



Dissociation of retinal and headcentric disparity signals in dorsal human cortex

David M. Arnoldussen^{1,2}, Jeroen Goossens¹ and Albert V. van Den Berg^{1*}

¹ Section Biophysics, Department of Cognitive Neuroscience, Radboud University Nijmegen Medical Centre, Donders Institute for Brain, Cognition, and Behavior, Nijmegen, Netherlands

² School of Psychology, University of Nottingham, Nottingham, UK

Edited by:

Mikhail Lebedev, Duke University, USA

Reviewed by:

Markus Lappe, Universität Münster, Germany

Arjun Ramakrishnan, Duke University, USA

*Correspondence:

Albert V. van Den Berg, Radboud University Nijmegen Medical Centre, Geert Grooteplein 21, 6525 EZ, Nijmegen, Netherlands
e-mail: a.vandenbergh@donders.ru.nl

Recent fMRI studies have shown fusion of visual motion and disparity signals for shape perception (Ban et al., 2012), and unmasking camouflaged surfaces (Rokers et al., 2009), but no such interaction is known for typical dorsal motion pathway tasks, like grasping and navigation. Here, we investigate human speed perception of forward motion and its representation in the human motion network. We observe strong interaction in medial (V3ab, V6) and lateral motion areas (MT⁺), which differ significantly. Whereas the retinal disparity dominates the binocular contribution to the BOLD activity in the anterior part of area MT⁺, headcentric disparity modulation of the BOLD response dominates in area V3ab and V6. This suggests that medial motion areas not only represent rotational speed of the *head* (Arnoldussen et al., 2011), but also translational speed of the *head* relative to the scene. Interestingly, a strong response to vergence eye movements was found in area V1, which showed a dependency on visual direction, just like vertical-size disparity. This is the first report of a vertical-size disparity correlate in human striate cortex.

Keywords: motion-in-depth, optic flow, self-motion, stereo, vergence eye movements, headcentric disparity, retinal disparity

INTRODUCTION

Optic flow, or the pattern of visual motion received by a moving observer, carries important information on the layout of the environment and self-motion (for reviews: Koenderink, 1986; Warren, 2004). Most of this work has dealt with the optic flow of a single vantage point, disregarding the potential contribution of binocular visual information to optic flow processing. The one-eyed pilot Wiley Post provides an intriguing example of the common notion that binocular information is *not essential* for complex navigation tasks (Regan and Gray, 2009), but that does not imply that the binocular contribution to optic flow processing is absent. Binocular vision could improve the judgment of direction and speed of self-motion from optic flow. Indeed, various perceptual studies using a complex mixture of rotation and forward self-motion have provided evidence for this notion; e.g., stereo signals help to recover the percept of *direction* of self-motion in the presence of noise (Van den Berg and Brenner, 1994a) or to prevent a biased heading percept due to parallel motion in a transparent overlay (Grigo and Lappe, 1998). The *feeling* of self motion (vection) is also faster and stronger for self-motion consistent than for self-motion inconsistent stereo signals (Palmisano, 2002).

Imaging studies found *overlapping responses* to optic flow and stereoscopic signals in the dorsal visual pathway (Tsao et al., 2003; Brouwer et al., 2005; Likova and Tyler, 2007) pointing to potential interactions between binocular and flow signals. More recently, interactions specific to *joint processing* of motion and stereoscopic signals were reported (Smith and Wall, 2008; Rokers et al., 2009;

Ban et al., 2012; Seymour and Clifford, 2012), however, *not* using self-motion stimuli.

The BOLD response in parietal-occipital sulcus (POS) increases with the presence of a gradient of stereoscopic depth, but only when this gradient is consistent with optic flow of forward self-motion (Cardin and Smith, 2011). Within anterior POS and the MT⁺-complex, BOLD correlates to the stereo-related increase of the vection percept to simulated forward motion were reported (Kovacs et al., 2008). Together, the latter studies suggest that medial and lateral motion areas may play a role in integration of disparity signals with optic flow for direction and/or forward speed of self-motion.

Traveled distance over time cannot be recovered solely from optic flow (Koenderink and Doorn, 1987). Yet, humans are quite accurate at such judgments when optic flow is augmented with stereo signals (Campos et al., 2012). Could binocular signals somehow provide a distance scale to enable distance judgments? By itself, the change in retinal disparity of an object that accompanies self-motion cannot provide this, because retinal disparity depends on the object's depth relative to fixation not on egocentric distance. By contrast, an object's headcentric disparity (H) provides egocentric distance which could provide the missing information. Headcentric disparity of an object is defined as the difference angle between the brain's representation of the two eyes' headcentric directions toward that object (Erkelens and van Ee, 1998). That theoretical study offered an interesting analysis how the vertical component of headcentric disparity could allow the visual system to register and correct errors in signaled eye-in-head posture, while the horizontal component of headcentric

disparity would reveal the distances in the visual scene. Because headcentric direction of each eye is not given of the retinal image alone, the notion implies convergence of retinal signals with extra-retinal sources like efference copy or proprioceptive signals about the eyes' orientations in the head. To our knowledge there is no physiological evidence for the concept of headcentric disparity but a recent behavioral study intriguingly reports depth percepts consistent with non-zero headcentric disparity just prior to a saccade when retinal disparity is still zero (Zhang et al., 2010). Here we study percepts and hemodynamic signals that distinguish between changes of retinal and headcentric disparity over time.

We report psychophysical evidence that motion in depth, perceived from identical optic flow, increases in proportion to the amount of changing headcentric disparity. BOLD responses in medial motion areas also increased in proportion to the added headcentric disparity. On the other hand, the anterior part of the MT⁺-complex responded to the optic flow and retinal disparity amplitude. These results add to accumulating evidence that medial and lateral motion areas contribute distinctly to self-motion perception.

MATERIALS AND METHODS

SUBJECTS

Eight healthy subjects participated in the fMRI experiment (3 female). Four of them participated in the perceptual study (1 female); five of them participated in the vergence eye movement measurements (2 female). All subjects passed a stereo-acuity test (TNO, Netherlands Scientific Organization). We obtained written informed consent from all subjects prior to scanning, and procedures were approved by the Radboud University Medical Centre. Five subjects were experienced with optic flow stimuli and participated previously in other fMRI studies. All subjects but one (one of the authors) were naïve to the purpose of the study.

DISENTANGLING BINOCULAR CUES TO SELF-MOTION

Natural self-motion in depth offers simultaneous modulation of 3D motion, retinal disparity, and headcentric disparity. To assess their individual contributions to the self-motion percept and the evoked BOLD signals, we developed a novel stimulus that allowed full 3D stereo flow presentation with preserved motion parallax, while retinal disparity-, headcentric disparity- and vergence amplitude could be varied independently.

Our stimulus started out with simulation of self-motion on a linear track with a sinusoidal speed as a function of time (Figure 1A). Seen from the front on the screen, this resulted in typical self-motion associated flow patterns, alternating expansion flow (forward motion) and contraction flow (backward motion, Figure 1A bottom panels).

Figure 1B illustrates a top view of an observer's converged eyes, during a forward self-motion that was simulated by a collection of approaching point targets (gray arrows). Red dashed lines illustrate the motion trajectory of different flow points as seen from the cyclopean eye, for a given time lapse. The apparent angle difference between the two flow points illustrates motion parallax, which signals monocular self-motion in depth. The motion

parallax naturally is concomitant with different changes in disparity. Our goal was to maintain motion parallax, but remove differences in disparity.

This was achieved by using a virtual screen (Figure 1C), on which all flow were projected (Figure 1D). The virtual screen has the specific property that fixation on any point on its surface requires the same eye vergence. Seen on top (Figure 1C), it describes a circle (a Vieth-Muller (VM) circle), connecting one target point in front of the head and the centers of both eyes. In 3D, the virtual screen has the shape of a toroidal surface—the VM surface—which is the extension of the Vieth-Muller circle to a surface, by rotation of that circle about the inter-ocular axis. It can be thought of as an oculomotor measure of egocentric distance. Hence, after projection of the flow points on the screen, required vergence angle for fixation on any point was equated (Figure 1D).

The center of projection was the cyclopean eye, located midway between the subject's eyes (Figure 1D). Each projected point on the VM surface was presented at a distance “d” from the cyclopean eye:

$$d = \frac{1}{2} \text{IOD} \left(\cos(\alpha) \cot(V_{VM}) + \sqrt{\csc^2(V_{VM}) - \cot^2(V_{VM}) \sin^2(\alpha)} \right) \quad (1)$$

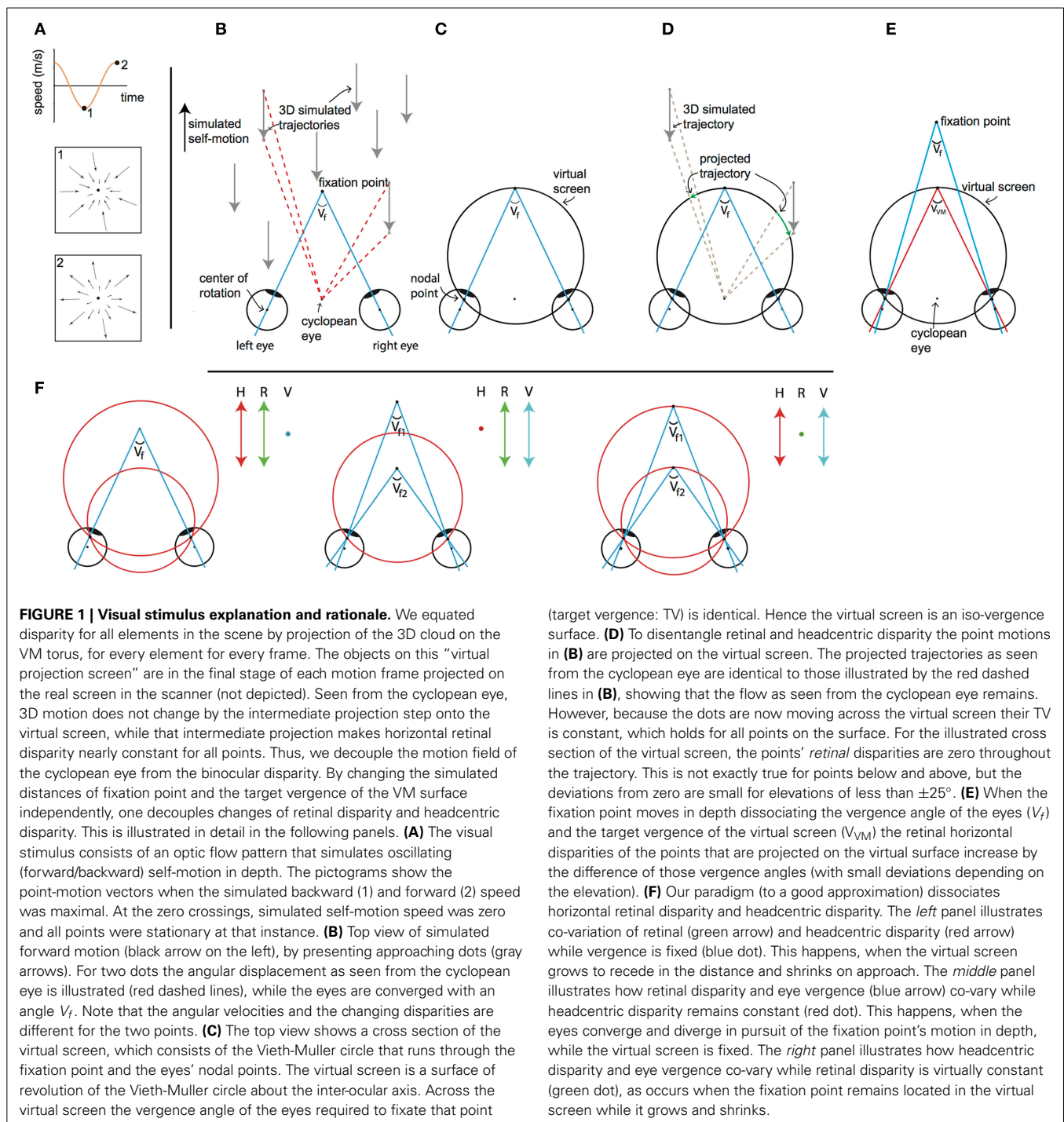
here IOD denotes the interocular distance and α the azimuthal direction of the point.

We define the required eye vergence to fixate a point (irrespective of the current eye vergence) as the *target vergence* (TV) of that point. As a result of the projection on the VM surface, all projected points maintain throughout their motion a constant TV that equals the target vergence level of the VM surface (V_{VM}). Note that V_{VM} and the azimuth angle determine the distance from the cyclopean eye (d) of the point that is projected on the virtual screen as defined in Equation (1).

Naturally, the three dimensional positions of points on the VM surface are given in Helmholtz angles: azimuth (α) and elevation (θ) and V_{VM} . For fixation in the horizontal plane (i.e., without cyclovergence of the eyes), the horizontal *retinal disparity* of points on the VM surface is close to zero. It is *exactly* zero for points on a vertical cylinder through the Vieth-Muller circle (Howard and Rogers, 2002). Points of the fixated VM surface that are elevated above or below the fixation point therefore possess crossed horizontal disparities, which turned out to be small, especially when the extent of the display was limited to $\pm 25^\circ$ of elevation (see Supplementary Material—Methods).

