Human cortical and behavioral sensitivity to patterns of complex motion at eccentricity

Ryan T. Maloney,1,2 Tamara L. Watson,3 and Colin W. G. Clifford1,2
1Colour, Form and Motion Laboratory, School of Psychology, Faculty of Science, The University of Sydney, Sydney, New South Wales, Australia; 2Australian Research Council Centre of Excellence in Vision Science, The University of Sydney, Sydney, New South Wales, Australia; and 3School of Social Sciences and Psychology, The University of Western Sydney, Bankstown, New South Wales, Australia

Submitted 17 June 2013; accepted in final form 8 September 2013

Maloney RT, Watson TL, Clifford CW. Human cortical and behavioral sensitivity to patterns of complex motion at eccentricity. J Neurophysiol 110: 2545–2556, 2013. First published September 11, 2013; doi:10.1152/jn.00445.2013.—Complex patterns of image motion (contracting, expanding, rotating, and spiraling fields) are important in the coordination of visually guided behaviors. Whereas specialized detectors in monkey visual cortex show selectivity for particular patterns of complex motion, their representation in human visual cortex remains unclear. In the present study, functional magnetic resonance imaging (fMRI) was used to investigate the sensitivity of functionally defined regions of human visual cortex to parametrically modulated complex motion trajectories, coupled with complementary psychophysical testing. A unique stimulus design made it possible to disambiguate the neural responses and psychophysical sensitivity to complex motions per se from the distribution of local motions relative to the fovea, which are known to enhance cortical activity when presented radial to fixation. This involved presenting several small, separate motion fields in the periphery in a manner that distinguished them from global optic flow patterns. The patterns were morphed through complex motion space in a systematic time-locked fashion when presented in the scanner. Anisotropies were observed in the fMRI signal, marked by an enhanced response to expanding vs. contracting fields, even in early visual cortex. Anisotropies in the psychophysical sensitivity measures followed a similar pattern that was correlated with activity in areas hV4, V5/MT, and MST. This represents the first systematic examination of complex motion perception at both a behavioral and neural level in human observers. The characteristic processing anisotropy revealed in both data sets can inform models of complex motion processing, particularly with respect to computations performed in early visual cortex.

complex motion perception; early visual cortex; fMRI; optic flow; psychophysics

AS AN OBSERVER MOVES forward or backward through space, characteristic motion flow patterns stimulate the eyes. These patterns, known as “complex motions,” provide important information useful in guiding navigation and locomotion through the three-dimensional (3-D) environment in humans and animals (Gibson 1950; Koenderink 1986; Lee 1980; Srinivasan and Zhang 2004) and can also arise as smaller patches anywhere in the visual field as a result of object motion (Regan and Beverley 1978). Complex motions are parameterized by the flow angle, or the frame-by-frame change in the angular trajectory of an element as it moves across space, and its speed, which scales linearly with distance from the focus of radial motion (i.e., the origin). Different flow angles produce phenomenologically different impressions of motion, such as expanding, contracting, rotating, or spiraling patterns (see Fig. 1).

The dorsal aspect of macaque medial superior temporal cortex (MSTd) contains neurons selective for complex motions, with a preponderance tuned to expansion (Duffy and Wurtz 1991a, 1991b; Graziano et al. 1994; Lagae et al. 1994; Orban et al. 1992; Tanaka et al. 1989; Tanaka and Saito 1989). Accordingly, MSTd and a network of posterior parietal structures that also show complex motion selectivity, including area 7a (Siegel and Read 1997; Steinmetz et al. 1987) and the ventral intraparietal (VIP) cortex (Schaafsma and Duysens 1996), are reputed to perform multimodal computations important in the control of self-motion, such as the estimation of heading and time-to-contact of obstacles (Britten and Van Wezel 1998, 2002; Chen et al. 2013; Mineault et al. 2012; Perrone and Stone 1998; Raffi and Siegel 2007). Meanwhile, the smaller receptive fields of early visual cortex, tuned to local translational motions, are generally not considered appropriate for this type of analysis (Burr et al. 1998; Vaina 1998); meaning little is understood about the representation of complex motion in early visual cortex, especially that of the human.

Although psychophysical work agrees that anisotropies exist in complex motion perception, the conclusions are mixed, with some studies demonstrating a bias for expansion (Ball and Sekuler 1980) and others for contraction (Edwards and Badcock 1993). Other systematic studies of complex motion perception have generally found similar sensitivity for expansion and contraction, albeit still greater than that of rotational (Beardsley and Vaina 2005) or spiral patterns (Burr et al. 2001; Morrone et al. 1999). Functional imaging suggests that subparts of the putative human homolog of the macaque middle temporal complex, V5/MT+ (Dumoulin et al. 2000; Zeki et al. 1991), respond more to certain complex motions than translational ones (Holliday and Meese 2005, 2008; Koyama et al. 2005; Morrone et al. 2000; Wall and Smith 2008); yet, a consensus remains elusive, with some studies showing a preference for expansion and rotation (Holliday and Meese 2005; Wall et al. 2008) and others for contraction (Giaschi et al. 2007).

