Decoding conjunctions of direction-of-motion and binocular disparity from human visual cortex

Kiley J. Seymour1,2,3 and Colin W. G. Clifford1,2

1School of Psychology and 2Australian Center for Excellence in Vision Science, University of Sydney, Sydney, New South Wales, Australia; and 3Department of Psychiatry and Psychotherapy, Charité - Universitätsmedizin Berlin, Berlin, Germany

Submitted 6 December 2011; accepted in final form 3 February 2012

Seymour KJ, Clifford CWG. Decoding conjunctions of direction-of-motion and binocular disparity from human visual cortex. J Neurophysiol 107: 2335–2341, 2012. First published February 8, 2012; doi:10.1152/jn.01103.2011.—Motion and binocular disparity are two features in our environment that share a common correspondence problem. Decades of psychophysical research dedicated to understanding stereopsis suggest that these features interact early in human visual processing to disambiguate depth. Single-unit recordings in the monkey also provide evidence for the joint encoding of motion and disparity across much of the dorsal visual stream. Here, we used functional MRI and multivariate pattern analysis to examine where in the human brain conjunctions of motion and disparity are encoded. Subjects sequentially viewed two stimuli that could be distinguished only by their conjunctions of motion and disparity. Specifically, each stimulus contained the same feature information (leftward and rightward motion and crossed and uncrossed disparity) but differed exclusively in the way these features were paired. Our results revealed that a linear classifier could accurately decode which stimulus a subject was viewing based on voxel activation patterns throughout the dorsal visual areas and as early as V2. This decoding success was conditional on some voxels being individually sensitive to the unique conjunctions comprising each stimulus, thus a classifier could not rely on independent information about motion and binocular disparity to distinguish these conjunctions. This study expands on evidence that disparity and motion interact at many levels of human visual processing, particularly within the dorsal stream. It also lends support to the idea that stereopsis is subserved by early mechanisms also tuned to direction of motion.

correspondence; feature binding; functional magnetic resonance imaging; stereopsis

MOTION AND BINOCULAR DISPARITY are two features in our visual environment that are commonly used by the brain to recover the three-dimensional (3-D) quality of a scene. In fact, the encoding of these two features relies on solving a common “correspondence problem”: a matching of object position over time (motion) or across the two eyes (binocular disparity). Thus it has been suggested that extracting these two features from retinal images may rely on a shared neural substrate (Anstis and Harris 1974; Bradley et al. 1995; Graham and Rogers 1982; Nawrot and Blake 1991). Psychophysical evidence suggests that conjunctions of motion and disparity are encoded within inseparable neural “units” before the computation of depth. For instance, experiments have demonstrated that a combination of these cues facilitates depth perception (Bradshaw and Cumming 1997; Johnston et al. 1994; van Ee and Anderson 2001). Furthermore, the well-known “motion aftereffect” has been shown to be contingent on the binocular disparities at which an adapting and test stimulus are presented (Anstis and Harris 1974; Nawrot and Blake 1989; Smith 1976; Verstraten et al. 1994). Similarly, disparity aftereffects are contingent on motion direction (Nawrot and Blake 1989). Although such findings are best explained by the adaptation of neurons coding for specific conjunctions of disparity and motion, the question of where in the human visual cortex these cells may arise remains unresolved.

Perhaps the strongest evidence for conjunction coding of motion and disparity has been provided by physiological experiments in nonhuman primates. These studies indicate that the initial sensory integration of motion and disparity occurs at, or before, area MT+ (Anzai et al. 2001; DeAngelis and Newsome 1999; DeAngelis and Uka 2003; Granter and Skoumbourdis 2004; Maunsell and Van Essen 1983; Pack et al. 2003; Poggio and Talbot 1981). However, only few studies confined to macaque MT actually present evidence for true conjunction coding (Bradley et al. 1995; DeAngelis and Newsome 2004; Dodd et al. 2001; Roy et al. 1992). For instance, although many neurons in macaque MT exhibit separable selective responses to motion and disparity information, a proportion also exhibit selectivity for motion that is “unfixed” and can be modulated by disparity (DeAngelis and Newsome 2004; Roy et al. 1992).