In some conditions we also simulated motion in depth of the (dichoptic) fixation ring, evoking eye vergence. Then, the difference of the target vergence of the fixation point (V_f) and the target vergence of the VM surface (V_{VM} , Figure 1E) determines the *horizontal retinal disparity*. Changes in horizontal retinal disparity by moving the fixation point out of the VM surface do not *exactly* generate horizontal retinal disparity ($V_f - V_{VM}$) across the entire VM surface, but spatially averaged deviations from this approximation are small (see Supplementary Material—Methods).

Finally, *headcentric disparity* of a point is defined as the vergence angle that would be required to fixate that point, irrespective of the current eye vergence. Hence, headcentric disparity of



a stimulus point equals its TV. We prefer to use this term in the remainder of the article to stress its perceptual attribute and its connection to the retinal disparity.

The levels of retinal and headcentric disparity could be varied independently by moving the fixation point in depth relative to the VM surface and by shrinking or growing the size of the VM torus.

This way, we presented an optic flow pattern (as seen from the ego-center) that was fully consistent with a sinusoidal movement

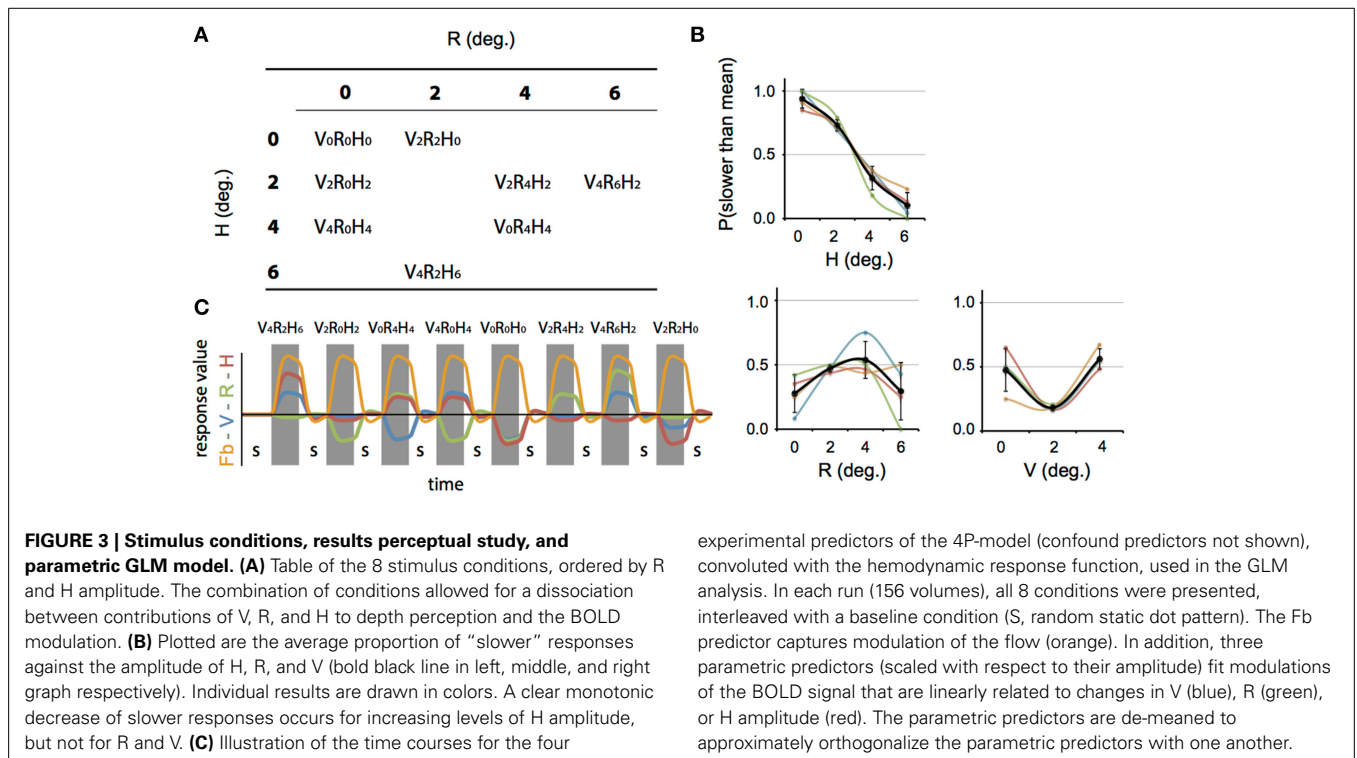
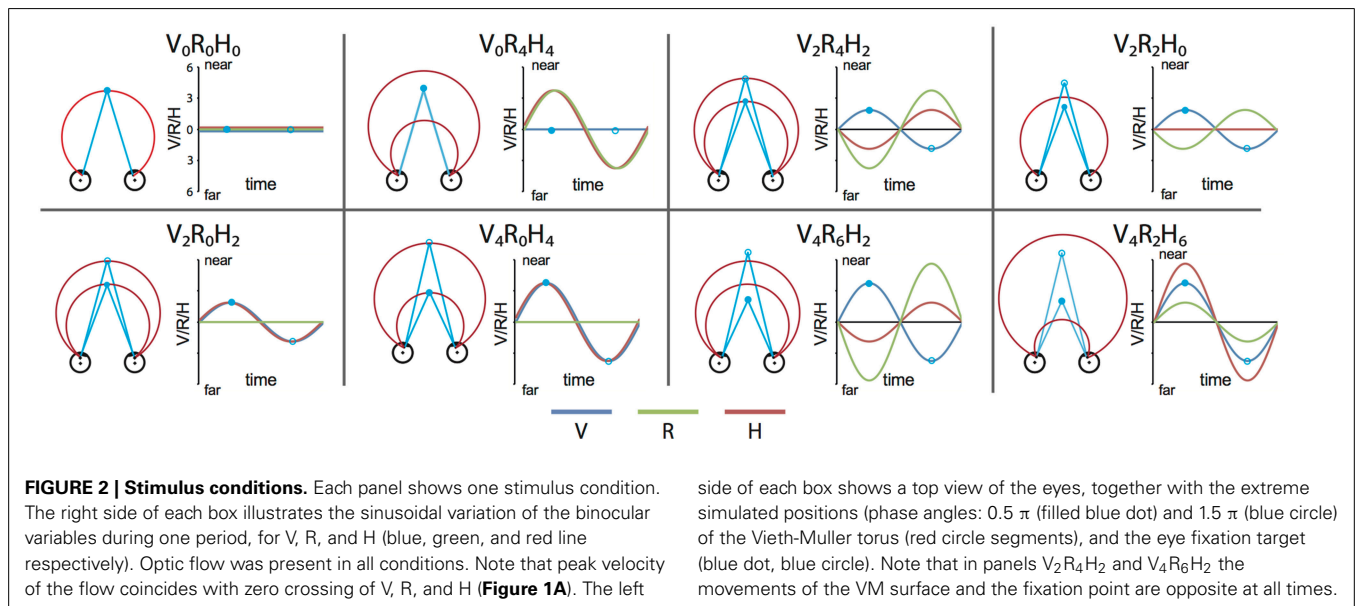
in depth including motion parallax cues to depth. Yet, for all points headcentric disparity was constant across the visual field, and horizontal retinocentric disparity was approximately constant. Both could be varied independently over time, by changing the target vergence angles of the virtual projection screen (V_{VM} , **Figure 1F** left), the fixation ring (V_f , **Figure 1F** middle), or both (**Figure 1F** right).

The flow speed was varied sinusoidally (frequency: 1/6 Hz) over time (**Figure 1A**). The parameters of the flow (see below)

were identical for *all* conditions. Just like the flow speed, the distance of the fixation point and the VM surface varied sinusoidally (1/6 Hz), always starting at the same target vergence of -6.5° .

We constructed a set of 8 stimulus conditions, in which we combined three amplitudes of eye vergence (V: 0, 2, and 4° of TV of the fixation ring) with 4 headcentric disparity amplitudes (H: 0, 2, 4, and 6° of TV of the virtual screen), which resulted in 4 amplitudes of retinocentric disparity

(R: 0, 2, 4, 6° ; **Figures 2, 3A**). Notice that we describe the amplitude of V, R, and H as the amplitude of the sinusoidal displacement. Subjects were instructed to rate the self-motion percept based on the velocity of the motion relative to their head (see below). Because we used a fixed frequency of 1/6 Hz, amplitude and peak velocity had a fixed ratio that was very close to 1. Labeling the levels of the stimuli by peak velocity or amplitude was therefore considered equivalent (**Figures 2, 3A**).



STIMULUS PRESENTATION

We used a custom-built wide-field visual projection system, previously described in more detail (Arnoldussen et al., 2011). Briefly, we projected the visual stimulus on a small projection screen very close to the subject's eyes (~ 3 cm). Subjects wore eye contact lenses (+25 to +33 D, dependent on individual refraction errors) in both eyes to allow for easy accommodation at this distance. After positioning into the head coil, the screen distance was adjusted until the subject reported sharp viewing for both eyes on the screen. Clay putty was put on the nose bridge to prevent cross viewing of the left and right eye projections. Next, an eye calibration procedure was performed inside the bore to assess the precise position of the eyes relative to the stimulus screen (Van den Berg, 1996). To check the calibration, a fixation mark was projected in stereo on a simulated distance of 5 meters. If the subject did not report easy fusion of the fixation spot in depth, the calibration procedure was repeated. Then, some of the stimulus conditions were shown prior to scanning to familiarize the subject with pursuit in depth. We ensured that the subject reported seeing the fixation point at eye level and reported the binocular fixation point moved in depth according to the subject's straight ahead.

VISUAL STIMULUS

All stimuli were generated on a Macintosh MacBook Pro Notebook, using OpenGL-rendering software. The same set of stimulus conditions was used for the psychophysical and the imaging studies. The simulated scene was a 3D stereoscopically presented cloud of dots (250 white dots against a black background) with one fixation ring. Dots were asynchronously refreshed every 1.0 s and distributed randomly within a volume in front of the eyes. Volume dimensions were 8600 (H) \times 6000 (V) \times 2500 (D) mm, with the nearest surface abutting the projection screen ($0.75 \times$ distance of the actual screen or about 2.3 cm). Pixel-size of the projection on the screen was $0.2\text{--}0.34^\circ$, depending on eccentricity (large at the center because of the nearer distance, smallest at largest eccentricity of 60°). Because we applied sub-pixel positioning by OpenGL software, 1/10 the of the pixel displacement resolution is offered, i.e., $0.02\text{--}0.034^\circ$. Oscillation frequency (f) of 1/6 Hz, dot-displacement amplitude (287 mm) and peak speed of dot movement ($A_{\text{Flow}} = 300$ mm/s) were fixed for all stimulus conditions.

In some conditions we also simulated motion in depth of the (dichoptic) fixation ring, evoking eye vergence. We note that the accommodation demand during the eye vergence conditions was smaller and opposite from normal viewing. Convergence was concomitant with *increased* rather than *decreased* distance between eye and the fixated position on the screen. However, that distance was nearly constant (about 1% variation during the largest simulated depth excursions of the fixation target).

PERCEPTUAL STUDY

In a psychophysics session, subjects ($n = 5$, all of whom participated in the fMRI experiment) were seated upright and with the head stabilized by a bite-board, viewing the screen at approximately the same distance as during the scanning session (~ 3 cm). Procedures and setup were similar to the fMRI setup described

below. Subjects viewed all stimulus conditions in random order without static conditions and for a shorter period than in the fMRI experiment ($T = 2$ s). Subjects were instructed to judge for each trial the speed at which the scene was oscillating in depth relative to their head. Subjects pressed keyboard buttons, to indicate the speed as slower or faster than an internal mean, which they built up during the experiment.

Because V, R, and H amplitudes were distributed unequally over the 8 stimulus conditions (Figure 3A), the number of repetitions for V, R, and H amplitudes could not all be balanced. We chose to balance the number of repetitions across different levels of H amplitude, offering an unequal amount of repetitions for each stimulus condition. This allows us to determine the psychometric function for each level of H with equal reliability but not so for R and V. Irrespective of that reliability we expect a monotonic relation between at most one of the predictors' stimulus level and "slower" choice-frequency because the levels can be arranged as an ascending ordered set only for one predictor at a time.

VISUAL STIMULI—FUNCTIONAL LOCALIZERS

Closely following our procedures in a previous study (Arnoldussen et al., 2011) we identified the main visual regions V1, V3ab, and V6 through wide-field retinotopy (Pitzalis et al., 2006), and distinguished sub-regions within the flow-responsive posterior portion of the medial inferior temporal sulcus as described before (MT⁺; Dukelow et al., 2001; Huk et al., 2002). We distinguished MT⁺/contralateral (MT⁺/c), which responds only to flow presented in the contra-lateral hemifield, and MT⁺/bilateral (MT⁺/b) which responds also to flow presented in the ipsi-lateral hemifield (MT⁺/bilateral, (MT⁺/b). We refer to the MT⁺ sub-regions in this way, because of growing evidence that they likely constitute several different functional regions (Kolster et al., 2010). The admittedly crude dissection of MT⁺ in bi-lateral and contra-lateral flow-sensitive sub-regions is useful nonetheless; a bi-laterally responsive sub-region is likely to have a distinctive functional contribution to wide-field flow processing and may show a different type of integration with binocular cues.