Although the inconsistency in the literature is probably due to variations in the stimuli and tasks involved, another possibility is that the measured sensitivity to complex motion has been confounded by the distribution of local motion biases relative to fixation, whereby topographic neural populations in V1–V3 show a preference for motions following a radial
trajectory toward or away from the fovea (Clifford et al. 2009; Raemaekers et al. 2009). In the present work, complementary behavioral and functional magnetic resonance imaging (fMRI) methods, with the use of a uniquely designed stimulus, were used to probe the mechanisms sensitive to a wide and systematic range of variations in complex motion per se, carefully decoupled from local motion biases centered at the fovea. This approach involved centering separate, small complex motion patterns in the visual periphery, in an array encircling fixation. As such, the patterns were not intended to simulate the movement of an observer through 3-D space as optic flow fields might. Primarily, this arrangement was designed to target responses in early visual cortex and not necessarily MST, where most neurons require relatively large stimuli to provoke a strong response (Lagae et al. 1994; Raiguel et al. 1997). Translational motions were not tested. Anisotropies marked by a higher sensitivity to expanding over contracting flow patterns were evident in both (the fMRI and psychophysical) data sets, even in early retinotopic cortex.

MATERIALS AND METHODS

Subjects

Six experienced psychophysical observers, including the authors, participated in both the psychophysical and functional imaging components of the study (ages 25–45 yr; 2 women). All had normal or corrected-to-normal visual acuity. Optical corrections in the scanner took the form of customized prescription goggles worn by the participants. All procedures were approved by the University of Sydney Human Research Ethics Committee; MRI scanning was done with the additional approval of the St Vincent’s Hospital ethics office (Darlinghurst, NSW, Australia).

Visual Stimuli

Separate stimulus sets were used in measuring sensitivity to complex motion patterns and in the definition of cortical regions of interest (ROIs) and retinotopic mapping procedures (see details below). The complex motion stimuli consisted of random dot kinemats (RDKs) that were matched in the fMRI and psychophysical settings.

Figure 2 provides an example of the stimuli used, which ranged in flow angle in 16 steps of 22.5° from a motion flow angle of 0° (expansion; see also Fig. 1). RDKs (saved as AVI files) were generated, loaded, and presented using MATLAB (7.10 R2010a; The MathWorks, Natick, MA) and PsychToolbox 3.0.9 (Brainard 1997; Pelli 1997). As illustrated in Fig. 2, A and B, eight unique RDKs were centered at equal intervals along the circumference of an invisible circle centered at fixation and with a radius of 4°. The RDKs, which were set against a background of the mean luminance of the display, were centered in the periphery to deconfound local from global complex motion anisotropies in an approach similar to that used previously to study complex polar form (Mannion and Clifford 2011). Each RDK was presented within a circular aperture of radius 1.41°, with dot contrast ramped at the outer edges with a raised cosine profile of 0.09°. Stimuli thus subtended an annular region within an eccentricity of 2.59° to 5.41°.

The RDKs consisted of 100% peak contrast dots with a Laplacian of Gaussian (LoG) profile (Curran et al. 2009) with σ = 0.05° and a peak spatial frequency of 10.6 cycles/deg. The spatial profile of the isotropic LoG dots minimized the artificial orientation signals known as “motion streaks” resulting from the temporal integration of dot trajectories (Alais et al. 2011; Geisler 1999), which may have arisen with nonfiltered dot stimuli. The RDKs had a dot density of 5.9 dots/deg² and moved at an average speed of 1.78°/s. This average speed is possibly suboptimal for cells in V5/MT and MST (Lagae et al. 1994; Tanaka and Saito 1989), but it was chosen to further limit the potential influence of motion streaks (Geisler 1999). The contrast of each individual dot in the RDK was randomly assigned in the first frame to be either a full contrast increment or decrement from the background luminance, with equal numbers of each polarity. The “flow angle” of the RDKs described the frame-by-frame change in the angular trajectory of a dot, all relative to a 0° (radially expanding) baseline (Clifford et al. 1999; Graziano et al. 1994; Meese and Anderson 2002; Morrone et al. 1999). The complex motion fields of the RDKs were generated according to the algorithm given in detail in Clifford et al. (1999) such that the position of each dot on a given frame of the display was discretely sampled from a continuous trajectory according to its distance from the focus of radial motion (the origin). As such, dot speed was scaled by the distance from the origin.

In the first frame of a kinematogram, the randomly positioned dots had a uniform density per unit area of the stimulus aperture. In an expanding pattern, some of the dots would fall outside of the stimulus with succeeding frames. These dots were repositioned back into the stimulus aperture at a random location calculated such that a constant dot density was maintained across the stimulus for its entire duration, according to the “wraparound” procedure given in Appendix A.3 of Clifford et al. (1999). Contracting patterns were generated in the same manner as expanding patterns but with the order of the frames reversed. The number of dots appearing (disappearing) within the stimulus aperture is greatest for the purely expanding (contracting) motions, whereas the dots in the purely rotational motions never escape the inside of the stimulus aperture. Thus the “effective coherence,” or the correspondence between dots from one frame to the next, was set to the same value (99.3%) for all flow angles. This meant that the number of dots disappearing and being redrawn was balanced for all flow angles, including the rotations, and it served as an objective way to equalize across stimuli the number of local temporal transients accompanying dot appearance and disappearance.

Psychophysical Methods

Apparatus. Stimuli were viewed at a distance of 61 cm (stabilized via a chin rest) on a gamma-corrected 46-cm Diamond Digital CRT monitor (1,024 x 768 resolution) operating at 85 Hz and with a mean luminance of 25 cd/m². The display was driven by a personal computer (running at 8 GB of RAM) with an Intel Core i7-2600 3.4-GHz processor and AMD Radeon HD6350 graphics card. Testing was conducted in a darkened testing cubicle (ambient luminance, ~2 cd/m²). During testing of complex motions (in both the fMRI and psychophysics milieus),
subjects’ eye position was monitored using the EyeLink1000 eye tracking system (SR Research, Mississauga, ON, Canada), in conjunction with the EyeLink Toolbox (Cornelissen et al. 2002) for MATLAB.

