The primary role of flow processing in the identification of scene-relative object movement

DOI:
10.1523/JNEUROSCI.3530-16.2017

Document Version
Accepted author manuscript

Link to publication record in Manchester Research Explorer

Citation for published version (APA):

Published in:
The Journal of Neuroscience

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
The primary role of flow processing in the identification of scene-relative object movement

Simon K. Rushton¹, Diederick C. Niehorster², Paul A. Warren³, & Li Li (李黎)²,⁴

¹ School of Psychology, Cardiff University, Tower Building, Park Place, Cardiff, Wales, UK
² Department of Psychology, The University of Hong Kong, Hong Kong SAR
³ Division of Neuroscience & Experimental Psychology, School of Biological Sciences, University of Manchester, Brunswick Street, Manchester, UK
⁴ Neural Science Program, NYU-ECNU Institute of Brain and Cognitive Science, New York University Shanghai, Shanghai, PRC

Corresponding authors:
Simon K. Rushton
School of Psychology
Cardiff University
Tower Building
Park Place
Cardiff
Wales
CF10 3AT
UK
rushtonsk@cardiff.ac.uk

Li Li
NYU Shanghai
1555 Century Avenue,
Pudong New Area
Shanghai
200122
PRC
ll114@nyu.edu

Abbreviated title: Optic flow and moving objects

Pages: 25
Figures: 4; tables: 0; multimedia and 3D models: 0
Abstract: 249 words; Introduction: 551 words; Discussion: 979 words
Note: Figures embedded (with captions) for review ms to enhance readability.

Conflict of Interest: The authors declare no competing financial interests.

Acknowledgements
This study was supported by grants from the Research Grants Council of Hong Kong (HKU 746013H), Shanghai Science and Technology Committee (17ZR1420100), and NYU-ECNU Joint Research Institute to LL, a PhD Fellowship from the Research Grants Council of Hong Kong (PF09-03850) to DCN, a research grant from UK Economic and Social Research Council (ES/M00001X) to SKR and a Wellcome Trust project grant (WT089934) to PAW and SKR. We thank Long Ni for his assistance in data collection. LL and DCN designed, ran and analysed Experiment 1. SKR designed, ran and analysed Experiment 2. The paper was written by SKR, LL, DCN and PAW.
The primary role of flow processing in the identification of scene-relative object movement

Abstract [249/250]

Retinal image motion could be due to movement of the observer through space or an object relative to the scene. Optic flow, form, and change of position cues all provide information that could be used to separate out retinal motion due to object movement from retinal motion due to observer movement. In Experiment 1 we used a minimal display to examine the contribution of optic flow and form cues. Human participants indicated the direction of movement of a probe object presented against a background of radially moving pairs of dots. By independently controlling the orientation of each dot pair, we were able to put flow cues to self-movement direction (the point from which all the motion radiated) and form cues to self-movement direction (the point all the dot pairs were oriented towards) in conflict. We found that only flow cues influenced perceived probe movement. In Experiment 2 we switched to a rich stereo display comprised of 3D objects to examine the contribution of flow and position cues. We moved the scene objects to simulate a lateral translation and counter-rotation of gaze. By changing the polarity of the scene objects (from light to dark and vice-versa) between frames, we placed flow cues to self-movement direction in opposition to change of position cues. We found that again flow cues dominated the perceived probe movement relative to the scene. Taken together, these experiments indicate the neural network that processes optic flow has a primary role in the identification of scene-relative object movement.
Significance Statement [118/120]

Motion of an object in the retinal image indicates relative movement between the observer and the object, but it does not indicate its cause: movement of an object in the scene; movement of the observer; or both. To isolate retinal motion due to movement of a scene object, the brain must parse out the retinal motion due to movement of the eye (“flow parsing”). Optic flow, form and position cues all have potential roles in this process. We pitted the cues against each other and assessed their influence. We found flow parsing relies on optic flow alone. These results indicate the primary role of the neural network that processes optic flow in the identification of scene-relative object movement.
Flow processing for the identification of object movement

Introduction [551/650]