DESIGN fMRI

The 8 stimulus conditions were presented in blocks of 18 s (3 periods), interleaved by 18 s of the baseline condition, resulting in runs of about 5 min. The conditions were presented pseudo-randomly and counter-balanced. The baseline condition was a pattern of stationary dots, placed randomly and independently for each eye (i.e., no binocular matches) with limited lifetime (1.0 s). Subjects were instructed to fixate the binocular fixation ring with TV angle of -6.5° during the static condition and to follow the binocular target when it moved in depth. Subjects performed 8 (subject 7 and 8) to 16 (subject 1 and 3) runs in total (average: 12.8 runs). All fMRI acquisition took place in three (subject 1 and 3) or two separate sessions.

MRI ACQUISITION

The MRI experiment was conducted on a 3 Tesla TIM trio Siemens scanner at the Donders Center for Cognitive Neuroimaging (Nijmegen, The Netherlands). For each subject, we obtained a high-resolution full-brain anatomical scan using a

32-channel head coil (T1-weighted MPRAGE, 192 slices, 256×256 matrix, resolution of 1 mm^3).

For the experimental scans, we used an 8-channel occipital surface coil. Scans were obtained with a resolution of 2 mm iso-voxel [T2*-weighted; single-shot echo-planar imaging; 32 slices; repetition time (TR): 2 s; echo time (TE): 30 ms]. The experimental runs consisted of 156 volume acquisitions. The initial 2 TR's were discarded to account for scanner drifts.

Subjects also performed 6 localizer runs (2 for polar angle, 2 for eccentricity, 2 for MT⁺/b-MT⁺/c localizer).

PREPROCESSING fMRI DATA

Brainvoyager QX (BVQX, version 2.6) was used for the analysis of functional and anatomical images (Brain Innovation, Maastricht, The Netherlands). All functional runs were corrected for motion and inter-slice scan time differences and were subsequently aligned with the high-resolution anatomical scan. For the retinotopy and MT/MST localizer scans, a high-pass filter was applied on each voxel's data time course by removing the fit of a General Linear Model (GLM) with predictors defined by a linear, a sine and a cosine function (the latter two with a period equal to the run duration). For the experimental runs, these three predictors were added as confound predictors to the GLM model, both for individual and multi-subject GLM analyses.

Using the high-resolution anatomical scan, white and gray matter boundaries were defined and used for the construction of the inflated and flattened representation of each subject's left and right hemisphere (using automatic segmentation by BVQX, followed by manual refinement).

ANALYSIS—FUNCTIONAL LOCALIZERS

The phase of the BOLD response to the retinotopic stimuli (rotating wedge, expanding ring) was determined by cross-correlation. Boundaries of visual areas V1 and V3ab were defined based on the phase reversal of the polar angle mapping stimulus by eye, on the flattened representation of each subject's left and right hemisphere (Sereno et al., 1995). In only a few hemispheres, we could distinguish the boundary between v3b and v3a on the basis of the retinotopy (Larsson and Heeger, 2006). Therefore, we decided to treat these regions as a single ROI, which we refer to as V3ab. V6 was defined as a region within the POS, medial to V3ab, containing an entire representation of the contra-lateral hemifield and an eccentricity map (Pitzalis et al., 2006).

For the MT⁺/b-MT⁺/c localizer, we used a GLM with separate predictors for full field, left-visual-field, and right-visual-field flow presentation. First, the region within the medial temporal sulcus that was highly responsive ($p < 0.01$, uncorrected) to a GLM contrast between full field optic flow and the static conditions was identified as MT⁺. MT⁺/b was then defined as an anterior sub-region of MT⁺ containing all contiguous voxels that responded to both ipsi-lateral and contra-lateral flow (with respect to the current hemisphere, $p < 0.05$ uncorrected; Dukelow et al., 2001; Huk et al., 2002). MT⁺/c was defined as a posterior sub-region of MT⁺ that consisted of all contiguous voxels responsive to contra-lateral flow ($p < 0.01$ uncorrected).

fMRI ANALYSIS

For the searchlight approach and subsequent multi-subject analyses, we used a parametric GLM model with 4 main predictors (4P-model; **Figure 3C**), described by:

$$Y = \beta_1 * Fb + \beta_2 * V + \beta_3 * R + \beta_4 * H + \beta_5 * L + \beta_6 * S + \beta_7 * C + \beta_8 \quad (2)$$

The first predictor captures the general BOLD response to the flow (Fb-predictor), which was identical for all conditions. The next three predictors capture parametric modulations of the BOLD signal by the amplitude of vergence (V-predictor), retinal disparity (R-predictor), and headcentric disparity (H-predictor; **Figure 3C**). Note that the β -values for the parametric predictors do not reflect level activation, but rather a region's linear sensitivity to amplitude changes in V, R, or H. β -values thus indicate % BOLD-signal per degree amplitude.

Predictor-values for the different conditions were offset to a mean of zero (Wood et al., 2008) to improve the orthogonality of the predictor functions. We included three confound predictors (a linear trend (L), a sinus (S), a cosinus (C), and a baseline predictor (β_8).

We applied the 4P-model on the data of every hemisphere separately, without voxel restrictions. We tested for a significant difference between baseline response (dichoptic static dots) and several joint responses: Flow baseline and R (R), flow baseline and H (H), and flow baseline and R and H (RH). A conjunction analysis refers to a logical AND operation for combinations of GLM contrasts. For every hemisphere, the results for R, H, and RH are plotted on a flat map representation, as depicted in **Figures 4, 5**. Per hemisphere, minimum p-level for each contrast was set to 75% of the p-level at which no more blobs were visible within any of the ROIs (except V1) with a cluster threshold of 25 mm^2 . If this value was lower than $p < 0.05$, it was set to $p < 0.05$. To prevent too restrictive statistics, minimum p(t)-levels beyond p(6.0) were set to p(6.0) (equal to $p < 2.7 e^{-9}$, uncorrected). P(t)_{min} values for each individual contrast are given in parentheses following subject name for R and H contrast, respectively.

Figure 8B shows the result of a GLM contrast between a conjunction of Fb and V (V), and baseline, for two hemispheres.

The multi-subject ROI results (**Figures 6, 8A**) were obtained by a random effects analysis on the averaged data of each ROI, using a multi-subject GLM with predictors separated for each subject. GLM contrasts were specified for each subject and the resulting mean values across subjects was tested for significant deviation from zero (two-sided *t*-test, $p < 0.05$, uncorrected). Contrasts tested include a deviation from zero for beta values of V, R, and H separately (**Figure 6** for R and H, **Figure 8A** for V); no conjunctions with the Fb predictor were included in these analyses.

We also analyzed the BOLD responses to each particular combination of binocular signals in our stimuli, using a GLM model with 8 main predictors (the 8P-model):

$$Y = \sum_{i=1}^8 \beta_i * E_i + \beta_9 * L + \beta_{10} * S + \beta_{11} * C + \beta_{12} \quad (3)$$

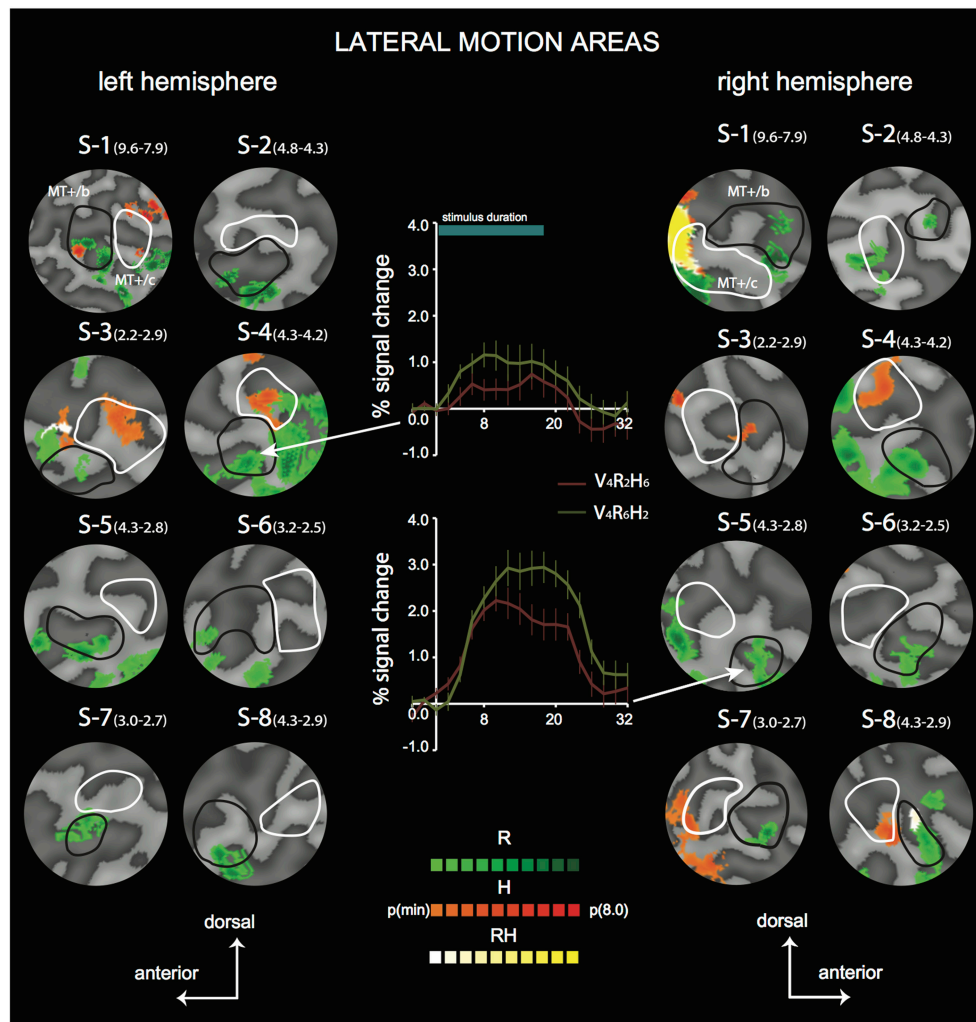


FIGURE 4 | Differential sensitivity to binocular signals in lateral motion areas. Sections of the flattened representation of all hemispheres, encompassing MT+/c (white outline) and MT+/b (black outline). On top, peak activations are drawn for R (green—dark green), H (orange—red), and RH (white—yellow). P(min) was set independently for each hemisphere to 75% of the p -value at which the most significant area of activation

surfaced $P(t)_{(min)}$ values are given in parentheses following subject name. For example, $P(t)_{(min)}$ for S-1 are 9.6 for R contrast, and 7.9 for H contrast. Center graphs show the average BOLD response for two illustrative conditions ($V_4R_2H_6$ and $V_4R_6H_2$), in the R peak activation area indicated by the white arrows for S4-LH (upper graph) and S5-RH (lower graph). Error bars represent SEM.

Including a separate predictor for each experimental condition (E_1 – E_8), three confound predictors [a Linear trend (L), a Sine (S), Cosine (C)], and baseline activity (β_{12}). This allowed us for example to distinguish the condition without temporal modulation of the binocular signals ($V_0R_0H_0$) from all the other conditions that add at least one temporally modulated binocular signal to the flow.

Vertical disparity—analysis

We observed in V1 a strong BOLD modulation by V, which was distributed unequally across the cortical surface (Figure 8B). We hypothesized that this activation is caused by the changing vertical retinal disparity during vergence eye movements.

Vertical disparity (χ) depends on visual direction [Helmholtz angles azimuth (α) and elevation (θ)] and the convergence of

the eyes (TV of the fixation point: V_f). This is captured by the following formula:

$$\chi(\alpha, \theta, V_f) = \frac{1}{2} \sin(2\theta) * \tan(\alpha) * V_f \quad (4)$$

This formula is based on Equation (1) of Read et al. (2009) with the constraints that the mean and difference of the eyes’ torsions, horizontal cyclopean eye orientation and the vertical vergence are zero (Read et al., 2009). This is appropriate for the experiments that we conducted because subjects only made slow vergence eye movement along a trajectory straight ahead at eye height.