Although research in the macaque brain supports the psychophysical evidence for conjunction coding of motion and disparity, no study has directly examined where in the human visual cortex these conjunctions are encoded. Functional MRI (fMRI) in combination with multivariate pattern classification has provided evidence for motion and binocular disparity selectivity throughout the visual cortex (Kamitani and Tong 2006; Preston et al. 2008), and a recent fMRI adaptation experiment (Smith and Wall 2008) showed that area MT responded differentially to motion when it was presented at different disparities. However, since voxels containing separate clusters of neurons selective for specific motion directions and specific binocular disparities could explain these findings, it remains unclear whether single functional units (i.e., voxels) respond to specific pairings of these two features.

In the current study, we examined whether fMRI blood oxygen level-dependent (BOLD) signals measured in human visual cortex could accurately discriminate stimuli differing exclusively by their specific pairing of two motion directions and two binocular disparities. A similar approach has previously been employed to examine conjunction coding of color-motion and color-form in the human visual cortex (Seymour et al. 2009, 2010). Here, we distinguished conjunction coding of...
disparity and motion (i.e., supralinear responses to distinct motion-disparity pairings) from joint selectivity, where a voxel could simply respond when either or both features were present (i.e., resulting in additive separable responses to the 2 features). Our results provide evidence that disparity and motion information are integrated early in human visual processing and are represented at many levels of the visual hierarchy, particularly within the dorsal stream.

METHODS

Five participants (3 male and 2 female) took part in the study, including both authors. All had normal vision and stereoeacuity (tested in a separate session within the scanner setup). Each subject was familiarized with the task during one preliminary psychophysics session outside of the scanner and gave informed written consent according to the ethics committee at the University of Sydney.

Stimuli

Basic stimulus parameters. A dichoptic stimulus display setup was designed for use in a Philips 3T MRI scanner with standard back projection (projector: Dell 5100MP, display resolution: 1024 × 768 pixels). This followed closely the design of Schurger (2009). A cardboard divider connected the viewing mirror and the projector screen (running through the bore of the magnet) allowed for each ocular image to be viewed separately. Two square stimulus frames were projected onto the screen (on either side of the vertical midline) and remained there for the entire scan session to help the subject maintain fixation of the images of the two eyes. The image of each eye also had its own set of nonius lines displayed at fixation (subtending 0.7°) so that, when fused, the two images were perceived as a single square frame subtending 15° of visual angle with a fixation cross at the center. To stabilize fixation further, subjects additionally fused a white fixation point (presented to the left eye) with a surrounding black ring of a bulls-eye (presented to the right eye). Throughout the entire scan session, each subject wore custom-cut prism lenses to adjust the viewing angle of each image of the eye and make fusion comfortable. The screen was viewed from a distance of 167 cm. Stimuli were presented using PsychToolbox 3.0.8 (Pelli 1997).

Specific conjunction stimuli. Since our experiment aimed to establish evidence for voxels tuned to specific conjunctions of motion and disparity, we relied on the use of multivariate pattern classifiers to distinguish between fMRI activation patterns associated with viewing of two specific conjunction stimuli. Both stimuli were composed of the same two directions of motion and two binocular disparity cues but differed exclusively by the way these cues were paired (i.e., their conjunctions). Hence, a classifier could not rely on activation associated with disparity-specific or motion-specific responses to distinguish the two stimuli. True conjunction coding would require a nonlinear (or unique) response to the combined features (i.e., 1 differing from the sum of the responses to each feature).

All stimuli were displayed on a gray background (37 cd/m²) within a square fusion frame. Each specific conjunction stimulus subtended 15° and consisted of translating black dots (size: 0.17° radius, speed: 2°/s) presented at disparities of +0.17° (crossed disparity) and −0.17° (uncrossed disparity). To ensure that depth was defined purely by binocular disparity and unaided by monococular cues, we removed all interocular unmatched dots (half-occlusions) from each image of the eye (Brooks and Gillam 2006). Figure 1 represents a schematic of the stimulus display for conjunction stimuli A and B. Note that both conjunction stimuli contained the same basic motion and disparity information (i.e., leftward motion, rightward motion, crossed disparity, and uncrossed disparity). The only difference between the two conditions was the way these features were paired. For example, where leftward motion was paired with crossed disparity in condition A, it was paired with uncrossed disparity in condition B, etc.

Avoiding low-level influences on classifier performance. We applied strict controls in our experimental design to avoid incorporating stimulus or task artifacts that could result in successful classification performance in the absence of conjunction-selective information.