**Design and procedure.** Perceptual sensitivity for the same set of 16 complex motions assessed with fMRI was also measured behaviorally. Sensitivity was defined as the inverse of the coherence threshold for detection of a particular motion flow angle, which was estimated in an eight-alternative forced choice (8-AFC) visual search task (Mannion and Clifford 2011). The task was performed in the form of a 30-trial run of the Psi adaptive staircase procedure (Kontsevich and Tyler 1999), where the coherence (i.e., the proportion of signal dots) of the RDKs was adjusted according to the posterior distribution of the threshold (and slope) parameters of the psychometric function at the latest point in the staircase. In the task, subjects made nonspeeded responses on the numeric keypad of a standard keyboard to indicate the spatial location of the single target motion flow angle (which they viewed at 100% coherence before the staircase began) among the eight-item array (Fig. 2, A and B). This location was assigned at random on each trial, and the distractors in the seven other positions of the array were unique RDKs of the “opposite” motion flow angle (i.e., 180° away from the target). This design was employed to avoid artifactual temporal cues to target location that were apparent when a single target was detected from among completely random motion (0% coherence) distractor RDKs. Setting the distractors to the opposite flow angle as the target allowed for the coherence (or motion energy) to be balanced for all stimuli in the display (target and distractors). In each new trial, all stimuli were reconstructed and redrawn at the updated coherence level. The RDKs were presented for 659 ms, with the contrast ramped on and off over the first and last 82 ms with a raised cosine profile. They were centered on a small fixation cross that appeared 1,000 ms before the arrival of the stimuli. Subjects were instructed to maintain fixation on the cross for the duration of the trial. Eye position was monitored using the EyeLink1000 gazetracker, and if a subject’s gaze shifted outside a fixation window of 1°/H11003° (for most subjects), the trial was aborted, the target position re-randomized, and the trial restarted. Each subject performed at least 4 repeats of the 30-trial staircase for each of the 16 motion flow angles. A single series of staircases for all 16 flow angles (performed in random order) was completed before a repeat series of staircases commenced (again in random order).

**Control for perceived size effects.** It has been reported that the perceived size of a stimulus aperture defined solely by kinetic signals can be influenced by the pattern of motion within it: expanding patterns can cause the aperture to appear larger, whereas contracting patterns can cause a “shrinking” of its perceived size (Ramachandran and Anstis 1990). To control for this effect and ensure that the subjects were basing their decisions on the target motion pattern per se and not on any perceived size differences, a control experiment was
also conducted wherein the radius of each of the eight RDKs in the array was randomly jittered by ±5% across trials, while dot density, speed, and all other parameters were kept constant (Takeuchi 1997). Subjects performed five runs of this task each for motion flow angles of 0° (expansion) and 180° (contraction).

Analysis. The mean coherence threshold for a given flow angle was computed from the final threshold parameter resulting from each repetition of the staircase (corresponding to 65.6% correct). Perceptual sensitivity was then specified as 1/mean threshold (sensitivity in the control task was computed in the same way). Thresholds were normalized for each subject by their mean threshold across the 16 flow angles before statistical analysis. Anisotropies were assessed by measuring the correlation between the mean (normalized) threshold across subjects with a cosine function. A strong positive correlation would indicate a bias for expanding patterns, whereas a strong negative correlation would indicate a bias for contracting patterns.

fMRI Methods

Apparatus. Functional MRI was performed with a Philips 3T Achieva scanner (Philips, Amsterdam, The Netherlands) with a whole head SENSE coil. Stimuli were back-projected through the Faraday shield with the use of a Dell 5100MP digital projector (Dell, Round Rock, TX) positioned behind the bore, with a resolution of 1,024 × 768 pixels, a refresh rate of 60 Hz, and a mean luminance of 275 cd/m². Images were viewed at a distance of 167 cm through a rear-facing mirror mounted on the head coil, giving a viewing angle of 19° × 14.3° (0.019° per pixel). Scanning was performed in a darkened chamber lit only by the light from the projector. The participants’ behavioral responses were collected via a LU400-PAIR Lumina response pad (Cedrus, San Pedro, CA). Reliable eye tracking data were obtained for two subjects using the MR-compatible camera for the EyeLink1000 system (SR Research). Unless noted otherwise, preprocessing and analysis of imaging data were performed using the Statistical Parametric Mapping package for MATLAB (SPM5; Friston et al. 2007) and custom MATLAB routines.

A boustrophedon, field-echo echo-planar imaging (FEEPI) pulse sequence was used in the acquisition of T₂*-weighted functional MR images of blood oxygenation level-dependent (BOLD) contrast. The FEEPI protocol was defined according to the following parameters: time to echo (TE) = 32 ms, time to repetition (TR) = 3,000 ms, flip angle = 90°, field of view = 192 × 69 × 192 mm, matrix = 128 × 128, voxel size = 1.5 mm (isotropic). The images were acquired in 46 interleaved ascending slices (1.5-mm thickness) in a tilted coronal plane that covered the entire occipital cortex as well as a portion of the posterior parietal and temporal lobes. An example EPI from a single subject overlaid on an anatomic scan is shown in Fig. 3F, whereas Fig. 3G shows example raw BOLD signal time courses for areas V1, V5/MT, and MST for the same subject and run as in Fig. 3F. In addition to the functional scans, a whole head structural MR image (voxel size = 1 × 1 × 1 mm) was obtained for each subject within each experimental scanning session, using a turbo field-echo protocol for optimal gray and white matter contrast.

