Implied Motion from Form in the Human Visual Cortex

Bart Krekelberg[1], Argiro Vatakis[2], Zoe Kourtzi[3,4]

[1]The Salk Institute for Biological Studies, La Jolla, CA, USA

[2]Department of Experimental Psychology, University of Oxford, Oxford, UK

[3]Max-Planck Institute for Biological Cybernetics, Tübingen, Germany

[4]School of Psychology, Birmingham University, Edgbaston, Birmingham, UK

Running Head : Implied Motion
Keywords: functional imaging, adaptation, MRI, orientation, structure

Contact Information:
Bart Krekelberg
The Salk Institute for Biological Studies
10010 N. Torrey Pines Rd
La Jolla, CA 92037
USA
E: bart@salk.edu

1

ABSTRACT

When cartoonists use speed lines - also called motion streaks - to suggest the speed of a stationary object, they use form to imply motion. The goal of this study was to investigate the mechanisms that mediate the percept of implied motion in the human visual cortex. In an adaptation functional imaging paradigm we presented Glass patterns that, just like speed lines, imply motion but do not on average contain coherent motion energy. We found selective adaptation to these patterns in the human motion complex, the lateral occipital complex as well as earlier visual areas. Glass patterns contain both local orientation features and global structure. To disentangle these aspects we performed a control experiment using Glass patterns with minimal local orientation differences but large global structure differences. This experiment showed that selectivity for Glass patterns arises in part in areas beyond V1 and V2. Interestingly, the selective adaptation transferred from implied motion stimuli to similar real motion patterns in dorsal but not ventral areas. This suggests that the same subpopulations of cells in dorsal areas that are selective for implied motion are also selective for real motion. In other words, these cells are invariant with respect to the cue (implied or real) that generates the motion. We conclude that the human motion complex responds to Glass patterns as if they contain coherent motion. This, presumably, is the reason why these patterns appear to move coherently. The lateral occipital complex, however, has different cells that respond to the structure of real motion patterns versus implied motion patterns. Such a differential response may allow ventral areas to further analyze the structure of global patterns.

INTRODUCTION

Humans perceive motion in some images that contain no real motion. A striking example of this observation is illustrated by the speed lines- also called motion streaks -used in cartoons to suggest the speed of a stationary object (Burr 2000; Burr and Ross 2002; Geisler 1999). This perceptual phenomenon is known as implied motion (form *implies* motion) and shows an interaction between motion and form processing, two aspects of visual analysis often assumed to be conducted by largely independent pathways (Mishkin et al. 1983).

Some implied motion cues operate at a cognitive level (e.g. an athlete about to throw a javelin), hence form and motion interactions could take place at a cognitive level. Recent neurophysiological (Jellema and Perrett 2003) and human fMRI studies, however, suggest that at least the end-result of this interaction is represented in the prototypical motion area of the human brain (hMT+/V5). This area is more active when presented with real-life images that imply motion than when similar images are shown that do not imply motion (Kourtzi and Kanwisher 2000a; Senior et al. 2000).

Other implied motion cues operate at a lower level. In a dynamic Glass pattern sequence (Glass 1969; Ross et al. 2000), oriented elements are aligned along a common trajectory. This alignment generates a global structure in these patterns (Figure 1) that evokes a percept of coherent motion (Ross et al. 2000). For instance, by orienting the elements along concentric trajectories, a percept of rotational motion is evoked. The direction of this rotation is ambiguously clockwise or counterclockwise, but the fact that a coherent

3

rotation is seen at all is surprising because, on average, the motion energy in these patterns is perfectly balanced in all directions (see *Appendix* and *Discussion*). Krekelberg et al. (2003) found a neural correlate of the motion percept in the superior temporal sulcus of the macaque, where a subpopulation of motion selective cells responds to Glass sequences as if they contain real coherent motion.

The goal of our study was threefold. First, we wished to determine which parts of human visual cortex show selective responses to the implied motion represented by Glass patterns. Second, we wanted to determine to what extent the subpopulations of cells that are selective for implied motion overlap with those selective for qualitatively similar real motion patterns. Third, we were interested in investigating to what extent the response to the global structure of the Glass patterns and its concomitant percept of motion was determined by local orientation detectors, possibly already at the level of V1, and how much of the response is due to the global organization of the oriented elements.

Our functional MRI data show that the human motion complex indeed contains subpopulations of cells that are selective for both implied and real motion. Selectivity for global patterns of implied motion and global patterns of real motion was also observed in ventral areas (VP, V4, and LOC). In contrast with the dorsal areas, however, the subpopulations selective for the structure of Glass sequences did not show significant fMRI selectivity for the structure of real motion patterns. Finally, only a small part of the implied motion selectivity could be explained on the basis of selectivity for local orientation changes in primary visual cortex (V1, V2).

4

These findings provide insight into the representation of form and motion in the human visual cortex. Dorsal areas have a large degree of cue-invariance (Albright 1992); in these areas real and implied motion patterns drive similar subpopulations of neurons.. The ventral areas – which presumably respond to the global structure rather than the motion in these sequences – do not show this cue-invariance. This allows ventral areas to discriminate between global structure generated by motion cues and that same structure generated by form cues. Dorsal areas, on the other hand, do not make this distinction and extract only the (implied or real) motion information.

MATERIALS AND METHODS

We used an event related adaptation fMRI paradigm. This section first describes our subjects, then the stimuli we used and how we adapted the stimuli for each subject. The last three sections are concerned with the logic and design of the experiments, and how we obtained and analyzed the fMRI data.

OBSERVERS

11 observers participated in Experiment 1 (Implied Motion), 14 in Experiment 2 (Real Motion), 16 in Experiment 3 (Implied and Real Motion Interactions) and 13 in Experiment 4 (Local versus Global Implied Motion). 3 observers were excluded from the analysis in Experiment 2, 3 in Experiment 3, and 2 in Experiment 4 due to excessive head movement. All observers had normal or corrected to normal vision, were paid for

participation and gave their informed consent. All procedures were in accordance with international standards for research involving human subjects (Declaration of Helsinki).

APPARATUS

An LCD projector (NEC GT950) displayed the visual stimuli on a tangent screen that the subjects viewed through a mirror. The refresh rate of the projector was set to 60Hz for all stimuli. The visible screen subtended 21 degrees of visual angle. A custom-made fiber optics button box allowed the subjects to communicate their decisions in the perceptual tasks.

STIMULI

Localizers

For the LOC localizer scans we used grayscale images of novel and familiar objects as well as scrambled versions of each set, as described previously (Kourtzi and Kanwisher 2001). We localized hMT+/V5 by using a moving dot-pattern (expanding and contracting for 9 s at a speed of 4º/s within a 21º aperture, reversal rate 1Hz) and a stationary random-dot-field (Huk and Heeger 2002; Watson et al. 1993). Area KO was localized by using kinetic boundaries and transparent motion stimuli (Dupont et al. 1997) that consisted of a field of random black (50%) and white (50%) dots (size: 3.1 arcmin, speed: 4.44 deg/s). To map the borders and the eccentricity of the retinotopic visual areas, we used rotating triangular wedge stimuli and concentric rings. These stimuli consisted of either gray-level natural images or black and white objects-from-texture images that were

presented at temporal frequency of 2Hz as described in previous studies (Kourtzi et al.

2003).

Random Dot Patterns

The random dot patterns used in the main experiments were created with in-house

OpenGL software. Each pattern consisted of 200 dots. A single dot subtended 0.24

degrees, and the whole stimulus pattern was contained within an 18 degree circular

aperture around the central fixation point.

Implied Motion

To generate implied motion, we used sequences of Glass patterns. In each Glass pattern,

the dots were arranged in pairs and all pairs in a given pattern were aligned either along

concentric circles around the fixation point, or along radial lines emanating from the

fixation point (Figure 1A). In a sequence of Glass patterns, a new pattern of the same

type, with a new set of randomly positioned pairs was presented every 83ms; this being

close to the optimal range to generate the impression of motion (Ross et al. 2000). The

duration of a single sequence was 300ms (see Procedure, below). The distance between

the dots in a pair, called the Glass-shift, was set per subject to maximize their impression

of motion (Ross et al. 2000) (see *Procedure*). When the pairs were aligned along

concentric circles, they gave an impression of clockwise or counter clockwise rotational

motion. They will be referred to as concentric Glass sequences. Orienting the pairs along

radial lines gave an impression of expanding or contracting motion; these will be referred

to as radial Glass sequences. As appendix A shows, the average motion energy in these

sequences is balanced. For each motion energy component in one direction, there is on

average an equal component with energy in the opposite direction. Hence, the distribution of motion energy in these sequences does not predict the coherent motion percept. In our terminology, stimuli with balanced motion have no *coherent* motion energy. In contrast, stimuli in which the motion energy does have a clear peak in some direction are referred to as *real* motion stimuli.

Real Motion

The motion sequences all had 200 dots that were identical to those in the Glass patterns. In the motion sequences, however, all dots were randomly positioned rather than arranged in pairs. Because Glass sequences implicitly contain two motion directions we used motion sequences in which the direction of motion reversed every 83ms. The motion trajectory could either be along concentric circles (concentric motion) or along lines emanating from the fixation point (radial motion). The speed of the dots in the motion sequences was chosen to perceptually match the (high) speed in the Glass sequences (see *Procedure*).

Local Orientation Controls

To investigate the influence of local orientation selectivity, we devised segmented Glass patterns. In a segmented Glass pattern, the 18° aperture was divided into a square grid of 8*8 segments (See Figure 5). We assigned a random pair orientation to each segment. The orthogonal Glass sequences were chosen such that the local orientation per segment was orthogonal to the orientation of the previously presented segmented Glass sequence.

Per segment, the change from a segmented sequence to the orthogonal sequence is locally the same as the switch from a radial sequence to a concentric sequence. This is most easily seen by comparing the orientation of the highlighted pairs in the example stimuli of Figure 1 and Figure 5. The difference between the transition from segmented to orthogonal sequences and the transition from concentric to radial Glass sequences lies solely in the global structure. Locally – at the scale of a single segment (~2º) – both transitions are from one orientation to the orthogonal orientation.

PROCEDURE

Observers participated in two LOC, one hMT+/V5, one KO localizer scan and two retinotopic mapping scans, and four scans for each of the four event-related experiments. Before the relevant scanning sessions the observers participated in a practice session. In this session the subjects were familiarized with the types of stimuli used in the experiment. In these sessions we used a simple adjustment procedure to determine the Glass shift that evoked the strongest coherent motion percept as well as the speed of the real motion sequences that matched the perceived speed in the Glass sequences.