First, we minimized the effects of systematic eye movements associated with the use of motion stimuli (Chukoskie and Movshon 2009) by displaying opposite motion-disparity conjunctions to the upper and lower hemifield of each stimulus. This ensured that each depth plane contained opposing motion directions, thus making it difficult for a subject’s eyes to follow one specific motion direction systematically in any particular condition. In addition, a 3° blank gap separated the entire upper and lower hemifields of the stimulus to restrict motion-induced eye movements. Furthermore, subjects were asked to fixate on the central fixation point for the entire duration of each experimental run.

Our aim was to ensure that depth in our stimuli was exclusively defined by binocular disparity. Therefore, inevitably, slightly different regions of each retina were stimulated by crossed vs. uncrossed depth and motion directions.
disparity-motion conjunctions. Since a classifier could in principle produce above-chance decoding performance by exploiting the differences in the direction of motion at the extreme edges of each eye’s image (width; 0.17°), we ran a control experiment to address this possibility. Two subjects were presented with identical stimuli to those displayed in the main experiment, however, rectangular patches, the same color and luminance as the background, occluded the aberrant monocular stimulus edges. This change effectively introduced half-octave changes to each image of the eye as an additional stereoscopic cue to depth (Brooks and Gilliam 2006) but resulted in a stimulus that no longer contained a potential monocular artifact that could aid classification of our two stimulus conditions.

Avoiding high-level influences on classifier performance. To prevent the effects of attentional biases to a particular depth plane, which could influence classifier performance (Neri and Levi 2008), subjects were engaged in a change-detection task that required attention to be divided across both depth planes and motion directions. In both depth planes, dots could change shape from circles to squares and vice versa over the duration of each block (3 randomly timed changes occurred per block). When both depth planes (and consequently both motion directions) displayed the same shape (i.e., circle-circle or square-square), the subject was required to press a button. When the two depth planes displayed different shapes, they were required to press an alternate button. Failure to divide attention across depth and motion direction would have made this task impossible to complete.

fMRI and Multivariate Analysis

Data acquisition. We used a Philips 3T scanner with a whole-head coil to collect data from 10 runs of our main block-designed experiment. Conjunction stimuli were presented in 15-s blocks and repeated 9 times over an entire run in an alternating fashion (i.e., in odd runs, block A followed block B). In addition, we collected data from 2 localizer runs, which consisted of presenting the same conjunction stimuli in blocks of 15 s interleaved by fixation blocks of equal duration. All functional images were collected using a T2*-sensitive, boustrophedon, field-echo, echo-planar imaging pulse sequence (repetition time, 3 s; time to echo, 30 ms; flip angle, 90°; field of view, 70.5, 192, and 192 mm; matrix, 128 × 128 pixels; voxel size, 1.5 mm isotropic). Images were acquired in 47 ascending interleaved slices in the coronal plane covering the occipital lobes. A within-session whole-head structural scan was collected using a turbo field-echo protocol. This was later used to aid optimally the alignment of the functional data to the subject’s high-resolution structural scan (see Retinotopic mapping and region-of-interest definition). All functional data were corrected for head movement before further analysis as implemented in SPM8 (http://www.fil.ion.ucl.ac.uk/spm).

Retinotopic mapping and region-of-interest definition. In a separate scan session using standard protocols (Engel et al. 1997), we performed retinotopic and MT localizer scans. First, a whole-head structural scan was collected using a turbo field-echo protocol for enhanced gray-white matter segmentation (voxel size: 1.0 mm isotropic). In addition, a high-resolution, partial-head coronal scan (voxel size: 0.75 mm isotropic) was acquired to allow for maximum detail in the occipital lobes and therefore good retinotopic mapping of this region. The functional data were collected as per the main experimental scan sequence. These data were transformed onto a flattened representation of the cortical surface using mrVista (http://white.stanford.edu/software). Manual delineation of the borders between visual areas V1, V2, V3, V3A/B, and hV4 was based on the phase of the responses of each voxel to standard polar angle (wedge) and eccentricity (ring) protocols. Area MT+ was localized as a region in the inferior temporal sulcus responding to moving vs. static random-dot stimuli presented at low contrast (Dumoulin et al. 2000). We did not distinguish MT from MST in this study. A mask defining each visual area was transformed from the flat map into the space of the subject’s anatomic scan. It was then smoothed (full-width at half-maximum: 1.5 mm) and resliced to the resolution of the experimental functional images using 4th degree B-spline interpolation.

