zeki et al., 1991). They used PET to localize an area at the temporo-parieto-occipital junction that had higher regional cerebral blood flow (rCBF) when observers viewed moving compared to stationary random squares. V5 is believed to be the human homolog of monkey areas MT, MST and adjacent motion selective cortex. Thus, it is frequently referred to as “the human MT complex”, “hMT+”, “MT+”, or sometimes just “MT” for short. Watson et al. (1993) further demonstrated that although MT+ could vary significantly in its stereotaxic location (Talairach & Tournoux, 1988), it was consistently located at the junction of the ascending limb of the inferior temporal sulcus and lateral occipital sulcus (Fig. 1, see also Dumoulin et al., 2000). MT+ is one of the most easily and consistently activated regions in human cortex, responding to a wide range of dynamic stimuli.
(e.g., flickering checkerboards, moving gratings, moving objects, moving dot patterns) in almost all observers, making it a prime target for neuroimaging research.

Motion responses can be generated by various stimulus types. Tootell and colleagues (Tootell, Reppas, Kwong et al., 1995) used fMRI to evaluate the response of MT+ to motion. Consistent with its place in the magnocellular stream, they showed that MT had much higher contrast sensitivity than V1 and that its activity was lowest (though still present) for chromatic stimuli at subjective equiluminance. They also demonstrated that MT was strongly driven by moving gratings, moving dots or incoherent flicker. Together, these findings confirm results from other techniques suggesting that MT is driven most strongly by luminance motion energy. This kind of motion is called first-order motion (Cavanagh & Mather, 1989). There are several other ways to generate a clear percept of motion, either in the case where luminance is balanced for opposite directions or when luminance information is absent, as in second-order motion (Cavanagh & Mather, 1989; Chubb & Sperling, 1988; Derrington & Badcock, 1985). Later in this paper, we consider areas that are sensitive to the second-order motion.

Recently, fMRI data from several imaging groups have suggested that it is possible to distinguish subregions within the MT+ complex. Given that in the macaque, MST cells have much larger receptive fields than MT cells (e.g., Raiguel et al., 1997), one would expect MT to respond only to stimuli within the contralateral hemifield, whereas, MST should respond to stimulation of the ipsilateral hemifield as well. Indeed, neuroimaging data suggest two regions of the MT+ cluster, a more posterior one with only a contralateral response and a more anterior one that responds to both contra- and ipsi-lateral stimulation (Dukelow et al., 2000, submitted; Khan et al., 1999). Responses within the MT+ complex also depend on the composition of the stimuli. Specifically, optic flow stimuli, such as rotation or expansion, produce activation in a region that is ventral to the one that responds to translation (although it appears essential that the flow pattern be continually changing; Morrone et al., 2000).

3.1.2. V3A

Retinotopic mapping has identified a dorsal extrastriate area V3A ("V3 accessory") which was later shown to be motion-selective (Tootell et al., 1997). As described elsewhere in this issue (Morland, Baseler, Hoffman, Sharpe, & Wandell, 2001), retinotopic mapping techniques use eccentricity and polar angle sweeping stimuli to delineate the borders between visual areas (DeYoe et al., 1995; Engel, Glover, & Wandell, 1997; Sereno et al., 1995). That is, visual areas can be separated by the boundaries between left and right visual fields (at the vertical meridian) or upper and lower visual fields (at the horizontal meridian) and by the fact that each area is mapped in a mirror image fashion to the adjacent areas (e.g., the V2 map is a mirror image of V1 and V3). These techniques have shown the existence of area V3A (Tootell et al., 1997), a region anterior to dorsal V3, near the junction of the intraparietal and transverse occipital sulci (at the superior occipito-parietal boundary). V3A has a representation of the entire contralateral hemifield (unlike lower-tier areas such as V3 which represent only a quadrant) and includes a foveal representation.
that is not contiguous with those of the lower-tier areas. In fact, some have suggested an additional area, V3B (anterior to V3 and lateral to V3A), that represents only the lower visual quadrant (Smith, Greenlee, Singh, Kraemer, & Hennig, 1998).