Introduction

Motion of an object in the retinal image indicates relative movement between the observer and the object, but it does not indicate its cause: movement of the object relative to the scene; movement of the observer; or a combination of the two. One way to resolve this ambiguity is to use information about self-movement: if the brain knows how the observer has moved it should be able to “factor out” motion due to self-movement thereby isolating motion due to object movement. Early work using dark environments demonstrated a limited role for extra-retinal information (copies of motor commands, vestibular cues, or felt position) about self-movement in this process (e.g. Gogel and Tietz, 1974; Wallach, Stanton and Becker, 1974). In lit environments, the global patterns of retinal motion that result from self-movement, known as “optic flow”, provide a very valuable source of information about self-movement. The human brain has a well-documented sensitivity to optic flow (Warren and Hannon, 1988). We suggested (the flow-parsing hypothesis, Rushton and Warren, 2005) that the brain identifies components of optic flow and then parses, or filters, them from the retinal flow field, isolating components of motion that are due to the movement of objects in the scene. In support of this hypothesis is evidence that shows that humans can make judgements of scene-relative object movement on the basis of retinal cues alone (e.g. Warren and Rushton 2007, 2008), and evidence of a subtractive process based on optic flow processing (Warren and Rushton, 2009a).

Although primate neurophysiological and human functional imaging work provides some suggestions of neural substrates involved in the flow-parsing process, e.g., MSTl (Eifuku and Wurtz, 1998; Tanaka, Sugita, Moriya and Saito, 1993), V6 (Pitzalis, Fattori and Galletti, 2013), MT (Kim, Angelaki and DeAngelis, 2015), V3a and V3B/KO (Bartels, Zeki and Logothetis, 2008; Calabro and Vaina, 2012) and V7/V3a (Bartels, Zeki and Logothetis, 2008), the exact underlying neural mechanisms remain unclear. An important step in understanding which sites and pathways are involved, and in understanding the process at a computational level, is to determine which cues are involved in flow parsing. In this paper, we explore the roles of optic flow, form, and change of position cues in the detection of scene-relative object movement.
In Experiment 1, we pitted optic flow and form cues to self-movement against one another, using an animated Glass pattern stimulus (Glass 1969; see Figure 1). Form cues and optic flow specified different directions of (forward) self-movement, and thus predicted different perceived scene-relative movement directions of a probe object placed into the display. We found that although form cues demonstrated a robust influence on the perceived direction of self-movement, only optic flow cues contributed to the perception of scene-relative movement of the probe object.

In Experiment 2, we pitted optic flow against change of position cues using the reverse-phi effect (Anstis, 1970; see Figure 2). By changing the luminance polarity (between light and dark) of scene objects between frames in a 10Hz animation, we created a stimulus in optic flow, and the change of object position cues, specified opposite direction of self-movement, and thus predicted different perceived scene-relative movement of a probe object. We found that optic flow dominated the perceived direction of self-movement and, in line, with the first experiment, optic flow cues also dominated in the perception of scene-relative object movement.

Materials and Methods

Experimental Design and Statistical Analyses

The experiments were within-subject designs. Data were analysed using t-tests. Exact p values are reported except when p <0.001. We report Cohen’s d as a measure of effect size.

Experiment 1: Optic flow vs. form cues

Glass patterns (Glass, 1969) are formed by randomly-positioned pairs of dots that are individually orientated to create a percept of global form. When the patterns are flashed, or independently generated on each frame, a percept of global motion parallel to the global structure arises (Ross,
Badcock, & Hayes, 2000; Burr & Ross, 2002). When the dot pairs have a common orientation, translational motion parallel to the orientation of the dot pairs results. When they are oriented in a radial pattern, a percept of movement towards or away from the middle or “focus” of the pattern results. Figure 1 shows a static image of the stimuli we used that contained randomly positioned pairs of dots oriented in a radial pattern towards a focus point (indicated by the blue cross).

A radial optic flow field is then generated by moving dot pairs outwards in a radial pattern. A percept of forward self-movement results. The focus point from which the dots radiate (indicated by the red cross) defines the direction of self-movement (i.e., heading, Warren & Hannon, 1988). Our animated stimuli thus combined both the form cues of the Glass pattern, and the motion cues of the flow field (See the movie at https://osf.io/pq5nh/). This allowed us to specify independent orientation-defined and motion-defined foci, and hence dissociate motion (optic flow) and form cues to self-movement direction. We have previously shown that when observers make judgements of the direction of self-movement while viewing such an animated display, the perceived direction is between the orientation- and motion-defined foci (Niehorster, Cheng, and Li, 2010).

Following earlier work (Warren and Rushton, 2009a), we used a probe dot placed within the display to measure characteristics of the flow parsing process (white arrow, Figure 1). After observing an animation, participants indicated the direction in which they had seen the probe move. We evaluated the respective role of form and optic flow cues in the identification of scene-relative object movement (flow parsing) by testing two display conditions: the animated Glass pattern display condition in which the motion-defined focus was in the middle of the screen and the form-defined focus was located 10° to the left or right, and the flow-only display condition in which dots were unpaired to remove the form cues. The difference in the perceived probe movement between these two display conditions indicated the contribution of form cues. We used the same manipulation to examine the role of the two cues in the perception of self-movement.