Feature (voxel) selection. As multivariate pattern classification is susceptible to overfitting, we implemented a feature selection strategy aimed at restricting voxels. Manual delineation of the borders between V1, V2, V3, and V3A/B. Data were coregistered in raw anterior commissure-posterior commissure space and were not transformed or smoothed. The time course of each voxel was shifted by 2 volumes (6 s) to compensate for the hemodynamic lag. In addition, the volumes corresponding to the 1st and last 2 blocks of each run were removed from the analysis. This resulted in 16 exemplars from each run and ensured that each conjunction condition had been equally preceded and followed by the other conjunction condition, eliminating any potential run-ordering confounds. The signal time course of each voxel was high-pass filtered (cutoff: 128 s) and normalized (z-scored) with respect to the remaining volumes of each run. Response vectors were then formed by averaging the signal over each 15-s block (5 volumes) for each voxel. These response vectors were used as training and test data sets for subsequent multivariate pattern classifier analysis.

Multivariate Pattern Classification

Rationale. We trained multivariate pattern classifiers to decode specific disparity-motion conjunctions from BOLD signals in functionally defined ROIs in visual cortex. This was achieved by using training data to train a linear support vector machine to assign a decision boundary that could be applied to a new independent test data set (Vapnik 1998). As linear classifiers integrate evidence provided (separately) by each voxel about category membership, the logic here was that a ROI classifier would only show above-chance decoding performance if some voxels in that ROI were individually sensitive to the dimension of interest, i.e., conjunctions of disparity and motion. As our stimuli differed only by their conjunctions of these features, a classifier could not rely on segregated disparity and motion signals to distinguish the two stimuli. In such a case, both conjunction stimuli A and B would elicit the same summed neural responses and thus result in chance decoding performance. This approach allowed us to interpret successful stimulus decoding as evidence for conjunction information being contained within voxel activation patterns.

Analysis

The pattern classifier analysis was conducted separately for each ROI of each subject. We implemented a 10-fold, leave-1-out strategy in which a classifier was trained on voxel responses from 9 runs and tested on the remaining run, cycling through all runs. This approach yielded 144 exemplars in the training set and 16 exemplars in the test set (an equal number of examples from the 2 conditions were used). The training set was submitted to the linear support vector machine algorithm (C parameter was set to 1.0) as implemented in SVMLight (Joachims 1998). This yielded a weight for each voxel, reflecting the strength of the bias in its response between the two conjunction conditions. This model was then applied to the independent test data.
set to produce a classification for each test exemplar as either conjunction stimulus A or B.

To optimize pattern classifier analysis further, we implemented an additional feature selection strategy that 1st ranked voxels in descending order by t values obtained from the independent localizer contrast (feature selection mask). On each separate ROI the pattern analysis was then performed by iteratively including additional voxels when forming the data set (i.e., starting with the top 5 voxels and continuing with a step size of 5 voxels on each iteration). The variation in classifier accuracy with an increasing number of voxels was fitted to an exponential growth function from which the fitted asymptote value was then interpreted as the accuracy of the ROI in discriminating between conjunction stimulus A or B patterns. In cases where the classifier accuracy did not conform to an exponential growth function, the mean accuracy over the range of voxels was taken as the accuracy measure for the ROI. Significance was determined using a 1-sample t-test performed across the classification accuracies of the 5 subjects (chance performance was always 50%).

RESULTS

This experiment sought evidence for conjunction coding of motion and disparity within the human visual cortex. To do this, we used multivariate pattern analysis to examine whether linear classifiers could discriminate two distinct conjunction stimuli using consistent patterns of BOLD signals distributed across voxels. By ensuring that the two stimuli contained the same basic motion and disparity information (i.e., left and rightward motion and crossed and uncrossed disparity) but differed exclusively in their conjunctions of these features, our analysis distinguished between joint selectivity (i.e., where a voxel could simply respond when either feature or both were present) and true conjunction coding. Figure 2 shows that a linear classifier could successfully discriminate responses to distinct disparity-motion conjunctions across early visual areas, starting in area V2 and extending throughout higher dorsal stream regions. In various ROIs, not all subjects’ data reached significance individually. However, mean prediction performance reached significance across subjects in V2 (64.8%), V3 (59.4%), V3A/B (59.0%), and V7 (60.1%) using a threshold of P > 0.01. Although some individual subjects’ data reached significance in V1 and MT+, classifier performance on data from these areas as well as hV4 was not found to be significantly higher than chance across all five subjects.