Interestingly, human V3A is more motion-selective than any other retinotopic areas including V3 (Tootell et al., 1997). This finding was somewhat surprising because in the macaque monkey, area V3 shows considerably more motion selectivity than V3A (Felleman & Van Essen, 1987). Furthermore, the response properties of human V3A appear to have more in common with macaque V3 than V3A (Singh, Smith, & Greenlee, 2000). These results suggest that interspecies homologies may break down for higher-tier visual areas and that fMRI may reveal new regions in human cortex which are unanticipated by monkey neurophysiology. Similar issues are suggested in the ventral visual areas by the controversial debates about V4 and V8 in human and monkey cortex (Hadjikhani, Liu, Dale, Cavanagh, & Tootell, 1998; Zeki, McKeefry, Bartels, & Frackowiak, 1998).

3.1.3. Other retinotopic areas

Motion-selective activation has been less consistent in lower-tier visual areas (e.g., Howard, Brammer, Wright, Woodruff, & Bullmore, 1996; McCarthy et al., 1995). Tootell and colleagues (Tootell et al., 1997; Tootell, Reppas, Dale et al., 1995) suggested that motion responsibility in visual areas could be ranked in the following order: V1 = V3 < VP < V2 < V3A < MT+. Smith et al. (1998) examined first- and second-order motion and found increasing motion specificity in the progression from lower- to higher-tier visual areas (V1–V2–V3/VP–V3B–V3A–MT). V3 and VP seem more selective for second-order motion than for first-order motion, a property that is not observed in MT+ (Smith et al., 1998).

3.1.4. Kinetic occipital cortex

Orban and his colleagues (Dupont et al., 1997; Orban et al., 1995; Van Oostende et al., 1997) have suggested a new area called “kinetic occipital” (KO) that responds to motion-defined contours. They used a “kinetic grating” stimulus in which random dots in bands of the grating moved in opposite directions. KO responded more to the kinetic grating than to either random dots that moved coherently in one direction (Dupont et al., 1997) or to transparently moving dots (Van Oostende et al., 1997). In comparison, MT+ showed a strong response to all motion stimuli regardless of the presence of kinetic boundaries. However, the response of KO to motion contours has not been reliably observed (Reppas, Niyogi, Dale, Sereno, & Tootell, 1997; Shulman, Schwarz, Miezin, & Petersen, 1998; Wang et al., 1999), possibly due to differences in stimulus size (Shulman et al., 1998; Van Oostende et al., 1997). KO is posterior to MT+ and lateral to V3A and is distinct from the other two areas (Van Oostende et al., 1997). KO appears to be adjacent to, and occasionally overlapping with, the lateral occipital area (LO) which is activated by static visual objects and may be involved in extracting surfaces (Mendola, Dale, Fischl, Liu, & Tootell, 1999) and objects (Malach et al., 1995). KO provides an example where neuroimaging has progressed beyond the identification of known homologs expected from neurophysiology to discover new areas with complex responses that would be difficult to
localize using conventional methods. It also provides an example of how neuroimaging can suggest new experiments using other methodologies. Specifically, Nawrot and colleagues (Nawrot, Rizzo, Rockland, & Howard, 2000) recently tested a patient with a lesion to both KO and MT. She had a transient deficit in extracting two-dimensional shape from motion, although she could determine shape from other (non-motion) cues.

3.1.5. Additional motion-sensitive areas

Numerous other motion responsive areas have been reported (Cheng, Fujita, Kanno, Miura, & Tanaka, 1995; Dupont et al., 1994; Sunaert et al., 1999). Although little has been done to characterize their functional properties, in some cases there are suggestive preliminary data, and in others, possible speculations based on related functions in specific brain regions. In particular, it is likely that the division of labor into ventral “what” pathway for recognition and a dorsal “where” or “how” pathway for guiding actions in space (Creem & Proffitt, 2001; Milner & Goodale, 1995; Ungerleider & Mishkin, 1982) can provide hints as to the function of motion-selective regions. MT is usually placed in the dorsal stream and it is common to consider motion processing as a “dorsal function”. However, fMRI studies suggest that motion and form processing each produce activation in both streams (Braddick, O’Brien, Wattam-Bell, Atkinson, & Turner, 2000) and information is obviously shared between the two streams.