We selected a sub-group of all V1 voxels for this analysis, i.e., those voxels that showed significance ($p < 0.05$, uncorrected) on the polar angle mapping and the eccentricity mapping, and on

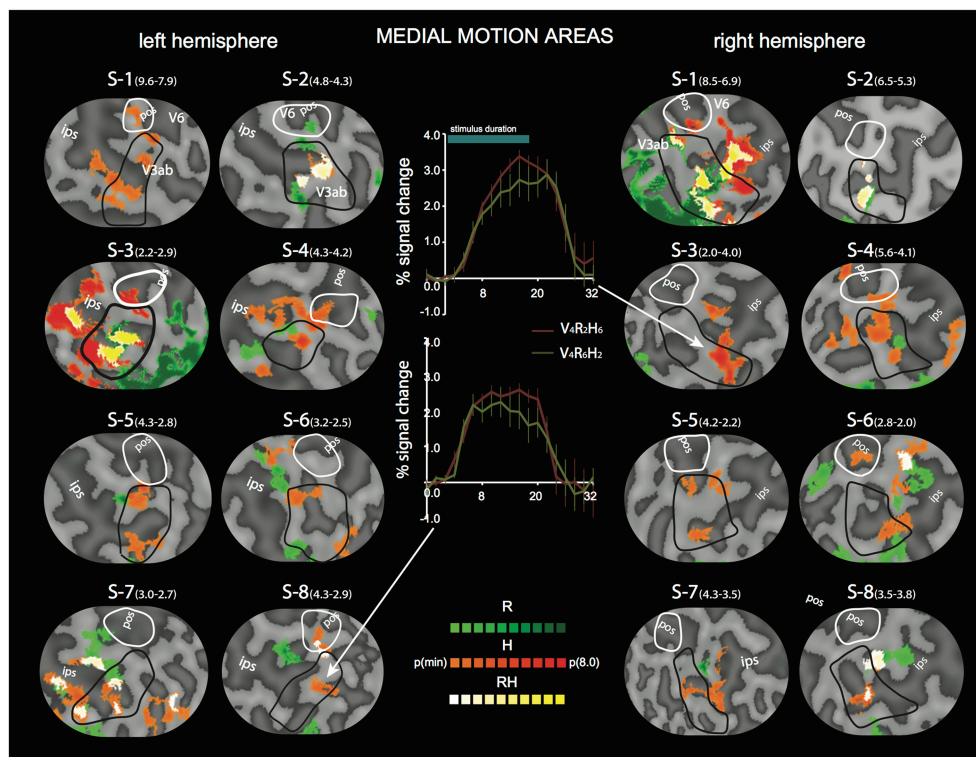


FIGURE 5 | Differential sensitivity to binocular signals in medial motion areas. Sections of the flattened representation of all hemispheres, encompassing V6 (white outline) and V3ab (black outline). On top, peak activations are drawn for R (green—dark green), H (orange—red), and RH (white—yellow). $P(t)_{(min)}$ was set independently for each hemisphere to 75% of the p -value at which

the most significant area of activation surfaced. $P(t)_{(min)}$ values are given in parentheses following subject name. Center graphs show the average BOLD response for two illustrative conditions ($V_4R_2H_6$ and $V_4R_6H_2$), in the H peak activation area indicated by the white arrows for S3-RH (upper graph) and S8-LH (lower graph). Error bars represent SEM.

the GLM contrast of β_V against static baseline. (For the polar angle mapping and eccentricity mapping, significance defines a significant correlation between the voxel's time course and the optimal lag value.) We converted each voxel's polar angle (μ) and eccentricity (ecc) fit from retinotopic mapping to the corresponding azimuth and elevation angles (see Supplementary Material—Methods). Then we examined if β_V was dependent on the voxel's receptive field location as predicted by Equation (4).

To test for a genuine relation between β_V and vertical disparity, the test was repeated with the retinal disparity beta value (β_R). Here, the voxel sub-group contained all voxels that showed significance ($p < 0.05$, uncorrected) on the polar angle mapping, and the eccentricity mapping, and on the GLM contrast of β_R against static baseline. Voxels' gain values (β_V or β_R) were normalized with respect to each subject's peak response in V1 in order to compensate for general BOLD level differences between subjects.

For each voxel, the predicted RF location dependency was computed as explained above, and the Beta values (either β_R or β_V) were obtained by application of the 4P-model on the each-subject's V1. For statistical testing and plotting, voxels were distributed in 10 equidistant bins, covering the range of the predicted RF location dependency value. Because the distribution of voxels' β_V and β_R response within each bin was highly skewed,

the median was taken as a measure of central tendency, on which the linear regression analysis was applied.

MODEL COMPARISON

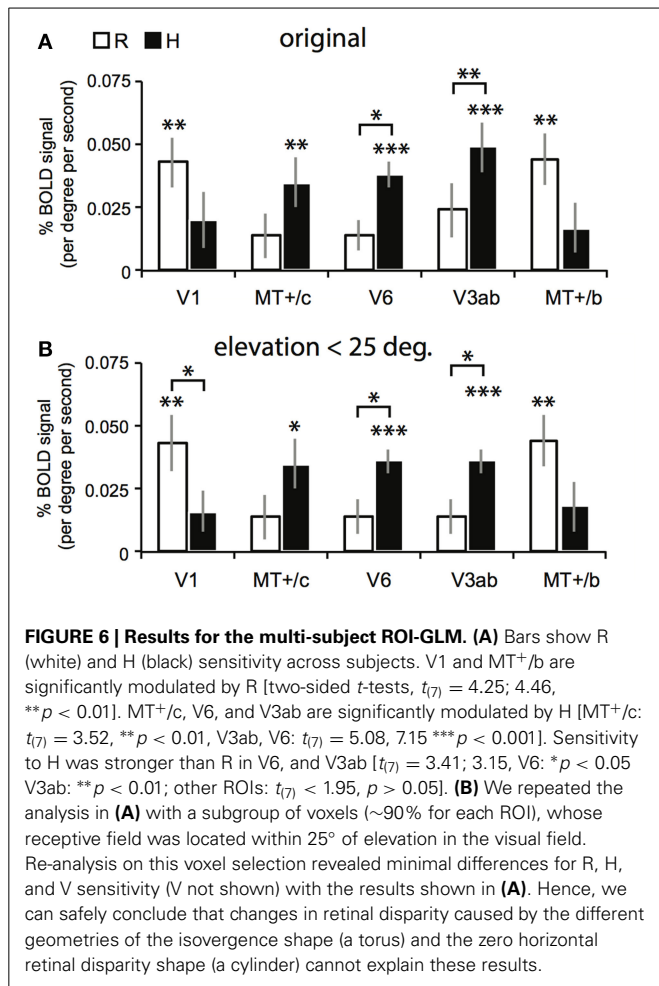
We assessed quality of fit of the 4P-model (Equation 2) on the fMRI data, by comparing it with the fit of the 8P-model (Equation 3). In the latter model, the amplitude of the BOLD response is fitted for each stimulus condition separately.

For each subject, an amplitude estimate was obtained for both models. Next, these values were correlated for each ROI for each subject, as plotted in **Figure 7A** for area V6, and the average of the correlation value across subjects is plotted in **Figure 7B**.

For all subjects, comparisons of the beta values of the three confound predictors revealed no significant difference between the two models for any ROI [two-sided t -test, all ROIs $t_{(7)} < 0.60$, $p > 0.56$]. Hence, the confound predictors were excluded from the comparison, assuming equal estimates.

Inter-subject variability

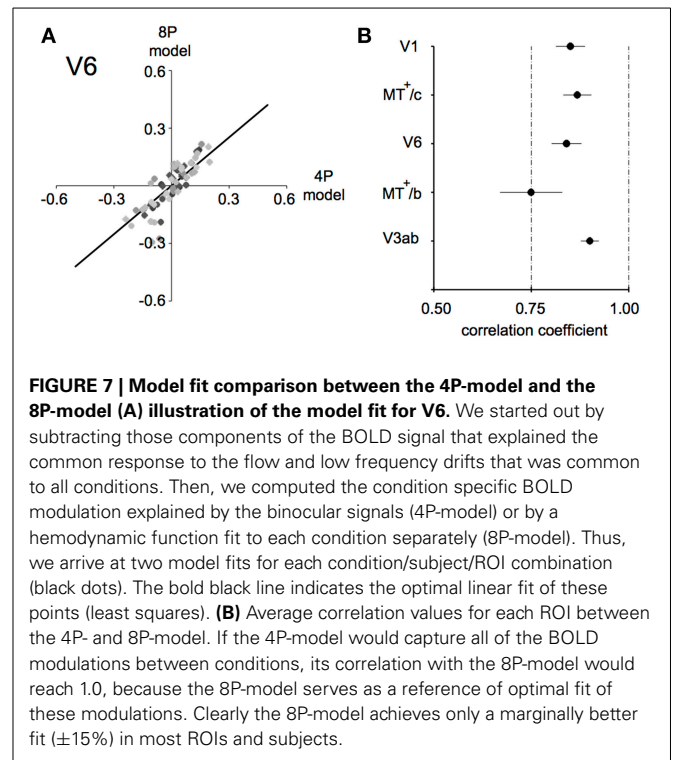
To assess spatial congruence across subjects regarding V, R, and H sensitivity across cortex, we performed a Cortex-Based Alignment (CBA) analysis for both hemispheres. The inflated cortical representation of subject's left and right hemispheres separately were morphed into a spherical structure in which original



curvature information of each vertex was maintained, which allowed for across subject analyses. Subsequently, all hemispheres were aligned by minimizing the curvature information between spheres, using the standard minimizing procedures of the CBA tool built-in in BrainVoyager QX (Goebel et al., 2006). The resulting curvature information of the averaged sphere, which also functioned as the target sphere for alignment, was used to reconstruct an inflated representation of left and right hemisphere. The single-subject GLM results of each subject, as depicted in **Figures 4, 5**, were then transformed to match the averaged hemisphere. This allowed us to establish for each contrast (V, R, and H) the proportion of overlap between the individual activation maps (at $p < 0.05$, uncorrected for each individual contrast), as plotted in **Figure 11**.

MEASURING VERGENCE EYE MOVEMENTS

Vergence eye movements were assessed in a separate session, in a dummy MR set-up, using an Eyelink II system (Eyelink® II, version 1.11, SR research, Canada). The experiment was performed in a dummy scanner with the same projector and very similar projection setup as during scanning. Contrary to the Eyelink® 2000 system available at the scanner, the Eyelink® II system uses two small separate cameras, which gave us the flexibility needed



to measure eye movements in both eyes during stimulus presentation in the dummy scanner, with a screen close to the head.

Subjects ($n = 5$) wore contact lenses and underwent the same calibration procedure as in the main experiment, and lay in supine position in the dummy scanner bore. Subjects were instructed to attend to and look at the stereoscopic fixation ring at all times. Stimulus conditions were presented in exactly the same stimulus runs as in the main experiment, except for the static conditions, which were reduced to 4 s. Each subject completed four or five runs, and thus made 4 or 5 repetitions for each stimulus condition. Raw traces for left and right eye are illustrated for two conditions in **Figure 9B**, for one subject (S-1). Because the subject's eyes were very close to the projection screen, the Eyelink was calibrated only on a small portion of the actual screen and the Eyelink resolution was limited to about 0.4°, as can be observed in **Figure 9B**. Because sample frequency was high (500 Hz), gain and phase analysis of the low frequency vergence movement was hardly affected by these discrete steps (**Table 2**).

Saccades and blinks were removed from the left and right eye movement data and interpolated using a linear fit function. We ensured that the subject's eyes were positioned along the horizontal axis of the projection. Therefore, for each subject, and for each vergence condition separately, we defined the vergence signal as the difference of the left and right eye azimuth signal. For each subject, vergence signals were averaged over the 4/5 repetitions and fitted by a sine resulting in an amplitude and phase lags, as shown in **Table 2**. Horizontal drift values were assessed by fitting a linear function on the residuals of the sinusoidal fit (for the vergence conditions), or on the raw data after blink and saccade removal (for the non-vergence conditions).

The effect of inaccurate vergence eye movements on the GLM results were assessed by comparing the GLM results for a $gain_{1.0}$ model (the original 4P-model with V_o and R_o), a model with hypothetical gain values for each condition of 0.5 ($gain_{0.5}$ -model), and a $gain_{eye}$ model. Here, V and R were adapted based on the measured gain in each condition (V_{eye} and R_{eye}). Specifically:

$$V_{eye} = gain * V_{target}$$

and

$$R_{eye} = |(gain * V_{target} - direction * H)|$$

Where *direction* refers to whether the direction of H is consistent (1) or opponent (-1) to the direction of V . Flow baseline and H are independent of V and thus are unchanged in the $gain_{eye}$ model and $gain_{0.5}$ model. For one subject with the lowest vergence gain, the described searchlight procedure was repeated for the $gain_{eye}$ and $gain_{0.5}$ model, and GLM contrast for R and H and RH were plotted on a flatmap representation of the cortex in both left and right hemisphere, using the same p -values as in the original model (Figure 10).

Goodness of fit (R^2) for both models was examined by performing a single-subject ROI-GLM on each ROI for each subject in whom we measured pursuit in depth accuracy.

RESULTS

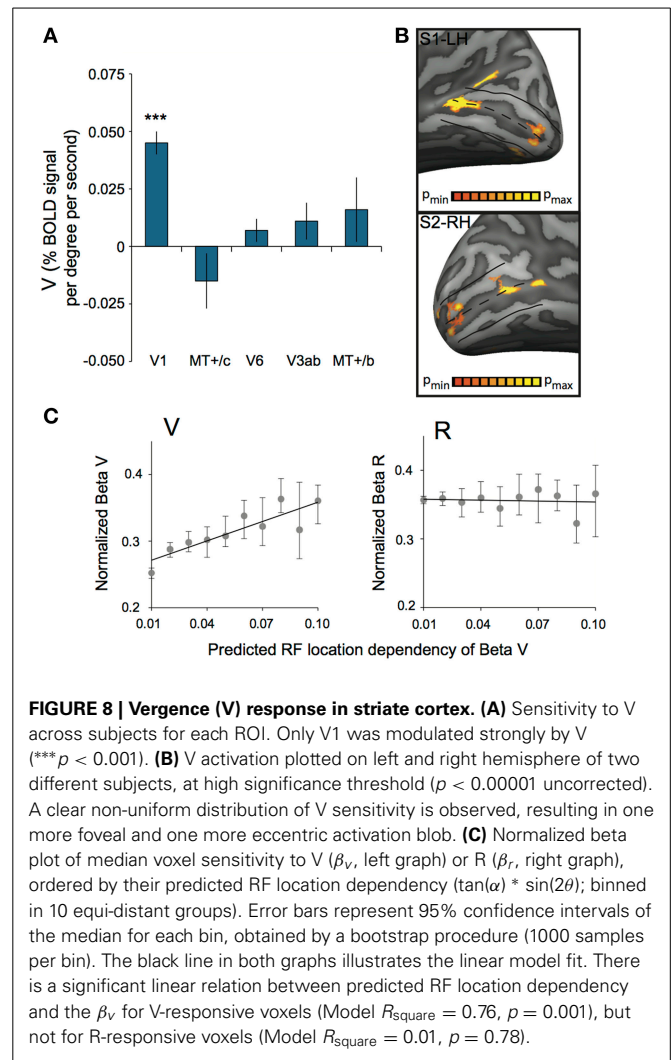
PERCEPTION OF DEPTH MODULATION

All stimuli evoked a vivid percept of a rigid scene that approached and receded from the subject periodically. The magnitude of that percept varied across conditions in relation to the *binocular* components of the stimulus, because the optic flow as seen from the ego-center was identical. Subjects viewed multiple repetitions of all stimulus conditions in a random order, and rated for each stimulus condition the relative speed between themselves and the cloud of dots. Ratings were expressed as “slower” or “faster” relative to an internal mean, which subjects acquired over the course of the experiment.