For retinotopic mapping of individual subjects and realignment of functional data, an average anatomic image of three further structural scans was prepared consisting of whole head sagittal and transverse images (voxel size = 1 × 1 × 1 mm) and a higher-resolution partial coronal image (voxel size = 0.75 × 0.75 × 0.75 mm) of the caudal brain to maximize anatomic detail in the occipital lobes. Before averaging, the images were aligned using normalized mutual information-based coregistration, corrected for inhomogeneity (Manjón et al. 2007), normalized according to their peak white matter intensities, and resampled (where necessary) to a voxel size of 0.75 mm (isotropic). Each subject’s average anatomic image was then segmented using the automatic algorithms of ITK-SNAP (Yushkevich et al. 2006; http://www.itksnap.org/) and mrGray (Teo et al. 1997), supplemented with careful manual editing.

Fig. 3. Single subject example of functional echo-planar imaging (EPI) and raw blood oxygen level-dependent (BOLD) signal time courses. A: a single functional EPI overlaid on a coregistered T₁-weighted anatomic scan through the right hemisphere (z = 6, x = 2). B: example raw BOLD signal time courses from primary visual cortex (V1), middle temporal visual area (V5/MT), and medial superior temporal area (MST) for the same subject and scanning run as that in A. On the abscissa, the cycled stimulus flow angle is given along the bottom, whereas time is given along the top (each stimulus flow angle was presented at full contrast for 2 s within each 3-s TR). The normalized time courses are shown corrected for the hemodynamic lag and averaged across all voxels within each visual area.
Design and procedure. A continuous presentation (phase encoded) paradigm was used in the presentation of stimuli (Freeman et al. 2011; Mannion and Clifford 2011; Mannion et al. 2010a, 2010b; Raemaekers et al. 2011; Swisher et al. 2010; Yacoub et al. 2008). See Engel (2012) for a recent review on the history and advantages of using phase-encoded (that is, temporally encoded) designs to recover the form of the BOLD response for continuously defined stimulus parameters. Motion at one of the 16 flow angles was tested with displayed with each 3,000-ms volume acquisition in an orderly, stepwise manner (i.e., every 22.5°) across 7.25 repeated cycles in a single scanning session, of which subjects undertook at least 6. The presentation of a full cycle of 16 flow angles was thus completed in 48 s (see Fig. 2C). The motions were cycled through in a clockwise (i.e., decreasing from 337.5°) or counterclockwise direction (i.e., increasing from 0°) on alternating sessions. Each of the eight RDKs (all at 100% coherence) in the stimulus array for a given motion flow angle was generated independently, and they were saved together as a single AVI file. These 16 AVI files were loaded into MATLAB using the PsychToolbox (Brainard 1997; Pelli 1997) before their presentation in the scanner. To avoid transients due to abrupt changes in the motion flow angles, the first and final 250 ms consisted of the mean luminance background only, meaning the different motion RDKs were separated by a blank period of 500 ms. Also, the contrast of the dots was ramped on and off over another 250 ms with a raised cosine profile, meaning each RDK was presented at full contrast for 2,000 ms at the center of the 3,000-ms fMRI volume acquisition.

Definition of ROIs. Each subject underwent separate scanning sessions before the main experiment to localize visual cortical ROIs, using the same scanning protocols. Retinotopic mapping was conducted to obtain eccentricity and polar angle maps of early visual areas V1, V2, V3, V3A/B, and hV4, with an additional localizer to identify the V5/MT+ complex and separate it into areas V5/MT and MST.

Retinotopic mapping. Functionally defined early visual cortex was delineated for each subject by using the nomenclature and criteria of Larsson and Heeger (2006) and Wandell et al. (2007), in the same manner as previous studies from this laboratory (e.g., see Goddard et al. 2010, Fig. 2; Mannion et al. 2010a, Supplementary Fig. 1). According to this scheme, areas V1–V3 and hV4 share a foveal representation at the occipital pole, whereas V3A and V3B (which were not separated in this analysis) share a dorsal foveal representation and border the dorsal portion of V3. Area hV4 was defined as a hemifield representation of the contralateral visual field bordering the ventral portion of V3 (Goddard et al. 2011). Subjects viewed slowly rotating wedge and expanding/contracting ring stimuli in separate scans. Wedges were high-contrast black and white radial checkerboard patterns that filled a sector of 45° of polar angle, extending from 0.4° to 7.15° eccentricity. The polar angle of the wedge shifted in stepwise 15° shifts around fixation every 1.5 s. The wedge itself was divided into three 15° sectors with checkerboard patterns of random phase that drifted smoothly (at random, with each 15° shift of the wedge) either inward or outward from fixation. The rotation direction of the wedges around fixation alternated between clockwise and counterclockwise on successive runs, of which subjects undertook no less than four. Ring stimuli were similar checkerboard patterns and filled an aperture of 1.2° of visual angle. The ring expanded outward (or on a separate run, contracted inward) from fixation in steps of 0.375° at a rate of 0.67 Hz and traversed the same space as the extent of the wedge stimulus (i.e., 0.4°–7.15° eccentricity). Like the wedges, the rings were divided into arc strips of 15°, with the checkerboard pattern of each strip at an initial random phase and moving smoothly inward or outward with each step in the cycle. The 24 stepwise shifts of both the polar wedge and rings cycles were thus completed in 36 s, and 10.75 cycles were completed per run. All stimuli were set against the mean luminance background of the display.