Ventral stream areas in occipitotemporal cortex are involved in perception and object recognition, processes which are frequently studied using stationary objects but which could potentially utilize motion cues. Neuroimaging studies using static objects have reported areas of ventral cortex specialized for objects (Malach et al., 1995), faces (Kanwisher, McDermott, & Chun, 1997) and places (Epstein & Kanwisher, 1998). Motion may also provide valuable cues to image segmentation and object grouping but few studies have examined its contribution (but see Wang et al., 1999). Area LO, a ventral area involved in object perception, responds to objects regardless of whether they are defined by luminance, texture, or motion (Grill-Spector, Kushnir, Edelman, Itzchak, & Malach, 1998), suggesting that motion input does indeed reach the ventral stream. Motion information may also be useful for motion grouping, particularly in interpreting the movements of others. One particularly intriguing area is in the superior temporal sulcus (STS; Allison, Puce, & McCarthy, 2000). Neuroimaging studies have reported activation in human STS in response to the motion of other individuals (see Allison et al., 2000, for a review). STS has been activated by body movements (Bonda, Petrides, Ostry, & Evans, 1996; Grossman et al., 2000; Howard et al., 1996), hand movement (Bonda et al., 1996), eye and mouth movement (Puce, Allison, Bentin, Gore, & McCarthy, 1998), movement of shapes with implied intentions (Castelli, Happe, Frith, & Frith, 2000), and implied body movement (Kourtzi & Kanwisher, 2000a; Senior et al., 2000). These properties suggest that human STS is a likely homolog of a fascinating area in the monkey, which also codes biological motion and other information about the actions of other individuals (Perrett et al., 1985). Such information may also be
useful in learning motor skills through the observation of others (di Pellegrino et al., 1992; Rizzolatti et al., 1996).

Dorsal stream areas in occipitoparietal cortex are involved in spatial processing and visuomotor control of actions. Many functions of the dorsal stream including spatial localization, spatial attention, and visuomotor control depend on stimulus location and presumably stimulus movement. Many areas of parietal cortex are activated by stimulus motion, including several regions along the intraparietal sulcus (anterior, posterior, intermediate IPS), postcentral sulcus and parieto-occipital fissure (Cheng et al., 1995; Dupont et al., 1994; Sonaert et al., 1999). Furthermore, activity in these areas is greatly enhanced when attention is used to track moving objects (Culham et al., 1998), suggesting that these areas are involved in the maintenance of attention across shifts (Corbetta, Miezin, Shulman, & Petersen, 1993) and/or the updating of target location for the purposes of guiding actions such as eye, arm and hand movements. Anatomically, these areas have potential homologs in motion-selective monkey areas such as VIP (Duhamel, Colby, & Goldberg, 1998), 7a (Siegel & Read, 1997) and V6/V6A in the parieto-occipital sulcus (Galletti, Fattori, Gamberini, & Kutz, 1999). Dorsal areas are likely to be cross-modal; for example, regions of the IPS are activated by both visual and auditory motion (Lewis, Beauchamp, & DeYoe, 2000) and some have shown that overlapping motion responses across modalities may be useful in identifying areas such as VIP (Bremmer et al., 2001). Dorsal areas are also likely to be involved in the monitoring of one’s own movement. For example, one motion-selective region in the intraparietal sulcus is activated more strongly by viewing motion of the hand than by biological or random motion (Bonda et al., 1996). This area is near areas involved in the control of grasping (Binkofski et al., 1998) and may use motion cues to guide the hand to a target using visual feedback.

The majority of studies from both physiology and neuroimaging have focussed on the initial stages of motion processing with less regard for the eventual uses of motion information. Although the data thus far are sketchy, the range of cognitive functions that could benefit from motion cues provides some justification for the large number of brain regions that demonstrate motion selectivity.

3.2. Neuroimaging has verified and extended findings from other techniques

Many of the experiments conducted on motion perception to date have been based on known findings from other techniques, primarily psychophysics and neurophysiology. Neuroimaging of motion has included topics such as speed sensitivity (Chawla, Phillips, Buechel, Edwards, & Friston, 1998), optic flow (Cheng et al., 1995; Greenlee, 2000), motion in depth (Orban, Sunaert, Todd, Van Hecke, & Marchal, 1999; Paradis et al., 2000), color-defined movement (Tootell, Reppas, Kwong et al., 1995; Wandell et al., 1999), sensitivity to flicker (Tootell, Reppas, Kwong et al., 1995), apparent motion (Goebel, Khorram-Sefat, Muckli, Hacker, & Singer, 1998; Seghier et al., 2000), and the perception of illusory motion (Zeki, Watson, & Frackowiak, 1993). Unfortunately we lack the space here for a com-
prehensive review of all of these findings. However, we consider one example in
detail, that of second-order motion, to illustrate how imaging can contribute to
theoretical models of motion perception.