Each stimulus condition is jointly characterized by its V , R , and H amplitude (Figure 3A). Which (combination) of these quantities matches the perceived motion in depth? Because the amplitude of V , R , and H were decoupled, one expects a monotonic (psychometric) relation between percentage slower scores and the level of those quantities contributing to the percept (Regan and Gray, 2009). We found, only for the level of H , a monotonic relation with the percentage slower (Figure 3B), indicating that the headcentric disparity amplitude modulated the perceived speed of motion in depth from flow.

LOCALIZERS

In 8 subjects, we measured Blood-Oxygenated Level-Dependent (BOLD) signals using a Siemens Tim-Trio 3T scanner, using a previously described wide-field stimulus presentation (Arnoldussen et al., 2011). We performed a set of functional localizers to independently establish the location of motion responsive areas in dorsal visual cortex: our Regions of Interest



(ROIs). These include V1, V3ab, and V6, which were identified using wide-field retinotopy (Sereno et al., 1995; Pitzalis et al., 2010), and also MT^+ -contralateral (MT^+/c) and MT^+ -bilateral (MT^+/b), by presentation of contra-lateral and ipsi-lateral flow (see also Materials and Methods). Average ROI Talairach coordinates are given in Table 1.

FLOW AND BINOCULAR CONTRIBUTIONS TO THE BOLD SIGNAL

All subjects were exposed to multiple functional runs that contained all 8 experimental conditions, alternated by a baseline condition, presented in a blocked design.

We first established the contributions to the BOLD signal of the temporal modulations of the flow and the binocular signals. We used the 8P-model instead of the 4P-model to make minimal model assumptions. The $V_0R_0H_0$ condition characterizes the BOLD response to binocular presentation of temporally modulated flow, without a *binocular* component to its modulation in time. Taking a grand average across subjects, across ROIs, we found a BOLD response to $V_0R_0H_0$ of about 0.45% (V1: 0.41%, V3ab: 0.56%, MT^+/c : 0.36%, MT^+/b : 0.51%, V6: 0.43%).

Next, we examined the contribution of temporal modulation of the binocular signals to the BOLD signal, by performing in each ROI separately a multi-subject GLM contrast between the BOLD response to the $V_0R_0H_0$ condition and the average of the remaining 7 conditions. The contrast (average of the 7 conditions $> V_0R_0H_0$ condition) was significant for V1, MT^+/b , V3ab, and V6 [mean modulation difference (mdf): 0.25%; 0.11%; 0.17%; 0.10%, two-sided t -test, $t_{(7)} = 4.66; 2.53; 3.59; 3.29$, $p < 0.05$], but not for MT^+/c [mdf: 0.11%, $t_{(7)} = 0.28$, $p = 0.78$]. Hence, the addition of binocular self-motion cues to the flow stimulus evoked considerable modulations in the BOLD response, in all ROIs but MT^+/c .

SEARCHLIGHT APPROACH—INDIVIDUAL SUBJECTS

To distinguish the different (V, R, H) contributions of binocular signals to the BOLD responses in flow sensitive areas, we used a searchlight approach, examining GLM contrasts (conjunctions of flow and any of the binocular components) from the 4P-model of individual data (Materials and Methods). For the *lateral* motion areas (MT^+/c and MT^+/b), we observed a strong R modulation that was localized within or at least encompassed MT^+/b in 14 out of 16 hemispheres (Figure 4, flat map section); peak modulation by H was found less frequently (9 out of 14 hemispheres), and was most often located within MT^+/c (7 hemispheres).

In most hemispheres, the R blobs were highly robust ($p_{\min} < 0.001$ uncorrected in 11/16). The strong R modulation is illustrated by the response preference for the stimulus with the highest speed of retinal disparity ($V_4R_6H_2$) rather than headcentric disparity ($V_4R_2H_6$) in R-sensitive regions when vergence was identical (Figure 4, graphs).

For the *medial* motion areas, we observed strong modulation by H in all subjects (V3ab, V6, Figure 5, flat map section). H modulation peaks were found in all hemispheres (16/16) in V3ab. Also in V6 (7/16 hemispheres) such peaks were found. In some subjects the contrast in H regions between responses to $V_4R_2H_6$ and $V_4R_6H_2$, showed a clear preference for the former (Figure 5, graphs), but in most subjects the contrast was less pronounced because of co-modulation of that ROI by R. Half of the hemispheres (8/16) showed also a peak R modulation in V3ab; one hemisphere showed peak R modulation in V6.

The searchlight approach suggested a stronger response to H compared to R within medial motion-responsive cortex. In contrast, a sub-region of MT^+/b responded more specifically to R. To find out whether the whole ROI had a response bias to one or the other type of binocular signal, we performed a random effects multi-subject GLM, for each ROI separately, using the 4P-model. Now, all voxels of the ROI were included without selection of voxels that responded significantly to flow. Congruent with the individual patterns of activation, MT^+/c , V3ab and V6 were significantly modulated by H [Figure 6A, two-sided t -test: MT^+/c : $t_{(7)} = 3.52$, $p < 0.01$; V3ab: $t_{(7)} = 5.08$, $p < 0.01$; V6: $t_{(7)} = 7.15$, $p < 0.001$]. H modulation was significantly larger than R modulation in V6 and V3ab [$t_{(7)} = 3.15; 3.41$, $p < 0.05$]. V1 and MT^+/b were significantly modulated by R [$t_{(7)} = 4.25; 4.46$, $p < 0.01$].

The use of a VM torus generates changes in disparity not captured by the 4P-model that are notable for eye elevation angles

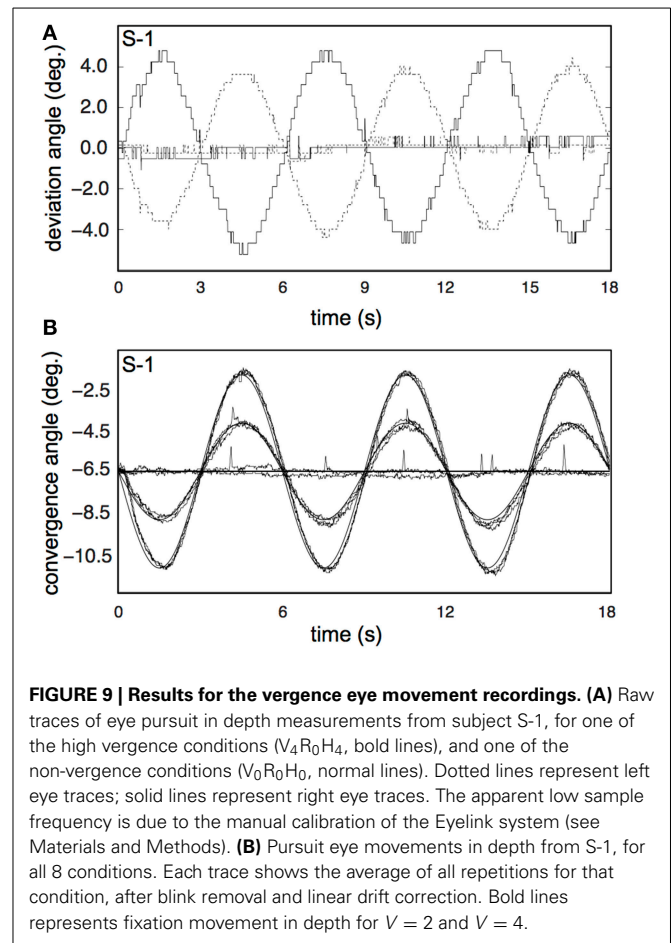


FIGURE 9 | Results for the vergence eye movement recordings. (A) Raw traces of eye pursuit in depth measurements from subject S-1, for one of the high vergence conditions ($V_4R_0H_4$, bold lines), and one of the non-vergence conditions ($V_0R_0H_0$, normal lines). Dotted lines represent left eye traces; solid lines represent right eye traces. The apparent low sample frequency is due to the manual calibration of the Eyelink system (see Materials and Methods). **(B)** Pursuit eye movements in depth from S-1, for all 8 conditions. Each trace shows the average of all repetitions for that condition, after blink removal and linear drift correction. Bold lines represent fixation movement in depth for $V = 2$ and $V = 4$.

larger than 25° (see Materials and Methods). To exclude possible confounding effects by these changes to the reported signal changes by V, R, and H, we repeated the multi-subject analysis, using only those voxels of which the preferred response was located within 25° of elevation in the visual field (calculated using the results of retinotopic mapping, see Materials and Methods). For all regions, this resulted in the exclusion of maximally 13 percent of voxels on average across subjects (for area V6). The results are very similar to the original analysis (Figure 6B). MT^+/c , V3ab and V6 were significantly modulated by H [two-sided t -test: MT^+/c : $t_{(7)} = 3.43$, $p < 0.05$; V3ab: $t_{(7)} = 7.79$, $p < 0.001$; V6: $t_{(7)} = 7.79$, $p < 0.001$]. H modulation was significantly larger than R modulation in V6 and V3ab [$t_{(7)} = 3.01; 3.01$, $p < 0.05$]. V1 and MT^+/b were significantly modulated by R [$t_{(7)} = 4.13; 4.51$, $p < 0.01$].

In sum, we found a notable difference in sensitivity to binocular signals in medial motion areas (high H sensitivity) and lateral motion areas (high R sensitivity).

VERGENCE RESPONSE IN AREA V1

Additionally, we assessed ROI responsiveness to vergence eye movements (V) across subjects, using the multi-subject GLM. We found significant activation to vergence eye movements only in V1 [$t_{(7)} = 8.86$, $p < 0.0001$, other ROIs: $t_{(7)} < 1.36$, $p > 0.21$, Figure 8A]. Inspection of the distribution of V activation across

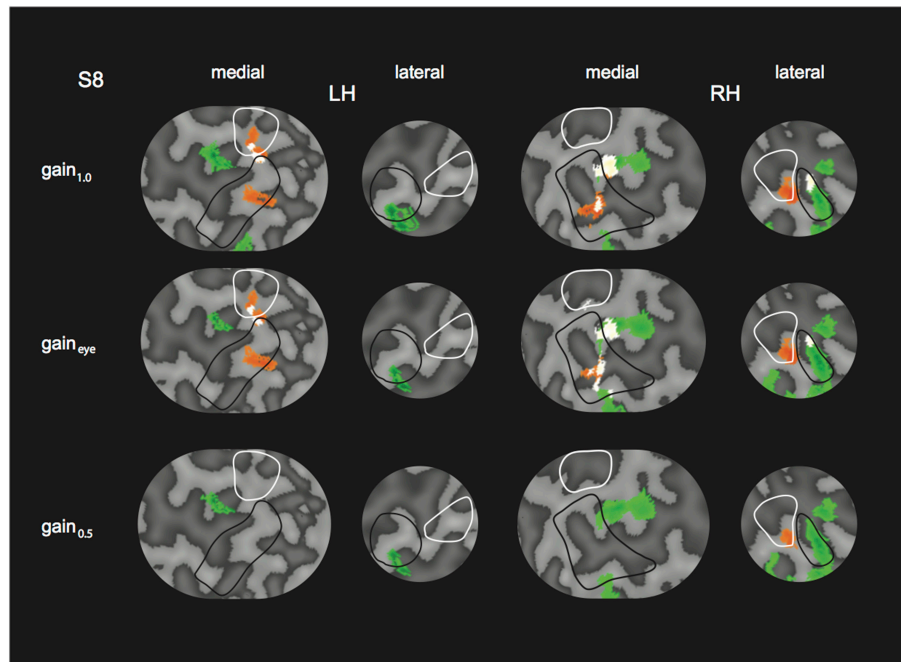


FIGURE 10 | Modulation by binocular signals for the 4P-model, with predictors adapted for different gain levels. Shown are sections of the flattened representation of both hemispheres in subject 8, encompassing medial and lateral motion areas. On top, peak activations are drawn for R (green—dark green), H (orange—red), and RH (white—yellow), for the

$gain_{1,0}$ model, the $gain_{eye}$ model, and the $gain_{0,5}$ model. $P(\min)$ was set to the same level as in the original model (Figures 4, 5). The $gain_{eye}$ model was adjusted using measured gain values in subject 8 (V2R2H0: 0.96; V2R4H2: 0.96; V2R0H2: 1.03; V4R0H4: 0.94; V4R2H6: 0.99; V4R6H2: 0.90).