We also ran a control experiment on two of our subjects to examine the possibility that small monocular differences at stimulus edges, inherent in stereo displays, could underlie the successful classification performance that we observed in our main experiment. Figure 3 shows that in the absence of monocular stimulus differences, the conjunctions of motion and disparity could be decoded with above-chance performance. Since a general increase in decoding performance was observed in this control condition, we believe the results of our main experiment reflect true conjunction coding of motion and disparity information. Furthermore, since it is known that V1 contains the highest proportion of monocular neurons and domains of ocular dominance at a spatial scale that might drive classifier performance (Hubel and Wiesel 1968), a failure in both experiments to classify successfully conjunctions from activity in area V1 strengthens this interpretation.

DISCUSSION

To date, evidence for conjunction coding of disparity and motion in the human brain has primarily come from human psychophysics (Anstis and Harris 1974; Bradshaw and Rogers 1996; Edwards and Badcock 2003; Graham and Rogers 1982, 1984a,b; Johnston et al. 1994; Nawrot and Blake 1989; Nerl and Levi 2008; Smith 1976; van Ee and Anderson 2001; Verstraten et al. 1994). However, no study has directly examined where in human visual cortex conjunction representations may exist. Although previous findings from physiology show that various regions within the visual hierarchy can exhibit joint selectivity for motion direction and binocular disparity (Anzai et al. 2001; DeAngelis and Newsome 1999; DeAngelis and Uka 2003; Grunewald and Skoumbourdis 2004; Maunsell and Van Essen 1983; Poggio and Talbot 1981), only a few studies in the macaque have actually presented evidence for true conjunction coding (Bradley et al. 1995; DeAngelis and Newsome 2004; Roy et al. 1992). Here, we investigated which areas of the human visual cortex encode distinct motion-disparity conjunctions. We employed a design that ruled out the possibility of separable additive fMRI responses underlying our results (i.e., joint selectivity) and sought direct evidence for inseparable (supraadditive) fMRI activations associated with viewing specific disparity-motion combinations (i.e., conjunction coding). Our findings indicate that conjunction coding of motion and disparity is distributed throughout the human dorsal visual stream, beginning in area V2. These results support physiological evidence for the specific tuning of conjunctions of motion and disparity within dorsal stream areas (Bradley et al. 1995; DeAngelis and Newsome 2004; Roy et al. 1992) and stand contrary to an account suggesting that before reaching MT+, motion and depth information are processed independently (Ponce et al. 2008). Furthermore, our data suggest that a mechanism for solving the correspondence problem may operate within extrastriate cortex at the level of V2. Such a finding provides support for theories concerning early mech-
nerisms also tuned to direction of motion subserving stereopsis (Anzai et al. 2001; Bradshaw and Rogers 1996; Nawrot and Blake 1989; Qian 1994; van Ee and Anderson 2001).

Although our main findings add to an existing view that conjunctions of motion and disparity are encoded within distinct neural units early in visual processing (Anzai et al. 2001; Lennie 1998; Nawrot and Blake 1991; Pack et al. 2003; van Ee and Anderson 2001), the weak classifier performance observed for area MT+ is in clear contrast to the compelling physiological evidence indicating that this region exhibits conjunction selectivity (Roy et al. 1992). However, since one must restrict any inference from our results to the level of voxels and not neurons (Logothetis 2008), our finding of a null result in MT+, V1, and hV4 should not be overstated as evidence for a lack of conjunction information. Importantly, factors such as the spatial arrangement of conjunction information within voxels as well as the number of voxels comprising a region can affect the decoding performance of a classifier. Indeed, a post hoc correlation analysis revealed that ROI size could account for 17% of the variation in decoding performance observed in our data (P < 0.05). Moreover, since MT+ is relatively small in size and is known to exhibit a high density of direction and disparity maps in close proximity to one another (DeAngelis and Newsome 1999), it is probable that the arrangement of conjunction information within this small visual region was not favorable for detection with the methodology employed here.