During the scan sessions, the subjects performed one of two behavioral tasks that ensured that an equal amount of attention was allocated to the stimulus in all conditions. In the first task ('matching task'), the subjects pressed a key to report whether the two stimuli in a trial were the same or different. We analyzed the percentage of trials in which the decision was correct. In the second task ('change detection task'), the central fixation point briefly (250ms) changed its color from red to blue at unpredictable times during the

9

trial. Subjects pressed a key to indicate that they detected this change. We analyzed the percentage correct detections as well as the reaction time for each correct detection.

DESIGN

Localizers

For the LOC localizer scans each stimulus condition was presented in a 16-second stimulus epoch (blocked design), as in previous studies (Kourtzi and Kanwisher 2000b). Each condition was repeated four times in a balanced order and with interleaved fixation periods. Twenty images were presented in each block, each for 300 ms with a blank interval of 500 ms between images. The observers fixated and performed a one-back matching task. In the hMT/V5+ localizer, a stationary dot-pattern was shown for 27 seconds and was then replaced by a moving (expanding and contracting) random dot pattern. Each condition was repeated 9 times. In the V3B/KO localizer scans, each stimulus condition (kinetic boundaries, transparent motion) was presented for seven 16-second epochs with interleaved fixation periods similar to the LOC localizer scan. For the retinotopic mapping scans, eight wedge positions and eight eccentricity rings were presented for 8s each and repeated eight times. During the hMT/V5, V3B/KO and retinotopic scans observers performed the change detection task.

Event-Related Adaptation Scans

Each scan started with a 16-second fixation epoch and ended with an 8-second fixation epoch. After the initial fixation epoch, the experimental trials started. We used an event-related adaptation paradigm (Buckner et al. 1998; Grill-Spector and Malach 2001;

Kourtzi and Kanwisher 2000b, 2001) in which two stimuli (e.g. A and B) were presented sequentially in 3s trials. Each stimulus was presented for 300ms with a 100 ms blank interval between stimuli and an inter-trial interval of 2300ms (See Figure 1B). Thus there were four experimental conditions: 1) A followed by A 2) B followed by B 3) A followed by B 4) B followed by A, and one fixation condition in which only the fixation point appeared throughout the trial. As in previous studies (Kourtzi and Kanwisher 2000b, 2001), the order of trials was counterbalanced across subjects and runs so that trials from each condition, including the fixation condition, were preceded (2 trials back) equally often by trials from any of the other conditions. In each of the experiments each condition was repeated 25 times per scan (total of 125 trials across conditions per scan) and subjects were run in 4 scans in one scanning session.

Experiment 1 ('Implied Motion')

The standard event-related adaptation design was used with A=concentric Glass and B=radial Glass resulting in four experimental conditions: 1) Concentric-Concentric, in which two concentric sequences were presented in a trial, 2) Radial-Radial, in which two radial sequences were presented in a trial, 3) Concentric-Radial, in which a radial sequence followed a concentric one, and 4) Radial-Concentric, in which a concentric sequence followed a radial one.

Experiment 2 ('Real Motion')

The same design as in experiment 1 was used with A = real concentric motion, B= real radial motion.

Experiment 3 ('Implied and Real Motion Interactions')

The first stimulus in a trial was always a Glass sequence, the second was always a real motion sequence. The conditions were: 1) Concentric-Concentric: concentric Glass followed by concentric motion, 2) Radial-Radial: radial Glass followed by radial motion, 3) Concentric-Radial: concentric Glass followed by radial motion, 4) Radial-Concentric: radial Glass followed by concentric motion.

Experiment 4 ('Local versus Global Implied Motion')

The first stimulus in a trial was always a segmented Glass sequence, and we had the following four conditions: 1) Random-Random: segmented Glass followed by segmented Glass 2) Random-Concentric: segmented Glass followed by concentric Glass, 3) Random-Radial: segmented Glass followed by radial Glass, 4) Random-Orthogonal: segmented Glass followed by orthogonal Glass.

IMAGING

The experiments were recorded in a 3 Tesla Siemens scanner at the University Clinic, Tübingen. Data were collected with a head coil from 11 axial (3 x 3 x 5 mm^3) slices that covered occipitotemporal regions using Gradient Echo pulse sequences (Localizer scans: TR=2s, TE = 90ms. Event-related scans: TR=1s, TE =40ms. Block-design scans: TR =2s, TE =90ms.).

DATA ANALYSIS

The fMRI data were processed using the Brain Voyager software package. Preprocessing of all functional data included head movement correction, temporal high-pass filtering

(cut-off frequency 0.0468Hz), and removal of linear trends. The 2D functional images were aligned to 3D anatomical data with 1x1x1mm resolution and the complete data set was transformed to Tailarach coordinates. Anatomical data were additionally inflated and unfolded.

Regions of interest

For each individual observer, early visual areas (V1, V2, Vp, V3, V3a, V4) were identified based on standard retinotopic mapping procedures (DeYoe et al. 1996; Engel et al. 1994; Sereno et al. 1995). The motion complex (hMT+/V5) was defined as the set of contiguous voxels in the ascending limb of the inferior temporal sulcus that showed significantly ($p<10^{-4}$, corrected) stronger activation for coherently moving (expanding, contracting) than stationary dots. The LOC was defined as the voxels in the ventral occipitotemporal cortex that showed significantly stronger activation ($p<10^{-4}$, corrected) to intact than scrambled images based on the averaged data of the two localizer scans. Area KO was defined as the set of voxels anterior to V3a and posterior to hMT+/V5 that showed significantly stronger activation ($p<10^{-4}$, corrected) for kinetic boundaries than transparent motion. All regions of interest are shown on a flattened representation of a single subject's cortex in Figure 2.

Event-related scans

For each observer, we extracted fMRI responses by averaging the data from all the voxels within each of the independently defined ROIs in the event-related scans. In each scan,

we averaged the signal intensity across all the trials in each condition. We then calculated the percent signal change for each condition in relation to the fixation baseline as described in previous studies (Kourtzi and Kanwisher 2000a, 2001). Finally, we averaged these time courses across scans and observers.

The hemodynamic response function peaks several seconds after the onset of the stimulus (Boynton et al. 1996). To identify the peak of the fMRI responses in a ROI we fitted a Gaussian model (Kruggel et al. 1999) to the average fMRI responses for each condition across observers. This analysis showed average peak responses within the same time window (3-5 sec after trial onset) across ROIs and experiments. Based on this analysis the average fMRI response between 3 and 5 sec after trial onset was taken as the measure of response magnitude for each condition in subsequent analyses. That is, all comparisons among conditions and fMRI response measures in the figures are based on this averaged signal.

From these averaged fMRI signals we derived a measure of fMRI selectivity to stimulus changes per ROI and experiment. For instance in an experiment with the four conditions (and corresponding fMRI signals) AA, BB, AB and BA, the selectivity index (or rebound index) was defined as: $SI = [(AB+BA)/(AA+BB)]-1$. This represents the enhancement of activity obtained when two different sequences are shown compared to when two sequences of the same type are shown. This measure quantifies a rebound effect; that is, the release from adaptation. Note that, even though we refer to SI as a *selectivity* index, we are aware that its relationship with true neuronal selectivity has not yet been

demonstrated conclusively (see *Discussion*). Nevertheless, it has proven to be a sensitive tool to investigate the BOLD signal selectivity at a spatial resolution below that of the imaged voxels (Buckner et al. 1998; Grill-Spector and Malach 2001; Henson and Rugg 2003; Kourtzi and Kanwisher 2000b, 2001).

RESULTS

We first present the details of the analysis and findings in the motion complex. Second, we describe the main findings in the other visual areas of interest shown in Figure 2. Third, we present the results of an analysis that determines whether the selectivity for Glass sequences can be explained on the basis of local orientation selectivity. The final part of the results documents that neither attentional confounds, nor eye-movement artifacts could explain our results.

THE HUMAN MOTION COMPLEX

In this section we present the analysis of the BOLD signal in the human motion complex (hMT+/V5) in some detail. This includes some of the intermediate steps in the analysis necessary to arrive at an index of selectivity for an ROI. In subsequent sections that report the results in other regions of interest we will skip these intermediate results.

Implied Motion

In this experiment we presented either two Glass sequences with implied motion of the same type (e.g. concentric-concentric) or two Glass sequences with implied motion of different type (e.g. concentric-radial), see Figure 1B.

Figure 3A shows the fMRI time courses (averaged over all subjects) in the motion complex. At time zero the two implied motion stimuli were presented sequentially. The BOLD signal responded with its typical delay of a few seconds and showed an undershoot after the peak of the response. Note that due to the slow time course of the BOLD response, the separate presentation of the two sequences (Figure 1B) cannot be resolved. A comparison of the two time courses in this panel, however, shows that the response to two successive Glass sequences of the same type (concentric-concentric) was lower than the response to two successive Glass sequences of different type (concentric-radial). This was a statistically significant effect (Repeated Measures ANOVA ($F(1,10) = 5.06$, $p<0.05$)). Figure 3B shows the same analysis for the two conditions that started with a radial pattern. The BOLD response was lower when two Glass sequences of the same type (radial-radial) were shown than when two Glass sequences of different type (radial-concentric) were shown.

To simplify the data presentation and for easy comparison across areas and experiments, we quantified the average response for each of the four conditions shown in panel Figure3A and B by averaging the fMRI signal in the peak time points of the fMRI time courses (gray time window; see *Materials and Methods*). This average peak response for the four conditions is shown in Figure 3C. The bar plot summarizes what is clear from panel A and B; the average peak BOLD response is lower in the conditions where the same pattern is shown twice than in the conditions where two different sequences are shown.

We interpret this as pattern selective adaptation. That is, when a second sequence of the same type is presented, the response is reduced due to adaptation. When a second sequence of a different type does not show this reduction, it must have stimulated a different (non-adapted) set of neurons. Hence, we infer from this so-called 'release from adaptation' that separate subpopulations of cells in hMT+/V5 respond to radial and concentric Glass sequences (see *Discussion*). In a control experiment described below we tested whether the local or the global differences between these categories could explain this selectivity.

To quantify the pattern selectivity we calculated an index that contrasts the BOLD signal in trials in which two sequences of the same type were presented with the BOLD signal in trials in which two sequences of different type were presented. This selectivity index (SI; see *Materials and Methods*) condenses the analysis to a single number per experiment and region of interest. This index is zero – reflecting no selectivity – when the response to two patterns of the same type equals the response to two patterns of different types. Index values significantly above zero correspond to the situation where the two patterns of different types evoke larger responses than two patterns of the same type. I.e. a positive index reflects release from adaptation and hence pattern selectivity in the underlying population. The implied motion selectivity for the human motion complex is shown as the white bar in Figure 3F.