3.2.1. Example: Second-order motion

Motion can be seen not just from luminance- and color-defined stimuli, but from
a wide range of stimuli defined by many other attributes, such as texture, contrast, or
relative motion. This so-called second-order motion has been a subject of intensive
study by psychophysicists in the last two decades (e.g., Smith, 1994). One dominant
idea in understanding the mechanism of second-order motion is that the second-
order signal first undergoes a rectification process, and the resulting signal can be
used to compute motion in more or less the same way as the conventional luminance-
defined motion energy model. The psychophysical model does not specify where in
the neural pathway rectification occurs. However, based on the visual attributes used
to define the second-order pattern, if we are willing to assume the likely site of the
rectifying stage (e.g., V1/V2 for contrast defined patterns), presumably this second-
order motion will activate the ensuing visual motion system just like a luminance-
defined pattern. It is possible that another task involving second-order patterns (e.g.,
object recognition based on patterns defined by second-order statistics) will activate
the early stages including the rectification stage, but not motion specific stage. In this
sense, fMRI studies of the second-order motion have the potential to inform us on
the validity of the rectification model of second-order motion.

Using fMRI, Smith et al. (1998) examined the activity in many different visual
areas in response to luminance-based motion and three types of second-order mo-
tion. V1 and V2 responded well to all types of dynamic stimuli but lacked specificity
to motion. The most specific response to second-order motion was found in V3 and
VP, while area MT+ responded equally well to luminance and second-order motion.
They suggest that the results are consistent with the hypothesis that first-order
motion sensitivity arises in V1, that second-order motion is first represented ex-
plicitly in V3 and VP, and that V5/MT+ (and perhaps also V3A and V3B) is involved
in further processing of motion information, including the integration of different
types of motion signals.

measured the response to moving plaids (comprised of two superimposed gratings
with different orientations and motion directions) compared to motion of the
components presented one at a time. The plaid used in their experiment is essentially
a contrast modulated second-order pattern. It is thus reassuring to see that the plaids
activated V3 most selectively. Not surprisingly, MT+ did not show much differential
activation, presumably because the moving plaids and component gratings activated
the area equally well.

3.3. Neuroimaging has examined top-down influences on motion perception

Here the advantages of neuroimaging are most obvious in allowing researchers to
examine higher-order motion phenomena with human observers who can report
their perceptual experiences unambiguously. While physiology has focused on bottom-up processing of motion, with a few studies of attentional influences (Seidemann & Newsome, 1999; Treue & Maunsell, 1996, 1999), neuroimaging has discovered a number of surprising top-down effects on motion perception which we consider below.

3.3.1. Attention

The amount of information impinging on the visual system is far more than is ever registered in conscious awareness. Attentional processes can enhance certain aspects of the input based on properties of the stimulus (“bottom-up” influences) or on the observer’s demands or expectations (“top-down” influences). Motion is a highly salient feature which grabs one’s attention and keeps it locked on important features and objects. In one of the first neuroimaging studies of attention, Corbetta and his colleagues (Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1990, 1991) presented observers with identical stimuli but asked them to attend to different attributes of the display – shape, color, or motion – during different PET scans. Activation was enhanced in areas specialized for the feature which was attended; for example, attention to the speed of the stimulus increased activation in an area believed to be MT+, compared to attention to other attributes or to all three attributes simultaneously. O’Craven, Rosen, Kwong, Treisman, and Savoy (1997) demonstrated conclusively that MT+ was indeed attentionally modulated, even in the absence of eye movements. Rees, Frith, and Lavie (1997) found that even when motion was irrelevant to the task, motion processing still occurs if attentional load is otherwise low. Attentional modulation may also depend on the task used or the attribute attended. Watanabe et al. (1998) demonstrated that the regions showing attentional modulation were task-dependent, with MT+ responding across a wide range of attentional tasks, but V1 responding only when the task required attending to translation (but not expansion or rotation) or local motion direction. Beauchamp, Cox, and DeYoe (1997) demonstrated that MT+ activity was modulated both by spatial attention and by featural attention. That is, activation was highest when observers attended both to the location and speed of the stimuli, weaker when they attended to the location and color, and almost absent when they attended to the fixation point. These results are corroborated by another study in which speed discrimination induced greater MT+ activity than during contrast discrimination (Huk & Heeger, 2000a) but are contradicted by studies from another lab that have reported little or no modulation of MT+ for speed discrimination (Orban et al., 1998; Orban et al., 1996; Sunaert, Van Hecke, Marchal, & Orban, 2000). However, both the tasks and observers’ level of performance differed, possibly accounting for the discrepant results (Huk & Heeger, 2000a).