Participants

Eight students and staff (seven naïve as to the specific goals of the study and one author; four males, four females) between the age of 22 and 33 at the University of Hong Kong participated in the experiment. All had normal or corrected to normal vision and provided informed consent. The study
was approved by the Human Research Ethics Committee for Non-Clinical Faculties at The University of Hong Kong.

**Visual stimuli and apparatus**

Two display types were generated and crossed with two judgment tasks.

In the animated Glass pattern display, (Figure 1) similar to the displays of Niehorster et al. (2010), the display simulated the observer translating at 0.6 m/s toward a frontal plane consisting of 500 randomly placed white dot pairs with 1° centroid-to-centroid average separation (dots: 0.4° in diameter, 75% luminance contrast) positioned 1.5 m away. The centroid of each dot pair moved outwards in a radial pattern. The focus of the motion-defined radial pattern (i.e., the motion-defined self-movement direction) was in the middle of the screen. The dot pairs were re-oriented on each frame to maintain a radial form pattern with a form-defined focus 10° to either the left or the right of the motion-defined focus (see Figure 1). As dot pairs moved off the edge of the screen, they were randomly repositioned on the frontal plane so that the dot distribution on the plane remained uniform.

In the flow only display, the display was the same, but the dots were unpaired by hiding one of the two dots that made up the pair. To equate the density with the Glass pattern condition, we consequently doubled the number of virtual pairs of dots.

For the object-movement judgment task, a red probe dot (1° diameter) moved upwards at 2.5 °/s. The midpoint of the probe’s movement trajectory was 3° to the left or the right of the motion-defined focus. A fixation point (0.8° diameter) was positioned at the motion-defined focus (located in the middle of the screen). Observers were instructed to look at the fixation point. A nulling lateral movement component, controlled by a Bayesian adaptive staircase procedure (Kontsevich and Tyler, 1999), was added to the movement of the probe. The staircase was designed to find the speed of the added lateral movement component at which participants were equally likely to report the probe moving obliquely leftwards as obliquely rightwards (i.e., the induced lateral movement component due to flow parsing was nulled by the added lateral movement component and the participants perceived the probe moving vertically). At the end of each trial, a blank screen appeared, and observers were asked to use the left or right mouse buttons to indicate whether they perceived that the
probe moved obliquely leftward or obliquely rightward. Response data was subsequently pooled and
fit with a cumulative Gaussian to obtain an estimate of the nulling speed at which the probe was
perceived as moving vertically.

For the self-movement judgement task, the probe dot and the fixation point were removed. At the
end of each trial, a white horizontal line appeared at the middle of a blank screen, and participants
were asked to use a mouse to move a red vertical line, which appeared in a random position within
20° from the middle of the screen, along the horizontal line to indicate their perceived direction of
self-movement, their heading. The angle between the perceived self-movement direction and the
motion-defined focus, defined as the shift of the perceived self-movement direction, was recorded.

For both judgement tasks, at the start of each trial, a blank screen containing only the fixation
point appeared for 600ms. Next the first frame of the stimulus was shown for 500ms, and then a
500ms animation followed. For the object-movement task, the probe was visible and moving
throughout the animation period.

The displays were programmed in MATLAB using the Psychophysics Toolbox 3 (Brainard,
1997; Pelli, 1997) and were rendered using a Dell Studio XPS Desktop 435T/9000 with an NVIDIA
GeForce GTX 560Ti graphics card running Windows 7. The displays (83°H x 83°V) were rear-
projected on a large screen at 60 Hz with an Epson EMP-9300 LCD projector (native resolution: 1400
x 1050 pixels). Participants viewed the displays binocularly with their head stabilized by a chin rest
at the viewing distance of 56.5 cm.

**Procedure**

After each trial, participants responded with a button press. All participants completed four blocks of
80 trials (40 trials for each adaptive staircase x 2 offset directions between motion- and form-defined
foci), one for each display (Glass pattern or pure flow) and task (object-movement or self-movement
judgement) condition. In each block, the two adaptive staircases (one for each offset direction) were
interleaved in each block and the testing order of the blocks was counterbalanced. Participants
received 3-5 training trials at the beginning of each block. No feedback was provided on any trial.
The experiment took approximately 35 minutes to complete.
Experiment 2: Optic flow vs. change of position cues

If a light or dark dot on a mid-luminance background is displayed in one position and then shortly after shifts to a nearby position, the dot is perceived to move from first position to the second (Figure 2, left panel). This is apparent or regular “phi” motion (Wertheimer, 1912). If the dot changes from light to dark, or vice-versa, between the first and second positions, then the dot is perceived to move in the opposite direction. This is “reverse-phi” motion (Anstis, 1970). In this experiment, we extended the reverse-phi technique to dissociate motion and change of position cues in complex 3D displays (Figure 2, middle panel), and examined the contribution of the two cues to the perception of self-movement and scene-relative object movement.