Real Motion

In this experiment, we presented either two real motion sequences of the same type (e.g. concentric-concentric) or two real motion sequences of different types (e.g. concentric-radial). The analysis was the same as that of the implied motion data. Figure 3D shows the bar plot of the average peak BOLD signal for the four conditions in this experiment. Here – as in the implied motion data – the BOLD signal in the conditions with two different types of motion sequences was significantly higher than that in the conditions with two of the same types of motion sequences. $(F(1,10) = 42.76, p<0.001)$. Hence we infer that – not surprisingly and entirely consistent with single cell findings– the motion complex has subpopulations of cells that are selective for the type of real motion.

From these average peak BOLD responses we determined the selectivity index of the human motion complex for real motion. The selectivity for real motion in hMT+/V5 was approximately 50% larger than that for implied motion. It is represented by the black bar in Figure 3F.

Interactions between Real and Implied Motion

The analysis so far shows that hMT+/V5 has subpopulations selective for real motion and subpopulations selective for implied motion sequences. However, this does not necessarily mean that these subpopulations were the same. Experiment 3 was designed to test this hypothesis directly. If the same cells that respond to implied rotations also respond to real rotations, it should be possible to obtain pattern selective adaptation for a concentric real motion pattern after the presentation of an implied rotation, and similarly

for expansions. Hence, in this experiment we presented concentric implied motion followed by concentric real motion and compared it to trials in which concentric implied motion was followed by radial real motion. We also compared radial implied motion followed by radial real motion to radial implied motion followed by concentric real motion.

Figure 3E shows the average peak BOLD responses for the four conditions. A real motion pattern following an implied motion pattern of the same type led to lower fMRI responses than a real motion pattern following an implied motion pattern of a different type ($F(1,10) = 6.28$, $p<0.05$). This was the case for both radial and concentric implied motion sequences. In terms of pattern selective adaptation this suggests that the adaptation caused by an implied motion sequence affects the response to a real motion sequence, but only if it is of the same motion type (radial or concentric). This in turn suggests that the same subpopulations of cells respond (and therefore adapt) to both implied and real motion sequences of the same type.

From the activation measures in Figure 3E we determined the selectivity index; it is shown as the gray bar in Figure 3F. It measures the extent to which pattern selective adaptation in the motion complex is invariant to changes in the sequence from implied motion to real motion. The fact that the index is non-zero shows that some of the cells selective for real motion were also selective for implied motion. If the gray bar were as tall as the black bar, that would mean that adapting with real motion had the same effect as adapting with implied motion (when tested with real motion). The fact that the

19

selectivity index appears lower in this experiment than in the experiments in which only real motion was used, suggests that not all cells selective for real motion are also selective for implied motion. From the relative sizes of the real motion selectivity index and interaction selectivity index we infer that about 45% of cells selective for real motion were also selective for implied motion.

The main points of our findings in the motion complex can be illustrated by panel F in Figure 3. The selectivity indices shown in this graph (together with their associated statistical tests of significance) make three points 1) hMT+/V5 is selective for the type of implied motion (white bar), 2) hMT+/V5 is selective for the type of real motion (black bar), 3) the selectivity for implied and real motion sequences is subserved by overlapping subpopulations (gray bar). For the other visual areas, we will only present these final steps of the analysis.

RETINOTOPIC VENTRAL AND DORSAL VISUAL AREAS AND THE LOC

The same analysis shown in detail for hMT+/V5 in Figure 3 was applied to the data from the other ROIs in the ventral and dorsal pathway.

Implied Motion

The responses for two Glass sequences of the same type were significantly lower than the responses to two Glass sequences of different type in all early areas (V1, V2), dorsal retinotopic areas (V3, V3a, V3b/KO) as well as ventral retinotopic areas (Vp,V4) and the higher occipitotemporal area (LOC). Hence all areas showed some selectivity for these

sequences; their selectivity indices are represented by the white bars in Figure 4. The statistical analysis of selectivity gave the following results: V1: $F(1,10) = 43.36$, $p<0.001$; V2: $F(1,10) = 37.15$, $p<0.001$; V3: $F(1,16) = 20.86$, $p<0.01$; V3a: $F(1,16) = 30.33$, $p<0.001$; V3b/KO: $F(1,16) = 53.76$, $p<0.001$;Vp: $F(1,20) = 68.61$, $p<0.001$; V4: $F(1,20) = 72.41$, $p<0.001$; LOC: $F(1,20) = 70.21$, $p<0.001$.

Real Motion

We also analyzed the responses to the real motion sequences (Experiment 2) and again found that there was significant - albeit small- selectivity in early visual areas (V1, V2, V3) ventral areas (Vp, V4, LOC) and the dorsal stream (V3a, V3B/KO). The statistical analysis gave the following results: V1: $F(1,10) = 122.12$, $p<0.001$; V2: $F(1,10) = 97.35$, $p<0.001$; V3: $F(1,12) = 5.79$, $p<0.05$; V3a: $F(1,12) = 20.88$, $p=0.001$; V3b/KO: $F(1,12) = 9.02$, $p=0.01$; Vp: $F(1,20) = 12.76$, $p<0.01$; V4: $F(1,20) = 6.89$, $p=0.01$; LOC: $F(1,20) = 6.19$, $p<0.05$. The selectivity index per area is represented by the black bars in Figure 4. Note that this selectivity does not necessarily imply motion selectivity; the direction of motion is one feature that distinguishes radial real motion patterns from concentric motion patterns, but each of these patterns also has a distinct structure or form. Given the known properties of cells in the ventral stream, the selectivity observed in V4 and the LOC is most likely attributable to selectivity for the form suggested by the motion, consistent with previous studies that show enhanced fMRI responses in ventral visual areas for moving compared to static forms (Grill-Spector et al. 1998; Kourtzi et al. 2002).

Interactions between Real and Implied Motion

As in hMT/V5, we then tested whether pattern selective adaptation to an implied motion pattern transferred to a reduced response in a real motion pattern. The aim of this experiment was to determine whether the neural subpopulations selective for implied motion sequences were also selective for real motion sequences. We found (Figure 4) that the selectivity indices in the dorsal areas (V3, V3a, V3b/KO) were typically at least twice as large as those in V1, V2 and the ventral visual areas (Vp, V4, LOC). This suggests that the overlap between the neural populations responding to implied and real motion stimuli was much larger in the dorsal areas than in early or ventral areas.

Specifically, the BOLD response for a real motion sequence following an implied motion sequence of the same type was lower than the response to a real motion sequence following an implied motion sequence of a different type in the dorsal visual areas (V3: $F(1,14) = 27.55$, $p<0.001$; V3a: $F(1,14) = 52.07$, $p<0.001$; V3b/KO: $F(1,14) = 2.64$, $p=0.05$). However, no significant differences were observed between the fMRI responses for same or different types of sequences in V1 and V2 ($F(1,10) = 1.11$, $p=0.31$), or ventral visual areas (Vp: $F(1,20) = 1.61$, $p=0.21$; V4: $F(1,20) < 1$, $p=0.85$; LOC: $F(1,20) = 1.68$, $p=0.21$).

SENSITIVITY TO LOCAL VERSUS GLOBAL CHANGES

Figure 4 shows that some selectivity for implied motion sequences is already present at the level of V1. This leads to two questions. First, how can area V1, with its small receptive fields be selective for such large patterns? And second, if V1 is selective, does

that mean that all selectivity in higher areas is simply inherited from V1, or can selectivity be observed in higher areas with patterns for which no selectivity is observed in V1?

The first question is easily answered; the sequences (concentric versus radial) differ not only at the global scale, but they also have systematic differences at the scale of a V1 receptive field. In fact, if a given part of the screen contains a pair of dots oriented one way in a concentric pattern, that same part of the screen will contain the orthogonal orientation in a subsequent radial pattern. The highlighted pairs in Figure 1A illustrate this. Hence, (local) orientation selectivity in an area could be enough to lead to the selective adaptation we observed. Experiment 4 was designed to test this hypothesis (Figure 5). We divided the screen into 64 segments and assigned a random orientation to each segment (See Figure 5 and *Materials and Methods*). In one condition, two segmented Glass sequences, with the same random orientation per segment were shown in succession. In the other condition, a segmented Glass sequence was followed by another segmented sequence in which the orthogonal orientation was assigned to each segment. Locally (at the scale of the segments), the difference between these two sequences was a 90-degree orientation change, just like the transition from a concentric to a radial Glass pattern.

Figure 6 shows that the selectivity index in V1 was about as large for these local changes as it was for the global changes documented in Figure 4. In particular, fMRI responses were significantly stronger for segmented Glass sequences with orthogonal orientations

than the same orientation (V1: $F(1, 30) = 7.16$, $p=0.01$; V2: $F(1, 30) = 2.13$, $p=0.05$)

Hence it seems likely that the selectivity to the global changes of Experiment 1 in V1 was

in fact a result of local orientation selectivity. This is in agreement with single cell data

demonstrating that the orientation signals present in Glass sequences are enough to drive

orientation selective V1 cells (Smith et al. 2002). As Figure 6 shows, selectivity for local

orientation was also found in other areas, both dorsal (V3: $F(1,54) = 6.59$, $p=0.01$; V3a:

$F(1,54) = 6.27$, $p=0.01$; V3b/KO: $F(1,54) = 2.98$, $p=0.05$; hMT+/V5: $F(1,10) = 8.49$, $p<$

0.01) and ventral (Vp: $F(1,60) = 16.72$, $p<0.01$; V4: $F(1,60) = 22.54$, $p<0.001$; LOC:

$F(1,60) = 12.21$, $p<0.001$). It is possible that cells in areas such as hMT+/V5 picked up

some of the implied motion signals at the scale of the segments, or that this reflects true

local orientation tuning in these areas (Albright 1984). Nevertheless, the most

parsimonious explanation of this selectivity is that it was inherited from the differential

responses in V1.

Finally, we tested whether there is *any* selectivity that can only be due to global and not

the local organization of the stimuli. The perfect control would be to show two stimuli

that are identical locally but different globally. Such a control does not exist because the

global structure is entirely defined by the local elements. Hence, the next best control

uses two sequences with minimal differences at the local scale but large differences at the

global scale. We used trials in which a segmented Glass sequence was followed by a

concentric or radial Glass sequence. In these trials the local orientation change is 45

degrees on average; this is on the order of the bandwidth of V1 orientation cells (Ringach

et al. 2002), hence we expect these two sequences to be minimally different to V1.