3.3.2. Motion imagery

Observers can “perceive” motion simply through the power of imagination. Mental imagery can activate some of the same regions engaged by perception alone (Kosslyn, Thompson, Kim, & Alpert, 1995), leading one to wonder whether imagining motion activates the same regions as perceiving it. Goebel et al. (1998) sug-
gested that indeed it does. Observers viewed several moving stimuli and then after a delay, they were asked to fixate a cross and imagine seeing the previous motion stimuli. In a second experiment, imagery of motion was compared to imagery of a static stimulus. Activation during imagery was seen in MT+ and several other motion-related areas, including V2 and V3A.

3.3.3. Implied motion

Recently, two groups have published the somewhat surprising result that motion areas can be activated merely by static images that imply motion (David & Senior, 2000; Kourtzi & Kanwisher, 2000a,b; Senior et al., 2000). For example, such images included an image of an athlete in motion compared to the same athlete at rest; or a cup falling off of a counter compared to a cup resting on the counter. MT+ responded more to the images that implied motion than those that did not. Areas near MT+ have also been implicated in tool naming and generating action words (Martin, Haxby, Lalonde, Wiggs, & Ungerleider, 1995; Martin, Wiggs, Ungerleider, & Haxby, 1996). Such results suggest top-down effects of object recognition and perceptual inference on the motion complex.

3.4. Neuroimaging has examined the neural correlates of awareness using motion perception

Although behavioral measures in monkeys may correlate closely with activity in single units (Logothetis & Schall, 1989; Salzman et al., 1990), extrapolating from behaviour to awareness in a non-verbal species requires careful training and a leap of faith. Neuroimaging studies in humans offer greater potential to examine the neural correlates of consciousness (Crick & Koch, 1995).

As we have seen earlier, MT+ can be active even when no physical motion is present (as in imagined or implied motion) and the strength of activation depends strongly on the attention of the viewer, which also affects the salience of the motion. This leads one to wonder whether MT+ activity and motion perception are equivalent. If so, MT+ should be active if and only if percept of motion is experienced. If not, it may be possible to observe MT+ activity in the absence of a motion percept or to observe a motion percept in the absence of MT+ activity.

3.4.1. Motion aftereffects

The motion aftereffect, also known as the movement aftereffect or the waterfall illusion, is probably one of the best-known visual illusions (Mather, Verstraten, & Anstis, 1998; Wohlgemuth, 1911). It typically occurs after an observer has watched a moving pattern (like a waterfall) for some time. If, after adaptation, one looks towards a stationary scene (like the rocks beside a waterfall), it appears to move opposite to the previously seen motion direction (e.g., upwards after viewing the falling water).

The effect is generally attributed to a transient imbalance between populations of motion sensors tuned to different directions of motion. For example, in the case of the waterfall, motion sensors tuned to downward motion become adapted over time
and as soon as the gaze direction is changed towards something stationary, the activity of these neurons falls below the baseline. Since the opposite direction was not stimulated and neurons coding this direction are not adapted, their activity is relatively higher and as a result the opposite direction is perceived. This idea underlies popular models like the ratio-model or distribution-shift type models (Mather & Harris, 1998; Mather & Moulden, 1980).