The wedge and ring stimuli produced a “traveling wave” of activation that allowed for a voxel’s preferred polar angle (wedges) and eccentricity (rings) within the stimulated field of view to be determined, based on the temporal phase of the best-fitting sinusoid (Engel et al. 1994, 1997; Larsson and Heeger 2006; Wandell et al. 2007). The resulting topographically organized map of angular and eccentricity preferences allowed for the manual definition of the borders of the visual areas using the mrVista software suite in MATLAB developed by the Stanford Vision, Imaging Science and Technology Laboratory (http://white.stanford.edu/software/). To obtain this map, the first half-cycle of each data set was discarded, and the hemodynamic delay was compensated for by shifting the time course of each run by +2 volumes (6 s). The contracting ring run and the clockwise-rotating wedge runs were then temporally reversed for averaging with the expanding ring and counterclockwise-rotating wedge runs, respectively (Larsson and Heeger 2006). In mrVista, the averaged fMRI time series from the wedge and ring scans were smoothed with a Gaussian kernel (full width at half-amplitude = 1.5 mm) and projected onto a computationally flattened representation of the two cerebral hemispheres derived from the anatomic segmentation image (described above).

Segregation of V5/MT and MST. For each subject, up to four runs of a separate localizer were also conducted. The localizer was designed to identify area V5/MT+, a strongly motion-responsive set of voxels in lateral occipital cortex near the ascending limb of the inferior temporal sulcus (Dumoulin et al. 2000; Kolster et al. 2010; Tootell et al. 1995; Zeki et al. 1991), and segregate it into separate V5/MT and MST ROIs. The localizer was based on the approach of previous studies (Fischer et al. 2012; Huk et al. 2002; Smith et al. 2006) and involved blocks of coherent radial motion restricted to the left or right 120°, with a static pattern filling the remaining two-thirds of the display. The stimulus extended within an annulus from 1.2° to 6.4° eccentricity and consisted of a RDK with dots of low (10%) contrast (Dumoulin et al. 2000) and a dot density of 9.9 dots/deg². The dots were smoothed with a Gaussian envelope (σ = 0.04°) and were randomly assigned to be of a positive or negative luminance increment (with a probability of 0.5), set against the mean luminance background of the display. Within the motion conditions, dots moved at an average speed of 5.3/°. The motion conditions were interleaved with static dot patterns and a fixation-only condition, where only the blank background was displayed. The static pattern was simply a snapshot of a random frame from the motion condition and was updated at a rate of 0.33 Hz. Each block was 5 volumes long (15 s), and blocks for each condition were repeated 6 times within a run. V5/MT+ complexes in the two hemispheres were identified using parametric maps of t values obtained from general linear model contrasts of the responses to coherent motion vs. the static dots, determined using SPM5 (P < 0.05; adjusted for multiple comparisons). Figure 4 provides a single-subject example. MST was defined as a cluster of voxels within V5/MT+ showing significant ipsilateral activation, whereas V5/MT was defined as the contralateral activation, excluding those voxels assigned to MST (Fischer et al. 2012; Huk et al. 2002; Smith et al. 2006). Finally, V5/MT and MST clusters in both hemispheres were exported for analysis using the MarsBaR ROI toolbox for SPM (Brett et al. 2002; http://marsbar.sourceforge.net/). Further examples from this laboratory of V5/MT+ ROIs obtained using this method, positioned in situ on a flattened map of cortex adjacent to visual areas defined using the traveling wave method, can be found in Goddard et al. (2010, Fig. 2).

Fixation task. During all fMRI scans conducted (the experimental, retinotopic, and localizer scans), subjects performed a demanding fixation task to promote stable gaze and control for the allocation of attention (for the experimental scans, this was in addition to the tracking of eye position using the EyeLink1000 system). This involved monitoring a small fixation point at the center of the display (0.19° in diameter) and responding by holding down a button on the response pad whenever it underwent a small increment in luminance. The luminance change occurred on average every 1,750 ms, with a random jitter around this value of ±250 ms. During the retinoptic
scans, a fixation grid made up of a series of isopolar and isoeccentric lines was continuously displayed against the mean luminance background, which is known to further encourage stable fixation during retinotopic mapping procedures (Hansen et al. 2007; Schira et al. 2007).

Reliable eye tracking data in the MRI scanner was obtained for two subjects for eight runs each (the duration of each run was 5 min 57 s). This allowed for fixation to be monitored and checks made for excessive blinking. The subjects made an average (per run) of 2.4 saccades of an amplitude >1° from fixation. Across all 16 runs, only 7 saccades of an amplitude >1.5° from fixation occurred. The mean interblink interval (IBI) across all 16 runs was 10.4 s, whereas the minimum mean IBI for any run was 6.3 s, which is still greater than the mean IBI for normal adults performing a visual task (Johnston et al. 2013). For the remaining subjects, behavioral performance on the fixation task was quantified by the mutual information (MI) between the state of the fixation spot and the subject response. The mean (across subjects/runs) maximum MI between the state of the fixation spot (an entropy of 1 bit) and the subject response was 0.64 bits (with a range of 0.44–0.81 bits), and the mean response latency at which this occurred was 376.4 ms (with range 300–500 ms). Taken together, these measurements suggest that the subjects were consistently fixating on the fixation spot and attending to the dimming task for the duration of each scan.