Participants

Eight students and staff (six naïve as to the specific goals of the study and the first and third author; 6 male, 2 female) between the age of 22 and 46 at Cardiff University participated in the experiment. All had normal, or corrected to normal, vision. The study was run in accordance with the requirements of the School of Psychology Ethics Committee at Cardiff University.

Visual stimuli and apparatus

As in Experiment 1, two display types were generated and crossed with two task conditions. In the reverse-phi display (Figure 2), the background scene consisted of 24 red wireframe objects of approximately 4 cm diameter, each randomly oriented and squashed, arranged in a volume of 26x22x70 cm, with the midpoint of the volume 87 cm from the observer. In the middle of the screen was a small (1 cm radius) target sphere. The background scene and target sphere were rendered stereoscopically. The position of the cameras (and hence the observer’s view of the scene) was updated at 10Hz. The screen background was a mid-red, and the scene objects (but not the target sphere) alternated between light and dark red each time the camera position was updated.
In the phi display, all was identical to the reverse-phi display except that the polarity of the scene objects did not change. In the reverse-phi display the change in polarity produced an obvious flicker. We therefore added comparable flicker to the phi display by alternating, at 10Hz, the luminance of the scene objects between two levels of light red (both lighter than the mid-red background).

In both display conditions, the background objects that made up the 3D scene were moved and transformed to simulate leftwards or rightwards lateral observer translation together with a counter-rotation of gaze to keep the middle of the volume straight ahead (Figure 2, middle and right panels). The translation was sinusoidal (amplitude 3cm, frequency, 0.35Hz, initial phase -0.75 radians). The target sphere was not moved or transformed. It remained at the same location relative to the observer.

In the phi motion display, the motion and change of position cues indicated the same direction of self-movement (Figure 2). In the reverse-phi motion display, the motion and change of position cues indicated opposite directions of self-movement.

For the self-movement judgment task, the target sphere was located at 87cm from the participant which coincided with the middle of the volume. The scene effectively rotated around the target sphere, and thus the target sphere appeared stationary relative to the observer and the scene. Participants used a mouse button to indicate their perceived direction of self-movement relative to the scene (or equivalently the direction of movement of the scene relative to them).

For the object-movement judgement task, the target sphere was located at a fixed location 64 or 110cm in front of the participant (23cm in front or behind of the middle of the volume). When the scene moved and the target sphere remained stationary, the target sphere would appear to translate relative to the scene. The geometric relationship between scene movement and object position that generates a percept of scene-relative object movement is the same relationship exploited by Gogel in his experiments on the perception of distance (Gogel and Tietz, 1974), and by previous work on the perception of object movement (Rushton and Warren, 2005; Rushton, Bradshaw and Warren, 2007; Warren and Rushton, 2009b). The participant judged the perceived direction of scene-relative object movement. On each trial, participants pressed a mouse button to indicate their perceived direction of movement (left or right) of the target relative to the scene.
Given that viewing large uncrossed disparities on a nearby monitor typically causes double-vision (diplopia), the scene dimensions and the speed of simulated self-movement (~0.7m distance and ~0.2m/s lateral translation speed) were equivalently scaled down from standard values. This allowed us to use angular retinal speeds that are comparable to movement in larger scenes at more typical speeds (~1m/s), while minimising the range of crossed and uncrossed disparities, and hence aiding stereoscopic fusion.

The displays were rendered at 120Hz using the OpenGL graphics library, programmed in Lazarus, a public domain Pascal compiler, and run under Windows 7 on a desktop (i3) computer with a Quadro 600 graphics card. The displays were presented on a 22” (40x30cm) Viewsonic p225f CRT monitor fronted with a red gel filter, and viewed through CrystalEyes stereo shutter glasses at a distance of 57cm with the participant’s eyes aligned with the middle of the screen and their head stabilised using a chin rest. The scene objects were rendered in red because red phosphor has the shortest decay time and hence minimises cross-talk with the stereo glasses. The CRT monitor had a refresh rate of 120Hz and a spatial resolution of 1280x960 pixels. Left and right eye images were shown on alternate frames and the opening of the left and right filters in the stereo glasses was synchronised to the display so that the stimuli were seen with stereoscopic depth.