Tootell and his colleagues (Tootell, Reppas, Dale et al., 1995) measured activation in MT+ during the perception of the motion aftereffect when a stationary pattern was presented after observers had viewed continuous rotation. MT+ activation remained elevated during the MAE and decayed at a rate that strikingly mirrored the decay of the illusory motion percept. In comparison, when the stationary test was preceded by an oscillating grating that produced no MAE, the MT+ signal dropped rapidly. These results confirmed the presence of direction-selective cells and provided evidence that MT+ activity was closely correlated with the perception of motion.

The fact that the fMRI MAE decays at a rate comparable to the perceptual MAE suggests a close relationship between the MT+ signal and the perception of motion. However, the phenomenon of MAE storage provides a more rigorous test of the relationship. Storage was, to our knowledge, first reported by Wohlgemuth (1911) in his seminal work on movement aftereffects. The effect is simple. Suppose one adapts to a moving pattern and then a static test is immediately presented. Assume that the MAE lasts x seconds. Now repeat the experiment, but as soon as the adapting motion is stopped, the observer’s eyes are closed for x seconds. If the observer opens the eyes after this time, the MAE can still be perceived. In other words, the MAE was ‘stored’ or at least the recovery from adaptation was delayed or slowed down.

Two groups independently used fMRI to investigate MT+ activity during storage of the MAE (Culham et al., 1999; He, Cohen, & Hu, 1998; see also Moore & Engel, 1999). Observers either adapted to continuous motion that produced an MAE or to oscillating motion that did not. Afterward, the adapting stimulus was turned off and no MAE was perceived. Both groups reported an elevation in MT+ activity during storage of the MAE, even though no illusory motion was seen. However, Culham et al. (1999) showed that the MAE-specific activity in MT+ (MAE condition compared to control condition) was considerably higher when subjects were perceiving the illusory motion from the MAE than during the preceding storage interval. He et al. (1998) showed that MT+ activity following adaptation was higher when the test appeared at the adapted location relative to an unadapted one. Thus, both studies found that it was possible to obtain MT+ activation in the absence of a motion percept; however, the size of this activity was much reduced compared to activation during a motion percept.

3.4.2. Neuropsychological patients with motion deficits

The correlation between motion perception and MT+ activation has also been addressed in neuropsychological patients. One study examined a patient with bilateral damage to MT+ (Zihl et al., 1983) who had cortical motion blindness (Shipp, de Jong, Zihl, Frackowiak, & Zeki, 1994). MRI anatomical scans confirmed that MT+ was damaged in both hemispheres, and PET imaging verified that no activa-
tion was observed in the region. However, other motion-selective regions were active, including superior parietal cortex, “V3” (though probably V3A), and the lingual and fusiform gyri.

Another intriguing study examined a patient with Riddoch phenomenon (Riddoch, 1917), a syndrome related to blindsight (see Lamme, 2001). Riddoch patients have damage in the visual cortex resulting in loss of conscious vision in parts of their visual field (scotomata) but nonetheless have some residual motion processing in their blind regions, particularly for fast moving stimuli. Using fMRI, Zeki and Ffytche (1998) tested G.Y., a well-studied Riddoch patient, by presenting motion stimuli at fast or slow speeds and asking him to report whether or not he was aware of the motion. They then compared trials in which G.Y. was aware vs unaware of the motion, showing that MT+ was more active during trials in which he was aware of the motion, even once speed had been factored out. This is perhaps not too surprising given that attention, which is closely related to awareness, can strongly modulate neural activity. The more surprising result was that the unaware motion conditions produced activity even though G.Y. had no report of “seeing motion”.

Taken together, results from normal observers and the cortically blind patient G.Y. offer insights into the relationship between MT+ activity and motion perception. That is, several experiments have shown that physical motion is not necessary for MT+ activity; illusory motion, imagined motion or implied motion will suffice. In all cases when motion is perceived, MT+ is active. However, the reverse may not be true. MT+ activation and the perception of motion do not necessarily always co-exist, as demonstrated by weak but significant enhancements of the MT+ signal during storage of the MAE in normal observers and during the presentation of undetected motion stimuli in a cortically blind patient.

4. Discussion and future

Human vision in general, and motion perception in particular, are disciplines that are relatively well established by traditional methodologies and have been extensively studied with new imaging techniques developed in the “Decade of the Brain”. Thus even if one is not interested in these topics per se, it may be worthwhile to consider them as a test case for the capacity, limitations and potential of functional neuroimaging.