Analysis. DATA PREPROCESSING. With the use of SPM5, fMRI data sets were corrected for between- and within-run head movement artifacts, averaged to produce a mean functional image across runs, and coregistered with the whole head structural MR acquired during the same session. Functional images were then resliced to the space of the mean anatomic image used to perform retinotopic mapping (described earlier) using fourth-degree B-spline interpolation. ROIs of the identified visual areas were also resliced to the space of the mean functional image from the experimental runs. For visual areas V1–V3 and V3A/B, the eccentricity phase maps were used to restrict the analysis of the experimental data to just those voxels sensitive to an annular region of the visual field covering the same range of eccentricities as the stimuli (i.e., 2.59°–5.41° eccentricity). To avoid problems of interpolation associated with the discontinuity in phase at 2π (i.e., neighboring voxels with phase close to 0 or 2π might be averaged to π), this eccentricity mask was generated by first taking the separate sine and cosine components of the (wrapped) phase maps, reslicing them to the resolution of the experimental functional images, and then recombining.

COMPUTING THE CORTEX SELECTIVITY TO COMPLEX MOTION FLOW ANGLES. The BOLD signal time course to the complex motion stimulus cycle was calculated for each voxel in each ROI. To eliminate the influence of onset transients and allow the BOLD signal to reach a steady state, the very first quarter-cycle (4 volumes; 12 s) in each run was discarded, and the response time course was advanced by +2 volumes (6 s) to compensate for the delay in the hemodynamic response. After this, the data from runs in which the motion flow angles advanced in a clockwise manner were temporally reversed before being combined with those that advanced counterclockwise. Low-frequency confounds were eliminated by high-pass filtering the data with a cutoff period of 128 s. The mean signal time course was computed by averaging across the voxels within a ROI, resulting in an (n × 112)-element matrix, where n was the number of runs (usually 8) and 112 was the length of the stimulus cycle (in volumes) after correction for the hemodynamic response (i.e., 16 flow angles × 7 full cycles). It was then normalized by taking the mean value across all 112 volumes, and subtracting and then dividing by that value, thereby centering the mean signal at zero (such that normalized BOLD responses below the across flow angles are negative). Finally, data were averaged across runs and stimulus cycles to produce a single vector describing the mean signal change (expressed as a percentage) across flow angles for each ROI (Mannion and Clifford 2011; Mannion et al. 2010a, 2010b).

RESULTS

Psychophysics

Detection thresholds for the 16 complex motion flow angles were converted to perceptual sensitivity for each subject. The mean sensitivity is given in Fig. 5 (note that data for contracting patterns are wrapped such that they appear at both −180° and 180° along the abscissa). The most notable feature of the data is the asymmetry in sensitivity to expanding over contracting patterns: subjects had greater difficulty in detecting a contracting pattern from an array of expanding distractors than they did detecting an expanding pattern from an array of contracting distractors. This result makes sense with respect to the distribution of selectivities for complex motions in MSTd, where cells tuned to expansion far outnumber those tuned to contraction (and most other flow angles, for that matter; Duffy and Wurtz 1991a, 1991b; Graziano et al. 1994; Heuer and
functionally defined regions of visual cortex. Figure 6, A–G, shows the mean response across subjects for each ROI (note that data for contracting patterns are wrapped such that they appear twice, at −180° and 180° on the abscissa). In areas V1, V2, V3, V3A/B, and hV4, there was a marked peak in response to expanding patterns. Generally, this was also coupled with a weak response to contracting patterns. The shapes of the anisotropies in the two anatomically smallest areas, V5/MT and MST, are less clear than in the other areas, but the modulations are nevertheless characterized by a weak response to contracting patterns.

As with the behavioral data, anisotropies in the mean BOLD response in each visual area as a function of complex motion flow angle were assessed by measuring the correlation between the data and a cosinusoid. The analyses measured the correlation between the 16-item normalized mean BOLD response across subjects (the vectors of data presented in Fig. 6) and a cosinusoid of the same length. To minimize the potential for spurious correlations due to the temporally low-pass nature of the hemodynamic response function, the statistical significance of the resulting $r$ values was assessed using the following bootstrapping procedure (Efron and Tibshirani 1993). First, a set of 10 independent random values was drawn from the standard normal distribution and convolved with the canonical hemodynamic response function for a TR of 3,000 ms, from SPM5. This was repeated 1,000 times to yield a distribution of noisy “surrogate” data sets to serve as proxy for the actual BOLD response to the 16 motion flow angles. Half of these data sets were then “temporally” reversed, just as was done with the actual data on runs where the complex motion cycle advanced in a clockwise direction. The two halves were added, and the correlation between each set of surrogate data and a cosinusoid was calculated, producing a distribution of 1,000 correlation coefficients. This process was repeated 10 times, and the value at the 95th percentile for each resulting distribution was taken. If the actual correlation coefficient given by the real data exceeded the mean 95th percentile of the 10 surrogate correlations, then the correlation with the cosinusoid was considered significant ($P < 0.05$, one-tailed test). A one-tailed test was used here because this procedure was considered sufficiently conservative to justify its use while there was also a priori reason, based on the existing neurophysiological data, to expect the BOLD signal to exhibit a greater increase at expanding patterns (and hence show a positive correlation with a cosinusoid).