Globally on the other hand, this change is a very clear transition from a random jumble of orientations to a structured pattern that appears to rotate or expand. Figure 6 shows that if this stimulus change did lead to changes in V1 or V2, those changes were too small to be resolved with BOLD signals. In particular, no significant differences were observed in the fMRI responses between two successive segmented sequences of the same orientation and sequences in which a Glass pattern followed a segmented pattern (V1: $F(1, 30) < 1$, p=0.88; V2: $F(1, 30) < 1$, p=0.87). In dorsal visual areas (V3,V3a, V3b/KO, hMT+/V5), however, two successive segmented Glass sequences did lead to significantly less activation than a segmented Glass pattern followed by a radial or concentric Glass pattern. The statistical analysis resulted in: V3: $F(1,54) = 13.01$, p<0.001; V3a: $F(1,54) = 5.76$, p<0.05; V3b/KO: $F(1,54) = 8.94$, p<0.01; hMT+/V5: $F(1,30) = 10.27$, p<0.01. For the ventral visual areas, we observed similar selectivity in V4 ($F(1,60) = 6.98$, p=0.01) and LOC ($F(1,60) = 20.83$, p<0.001) but not in Vp ($F(1,60) = 1.81$, p=0.18).

This shows that higher areas in both the dorsal and the ventral stream but not the early areas V1 and V2 were selective (as measured by the BOLD signal) to sequences that have similar local properties but distinct global structure. These findings suggest that at least part of the selectivity for the global organization of Glass sequences arises in areas beyond V1 and V2.

CONTROL EXPERIMENTS & ANALYSIS

It is conceivable that attention could be engaged more during trials in which two different stimuli were presented than trials in which the same stimulus was shown twice. Because attention is known to modulate responses in visual areas, this could then have influenced our results. To control for this possibility, subjects performed a matching task during all scans (see *Materials and Methods*). This task drew the subject's attention towards the stimuli, even in trials in which the same stimulus was shown twice. A Kruskal-Wallis ANOVA on ranks of the subjects' performance on this task showed that there was no significant difference in the percentage correct across conditions in any of the experiments (Experiment 1: $p=0.92$; Experiment 3: $p = 0.32$, Experiment 4: $p=0.98$; due to a software error, the behavioral responses from Experiment 2 could not be analyzed). The constant level of performance in these experiments indicates that a similar amount of attention was always devoted to the stimuli, regardless of condition. Moreover, it is highly unlikely that observers could selectively choose to attend to particular conditions as trials were presented in quick succession and were randomly interleaved. However, the matching task was not particularly difficult as witnessed by the high average level of performance (96% correct). This implies that some attentional capacity may have been left for the subject to allocate differently in different conditions.

We therefore performed an additional control experiment. We repeated the implied motion experiment, with the instruction to the subjects to detect a change in the color of the fixation point (see *Materials and Methods*). This task was subjectively much more difficult than the matching task. Analysis of the behavioral data showed that the number

of undetected changes in the fixation point did not vary significantly with condition ($F(2,4) = 0.93$, $p=0.49$). Moreover, the reaction times for the correctly detected changes in the fixation spot were also not significantly modulated by the stimulus condition. ($F(2,4) = 2.28$; $p>0.05$). These behavioral data suggest that attention was allocated similarly (to the fixation point) in all five conditions. Analysis of the fMRI data and the corresponding selectivity indices obtained in these sessions confirmed the adaptation effects that we reported above for implied motion in Experiment 1 across all areas (V1: $F(1,3) = 58.08$, $p<0.01$, V2: $F(1,3) = 104.97$, $p<0.01$, V3: $F(1,3) = 16.46$, $p<0.05$, V3a: $F(1,3) = 12.24$, $p<0.05$, Vp: $F(1,3) = 134.04$, $p<0.001$, V4: $F(1,3) = 119.92$, $p<0.001$, LOC: $F(1,3) = 62.98$, $p<0.001$). This control suggests that our findings were not confounded by a differential allocation of attention across conditions.

Eye Movements

During the scans, observers were instructed to fixate the central fixation point. Eye movements of three subjects in each experiment were recorded (Eye-Link video based system, 250Hz sample rate). We compared eye position and saccades across conditions in each experiment. This analysis showed no significant differences in the average saccade number, the vertical or horizontal eye position, or the vertical or horizontal saccade amplitude between experimental conditions and the fixation condition. This applied to all four experiments and none of the statistical tests reached a p-value below 0.11 ($F[2,8]>=2.58$). This shows that the subjects were able to fixate for long periods of time and that it is unlikely that our findings could be confounded by differential eye movements across conditions.

27

Block design analysis

Region of interest analyses are sometimes criticized for pre-selecting areas and ignoring the rest of the brain. Our study tested very specific hypotheses about the involvement of typical motion and form areas in implied motion perception. In that context we believe the ROI approach to be appropriate. Nevertheless, to investigate whether other brain areas beyond the measured ROIs respond to implied motion sequences we ran block design scans that presented real, implied and random motion. Each scan consisted of 16 second blocks with a given visual pattern that were presented in a counterbalanced order. The 7 conditions were: fixation only, concentric Glass, radial Glass, segmented/random Glass, concentric motion, radial motion. Thus, all stimuli that were used in the event-related paradigms were presented here in separate blocks. In each block 20 sequences were presented each for 300 ms followed by 500 ms blank. During this experiment, subjects performed the fixation point color-change detection task (5 color changes of the fixation point per block).

Confirming our ROI analysis, we observed significantly stronger activations for Glass sequences than segmented/random (p<0.01) in hMT+/V5, V3b/KO and LOC. These activations did not cover the full extent of these areas as defined by independent localizers. Moreover, consistent with previous studies of motion-related areas, we also observed significantly stronger (p<0.01) activations for real and implied than random

motion sequences anterior to hMT+/V5 (Senior et al. 2000; Zeki et al. 1993) and along the intra parietal sulcus (Claeys et al. 2003).

DISCUSSION

Our study demonstrates the following main findings that advance our understanding of interactions between form and motion processing in the human visual cortex. First, subpopulations in the human motion complex were shown to be selective for different types of global Glass pattern sequences that imply motion. Second, this selectivity in hMT+/V5 to implied motion was about 50% of the selectivity for real motion sequences. Importantly, we provide evidence for significant overlap between the subpopulations that encode implied motion and those that encode real motion. A similar pattern of selectivity was found in other dorsal areas, notably, V3, V3a and V3b/KO. Taken together we interpret the selectivity for global implied motion patterns and the significant overlap of the populations that are selective for real and implied motion as true selectivity for implied motion. Third, in early visual areas (V1, V2,) we observed selectivity for local orientation, but no selectivity for the global structure of Glass sequences. Thus, a significant part of the selectivity observed in the dorsal areas could not be attributed to selectivity for local orientation changes inherited from V1, but were due to the global arrangement of the sequences. Finally, ventral areas – V4 and LOC – were also selective for the global structure of Glass sequences and real motion sequences. The main difference with the dorsal areas was that in the ventral areas there was much less overlap between the populations selective for the global structure of Glass sequences and the populations selective for the global structure of real motion sequences (Figure 4). In sum,

29

our studies provide novel evidence that both dorsal and ventral visual areas in the human brain contribute to the selective processing of global motion from form. Dorsal motion-related areas represent global pattern sequences independent of the motion cue that defines them (real or implied) and thus may mediate the perception of implied motion from static forms. In contrast, this cue-invariance was not evident in ventral areas suggesting that these areas are better able to distinguish real from implied sequences; this may support the fine discrimination of the global structure in such patterns.

fMRI ADAPTATION

The fMRI adaptation paradigm (Grill-Spector and Malach 2001) is being used in an increasing number of studies to uncover selective subpopulations of neurons at a resolution below that of the typical human fMRI voxel. The main assumption behind this paradigm is that if a neuron responds to a pattern (pattern selectivity) it will respond less to the second presentation of a pattern than to the first. This assumption has considerable support in early sensory areas (e.g. V1: (Movshon and Lennie 1979), MT: (Petersen et al. 1985)). In psychophysics (Blakemore and Campbell 1969) and functional imaging (Grill-Spector and Malach 2001) a non-invasive measure of pattern selective adaptation is used to infer neuronal selectivity. Specifically, to infer whether a population can distinguish pattern A from B, one presents two identical patterns successively (AA, as well as BB) and measures whether this leads to a smaller response than presenting two different patterns successively (AB and BA). If both the AA and BB responses are smaller than the AB and BA responses, then there must be two separate mechanisms that respond and adapt to A and B.

The validity of adaptation fMRI has not yet been confirmed by detailed comparisons of adaptation effects at the neuronal and imaging level. Initial reports (Sawamura et al. 2004), however, suggest that although successive presentation of identical stimuli does typically reduce responses, there are more complicated sequence effects. Some of these properties may be area-specific as witnessed by the qualitatively different adaptation effects found in, for instance, area MT compared to area V1 (Kohn and Movshon 2004; Krekelberg et al. 2005). Moreover, there are still many uncertain steps relating even neuronal activity to the BOLD response (Logothetis and Wandell 2004). Hence, even if a reduced BOLD response is observed in an adaptation paradigm, the underlying mechanism need not be neuronal pattern selective adaptation.

These caveats notwithstanding, adaptation fMRI has been successful in various contexts in which the results could be verified at least indirectly with intracortical recordings. Examples are orientation selectivity in visual areas (Boynton and Finney 2003; Kourtzi et al. 2003; Tootell et al. 1998), and motion selectivity in area MT (Huettel et al. 2004; Huk et al. 2002; Huk et al. 2001; Tolias et al. 2001). Our study is another example; monkey single cell recordings showed implied motion selectivity in MT (Krekelberg et al. 2003) and our adaptation fMRI revealed selectivity in the human motion complex. However, while some studies, including the current one, could demonstrate orientation selectivity in V1 with adaptation techniques (Kourtzi et al. 2003; Tootell et al. 1998) others could not (Boynton and Finney 2003). These discrepancies may be due to differences in the stimuli (Boynton and Finney used low spatial frequency stimuli that were non-optimal for typical

V1 cells), but it also points to possible area differences in adaptation and the effect it has on the BOLD signal. Hence until a fuller validation of the adaptation fMRI paradigm has been obtained with intracortical recordings, statements regarding underlying neuronal selectivity made on the basis of imaging data alone should be treated with caution. Imaging data, however, certainly can be suggestive of and consistent with neuronal selectivity. Treated as a piece of evidence in favor of such selectivity - not the final and conclusive answer - they can be highly valuable.