Through the examples of motion perception discussed above, we have suggested several instances where the application of modern neuroimaging techniques, especially fMRI, has increased our knowledge of the functional organization of the brain. Specifically,

- Cortical areas are mapped more accurately. Using characteristics like mirror image phase relationships of the retinotopic areas, borders between them can be charted and the representation of visual space within these areas is now well established.
- Similarities and differences in the relationship between macaque monkey brain and human brain have been identified. Much knowledge acquired from monkey
physiology can be applied to human brain with the guidance of functional imaging, but one should also be careful in generalizing results obtained using electrophysiology.

- New areas have been discovered. Unlike neurophysiology, neuroimaging has made it much more efficient to measure cortical responses to certain motion stimuli for the whole brain with no a priori assumptions about where to look.
- The effect of higher-order motion can be studied with greater confidence and with the full cooperation of conscious observers. Questions that are extremely difficult to study with non-human primates (e.g., imagination, awareness, etc.) are all within the scope of functional imaging studies.

While it is clear that neuroimaging has been a useful tool in the study of motion perception, its greatest contributions may be yet to come. The studies to date have largely focussed on the MT+ complex which is only one of the many regions and which itself likely consists of several subregions that only begin to be dissociated (Dukelow et al., 2000; Khan et al., 1999; Morrone et al., 2000). Given the range of reliable physiological differences between MT and MST (and perhaps other poorly understood regions like FST), one hopes to find similar divisions between corresponding subregions of the MT+ complex. Neuroimaging may eventually also be able to push the limits of spatial resolution to discern columnar organization within MT/MST for direction (Albright, Desimone, & Gross, 1984), disparity (DeAngelis & Newsome, 1999) or global/local motion (Born & Tootell, 1992). Such experiments would require high spatial resolution which has recently been used to visualize orientation and ocular dominance columns in V1 with fMRI (Kim, Duqong, & Kim, 2000; Menon, Ogawa, Strupp, & Ugurbil, 1997). The improved temporal resolution of single-event designs (e.g., Buckner et al., 1996) permits more sophisticated designs and greater capacity to correlate behaviour or awareness with neural activity on a trial-by-trial basis. Recently, a single-event design found that a directional cue (a stationary arrow) increased activation in MT+ and parietal areas in the cueing period and during a subsequent motion stimulus (Shulman et al., 1999). Single-event designs also allow one to use more sophisticated designs such as fMRI adaptation (Grill-Spector et al., 1999; Grill-Spector & Malach, 2001) to address motion processing. For example, preliminary data suggest that following adaptation to one direction of motion, the MT+ response is weaker for a brief movement in the same direction compared to the opposite direction (Huk & Heeger, 2000b), another measure of motion adaptation in addition to the motion aftereffect.

One particularly thorny problem for the understanding of motion is the contribution of top-down influences on neural activity. That is, as we have discussed, high-level factors such as attention and motion imagery can substantially modulate activity in MT+. In fact, some have argued that factors like attention that affect the baseline neural firing rates of many neurons make a disproportionate contribution to the population response measured by fMRI (compared to factors that are encoded within a small subpopulation of the neurons; Scannell & Young, 1999). If indeed this is true, the attentional state of the observer may be crucial to the outcome of the data. For example, Huk and Heeger (2000b) have recently suggested that the fMRI correlate of the MAE disappears when observers are forced to attend to a small
Although we have focused here on how neuroimaging informs the study of motion perception, the study of motion perception may also elucidate our understanding of how the brain processes information. Neuroimaging studies often utilize task-based paradigms to elicit specific patterns of neural activation. These tasks can range from simple motor movements to more complex cognitive processes, such as decision-making or language production. By manipulating these tasks within a neuroimaging paradigm, researchers can gain insights into the neural mechanisms underlying these processes.