The correlations were significant in areas V2, V3, V3A/B, hV4, and V5/MT (and came very close to significance in V1). The magnitude of each correlation for each ROI (and the correlation between each set of surrogate data and a cosinusoid was calculated, producing a distribution of 1,000 correlation coefficients. This process was repeated 10 times, and the value at the 95th percentile for each resulting distribution was taken. If the actual correlation coefficient given by the real data exceeded the mean 95th percentile of the 10 surrogate correlations, then the correlation with the cosinusoid was considered significant ($P < 0.05$, one-tailed test). A one-tailed test was used here because this procedure was considered sufficiently conservative to justify its use while there was also a priori reason, based on the existing neurophysiological data, to expect the BOLD signal to exhibit a greater increase at expanding patterns (and hence show a positive correlation with a cosinusoid).

The correlations were significant in areas V2, V3, V3A/B, hV4, and V5/MT (and came very close to significance in V1). The magnitude of each correlation for each ROI (and the uncorrected test alpha) is provided in Table 1. Correlations between the mean psychophysical data and the mean BOLD response in each ROI were also measured, and the resulting $r$ values are also given in Table 1. The statistical significance of these correlations was again computed using the same procedure as for the correlations of the neural data with a cosinusoid, except this time a distribution of correlations of the actual behavioral data set with surrogate fMRI data sets was calculated. The correlation between the behavior and the fMRI measures was significant for areas hV4, V5/MT, and MST ($P < 0.05$, corrected, one-tailed test).

**Perceived Size Effects**

In a control condition, the aperture size of targets and distractors at the expanding and contracting flow angles was jittered in size by ±5% to rule out the possibility that subjects were basing their decisions on perceived size differences between the two motion patterns. The results were very similar to the results of the main psychophysics experiment. There was again a greater sensitivity to expansion over contraction (paired $t$-test, $t_5 = 4.54$, $P < 0.01$). This indicates that the subjects were basing their decisions on the patterns of motion per se and not on any illusory size differences.

**Functional MRI**

The BOLD fMRI signal to full coherence complex motion stimuli centered in the periphery was also measured for seven
DISCUSSION

This study used matching stimulus arrangements to explore the sensitivity of human vision to a wide range of complex motion patterns, measured both psychophysically and at the level of the fMRI BOLD signal in functionally defined regions of visual cortex. The stimuli were carefully arranged so that sensitivity could be assessed for the motion patterns per se, decoupled from the local biases radial to the point of fixation evident in the activity of early visual cortex (Clifford et al. 2009; Raemaekers et al. 2009). Characteristic anisotropies to complex motion patterns were revealed in both data sets. In all cortical regions of interest examined, motion flow angles...
Table 1. Correlations of mean fMRI results with a cosinusoid and mean behavioral data

<table>
<thead>
<tr>
<th>ROI</th>
<th>Pearson Correlation With Cosinusoid</th>
<th>Pearson Correlation With Mean Psychophysical Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
<td>$P$ (uncorrected)</td>
</tr>
<tr>
<td>V1</td>
<td>0.64</td>
<td>0.008</td>
</tr>
<tr>
<td>V2</td>
<td>0.69*</td>
<td>0.003</td>
</tr>
<tr>
<td>V3</td>
<td>0.72*</td>
<td>0.002</td>
</tr>
<tr>
<td>V3A/B</td>
<td>0.85*</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>hV4</td>
<td>0.65*</td>
<td>0.007</td>
</tr>
<tr>
<td>V5/MT</td>
<td>0.67*</td>
<td>0.005</td>
</tr>
<tr>
<td>MST</td>
<td>0.57</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Correlation analyses were conducted across the vector of normalized mean blood oxygen level-dependent (BOLD) modulation across the 16 complex motion flow angles (data shown in Fig. 6) with a cosinusoid of the same length or across the vector of normalized mean psychophysical sensitivity measured across the same 16 stimuli (Fig. 5). $P$ values indicate the statistical significance of the initial, uncorrected correlation analyses performed before any bootstrapped correction procedure with surrogate functional MRI (fMRI) data. *Correlation coefficients ($r$) significant according to the bootstrapped distribution of correlations with surrogate data ($P < 0.05$, 1-tailed). Regions of interest (ROI) include early visual areas (V1, V2, V3, V3A/B, and hV4), middle temporal visual area (V5/MT), and medial superior temporal area (MST).