GLASS PATTERNS AND IMPLIED MOTION

The percept of motion in a sequence of Glass patterns is so convincing that many observers find it difficult to believe that the motion energy in these sequences is just as balanced as the motion energy in a sequence of random dot patterns. Appendix A provides a mathematical proof of this theorem, which states that in both random dot sequences and in Glass sequences, the average motion energy spectrum is symmetric: The motion energy in any direction is on average matched by the motion energy in the opposite direction.

Motion energy detectors rely on asymmetries in the stimulus power spectrum (Adelson and Bergen 1985). Because the average power spectrum of a Glass pattern is symmetric, a motion energy detector will not assign a consistent direction of motion to it. On the basis of the motion energy distribution alone such a stimulus is therefore not expected to lead to a consistent coherent motion percept. This is our reason for referring to these balanced motion energy patterns as containing no coherent motion. When we speak of

implied motion in Glass patterns, we refer to the percept of coherent motion generated in the absence of coherent motion signals. By contrast, when we speak of real motion, we refer to the percept generated by stimuli whose average power spectrum is asymmetric and therefore contain unambiguous coherent motion signals.

The theorem in Appendix A, however, applies to the *average* motion energy in all Glass pattern sequences of a particular type. Due to the stochastic nature of the placement of the dots, the motion energy at any given time can be larger in any given direction, just as in random dot sequences. This leaves open the possibility that any individual Glass sequence does contain an asymmetry and that this asymmetry is detected by a simple motion energy detector. However, if these accidental coherent motion components were the underlying cause of the perceived direction of motion, one would expect that a reversal of the sequence would lead to a reversal of the perceived direction of motion. Krekelberg et al (2003) demonstrated that this prediction is not borne out by the data. We conclude that the implied motion percept is not generated by stochastic fluctuations in the motion energy of the stimulus.

Some stage of processing beyond motion energy extraction is therefore required to explain why Glass patterns appear to move coherently. In this context it is instructive to note that some stimuli with balanced motion energy fail to evoke a coherent motion percept. In a random dot sequence, for instance, the balanced motion energy does not lead to the percept of a globally coherent direction of motion. Instead, it results in a percept of directed motion that changes rapidly over time, and is presumably driven by

the stochastic fluctuations in dot placement. In the Glass pattern sequences, on the other hand, a globally coherent motion percept is evoked that alternates between only two directions over time (e.g. clockwise and counterclockwise for concentric Glass patterns). This perceptual difference is the phenomenon whose neural mechanisms we wish to understand. The difference between random sequences and Glass sequences is one of form: both global form (concentric patterns) and local form (oriented elements). Hence, in this descriptive sense, the form of the Glass patterns generates the implied motion. Whether neural mechanisms make explicit use of this form information to enhance motion processing (Geisler 1999) is a topic of ongoing investigation.

IMPLIED MOTION SELECTIVITY IN THE HUMAN MOTION COMPLEX

A recent imaging study has reported not to find selectivity for Glass sequences in hMT+/V5 (Wade et al. 2004). We believe this was due to mainly two factors. First, the authors used a restrictive definition of hMT+/V5; they choose voxels that responded better to coherent motion than to random motion. Many MT cells in the macaque, however, respond vigorously to random motion (Churan and Ilg 2002; Krekelberg and Albright 2005; Thiele et al. 2000). Our definition of hMT+/V5 included these cells. Secondly, the block design of Wade et al. looked for voxels that responded better to Glass sequences than to random motion. However, if cells responsive to random motion and implied motion are spatially interleaved, no difference in activation between the blocks would be expected. The adaptation technique, on the other hand, can resolve selectivity at a spatial resolution below that of a single voxel. Hence Wade et al.'s findings do not contradict ours; at the scale of MRI voxels there is no part of hMT+/V5

specialized for the detection of implied motion, but our study shows that hMT+/V5 does contain spatially intermingled subpopulations of neurons selective for implied motion.

MECHANISMS FOR IMPLIED MOTION PERCEPTION

Our data confirm and extend the findings of Krekelberg et al. (2003) who showed that a subpopulation of cells in the macaque superior temporal sulcus (MT and MST) responds to implied motion as if it is real motion. These cells responded best to those Glass patterns that evoked the strongest sense of motion and the cells' selectivity for real motion carried over to implied motion. This generalization suggests that these cells extract motion signals independent of the real or implied motion cue that delivers the signal. This cue-invariance was also evident in the imaging data obtained from dorsal motion areas, but not ventral areas.

The implied motion selectivity in the dorsal stream could arise from direct judicious sampling of neurons in the early visual areas (V1, V2). Both the single cell data (Smith et al. 2002) and our imaging data show that these areas contain the necessary local orientation information. Moreover, in MT, a subset of cells responds to oriented features that are aligned along the preferred direction of motion of the cell (Albright 1984; Maunsell and Van Essen 1983). The oriented form features of Glass patterns could activate these cells and thereby signal motion along those oriented features. In line with the percept of motion in plaids, these type-II MT cells respond to plaid stimuli containing two directions of motion as if they contain a single direction of motion (Rodman and Albright 1989). Hence, they are promising candidates for the computation needed to

extract coherent motion from the multiple balanced motion signals in Glass patterns. At the same time, however, it is important to note that the majority of MT cells (~80%) respond preferentially to oriented features orthogonal to their preferred direction of motion (Albright 1984). Activation of these cells by Glass patterns would signal motion orthogonal to the motion implied by the pattern. Hence, even though some MT cells are expected to respond simply to the oriented features in the Glass patterns, this by itself does not explain why motion is perceived along and not orthogonal to the oriented features.

Alternatively, implied motion selectivity in dorsal areas could arise from interactions with early form areas (V4, LOC). Our data support this view in that V4 and LOC contain subpopulations selective for the global structure in these patterns. Such an explicit use of orientation information could improve motion processing at high velocities (Geisler 1999). The temporal resolution of current functional imaging is not high enough to test whether selective responses in form areas precede those in the motion areas, hence we cannot resolve whether implied motion selectivity arises directly from V1 and V2 or through an interaction with V4 and the LOC. However, our study does provide electrode guidance for intracortical recordings that could address this issue. Area V4 and the inferotemporal cortex (IT) of the monkey have form selectivity that is comparable to that of V4 and the LOC in humans (Denys et al. 2004; Gallant et al. 2000; Kobatake and Tanaka 1994). Using simultaneous single cell recordings from areas that respond to complex form (such as V4 or IT) and MT, it should be possible to test the hypothesis that an interaction between form and motion areas underlies our perception of implied

motion. It would be particularly interesting to determine whether that same interaction also underlies the perception of motion in the more cognitive implied motion images, such as a cup about to fall from a table (Kourtzi and Kanwisher 2000a; Senior et al. 2000).

MECHANISMS FOR GLOBAL FORM PERCEPTION

We concentrated on the implied motion percept generated by Glass pattern sequences. These patterns, however, also generate a strong perception of global form. In fact, most of the work on Glass patterns has concentrated on how the local elements that carry the orientation information are bound together to generate the global form percept. Behavioral (Wilson and Wilkinson 1998) and recent ERP studies (Pei et al. 2005) show that the detection of structure in radial or concentric Glass patterns involves pooling of local information beyond the scale of typical V1 receptive fields. A case study by Gallant et al.(2000) strongly implicated V4 because a lesion involving V4 significantly disrupted a patient's ability to detect the global structure in Glass patterns. Single cell studies in monkeys also point to areas beyond V1 because selectivity for so-called non-Cartesian gratings and complex object features arises in V4 (Gallant et al. 1993; Gallant et al. 1996; Kobatake and Tanaka 1994). Moreover, a recent psychophysical study (Clifford and Weston 2005) provides evidence that adaptation to the global structure of Glass patterns also has two components, one likely to originate in the local orientation detectors of V1, the other likely to originate in cells with much larger receptive fields and therefore presumably in extrastriate areas. Our data are compatible with this view. Local

37

orientation selectivity – even for the noisy oriented elements in a Glass pattern – was evident in V1, but selectivity for the global organization of a Glass pattern was only observed in later ventral areas (V4, LOC). In agreement with previous models this suggests that global form selectivity starts to arise in V4 from appropriate pooling of V1 orientation detectors.

Interestingly, ventral areas were also selective for the structure of global *motion* patterns. Previous work has also documented this overlap of sensitivity for motion and form (Braddick et al. 2000; Denys et al. 2004). Our data suggest, however, that the selectivity for real motion is largely carried by a different subpopulation of cells than those selective for the structure of Glass patterns. In other words, while the representation of motion information in the human motion complex has a significant degree of invariance with respect to implied and real motion cues, the representation of the form of those same patterns in V4 and LOC is not cue-invariant. The reason for this may be that while motion processing is essentially complete once the velocity is estimated, form information may have to be processed in much more detail to extract further information on the objects that generate it. For such a detailed analysis an early stage of cue-invariant processing would be detrimental.

APPENDIX

Let P be a sequence in which a new random dot pattern is chosen every time step. P is a function of two spatial dimensions (x, y) and time (t). The Fourier transform of the sequence is denoted by F(P), it is a function of spatial frequencies (k_x, k_y) and temporal frequency (ω). The power spectrum of P is written as $F(P)F^*(P)$, where * denotes the complex conjugate. Because dot patterns are chosen randomly in each frame of the sequence, there are, on average, no space-time correlations in P. Averaged over time, each direction of motion is equally likely to occur, and the motion signals are balanced. In Fourier space, the balance in directional signals results in a stimulus power spectrum that is symmetric around the ω-axis in k-ω space. We refer to such a stimulus as one that has no coherent motion energy. (By contrast, a stimulus whose power spectrum is consistently oriented in k-ω space is referred to as a real motion stimulus.)