For instance, in this study, MT+ comprised an average of 104 voxels across all subjects, whereas area V2, for example, averaged 956 voxels. Furthermore, 2 individual subjects possessing the largest MT+ volumes exhibited above-chance decoding performance in this region. This suggests that poor classifier performance in MT+ compared with earlier visual areas may be attributed to substantially fewer selective features (or dimensions) being available to a classifier for successful assignment of a decision boundary. In addition, since the receptive field sizes of MT+ neurons are significantly larger than those in earlier retinotopic regions (Felleman and Kaas 1984), the spatial configuration of our stimuli (i.e., the presentation of different conjunction information to the upper and lower hemifields) may have also biased classifier performance in favor of areas exhibiting clearer dorsal-ventral retinotopy. For instance, if both the upper and lower hemifields of our stimuli fell within a single receptive field of an MT+ voxel, then both conjunction conditions would elicit indistinguishable responses since both our stimuli contained not only the same individual features, but also the same conjunctions. Furthermore, since MT neurons usually prefer zero disparity (DeAngelis and Newsome 1999), and small nonzero disparity tunings are common in V1, our stimulus parameters may have been biased against successful decoding performance within this region. Thus we believe our null result in MT+ is unlikely to reflect an absence of conjunction coding, per se, but more a spatial limitation of the current fMRI methodology.

The results of this study provide direct evidence for the coding of motion and disparity conjunctions early in human visual cortex. Nonetheless, the approach used here does not disentangle whether conjunction information observed across different areas of the dorsal stream relates to the sensory encoding of motion-disparity pairings or the subsequent perception of transparent surfaces separated in depth. Furthermore, a number of other roles have been suggested for conjunction neurons including an involvement in egocentric navigation (Lappe 1996; Roy et al. 1992) and the control of eye movements (Howard and Simpson 1989).

The dorsal stream in general (Sunaert et al. 1999) and area V3A in particular (Kamitani and Tong 2006; Tootell et al. 1997) have been shown to be heavily involved in motion processing. fMRI studies of human depth perception have consistently shown the strongest activity in dorsal areas, particularly V3A and V7 (Backus et al. 2001; Georgieva et al. 2009; Minini et al. 2010; Tsao et al. 2003). Areas V3A and V7 appear to code predominantly for absolute rather than relative disparity (Anzai et al. 2011; Neri et al. 2004). In contrast, it has been suggested that ventral areas split neural resources between absolute and relative disparity (Neri et al. 2004; Neri 2005). Indeed, a failure to detect conjunction information in V1 and hV4 may imply that absolute disparity signals, rather than relative disparity signals, are combined with motion information (Tyler 1990). This interpretation would favor a role for conjunction neurons in providing a rough estimate of the distance of an approaching object (Goodale and Milner 1992; Neri et al. 2004) as well as controlling vergence eye movements (Howard and Simpson 1989). However, in a study of the perception of bistable structure from motion, the highest levels of classification accuracy of the subjective motion-depth contingency was observed in areas V3A and V7 and the posterior intraparietal sulcus (Brouwer and van Ee 2007), indicating a role for these areas not just in the processing of disparity-motion conjunctions, but also in their perception.

Although our finding of conjunction information in V2 supports the evidence that this region may play a unique role in transforming and integrating depth-related information from primary visual cortex to facilitate 3-D surface perception (Bakin et al. 2000; Chen et al. 2008; den Ouden et al. 2005; Nienborg and Cumming 2006; Peterhans and van der Heydt 1993; Qiu and von der Heydt 2005; Tanabe and Cumming...
Anzai A, Chowdhury SA, DeAngelis GC. 2008; von der Heydt et al. 2000), future investigations employing stimuli containing disparities to which the perceptual system is blind (e.g., Cumming and Parker 1997; Preston et al. 2008), as well as stimuli that separate absolute disparity from relative disparity (Neri et al. 2004), may help to address the exact nature of conjunction representations observed in this region. Since dissociations between the physical encoding of low-level stimulus properties and the perception of depth have been observed across the visual hierarchy in both monkey and human (Bradley et al. 1995; Brouwer and van Ee 2007; Preston et al. 2008), it is probable that future studies will find a similar ascending gradient from sensory encoding of motion-disparity pairings to responses relating to the perception of 3-D surfaces defined by these features.

ACKNOWLEDGMENTS

We thank Kirsten Moffat and St. Vincent’s Hospital staff for assistance with fMRI data collection, Ciaran Baranasaorioy for assistance with the dichoptic viewing apparatus, and Damien Mannion, Kevin Brooks, Tatjana Seizova-Cajic, David Badcock, and Bart Anderson for helpful discussions.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

K. J. Seymour: experimental conceptualization and design, stimulus programming, creation of dichoptic apparatus, data acquisition and analysis, and manuscript production. C. W. G. Clifford: experimental conceptualization and design, data acquisition, manuscript production, and provision of funds.

REFERENCES


Qiu FT, von der Heydt R. Figure and ground in the visual cortex: v2 combines stereoscopic cues with gestalt rules. *Neuron* 47: 155–166, 2005.