Complex motions represent a particularly compelling class of motion stimuli, not only because of their role in behavioral processes of navigation and locomotion (Vaina 1998) but also because of what they can reveal about the mechanisms of motion perception. With complex motions, and in particular with the peripherally centered stimuli used in the present study, the distribution of local motion directions remains uniform across flow angles, making these stimuli highly informative when mechanisms of global, in contrast to local, motion perception are considered. In previous psychophysical work on complex motion, sensitivity was generally found to be maximal for the radial motions, contraction, and expansion, with disagreement as to whether sensitivity is poorest for rotations (Beardsley and Vaina 2005; Meese and Anderson 2002) or spirals (Burr et al. 2001; Morrone et al. 1999). In the present study, sensitivity was greatest for the rotations, whereas there was a marked asymmetry between contraction and expansion (with spiral motions being intermediate). Earlier studies have tended to use large stimuli centered at fixation, meaning that radial biases for local motion (Albright 1989; Cliffor et al. 2009; Raemaekers et al. 2009) may have artifactually inflated sensitivity to contraction while reducing it for rotation. In the present study, it was more difficult to detect a contracting target from among a field of expanding distractors, compared with the reverse situation. Similar asymmetries have also been reported in the visual search literature. Reaction times, error rates, and search slopes are all higher for contracting compared with expanding targets (Shirai and Yamaguchi 2004; Takeuchi 1997; Thornton and Gilden 2001). Stimuli in visual search paradigms (as in the present study) are usually not centered at fixation and thereby should avoid confounds due to radial biases that might otherwise obscure actual perceptual asymmetries. It is not entirely obvious at this juncture why the rotational motions should have yielded the greatest measures of sensitivity. One could speculate on the basis of torsional movements of the eye/head or perhaps a greater activation of higher-level object representations by the rotational motions. Even though attention and eye movements were controlled in both the fMRI and psychophysics settings, the deployment of attention toward different complex motions (whether voluntarily or involuntarily) is also worthy of consideration. The visual search literature, albeit often with conflicting results, suggests differences in the efficiency of search between and within the cardinal complex motions and translational ones (Braddick and Holliday 1991; Takeuchi 1997; Thornton and Gilden 2001). Furthermore, judgments of heading in expanding flow fields are mostly unaffected when subjects attend to a concurrent motion-discrimination task (Royned and Hildreth 1999), suggesting perception of expansion (at least) either occurs preattentively or does not compete for attentional re-
Most models of complex motion perception emphasize a special role for MST, along with interconnected regions of posterior parietal cortex that also exhibit similar selectivity, such as 7a (Raffi and Siegel 2007; Siegel and Read 1997; Steinmetz et al. 1987) and the ventral intraparietal area (Chen et al. 2013; Schaafsma and Duy sens 1996; Wall and Smith 2008). The primary source of projections to MST is V5/MT (Maunsell and Van Essen 1983; Ungerleider and Desimone 1986), whereas the large receptive fields, position invariance, and preference for large stimuli of MST neurons all suggest significant spatial summation in the computations underlying complex motion selectivity (Burr et al. 1998; Duffy and Wurtz 1991a; Graziano et al. 1994; Lagae et al. 1994; Morrone et al. 1995; Orban et al. 1992; Raiguel et al. 1997; Tanaka et al. 1989; Tanaka and Saito 1989). Integration of inputs from V5/MT cells tuned to translational motion directions is, accordingly, an important component of models of complex motion perception (Mineault et al. 2012; Morrone et al. 2000; Perrone and Stone 1994, 1998; Saito et al. 1986; Tanaka and Saito 1989; Yu et al. 2010). For example, Mineault et al. (2012) recently demonstrated that a nonlinear integration mechanism among the inputs to MST provided a successful account of the complex motion selectivity of their sample of MST cells. Such models are typically feedforward and hierarchical, and as such do not necessarily predict the anisotropies in complex motion selectivity in early visual cortex or V5/MT measured in the present study. It is tempting to speculate that feedback mechanisms, perhaps from MST or posterior parietal cortex (VIP or 7a), and possibly the cingulate cortex (Wall and Smith 2008), are responsible for these activations. After initial processing, reentrant signals might influence the responses of populations in earlier visual cortex, with smaller receptive fields and an appropriate spatial distribution of tunings and for local motion directions, thereby enhancing the response to particular complex motions. The present findings cannot resolve this question, particularly given that (as noted above) the stimuli were not necessarily designed to elicit a vigorous response from MST. Still, these findings indicate that future work, using stimuli optimized to drive MST, could be used to elucidate the role of reentrant mechanisms in complex motion perception.

In summary, the present study disentangled the response of early visual cortex to complex motion patterns per se from local biases for motion radial to the fovea, in an attempt to resolve some of the discrepancies in the existing literature. The unique approach involved presenting a wide range of small complex motion patterns centered in the periphery, tested in a complementary manner using both fMRI and psychophysical procedures. Both experiments revealed anisotropies in the sensitivity/response to the complex motions. This was marked by a significantly greater response to expanding over contracting patterns in areas V2–V3A/B, hV4, and V5/MT, and psychophysically. There were also two maximal peaks in the psychophysical sensitivity at the rotations. The anisotropies may reflect the activity of local looming detectors and/or feedback from higher centers such as MST or the posterior parietal cortex. The stimuli and design employed in this work should prove useful in further investigations of certain complex stimulus classes whose representation in early visual cortex might otherwise be overlooked.

ACKNOWLEDGMENTS

We thank Dr. Isabelle Mareschal for helpful discussions and Kirsten Moffatt and the MRI radiography team at St. Vincent’s Hospital, Darlinghurst.

GRANTS

This work was supported by National Health and Medical Research Council Grant APP1027258 (to C. W. G. Clifford), Australian Research Council (ARC) Future Fellowship Grant FT110100150 (to C. W. G. Clifford), ARC Post-doctoral Fellowship Grant DP11010217 (to T. L. Watson), and the ARC Centre of Excellence in Vision Science.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

J Neurophysiol • doi:10.1152/jn.00445.2013 • www.jn.org
AUTHOR CONTRIBUTIONS


REFERENCES


Regan D, Beverley KI. How do we avoid confusing the direction we are looking and the direction we are moving? *Science* 215: 194–196, 1982.