To construct a Glass pattern sequence from a sequence of random dot patterns, each element in the original sequence P is shifted along some one-dimensional coordinate. For translational Glass patterns this is a simple spatial translation, for concentric and radial Glass patterns this is a translation in polar coordinates. We will represent this translation/rotation/expansion operation by the operator S, which shifts a pattern by an amount s. With this notation, a Glass pattern (G) is simply the sum of a random dot pattern sequence and that same sequence after the operation S.

$$G = P + SP$$

To assess the motion energy in the Glass pattern, we determine its power spectrum:

$$|F(G)|^2 = |F(P+SP)|^2$$

Because the Fourier transform is a linear operation, we can write this as:

$$|F(G)|^2 = |F(P)+F(SP)|^2$$

$$= (F(P)+F(SP)) \, (\, F(P)+F(SP))*$$

Because the operator S is a simple translation in space-time, it can be written as a multiplication by e^{iks} in the Fourier domain. The spatial frequency k here refers to the dimension along which the operator S operates, s is the amount by which operator S shifts P (the Glass-shift). With the transformation $F(SP) = e^{iks}F(P)$., we can write the above formula as:

$$= (F(P) + e^{iks}F(P))(F(P)+ e^{iks}F(P))*$$

$$= F(P)F^*(P) + e^{iks}F(P) \, e^{-iks}F^*(P) + e^{iks}F(P)+ e^{-iks}F^*(P)$$

$$= |F(P)|^2 \; (2+ e^{iks} + e^{-iks})$$

$$= 2|F(P)|^2 \, (1+\cos(ks))$$

In words: the power spectrum of the Glass pattern is equal to the power spectrum of the underlying random dot pattern ($|F(P)|^2$) times a factor that depends only on the *spatial* frequency k. Importantly, because $\cos(ks) = \cos(-ks)$, this term is symmetric in k. The power spectrum of the random dot pattern $|F(P)|^2$ is symmetric in k-ω. The product of two symmetric functions, however, is also symmetric. Hence, the power spectrum of the Glass pattern sequence is symmetric in k-ω space and contains no coherent motion energy.

FIGURES

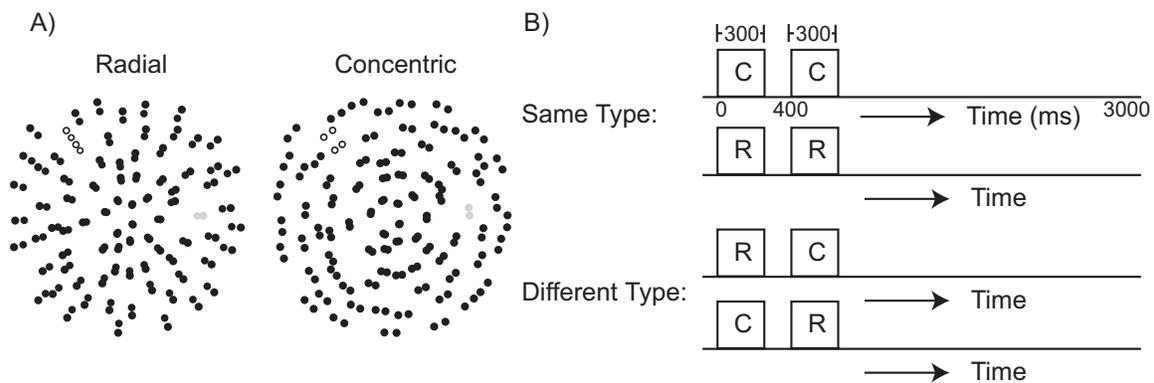Figure 1 Krekelberg et al.

Figure 1 Stimuli and design.

A) Glass patterns are created by placing pairs of dots aligned along a common pathway at random positions on the screen (Glass 1969). The examples show a common concentric pathway (concentric Glass) and a common radial pathway (radial Glass). The open and grey circle pairs highlight two examples of elements that are orthogonal to each other in the two types of patterns. In the experiment all dots were white on a black background. In a Glass sequence, randomly chosen patterns of the same type (radial or concentric) are presented in rapid succession. B) The temporal design of the trials. In each trial, two sequences with a duration of 300ms and an inter pattern interval of 100ms were shown. In the 'same type' conditions, two sequences of the same type were shown; concentric-concentric or radial-radial. In the 'different type' conditions, two

sequences of different type where shown. In these examples; a radial sequence followed by a concentric sequence or a concentric sequence followed by a radial sequence. After the visual stimuli had been extinguished, the fixation point remained on for another 2.3s after which the next trial started.

Figure 2   Krekelberg et al.
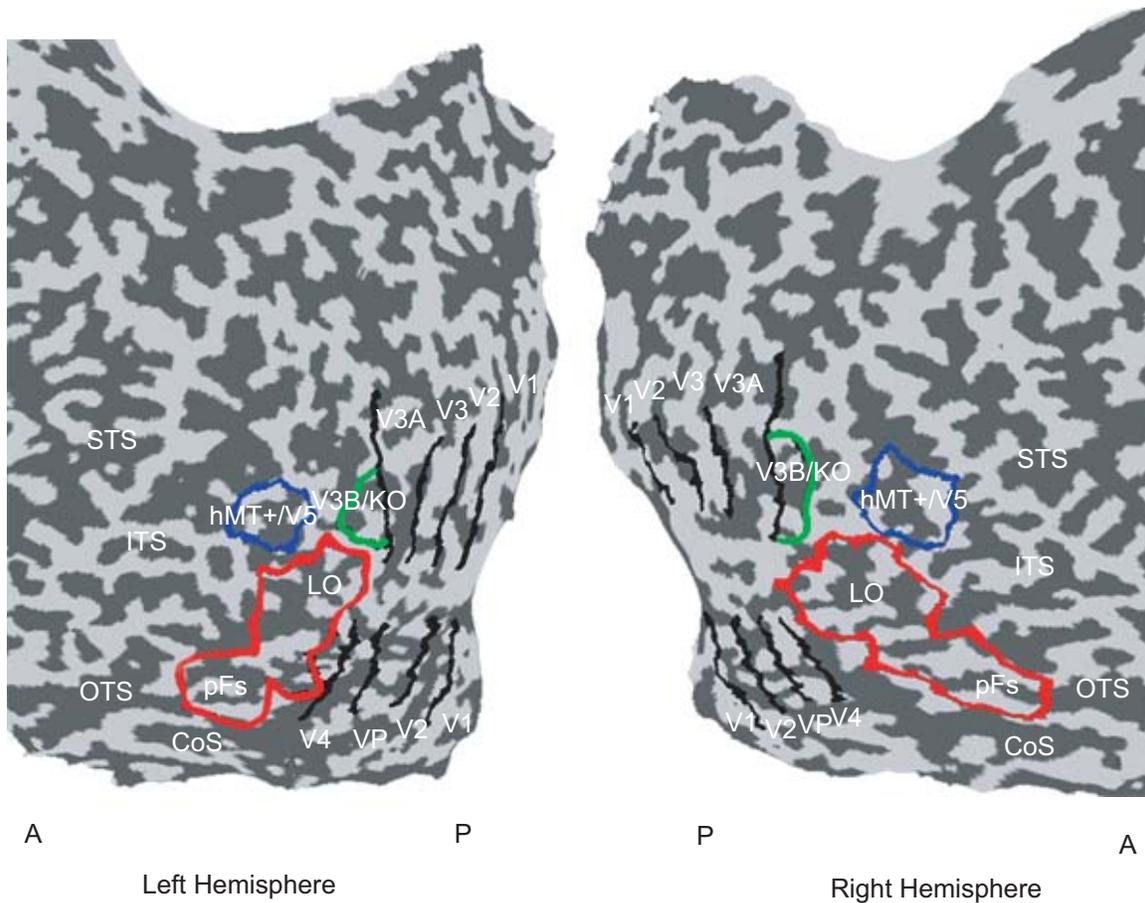
Figure 2 Regions of interest.

Functional activation maps for one subject showing the retinotopic ventral (V1, V2, VP, V4) and dorsal (V1, V2, V3, V3a) areas, V3b/KO, hMT+/V5, and LOC. Functional activations are superimposed on flattened cortical surfaces of the right and left hemispheres. A: anterior, P: posterior. STS: superior temporal sulcus, ITS: inferior temporal sulcus, OTS: occipitotemporal sulcus, CoS: collateral sulcus.
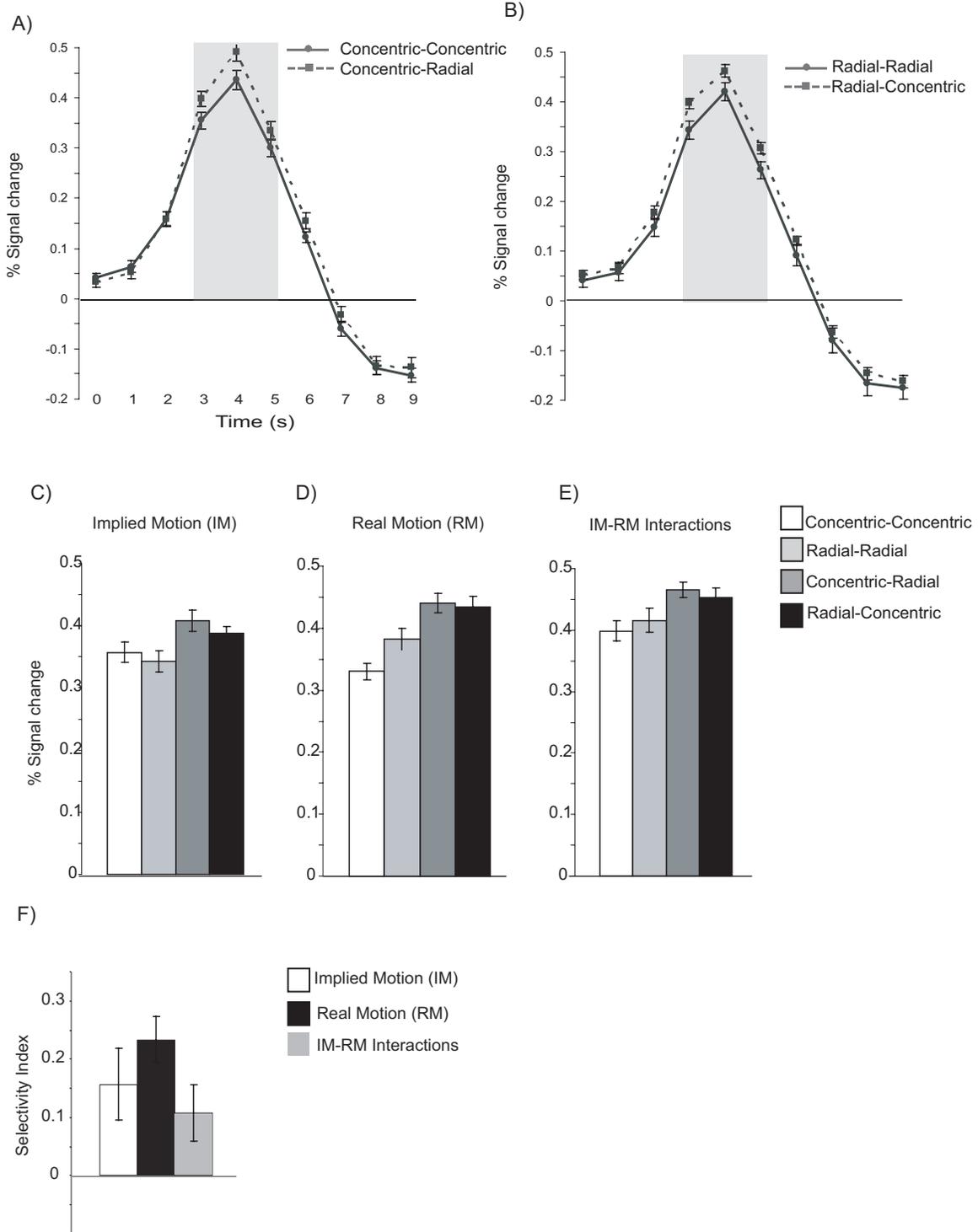
Figure 3 Implied and real motion selectivity in the human motion complex.