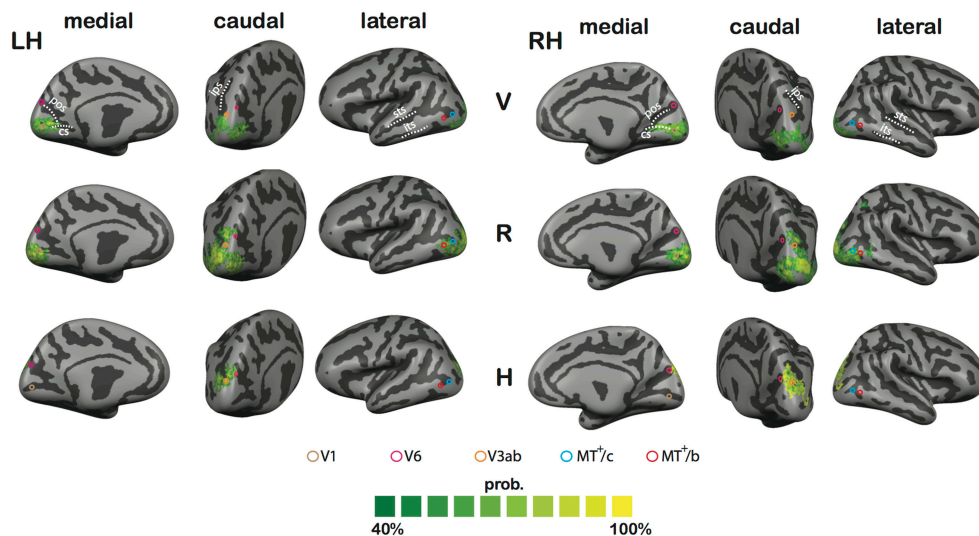


FIGURE 11 | Results of the CBA analysis. Shown is the proportional spatial overlap (green to yellow) between subject's functional maps for V (top row), R (middle row), and H (bottom row; $p < 0.05$, uncorrected for all individual contrasts). The proportion is projected on the averaged

representation of left (LH) and right hemisphere (RH), which was derived from the average of all curvature information after CBA across subjects. Circles define approximate ROI locations, based on the Talairach coordinates reported in Table 1.

V1 showed that vergence response was distributed unevenly across the cortical surface. In most subjects, an anterior and a posterior region with strong modulation by eye vergence could be distinguished (Figure 8B). Either sub-region's response could

be due to an efference copy signal or due to a visual correlate of eye vergence.

An efference copy signal of eye vergence should not depend on the visual direction. For vertical disparity however, we know that

Table 1 | Mean Talairach coordinates of the ROIs in this study (\pm SD across subjects, $n = 8$).

ROI	Left			Right		
	x	y	z	x	y	z
V1	-10 ± 3	-77 ± 2	0 ± 4	8 ± 2	-74 ± 4	0 ± 5
V3ab	-17 ± 3	-87 ± 5	20 ± 8	18 ± 5	-86 ± 3	19 ± 4
MT ⁺ /c	-42 ± 2	-76 ± 4	7 ± 7	40 ± 5	-71 ± 4	7 ± 4
MT ⁺ /b	-44 ± 3	-69 ± 5	3 ± 5	44 ± 3	-64 ± 5	4 ± 5
V6	-13 ± 4	-76 ± 4	26 ± 5	14 ± 2	-75 ± 5	26 ± 5
VIP	-26 ± 3	-59 ± 8	46 ± 4	24 ± 5	-59 ± 6	45 ± 6

Table 2 | Gain and phase lag of eye vergence averaged across subjects ($n = 5$).

	Gain (deg)	Phase (deg)	Drift (deg)	Gain S8
V2R2H0	1.01 (0.02)	4.62 (1.71)	-0.30 (0.07)	0.96
V2R4H2	1.01 (0.02)	3.50 (0.69)	0.21 (0.36)	0.96
V2R0H2	1.01 (0.02)	5.17 (1.22)	-0.21 (0.07)	1.03
V4R0H4	1.00 (0.02)	5.56 (0.51)	-0.06 (0.05)	0.94
V4R2H6	1.00 (0.02)	5.48 (0.82)	0.07 (0.20)	0.99
V4R6H2	0.94 (0.05)	6.64 (1.96)	0.07 (0.18)	0.90
V0R0H0			0.06 (0.17)	
V0R4H4			0.02 (0.08)	

Gain and phase lag are shown for each of the 6 vergence conditions, and average horizontal drift is shown for all 8 stimulus conditions. Gain is expressed as the ratio of the vergence amplitude to the fixation target vergence amplitude. Positive phase difference indicates a lag of the eye vergence position over time compared to the position of the fixation point over time. The horizontal drift expresses the total drift in degrees over the course of a trial. The final column shows the gain values for subject 8. SEM is given in parentheses.

it depends on horizontal vergence and cyclopean visual direction (Read et al., 2009). Thus, we looked in retinotopic area V1 for a visual direction dependent magnitude of the vergence response (as characterized by the GLM's Vergence Beta value: β_V). (see Materials and Methods).

To test this prediction, we selected for each subject, all significant voxels in area V1, both for the retinotopic localization fits (polar angle and eccentricity mapping) and β_V ($p < 0.05$ on each test). We limited the search to an eccentricity of 45° ; resulting in a grand total across subjects of 3928 voxels that passed all criteria. Subsequently, we computed for each voxel the magnitude of the local vertical disparity level. This was done based on its RF location within the visual field, and the geometric relations between vertical- disparity and horizontal vergence (see Materials and Methods).

We established a gain factor $g = 0.83$ and a constant $C = 0.27$ from the regression (Figure 8C, left graph), indicating a contribution of vertical-size disparity to the vergence response in area V1 and possibly a non-visual contribution.

We applied the same procedure to the voxels in area V1, but now with significant values of β_R and retinotopic localization fits (Figure 8C, right graph). This analysis involved about 5%

more voxels (4164), but yielded no significant relation with the predicted function for vertical disparity ($g = -0.05$, $C = 0.36$).

Finally, to assess potential effects of the non-ideal control of horizontal retinal disparity at large elevations, we repeated the above analysis for each subject selecting only those voxels with a preferred stimulus location within $\pm 25^\circ$ elevation. This selection removed for each subject less than 3% of the voxels. This had marginal effects (in the 3rd decimal) on the slopes and offsets reported above.

MODEL QUALITY ON CAPTURE OF PARAMETRIC MODULATIONS

We explain the differential responses to the different stimulus conditions by three binocular parametric predictors ($\beta_1 * V_i + \beta_2 * R_i + \beta_3 * H_i$) on top of the flow response ($\beta_4 * Fb$), which is common to all conditions (4P-model). How successful is this model compared to a model where each condition is fitted separately by a BOLD response template (8P-model)? To find out we subtracted from each model fit the common flow component from the 4P-model:

$$R_{4P,i} = \eta_1 * V_i + \beta_2 * R_i + \beta_3 * H_i (i \in [1, 8]) \quad (5)$$

$$R_{8P,i} = \beta_i - \beta_4 * Fb (i, \in [1, 8]) \quad (6)$$

In each ROI, we regressed the model predictions R_{4P} and R_{8P} against each other, for all subjects and conditions, as illustrated in Figure 7A for area V6. The β 's of the 4-parameter model captured the vast majority of the BOLD modulation over stimulus conditions (Pearson's correlation: lowest value: 0.74 in MT⁺/c, Figure 7B). This indicates that in our target ROIs a linear response to the amplitude of vergence, retinocentric- and headcentric disparity modulation explains 75–90% of the hemodynamic response on top of the BOLD response to flow.

MEASURING VERGENCE EYE MOVEMENTS

For the 4P-GLM-model, we based the scaling of the parametric predictors on the stimulus amplitude of the binocular components. The scaling assumes perfect fixation on the target ring that moved in depth. Hence, it is key to the interpretation of the BOLD signals that subjects make accurate vergence eye movements during all eye vergence conditions. Unfortunately, the wide-field projection system precluded measurement of eye movements in the scanner. In a separate session in a dummy MRI-setup, we assessed the accuracy of eye vergence movements for each stimulus condition using an Eyelink II system (Eyelink® II, version 1.11, SR research, Canada). Gain and phase lag for all 6 vergence stimulus conditions were assessed in five of the subjects that had participated in scanning. We also assessed linear horizontal drift during each trial.

Subjects' vergence eye movements showed minimal phase lag, nearly perfect vergence pursuit for all stimulus conditions (Table 2, Figure 9 for a single subject). Horizontal drift during each trial was minimal and likely reflects drifts of the head relative to the camera assembly rather than real eye movements (Table 2).

We tested the influence of inaccurate vergence eye movements on the GLM results, by changing the GLM predictor values from the original model (gain_{1,0} model, i.e., 4P-model) based on the gain values obtained from eye tracking (gain_{eye} model). We did

so, for the subject with the lowest gain values (S8, gain values in **Table 2**). Changes in R and H modulation showed minimal changes on the flat maps (**Figure 10**). We also assessed the influence of inaccurate gain by changing the GLM predictors based on a hypothetical gain of 0.5 (gain_{0.5} model). Here, R and H activity changes were clearly visible on the flap maps (**Figure 10**). For all subjects, we examined if the measured-gain model outperformed the perfect gain-model, by comparing goodness of fit (R^2). Across subjects for each ROI, R^2 -values did not differ significantly [paired two-sided t -test, $t_{(4)} = 0.50; 1.17; 1.00; 1.05; 1.04$, for MT, MST, V6, V3ab, and V1 respectively].

Overall, the similar results for the gain_{eye} model show that the observed vergence eye movement inaccuracies do not affect our conclusions.

INTER-SUBJECT VARIABILITY

The search light approach identifies locations with peak H, V or R sensitivity in different subjects. Although the ROI analysis confirms the separate sensitivity to H, V, and R tuning at the group level (4P-model, **Figure 6**) in different ROIs, it is of limited use for inferring spatial congruence between the subjects regarding loci with H or predominant R sensitivity across dorsal and ventral cortex. To examine spatial congruence, we projected the functional maps of each subject onto an averaged inflated representation of left and right hemisphere, which was computed using a Cortical-Based Alignment (CBA) analysis. CBA analysis projects single subject hemisphere curvature information on a common spherical representation and subsequently minimizes curvature differences across subjects, resulting in an alignment based on cortical curvature information (Goebel et al., 2006).

As a result, each subject's functional maps could be transformed to a singular target sphere, which consisted of an average of all 8 hemispheres. For V, R, and H separately, we assessed spatial congruence by computing the proportion of subjects that showed significance for the contrast for each vertex ($p < 0.05$, uncorrected, **Figure 11**). The highest proportion of V activation was found in V1, along the calcarine sulcus. A high proportion of R activation was found along both dorsal and ventral stream, covering all our ROIs. This seems in contrast with the results of the multi-subject analyses, where R did not reach significance in MT⁺/c, V3ab, and V6, pointing to weak but consistent R sensitivity in these areas. Among our ROIs, the highest proportion of H activation is restricted to the more foveal part of V6, while V3ab seems fully covered. The results also uncover an additional region of consistent H activation that falls outside our pre-defined ROI, which likely corresponds to putative area V7. Regarding R activation, a large consistent region of activation is seen lateral and posterior to V3ab. This region most likely coincides with the Lateral Occipital (LO) region, based on anatomical location (Larsson and Heeger, 2006), and previously reported sensitivity to changing retinal disparity (Rokers et al., 2009).

In sum, the CBA results further bolster the observation that H activation is confined to Posterior Parietal Cortex and Posterior Occipital Cortex. Lower R thresholds reveal consistent (but probably weaker) R activation along both lateral and medial motion areas.

DISCUSSION

Monocular optic flow provides important information for the direction of heading (Lappe et al., 1999). Many believe *binocular* information is superfluous. Yet, heading judgment can operate on purely stereoscopic cues (Macuga et al., 2006). Judgment of the speed of self-displacement in depth may profit even more from binocular signals than perception of self-motion direction; the inherently temporal information from optic flow (Koenderink, 1986) does not provide speed. It needs extension with a spatial scale that may be mediated through binocular information. Our observation that perceived speed of sinusoidal back-and-forth self-motion is enhanced by concurrent sinusoidal headcentric disparity, suggests that binocular signals may contribute to such a headcentric distance scale to the flow. Because retinal disparity refers to depth relative to the fixation point, modulation by R indicates an influence of speed by which the scene approaches the fixation point. In contrast, modulation by H indicates a true influence of the approach speed between the head and the scene. It remains to be seen, however, whether more veridical distance judgments arise from the combination of optic flow and headcentric disparity. Our observation extends previous psychophysical reports that speed judgments of approaching objects include changing eye vergence information and changing relative retinal disparity (Howard, 2007; Lutgheid et al., 2011). These studies did not distinguish between contributions of headcentric disparity and eye vergence as we do here.

Our observation bolsters an earlier observation (Arnoldussen et al., 2013b) that perceived self-rotation depends on the rotational flow relative to the head and not relative to the gaze line. That study and the present indicate that the visual flow relative to the *head* is paramount to the self-motion percept. The visual system, extended with eye movement signals, appears to gather information about the rotation and translation of the *head* relative to the scene. This shift from visual retinal flow to headcentric flow was proposed before (Beintema and Van den Berg, 1998; Arnoldussen et al., 2013b) as a step to align *visual information on self-motion* with the information from the *vestibular* system, which reports the head's motion relative to the world.

It is known that visual and vestibular information combine optimally for perceived heading direction (Gu et al., 2008), but so far comparable data for visuo-vestibular judgment of speed (translational or rotational) of self-motion are lacking as far as we know.

Only areas V3ab and V6 were modulated strongly by amplitude changes in H. In contrast, the BOLD modulation in MT⁺/b showed strong modulations to R, but none to H (**Figures 4, 5**). This distinction between V3ab/V6 and MST⁺/b with regard to the type of binocular sensitivity provides strong support that analysis of motion signals in the dorso-medial and dorso-lateral pathways (Galletti and Fattori, 2003; Kravitz et al., 2011; Pitzalis et al., 2012b) serve different functions.