However, it is important to note that neuroimaging studies are not without limitations. For example, the spatial and temporal resolution of current neuroimaging techniques may not be sufficient to fully capture the complexity of neural activity. Additionally, the interpretation of neuroimaging data can be challenging, as it requires careful consideration of factors such as task design, stimulus presentation, and individual differences in neural responses. Despite these challenges, neuroimaging remains an essential tool for advancing our understanding of the brain and its functions.
well to mathematical models based on the single-cell data, a result taken to mean that the BOLD response is directly proportional to average neuronal firing rate. The results suggest that when coherence is increased in one direction, say upward, the firing increase in upward-tuned neurons is larger than the decrease in downward-tuned neurons, otherwise no increase in BOLD would be observed (Heeger, Huk, Geisler, & Albrecht, 2000). Increased fMRI activation with higher motion coherence has been reported by other labs (Braddick et al., 1998; Braddick et al., 2000; Previc, Liotti, Blakemore, Beer, & Fox, 2000) but there are also reports indicating greater MT+ activity for incoherent than coherent motion (McKeefry, Watson, Frackowiak, Fong, & Zeki, 1997) or no difference between them (Paradis et al., 2000). These discrepancies imply that the BOLD response as a function of motion coherence may depend critically on the particular stimulus parameters employed. Interactions between neuronal populations have also been studied by Heeger and colleagues (Heeger, Boynton, Demb, Seidemann, & Newsome, 1999; see also Singh et al., 2000). Based on earlier results (Qian & Andersen, 1994) and their own results from neurophysiology, they predicted that when one direction of motion was superimposed on another, the response in MT+ would decrease. This prediction was confirmed with fMRI. Such results imply suppression between mutually opponent directions (otherwise one would predict greater activity for two vs one directions). Modeling may also be useful in understanding the interactions between areas within a neural network. For example, one study used structural equation modeling to demonstrate that attention to motion modulated the connectivity between MT+ and posterior parietal cortex (Chawla, Rees, & Friston, 1999).

Perhaps the greatest potential for neuroimaging is in conjunction with other cognitive neuroscience techniques. Many of the issues in understanding the transition from neuron to brain may eventually be resolved by combining neurophysiological recordings with fMRI recording in macaque monkeys. Several groups have independently reported fMRI activation of visual areas in macaques (Dubowitz et al., 1998; Logothetis, Guggenberger, Peled, & Pauls, 1999; Stefanacci et al., 1998). The MT complex is likely to be one of the first regions studied and modelled. Functional imaging of neuropsychological patients is also promising (for example, Morland et al., 2001), allowing one to determine not only what areas are damaged, but also what areas are intact and are likely to be subserving residual function. Neuroimaging may also be useful in studying patient populations with potential motion deficits such as dyslexics (Demb, Boynton, & Heeger, 1997; Eden et al., 1996), glaucoma patients (Joffe, Raymond, & Chrichton, 1997; Silverman, Trick, & Hart, 1990), hemianopes with blindsight/Riddoch phenomena (Bittar, Ptito, Faubert, Dumoulin, & Ptito, 1999), or schizophrenics (Chen et al., 1999). It may also be fruitful to combine fMRI, which has high spatial resolution, with other techniques such as electroencephalography (EEG) or magnetoencephalography (MEG), which have high temporal resolution, as two groups have done with motion perception (Ahlfors et al., 1999; Wang et al., 1999).

Neuroimaging has been accused of being modern-day phrenology (Blakeslee, 2000): simply localizing activations with little regard for actual function. To a certain extent, this has been true in the neuroimaging of motion perception. Many of the
initial imaging studies simply localized a host of regions that were activated by motion. However, a survey of neuroimaging studies of motion reveals how the techniques can progress beyond phrenology (Wandell, 1999). The robust and reliable localization of the human motion complex, MT+, allowed many neurophysiologists, psychophysicists and neuropsychologists to also use neuroimaging to address theoretically motivated questions about motion perception. In some cases, the answers verified those from other cognitive neuroscience techniques. In others, the answers were more surprising, such as indications of extensive top-down modulation of motion by factors such as attention and expectation, and the complex relationship between motion-related activity and the conscious perception of motion. Now we are left with clues to other motion-related areas and suggestive evidence about their potential function. Improvements in scanning technology and analysis methodology combined with clear theoretical goals and cross-talk with other neuroscience techniques should move us still further in this challenging endeavor to gain a full understanding of the neural mechanisms of visual motion. One can only hope that “The Decade After the Decade of the Brain” will resolve many of the intriguing issues posed by these initial neuroimaging results.

References