A) Averaged time courses showing the BOLD response in trials in which two concentric Glass sequences were shown (solid line) and the response in trials in which a concentric Glass sequence was followed by a radial Glass sequence (dotted line). Signal changes in all panels are relative to the fixation baseline. B) Averaged time courses showing the BOLD response to two successive radial Glass sequences (solid line) and the response to concentric Glass sequence followed by a radial Glass sequence (dashed line). C) The average BOLD response to four implied motion conditions (experiment 1). Calculated by averaging the BOLD signal around the peak response (gray window in panels A and B). Error bars represent standard errors in the mean. D) The average BOLD response to the four real motion conditions (experiment 2). E) The average BOLD response to the four interaction conditions (experiment 3; see main text). F) The selectivity index (SI:[(AB+BA)/(AA+BB)]-1) calculated from the response measures in panel B, C and D. Similar results were observed in the MT and MST subregions of hMT+/V5 that were identified based on retinotopic mapping techniques (Huk et al. 2002). Together, these data show that cells in the human motion complex that are selective for real radial versus concentric motion are also selective for implied motion of the same types.
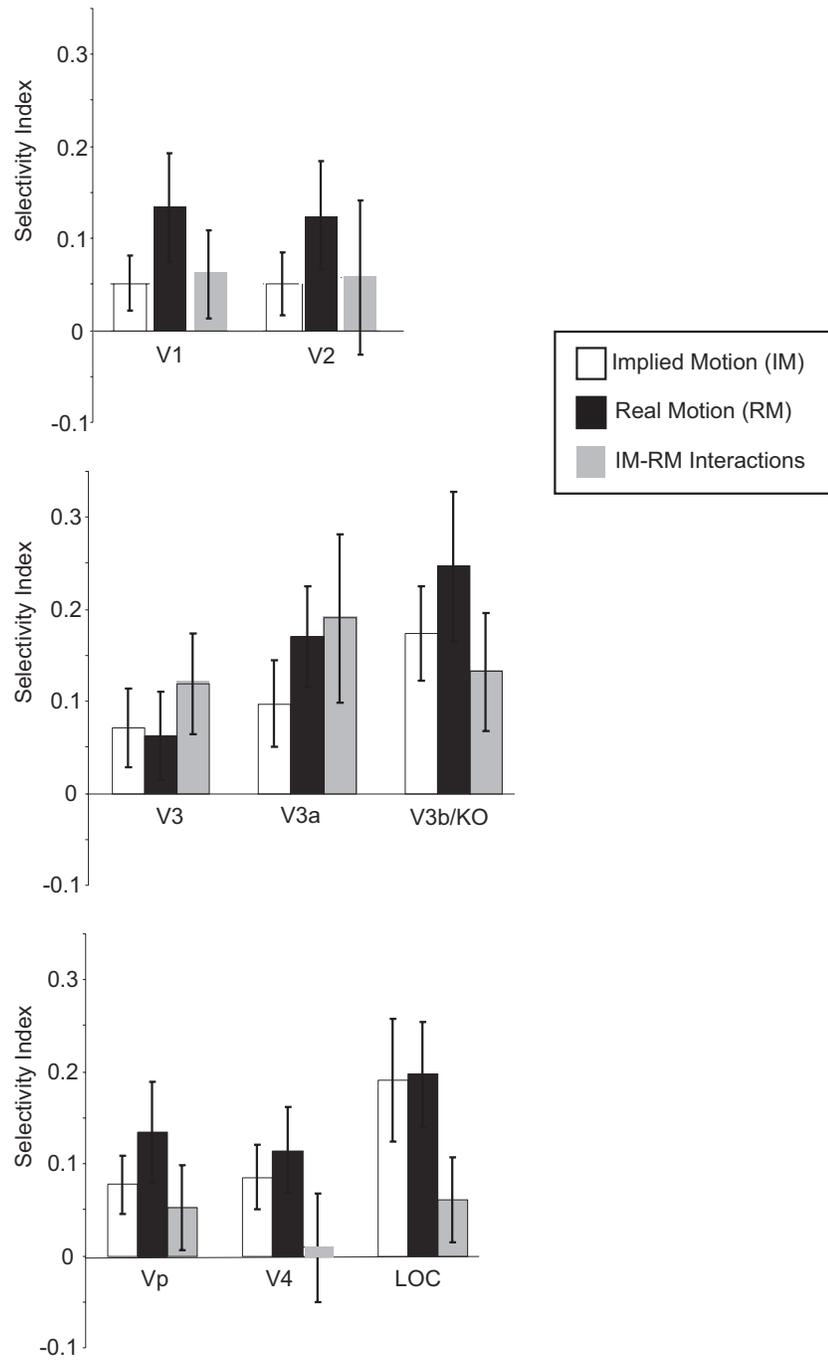
Figure 4 - Krekelberg et al.

Figure 4 Implied and real motion selectivity in early visual areas, V3b/KO and the LOC.

A) The selectivity index for implied motion (white), real motion (black) and interactions between real and implied motion (gray). These indices were calculated from the intermediate steps shown for the motion complex in Figure 3 and are based on the data from experiments 1, 2, and 3, respectively. Error bars represent SEM; they are large because we used a conservative method that incorporates the error estimates of both numerator and denominator. Similar results were observed in two subregions of the LOC: the LO (lateral occipital) at the posterior part of the inferior-temporal sulcus and the pFs (posterior Fusiform) in the posterior fusiform gyrus. These results show that all areas contained some subpopulations that were selective for concentric versus radial real and implied motion, but that only in V3, V3a, and V3b/KO the same subpopulations were selective for real and implied motion.
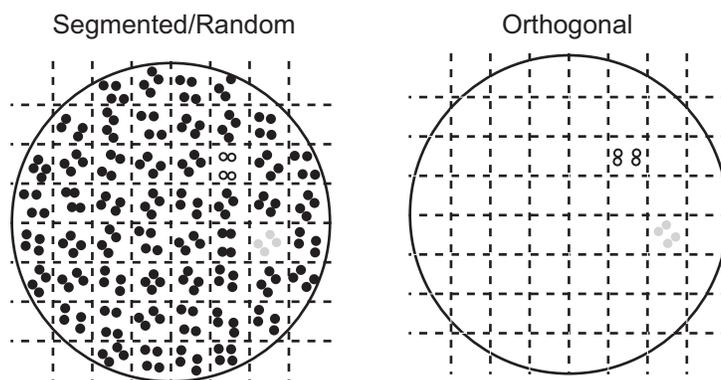
Figure 5 Krekelberg et al.

Segmented/Random   Orthogonal

Figure 5 Segmented/Random Glass sequences.

These sequences were constructed to generate a stimulus that has no global structure, but clear local orientation structure. The orientation of the pairs was assigned randomly to each of the 64 segments (indicated by the dashed grid that was not visible in the stimulus). All pairs within a patch have the same orientation. The orthogonal Glass sequence is matched to a segmented Glass sequence; in each segment, the orientation is orthogonal to the corresponding segment in the associated segmented Glass sequence. For clarity only two segments are shown, in the real stimuli all segments were filled with pairs of dots. The open circle and grey circle pairs highlight

two examples of elements that are orthogonal to each other. In the experiment all dots were white

on a black background.
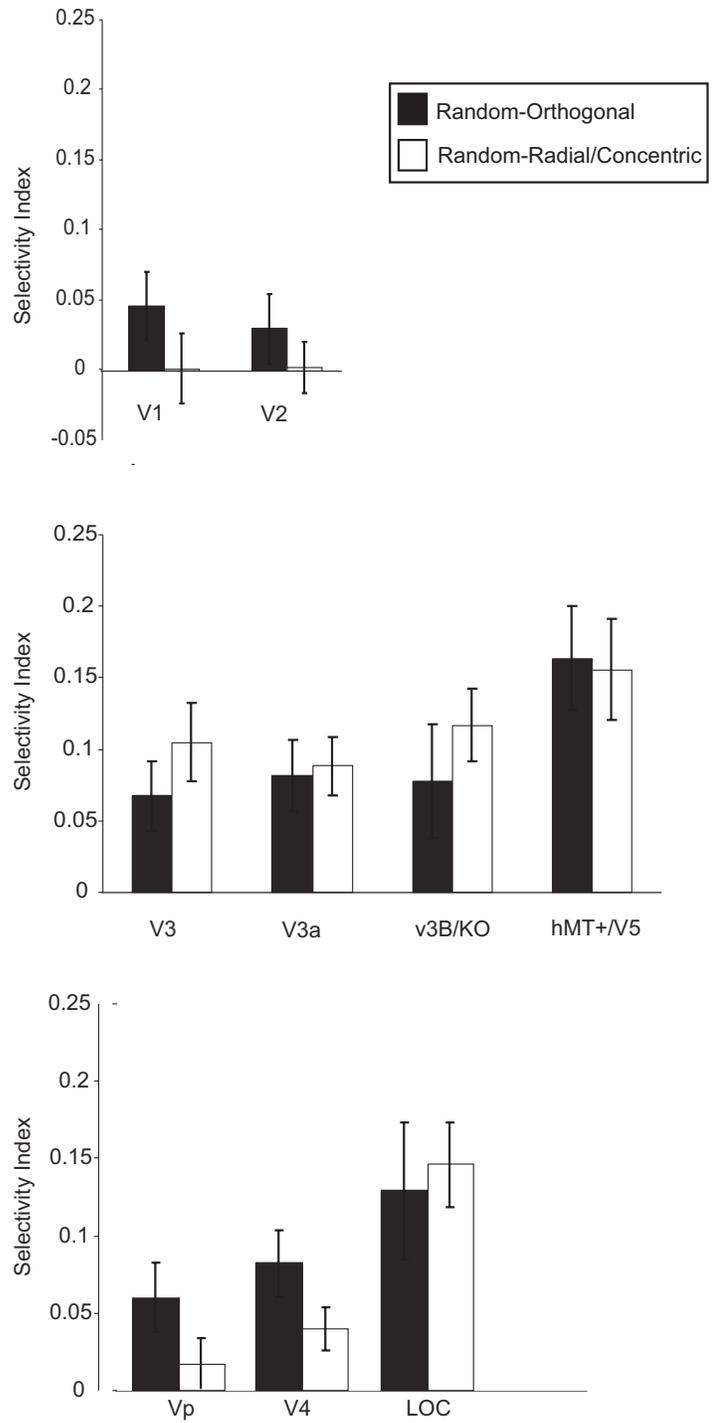
Figure 6   Krekelberg et al.