DORSO-MEDIAL RESPONSES

Human V6 shows strong BOLD responses to full field optic flow (Pitzalis et al., 2010), specifically patterns that contain depth by motion parallax (Pitzalis et al., 2013). Also, V6 integrates depth from motion parallax with depth by binocular disparity (Cardin

and Smith, 2011), possibly to rescue direction of heading perception from contamination by motion noise, and to distinguish the depth components of multiple moving objects (Arnoldussen et al., 2013a). These studies did not distinguish between retinal and headcentric disparity.

Headcentric disparity (H), more than retinal disparity is suited for the integration with motion parallax, because it shares with parallax a dependency on target distance to the head, whereas retinal disparity depends on distance to the fixation point. Thus, the strong modulation of optic flow responses by H suggests an important contribution of V3A/V6 to process the *speed* of the head's movement in depth. This lines up nicely with a previous study concluding that V6 combines retinal rotational motion and eye *pursuit* signals to indicate the speed of rotation of the head in space rather than rotation speed of the gaze line (Arnoldussen et al., 2011, 2013b).

Other findings suggest that the flow-sensitivity of human V6 may not point to a role in perception of the *direction* of self-motion; fMRI adaptation in V6 does not rise for repeated presentation of the same focus of expansion of optic flow (Cardin et al., 2012). Also, one might expect sensitivity to the vestibular modality to perceive the direction of heading correctly but this appears to be lacking in area V6 (Smith et al., 2012). Because these studies presented monocular flow one cannot exclude processing of the direction of heading for binocular stimuli for now.

Further functional considerations follow from the comparison with Macaque V6. The majority of V6 neurons are strongly motion-sensitive (Galletti et al., 1996, 1999), but its response to optic flow remains unstudied. Macaque V6 lies within the parieto-occipital stream and has strong connections to parietal area V6A (Galletti et al., 2001), an area strongly related to reaching (Fattori et al., 2001, 2005). Many neurons in V6 are “real-motion cells,” meaning they distinguish “real” retinal motion from eye-movement evoked retinal motion (Galletti and Fattori, 2003). These findings led to the idea that V6 plays a role in visually guided actions (Pitzalis et al., 2012a). Our findings support this notion, because the processing of the speed of head rotation (Arnoldussen et al., 2011) and head translations (current study) from optic flow would serve proper control of motor actions, especially in peri-personal space, such as reaching, grasping, and repelling objects that approach the head.

The virtually lacking modulation by headcentric disparity in the MT⁺/b complex (Figure 4) suggests that the speed of the head's movement in depth is not processed there. What might be the reason for the strong modulation by retinocentric disparity? This region shares certain properties with activity evoked in the medial temporal cortex, area MST, of the monkey: sensitivity to self-motion specific patterns of optic flow in contra and ipsi-lateral parts of the visual field. Such patterns provide direction of self-motion and depth information but the informative component must be separated from the uninformative rotational component of the flow, which does not carry depth information (Warren and Hannon, 1988). To support that separation, primate self-motion perception relies on non-visual information, like efference copy signals of eye movements (for reviews: Lappe et al., 1999; Britten, 2008) and vestibular signals of head movements in MST (Angelaki et al., 2009, 2011). It also uses

independent signals on depth order (Van den Berg and Brenner, 1994b). Depth order can be derived from many sources including both types of binocular disparities. Perhaps retinal disparities are preferred to establish depth order as they are less computationally expensive.

VERGENCE EYE MOVEMENT MODULATION IN V1?

The ocular convergence angle required for fixating an object, signals a target's distance relative to the head (Von Hofsten, 1976; Foley, 1980; Viguier et al., 2001). Neuro-physiological work points to the integration of the eye vergence signal with retinal disparity for reach distances in various regions within PPC, for example the parietal reach region (PRR; Bhattacharyya et al., 2009). Early neurophysiological work reported modulation of disparity-selective neurons by eye position signals as early as V1 (Trotter et al., 1992, 1996). Also V1, V2, and V4 neurons' responses to a contrast target are up regulated by increasing eye vergence angle (Rosenbluth and Allman, 2002). However, both studies have been criticized on methodological grounds (Cumming and DeAngelis, 2001).

Here, we show strong modulation of the BOLD signal by V and R in area V1 (Figures 6, 8A). Geometrically, changes in retinal *vertical* disparity correlate with eye vergence, and depend only on visual direction. Therefore, *changes* in V and *changes* in vertical disparity take place concomitantly. Interestingly, that visual direction dependency was reflected in the magnitude of the BOLD response to vergence within area V1 (Figure 8C, left graph), but not at all in the response to the amplitude of the horizontal retinal disparity in V1 (Figure 8C, right graph). Vertical disparity alternates polarity between visual field quadrants, but polarity is unlikely to be represented in the BOLD signal. Hence, the nature of the BOLD signal prevents us to establish the presence of that polarity change and we cannot distinguish a response to vertical disparity from the mere vertical-size mismatch of the image. Nevertheless, our data strongly suggest that the modulation by V in cortical area V1 at least reflects sensitivity to vertical-size disparity. To our knowledge, this is the first human evidence for vertical-size disparity sensitivity in V1, already demonstrated in Macaque striate cortex (Cumming, 2002). It appears then that V1 may represent eye vergence from the global pattern of vertical size mismatch between the eyes' images.

CAVEATS—VECTION

Unequal attentional load between stimulus conditions can alter the BOLD modulations (Huk and Heeger, 2000), also in MT⁺/b and MT⁺/c (Treue and Maunsell, 1999). To prevent differences in attentional load due to vection, we instructed subjects to keep attention on the fixation point in depth and to make vergence eye movements as accurately as possible. Also, peak forward speed of the simulated motion was slow (300 mm/s), and *post-hoc* reports by the subjects indicated no strong percepts of vection (not quantified). Importantly, the optic flow was similar for all stimulus conditions. Because the optic flow remains the primary and strongest visual cue to self-motion, irrespective of additional binocular cues, vection differences between stimulus conditions were minimized. Overall, the effects described are unlikely due to attention or vection.

CAVEATS—INDEPENDENT CONTROL OF RETINAL AND HEADCENTRIC DISPARITY

Our study relies on independent control of retinal headcentric disparity and eye vergence. As we mentioned before, geometric properties of the disparity field preclude a full dissociation of retinal horizontal disparity and headcentric disparity across the VM surface. We showed by simulations that changes in retinal horizontal disparity associated with the changes of the position of the virtual screen (VM surface) and the fixation point in depth were very close to the intended values and even better so when the field of view was limited to the central range ($\pm 25^\circ$) of elevation. We found that only for a small minority (<13%) of the voxels in our ROIs the receptive field (RF) was located at elevations beyond 25° . Accordingly, we found little change in the results (Figures 6A,B) when we limited the analysis of the 4P-model to the voxels with RFs within the central $\pm 25^\circ$ of elevation. Also, we note that when we entertain the hypothesis that voxels responsive to headcentric disparity would actually respond to spurious horizontal retinal disparities we should be puzzled why such voxels are not found in area V1? Hence, we conclude that our choice of the VM surface for the virtual screen was appropriate to distinguish headcentric and retinal disparity tuning in the ROIs.

CAVEATS—VERGENCE EYE MOVEMENTS

Can inaccurate vergence eye movements explain our results? The wide-field projection system did not allow for online registration of eye movements. Vergence eye movement accuracy for the different stimulus conditions was measured in 5 of the subjects in dummy scanner setup. Subjects made accurate vergence eye movements for all stimulus conditions (Figure 9, Table 2), with gain and phase values comparable to previously published data (Erkelens and Collewijn, 1985a).

An adaptation of the 4P-model that reflected adapted V and R predictor values given non-perfect gain (gain_{eye} model) showed minimal differences in the single subject results compared to the original model (Figure 10). Across subjects, the gain_{eye} model did not improve on the original model. We conclude that inaccurate vergence eye movements cannot explain our fMRI results.

CAVEATS—RELATIVE RETINAL DISPARITY

We note that our projection method on the virtual screen removed relative retinal disparity associated with the positions within the cloud of individual dots. Therefore, the changing relative retinal disparity normally associated with self-motion through a 3D cloud was not presented. We cannot rule out the possibility that inclusion of relative disparity information would contribute to the percept motion in depth, as has been found experimentally (Erkelens and Collewijn, 1985b; Brenner et al., 1996; Welchman et al., 2009). In light of this work, our results show at least that by minimizing relative disparity information, the contribution of other binocular cues to perceived motion in depth is revealed. Two other reasons make us doubt a strong contribution of relative disparities to the percept of motion depth. First, we note that headcentric disparity, absolute retinal disparity, and relative retinal disparity provide decreasing headcentric information in that order. Secondly, relative retinal disparity was present between the fixation point and the dots, in all conditions

in which the binocular fixation point was not placed on the virtual projection screen (i.e., $R > 0$). We did not find evidence that these conditions provided especially large BOLD signals. Nevertheless, to clarify, a potential contribution of relative disparity signals to the perceived speed of self-motion needs further investigation within the context of our type of stimuli.

CAVEATS—CYCLOVERGENCE AND TILTED HOROPTER

Our stimulus presentation did not take into account cyclovergence of the eyes. Cyclovergence that is coupled to horizontal vergence occurs when the primary position is not located in the plane of regard. We did an analysis of the effect of adding elevation dependent cyclovergence (Mok et al., 1992; Van Rijn and Van den Berg, 1993) to the model, because we did not know the primary position of our subjects' Listing planes. We varied the elevation of the assumed primary position by 10° , which then results in cyclovergence coupled to the imposed horizontal vergence, and associated vertical disparities (using full equation 1 from Read et al., 2009). We found only marginal changes of fit of the linear regression by the inclusion of that cyclovergence component (not shown). Hence, our data suggest that the tilt of the eyes in opposite directions has not been an important contribution to the BetaV fits of Figure 8C.

The horopter shows a backward tilt in the vertical dimension (von Helmholtz, 1962). We did not take into account this tilt, because our hypothesis involves a change of the vertical size disparity over time. This would not be affected by an offset like the tilt of the horopter.

FUNDING

This work was supported by Netherlands Organization for Scientific Research NWO-ALW (grant number 818.02.006) to A. V. van Den Berg.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fnsys.2015.00016/abstract>

REFERENCES

- Angelaki, D. E., Gu, Y., and DeAngelis, G. C. (2011). Visual and vestibular cue integration for heading perception in extrastriate visual cortex. *J. Physiol.* 589, 825–833. doi: 10.1113/jphysiol.2010.194720
- Angelaki, D. E., Klier, E. M., and Snyder, L. H. (2009). A vestibular sensation: probabilistic approaches to spatial perception. *Neuron* 64, 448–461. doi: 10.1016/j.neuron.2009.11.010
- Arnoldussen, D. M., Goossens, J., and Van den Berg, A. (2013a). Differential responses in dorsal visual cortex to motion and disparity depth cues. *Front. Hum. Neurosci.* 7:815. doi: 10.3389/fnhum.2013.00815
- Arnoldussen, D. M., Goossens, J., and van den Berg, A. (2013b). Visual perception of axes of head rotation. *Front. Behav. Neurosci.* 7:11. doi: 10.3389/fnbeh.2013.00011
- Arnoldussen, D. M., Goossens, J., and van den Berg, A. V. (2011). Adjacent visual representations of self-motion in different reference frames. *Proc. Natl. Acad. Sci. U.S.A.* 108, 11668. doi: 10.1073/pnas.1102984108
- Ban, H., Preston, T. J., Meeson, A., and Welchman, A. E. (2012). The integration of motion and disparity cues to depth in dorsal visual cortex. *Nat. Neurosci.* 15, 636–643. doi: 10.1038/nn.3046
- Beintema, J. A., and Van den Berg, A. (1998). Heading detection using motion templates and eye velocity gain fields. *Vision Res.* 38, 2155–2179. doi: 10.1016/S0042-6989(97)00428-8