Figure 6 Local and global pattern selectivity.

Black bars show selectivity for a change from a segmented Glass pattern to an orthogonal Glass pattern. This selectivity is most likely attributable to simple orientation selectivity. White bars show the selectivity for a change from a segmented Glass sequence to a rotation Glass sequence. Error bars represent the SEM. These results show that local orientation selectivity in V1 may have contributed to the selectivity observed in higher areas (black bars). Some selectivity in higher areas, however, occurs without a measurable selectivity in V1 and V2 (white bars), suggesting that this global pattern selectivity arises in areas beyond V1 and V2.

REFERENCES

Adelson EH and Bergen JR. Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A* 2: 284-299, 1985.

Albright TD. Direction and orientation selectivity of neurons in visual area MT of the macaque. *J Neurophysiol* 52: 1106-1130., 1984.

Albright TD. Form-cue invariant motion processing in primate visual cortex. *Science* 255: 1141-1143., 1992.

Blakemore C and Campbell FW. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *J Physiol* 203: 237-260, 1969.

Boynton GM, Engel SA, Glover GH, and Heeger DJ. Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16: 4207-4221, 1996.

Boynton GM and Finney EM. Orientation-specific adaptation in human visual cortex. *J Neurosci* 23: 8781-8787, 2003.

Braddick OJ, O'Brien JM, Wattam-Bell J, Atkinson J, and Turner R. Form and motion coherence activate independent, but not dorsal/ventral segregated, networks in the human brain. *Curr Biol* 10: 731-734., 2000.

Buckner RL, Goodman J, Burock M, Rotte M, Koutstaal W, Schacter D, Rosen B, and Dale AM. Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI. *Neuron* 20: 285-296, 1998.

Burr D. Motion vision: are 'speed lines' used in human visual motion? *Curr Biol* 10: R440-443, 2000.

Burr DC and Ross J. Direct evidence that "speedlines" influence motion mechanisms. *J Neurosci* 22: 8661-8664, 2002.

Churan J and Ilg UJ. Flicker in the visual background impairs the ability to process a moving visual stimulus. *Eur J Neurosci* 16: 1151-1162, 2002.

Claeys KG, Lindsey DT, De Schutter E, and Orban GA. A higher order motion region in human inferior parietal lobule: evidence from fMRI. *Neuron* 40: 631-642, 2003.

Clifford CW and Weston E. Aftereffect of adaptation to Glass patterns. *Vision Res* 45: 1355-1363, 2005.

Denys K, Vanduffel W, Fize D, Nelissen K, Peuskens H, Van Essen D, and Orban GA. The processing of visual shape in the cerebral cortex of human and nonhuman primates: a functional magnetic resonance imaging study. *J Neurosci* 24: 2551-2565, 2004.

DeYoe EA, Carman GJ, Bandettini P, Glickman S, Wieser J, Cox R, Miller D, and Neitz J. Mapping striate and extrastriate visual areas in human cerebral cortex. *Proc Natl Acad Sci U S A* 93: 2382-2386, 1996.

Dupont P, De Bruyn B, Vandenberghe R, Rosier AM, Michiels J, Marchal G, Mortelmans L, and Orban GA. The kinetic occipital region in human visual cortex. *Cereb Cortex* 7: 283-292, 1997.

Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky EJ, and Shadlen MN. fMRI of human visual cortex. *Nature* 369: 525, 1994.

Gallant JL, Braun J, and Van Essen DC. Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science* 259: 100-103, 1993.

Gallant JL, Connor CE, Rakshit S, Lewis JW, and Van Essen DC. Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. *J Neurophysiol* 76: 2718-2739, 1996.

Gallant JL, Shoup RE, and Mazer JA. A human extrastriate area functionally homologous to macaque V4. *Neuron* 27: 227-235., 2000.

Geisler WS. Motion streaks provide a spatial code for motion direction. *Nature* 400: 65-69, 1999.

Glass L. Moiré effect from random dots. *Nature* 223: 578-580, 1969.

Grill-Spector K, Kushnir T, Edelman S, Itzchak Y, and Malach R. Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron* 21: 191-202, 1998.

Grill-Spector K and Malach R. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol (Amst)* 107: 293-321, 2001.

Henson RN and Rugg MD. Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia* 41: 263-270, 2003.

Huettel SA, Obembe OO, Song AW, and Woldorff MG. The BOLD fMRI refractory effect is specific to stimulus attributes: evidence from a visual motion paradigm. *Neuroimage* 23: 402-408, 2004.

Huk AC, Dougherty RF, and Heeger DJ. Retinotopy and functional subdivision of human areas MT and MST. *J Neurosci* 22: 7195-7205, 2002.

Huk AC and Heeger DJ. Pattern-motion responses in human visual cortex. *Nat Neurosci* 5: 72-75, 2002.

Huk AC, Ress D, and Heeger DJ. Neuronal basis of the motion aftereffect reconsidered. *Neuron* 32: 161-172, 2001.

Jellema T and Perrett DI. Cells in monkey STS responsive to articulated body motions and consequent static posture: a case of implied motion? *Neuropsychologia* 41: 1728-1737, 2003.

Kobatake E and Tanaka K. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *J Neurophysiol* 71: 856-867, 1994.

Kohn A and Movshon JA. Adaptation changes the direction tuning of macaque MT neurons. *Nat Neurosci* 7: 764-772, 2004.

Kourtzi Z, Bulthoff HH, Erb M, and Grodd W. Object-selective responses in the human motion area MT/MST. *Nat Neurosci* 5: 17-18, 2002.

Kourtzi Z and Kanwisher N. Activation in human MT/MST by static images with implied motion. *J Cogn Neurosci* 12: 48-55, 2000a.

Kourtzi Z and Kanwisher N. Cortical regions involved in perceiving object shape. *J Neurosci* 20: 3310-3318, 2000b.

Kourtzi Z and Kanwisher N. Representation of perceived object shape by the human lateral occipital complex. *Science* 293: 1506-1509, 2001.

Kourtzi Z, Tolias AS, Altmann CF, Augath M, and Logothetis NK. Integration of local features into global shapes: monkey and human FMRI studies. *Neuron* 37: 333-346, 2003.

Krekelberg B and Albright TD. Motion mechanisms in macaque MT. *J Neurophysiol* 93: 2908-2921, 2005.

Krekelberg B, Dannenberg S, Hoffmann KP, Bremmer F, and Ross J. Neural correlates of implied motion. *Nature* 424: 674-677, 2003.

Krekelberg B, Van Wezel RJ, and Albright TD. Adaptation in Macaque MT reduces Perceived Speed and Improves Speed Discrimination. Submitted, 2005.

Kruggel F, von Cramon DY, and Descombes X. Comparison of filtering methods for fMRI datasets. *Neuroimage* 10: 530-543, 1999.

Logothetis NK and Wandell BA. Interpreting the BOLD signal. *Annu Rev Physiol* 66: 735-769, 2004.

Maunsell JH and Van Essen DC. Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *J Neurophysiol* 49: 1127-1147, 1983.

Mishkin M, Ungerleider L, and Macko K. Object vision and spatial vision: two central pathways. *Trends Neurosci* 6: 414-417, 1983.

Movshon JA and Lennie P. Pattern-selective adaptation in visual cortical neurones. *Nature* 278: 850-852, 1979.

Pei F, Pettet MW, Vildavski VY, and Norcia AM. Event Related Potentials show configural specificity of global form processing. *Neuroreport* In press, 2005.

Petersen SE, Baker JF, and Allman JM. Direction-specific adaptation in area MT of the owl monkey. *Brain Res* 346: 146-150, 1985.

Ringach DL, Shapley RM, and Hawken MJ. Orientation selectivity in macaque V1: diversity and laminar dependence. *J Neurosci* 22: 5639-5651, 2002.

Rodman HR and Albright TD. Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Exp Brain Res* 75: 53-64, 1989.

Ross J, Badcock DR, and Hayes A. Coherent global motion in the absence of coherent velocity signals. *Curr Biol* 10: 679-682., 2000.

Sawamura H, Vogels R, and Orban GA. Single cell investigation of stimulus adaptation in monkey inferior temporal cortex. *Society for Neuroscience Abstracts*: 824.821, 2004.

Senior C, Barnes J, Giampietro V, Simmons A, Bullmore ET, Brammer M, and David AS. The functional neuroanatomy of implicit-motion perception or representational momentum. *Curr Biol* 10: 16-22, 2000.

Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, and Tootell RB. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268: 889-893, 1995.

Smith MA, Bair W, and Movshon JA. Signals in macaque striate cortical neurons that support the perception of Glass patterns. *J Neurosci* 22: 8334-8345, 2002.

Thiele A, Dobkins KR, and Albright TD. Neural correlates of contrast detection at threshold. *Neuron* 26: 715-724, 2000.

Tolias AS, Smirnakis SM, Augath MA, Trinath T, and Logothetis NK. Motion processing in the macaque: revisited with functional magnetic resonance imaging. *J Neurosci* 21: 8594-8601, 2001.

Tootell RB, Hadjikhani NK, Vanduffel W, Liu AK, Mendola JD, Sereno MI, and Dale AM. Functional analysis of primary visual cortex (V1) in humans. *Proc Natl Acad Sci U S A* 95: 811-817, 1998.

Wade AR, Vildavski VY, Pettet M, Tyler CW, and Norcia AM. Glass patterns look like noise to human MT. *Society for Neuroscience Abstracts*: 935.910, 2004.

Watson JD, Myers R, Frackowiak RS, Hajnal JV, Woods RP, Mazziotta JC, Shipp S, and Zeki S. Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cereb Cortex* 3: 79-94, 1993.

Wilson HR and Wilkinson F. Detection of global structure in Glass patterns: implications for form vision. *Vision Res* 38: 2933-2947, 1998.

Zeki S, Watson JD, and Frackowiak RS. Going beyond the information given: the relation of illusory visual motion to brain activity. *Proc Biol Sci* 252: 215-222, 1993.