- Bhattacharyya, R., Musallam, S., and Andersen, R. A. (2009). Parietal reach region encodes reach depth using retinal disparity and vergence angle signals. *J. Neurophysiol.* 102, 805–816. doi: 10.1152/jn.90359.2008
- Brenner, E., Van den Berg, A., and Van Damme, W. (1996). Perceived motion in depth. *Vision Res.* 36, 699–706. doi: 10.1016/0042-6989(95)00146-8
- Britten, K. H. (2008). Mechanisms of self-motion perception. *Annu. Rev. Neurosci.* 31, 389–410. doi: 10.1146/annurev.neuro.29.051605.112953
- Brouwer, G. J., van Ee, R., and Schwarzbach, J. (2005). Activation in visual cortex correlates with the awareness of stereoscopic depth. *J. Neurosci.* 25, 10403–10413. doi: 10.1523/JNEUROSCI.2408-05.2005
- Campos, J. L., Butler, J. S., and Bulthoff, H. H. (2012). Multisensory integration in the estimation of walked distances. *Exp. Brain Res.* 218, 551–565. doi: 10.1007/s00221-012-3048-1
- Cardin, V., Hemsforth, L., and Smith, A. T. (2012). Adaptation to heading direction dissociates the roles of human MST and V6 in the processing of optic flow. *J. Neurophysiol.* 108, 794–801. doi: 10.1152/jn.00002.2012
- Cardin, V., and Smith, A. T. (2011). Sensitivity of human visual cortical area V6 to stereoscopic depth gradients associated with self-motion. *J. Neurophysiol.* 106, 1240–1249. doi: 10.1152/jn.01120.2010
- Cumming, B. (2002). An unexpected specialization for horizontal disparity in primate primary visual cortex. *Nature* 418, 633–636. doi: 10.1038/nature00909
- Cumming, B., and DeAngelis, G. (2001). The physiology of stereopsis. *Annu. Rev. Neurosci.* 24, 203–238. doi: 10.1146/annurev.neuro.24.1.203
- Dukelow, S. P., DeSouza, J. F. X., Culham, J. C., van den Berg, A. V., Menon, R. S., and Vilis, T. (2001). Distinguishing subregions of the human MT⁺ complex using visual fields and pursuit eye movements. *J. Neurophysiol.* 86, 1991.
- Erkelens, C., and Collewijn, H. (1985a). Eye movements and stereopsis during dichoptic viewing of moving random-dot stereograms. *Vision Res.* 25, 1689–1700. doi: 10.1016/0042-6989(85)90141-5
- Erkelens, C., and Collewijn, H. (1985b). Motion perception during dichoptic viewing of moving random-dot stereograms. *Vision Res.* 25, 583–588. doi: 10.1016/0042-6989(85)90164-6
- Erkelens, C. J., and van Ee, R. (1998). A computational model of depth perception based on headcentric disparity. *Vision Res.* 38, 2999–3018. doi: 10.1016/S0042-6989(98)00084-4
- Fattori, P., Gamberini, M., Kutz, D. F., and Galletti, C. (2001). Arm-reaching neurons in the parietal area V6A of the macaque monkey. *Eur. J. Neurosci.* 13, 2309–2313. doi: 10.1046/j.0953-816x.2001.01618.x
- Fattori, P., Kutz, D. F., Breveglieri, R., Marzocchi, N., and Galletti, C. (2005). Spatial tuning of reaching activity in the medial parieto-occipital cortex (area V6A) of macaque monkey. *Eur. J. Neurosci.* 22, 956–972. doi: 10.1111/j.1460-9568.2005.04288.x
- Foley, J. M. (1980). Binocular distance perception. *Psychol. Rev.* 87:411. doi: 10.1037/0033-295X.87.5.411
- Galletti, C., and Fattori, P. (2003). Neuronal mechanisms for detection of motion in the field of view. *Neuropsychologia* 41, 1717–1727. doi: 10.1016/S0028-3932(03)00174-X
- Galletti, C., Fattori, P., Battaglini, P., Shipp, S., and Zeki, S. (1996). Functional demarcation of a border between areas V6 and V6A in the superior parietal gyrus of the macaque monkey. *Eur. J. Neurosci.* 8, 30–52. doi: 10.1111/j.1460-9568.1996.tb01165.x
- Galletti, C., Fattori, P., Gamberini, M., and Kutz, D. F. (1999). The cortical visual area V6: brain location and visual topography. *Eur. J. Neurosci.* 11, 3922–3936. doi: 10.1046/j.1460-9568.1999.00817.x
- Galletti, C., Gamberini, M., Kutz, D. F., Fattori, P., Luppino, G., and Matelli, M. (2001). The cortical connections of area V6: an occipito-parietal network processing visual information. *Eur. J. Neurosci.* 13, 1572–1588. doi: 10.1046/j.0953-816x.2001.01538.x
- Goebel, R., Esposito, E., and Formisano, E. (2006). Analysis of functional image analysis contest (FIAC) data with Brainvoyager QX: from single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Hum. Brain Mapp.* 27, 392–401. doi: 10.1002/hbm.20249
- Grigo, A., and Lappe, M. (1998). Interaction of stereo vision and optic flow processing revealed by an illusory stimulus. *Vision Res.* 38, 281–290. doi: 10.1016/S0042-6989(97)00123-5
- Gu, Y., Angelaki, D. E., and DeAngelis, G. C. (2008). Neural correlates of multi-sensory cue integration in macaque MSTd. *Nat. Neurosci.* 11, 1201–1210. doi: 10.1038/nn.2191
- Howard, I. P. (2007). Vergence modulation as a cue to movement in depth. *Spat. Vis.* 21, 581–592. doi: 10.1163/156856808786451417
- Howard, I. P., and Rogers, B. J. (2002). *Seeing in Depth, Vol. 2: Depth Perception*. Ontario: I Porteous.
- Huk, A. C., Dougherty, R. F., and Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *J. Neurosci.* 22, 7195–7205.
- Huk, A. C., and Heeger, D. J. (2000). Task-related modulation of visual cortex. *J. Neurophysiol.* 83, 3525–3536.
- Koenderink, J., and Doorn, A. J. (1987). Facts on optic flow. *Biol. Cybern.* 56, 247–254. doi: 10.1007/BF00365219
- Koenderink, J. J. (1986). Optic flow. *Vision Res.* 26, 161–179. doi: 10.1016/0042-6989(86)90078-7
- Kolster, H., Peeters, R., and Orban, G. A. (2010). The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbors. *J. Neurosci.* 30, 9801–9820. doi: 10.1523/JNEUROSCI.2069-10.2010
- Kovacs, G., Raabe, M., and Greenlee, M. W. (2008). Neural correlates of visually induced self-motion illusion in depth. *Cereb. Cortex* 18, 1779–1787. doi: 10.1093/cercor/bhm203
- Kravitz, D. J., Saleem, K. S., Baker, C. I., and Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nat. Rev. Neurosci.* 12, 217–230. doi: 10.1038/nrn3008
- Lappe, M., Bremmer, F., and Van den Berg, A. (1999). Perception of self-motion from visual flow. *Trends Cogn. Sci.* 3, 329–336. doi: 10.1016/S1364-6613(99)01364-9
- Larsson, J., and Heeger, D. J. (2006). Two retinotopic visual areas in human lateral occipital cortex. *J. Neurosci.* 26, 13128–13142. doi: 10.1523/JNEUROSCI.1657-06.2006
- Likova, L. T., and Tyler, C. W. (2007). Stereomotion processing in the human occipital cortex. *Neuroimage* 38, 293–305. doi: 10.1016/j.neuroimage.2007.06.039
- Lutigheid, A. J., Brenner, E., and Welchman, A. E. (2011). Extra-retinal signals affect the perceived speed of 3D motion. *J. Vis.* 11, 323–323. doi: 10.1167/11.11.323
- Macguga, K. L., Loomis, J. M., Beall, A. C., and Kelly, J. W. (2006). Perception of heading without retinal optic flow. *Percept. Psychophys.* 68, 872–878. doi: 10.3758/BF03193708
- Mok, D., Ro, A., Cadera, W., Crawford, J., and Vilis, T. (1992). Rotation of Listing's plane during vergence. *Vision Res.* 32, 2055–2064. doi: 10.1016/0042-6989(92)90067-S
- Palmisano, S. A. (2002). Consistent stereoscopic information increases the perceived speed of vection in depth. *Perception* 31, 463–480. doi: 10.1068/p3321
- Pitzalis, S., Bozzacchi, C., Bultrini, A., Fattori, P., Galletti, C., and Di Russo, F. (2012b). Parallel motion signals to the medial and lateral motion areas V6 and MT⁺. *Neuroimage* 67, 89–100. doi: 10.1016/j.neuroimage.2012.11.022
- Pitzalis, S., Fattori, P., and Galletti, C. (2012a). The functional role of the medial motion area V6. *Front. Behav. Neurosci.* 6:91. doi: 10.3389/fnbeh.2012.00091
- Pitzalis, S., Galletti, C., Huang, R. S., Patria, F., Committeri, G., Galati, G., et al. (2006). Wide-field retinotopy defines human cortical visual area v6. *J. Neurosci.* 26, 7962–7973. doi: 10.1523/JNEUROSCI.0178-06.2006
- Pitzalis, S., Sdoia, S., Bultrini, A., Committeri, G., Di Russo, F., Fattori, P., et al. (2013). Selectivity to translational egomotion in human brain motion areas. *PLoS ONE* 8:e60241. doi: 10.1371/journal.pone.0060241
- Pitzalis, S., Sereno, M. I., Committeri, G., Fattori, P., Galati, G., Patria, F., et al. (2010). Human V6: the medial motion area. *Cereb. Cortex* 20, 411–424. doi: 10.1093/cercor/bhp112
- Read, J. C., Phillipson, G. P., and Glennerster, A. (2009). Latitude and longitude vertical disparities. *J. Vis.* 9, 11–37. doi: 10.1167/9.13.11
- Regan, D., and Gray, R. (2009). Binocular processing of motion: some unresolved questions. *Spat. Vis.* 22, 1–43. doi: 10.1163/156856809786618501
- Rokers, B., Cormack, L. K., and Huk, A. C. (2009). Disparity- and velocity-based signals for three-dimensional motion perception in human MT⁺. *Nat. Neurosci.* 12, 1050–1055. doi: 10.1038/nn.2343
- Rosenbluth, D., and Allman, J. M. (2002). The effect of gaze angle and fixation distance on the responses of neurons in V1, V2, and V4. *Neuron* 33, 143–149. doi: 10.1016/S0896-6273(01)00559-1
- Sereno, M., Dale, A., Reppas, J., Kwong, K., Belliveau, J., Brady, T., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268, 889. doi: 10.1126/science.7754376

- Seymour, K. J., and Clifford, C. W. G. (2012). Decoding conjunctions of direction-of-motion and binocular disparity from human visual cortex. *J. Neurophysiol.* 107, 2335–2341. doi: 10.1152/jn.01103.2011
- Smith, A. T., and Wall, M. B. (2008). Sensitivity of human visual cortical areas to the stereoscopic depth of a moving stimulus. *J. Vis.* 8, 1.1–1.12. doi: 10.1167/8.10.1
- Smith, A. T., Wall, M. B., and Thilo, K. V. (2012). Vestibular inputs to human motion-sensitive visual cortex. *Cereb. Cortex* 22, 1068–1077. doi: 10.1093/cercor/bhr179
- Treue, S., and Maunsell, J. H. (1999). Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas. *J. Neurosci.* 19, 7591–7602.
- Trotter, Y., Celebrini, S., Stricanne, B., Thorpe, S., and Imbert, M. (1992). Modulation of neural stereoscopic processing in primate area V1 by the viewing distance. *Science* 257, 1279–1281. doi: 10.1126/science.1519066
- Trotter, Y., Celebrini, S., Stricanne, B., Thorpe, S., and Imbert, M. (1996). Neural processing of stereopsis as a function of viewing distance in primate visual cortical area V1. *J. Neurophysiol.* 76, 2872–2885.
- Tsao, D. Y., Vanduffel, W., Sasaki, Y., Fize, D., Knutsen, T. A., Mandeville, J. B., et al. (2003). Stereopsis activates V3A and caudal intraparietal areas in macaques and humans. *Neuron* 39, 555–568. doi: 10.1016/S0896-6273(03)00459-8
- Van den Berg, A. (1996). Judgements of heading. *Vision Res.* 36, 2337–2350. doi: 10.1016/0042-6989(95)00247-2
- Van den Berg, A., and Brenner, E. (1994a). Why two eyes are better than one for judgements of heading. *Nature* 371, 700–702. doi: 10.1038/371700a0
- Van den Berg, A., and Brenner, E. (1994b). Humans combine the optic flow with static depth cues for robust perception of heading. *Vision Res.* 34, 2153–2167. doi: 10.1016/0042-6989(94)90324-7
- Van Rijn, L., and Van den Berg, A. (1993). Binocular eye orientation during fixations: listing's law extended to include eye vergence. *Vision Res.* 33, 691–708. doi: 10.1016/0042-6989(93)90189-4
- Viguiet, A., Clement, G., and Trotter, Y. (2001). Distance perception within near visual space. *Perception* 30, 115–124. doi: 10.1068/p3119
- von Helmholtz, H. (1962). *Helmholtz's Treatise on Physiological Optics*. New York, NY: Dover Publications.
- Von Hofsten, C. (1976). The role of convergence in visual space perception. *Vision Res.* 16, 193–198. doi: 10.1016/0042-6989(76)90098-5
- Warren, W. H. (2004). "Optic flow," in *The Visual Neurosciences*, eds L. M. Chalupa and J. S. Werner (London: The MIT press), 1257–1259.
- Warren, W. H., and Hannon, D. J. (1988). Direction of self-motion is perceived from optical flow. *Nature* 336, 162–163. doi: 10.1038/336162a0
- Welchman, A. E., Harris, J. M., and Brenner, E. (2009). Extra-retinal signals support the estimation of 3D motion. *Vision Res.* 49, 782–789. doi: 10.1016/j.visres.2009.02.014
- Wood, G., Nuerk, H.-C., Sturm, D., and Willmes, K. (2008). Using parametric regressors to disentangle properties of multi-feature processes. *Behav. Brain Funct.* 4, 38. doi: 10.1186/1744-9081-4-38
- Zhang, Z. L., Cantor, C. R., and Schor, C. M. (2010). Perisaccadic stereo depth with zero retinal disparity. *Curr. Biol.* 20, 1176–1181. doi: 10.1016/j.cub.2010.04.060

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 26 August 2014; accepted: 02 February 2015; published online: 24 February 2015.

Citation: Arnoldussen DM, Goossens J and van Den Berg AV (2015) Dissociation of retinal and headcentric disparity signals in dorsal human cortex. *Front. Syst. Neurosci.* 9:16. doi: 10.3389/fnsys.2015.00016

This article was submitted to the journal *Frontiers in Systems Neuroscience*.

Copyright © 2015 Arnoldussen, Goossens and van Den Berg. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.