Each trial began with a blank screen presented for 500ms. The background scene and target sphere then became visible. The first frame of the stimuli appeared statically for 500ms, and was followed by a 1500ms animation in which the viewpoint changed position at 10 Hz to simulate relative movement between the observer and the scene movement. The scene and target were then replaced with a blank screen.

Procedure
Each participant completed four blocks: one for each display type and task condition. Each block contained 80 trials (10 trials x 2 simulated scene movement directions x 2 target distances x 2 display types) presented in a random order. To make sure observers understood the task, they completed a short block of approximately 10 training trials with the phi motion display at the beginning of the experiment. No feedback was provided on any trial. The experiment took a little over 10 minutes to
Results

Experiment 1

There was no significant difference between the mirror-symmetric left/right versions of the stimuli for both the Glass pattern and pure flow conditions (object movement, $t(7) = 0.98, p = 0.36$, Cohen’s $d = 0.35$ and $t(7) = 0.71, p = 0.50$, Cohen’s $d = 0.25$, respectively; self-movement, $t(7) = 1.51, p = 0.17$, Cohen’s $d = 0.53$ and $t(7) = 1.24, p = 0.25$, Cohen’s $d = 0.44$, respectively). Consequently, the data were collapsed.

Self-movement judgments

In the flow only display condition, the mean shift of the perceived self-movement direction from the motion-defined focus was approximately zero and the variability between participants was low (mean±SE: $-0.13\pm0.14^\circ$, difference from zero shift, $t(7) = -0.98$, $p = 0.36$, Cohen’s $d = 0.35$). This indicates that participants perceived the motion-defined focus in the radial flow pattern as their self-movement direction. In the animated Glass pattern display condition, with the form-defined focus displaced by $10^\circ$ from the motion-defined focus, the mean shift of the perceived self-movement direction from the motion-defined focus was significantly larger than zero ($5.05\pm1.00^\circ$, $t(7) = 5.05$, $p = 0.001$, Cohen’s $d = 1.8$). The difference in the shift of the perceived self-movement direction from the motion-defined focus between the two display conditions was statistically significant ($t(7) = 5.72$, $p < 0.001$, Cohen’s $d = 2.0$). The $5^\circ$ shift of the perceived self-movement direction from the motion-defined focus in the animated Glass pattern display condition was consistent with an approximately equal weighing of motion and form cues (see also Niehorster et al., 2010).

Object movement judgments

For both the animated Glass pattern and the flow-only display conditions, the mean speed of the added lateral movement to null the induced lateral movement of the probe due to flow parsing was significantly larger than zero ($2.45\pm0.21$ cm/s and $2.48\pm0.19$ cm/s, both $t(7) > 11.5$, $p < 0.001$, Cohen’s $d > 4.0$). The direction of the added lateral movement was the same in both display conditions and was consistent with the fact that the vertically moving probe was perceived to be
moving on a diagonal toward the direction of the motion-defined focus located in the middle of the screen. There was no statistically significant difference in the speed of the added lateral movement between the animated Glass pattern and the flow-only display conditions \( t(7) = -0.32, p = 0.76 \), Cohen’s \( d = 0.11 \), supporting the conclusion that in contrast to self-movement judgements, motion cues in the flow field play the primary role in object-movement judgments.

Comparing self-movement and object movement judgements

To allow direct comparison of data from the self-movement and object movement judgement task conditions, we computed the difference scores in the judgement performance between the animated Glass pattern and the flow-only display conditions and converted the difference scores to z-scores by dividing the difference scores by their standard deviation. Figure 3 plots the z-scores for the self-movement and the object-movement judgment tasks.

The same participants performed both the self-movement and object-movement judgement tasks and the displays used for the object movement judgment task were identical to those used for the self-movement judgement task except for the presence of a moving probe. If the identification of scene-relative object movement occurs at a later stage in the processing hierarchy, relying on the output of the perception of self-movement system, then we might expect perceived self-movement and the perceived object movement results to show a similar pattern.

There was a significant difference between the self-movement and the object-movement judgment tasks \( t(7) = 4.06, p=0.005 \), Cohen’s \( d = 1.4 \). The mean of the normalized difference scores was significantly larger than zero for self-movement judgements \( (2.0\pm0.35, t(7) = 5.72, p < 0.001, \) Cohen’s \( d = 2.0 \) ), but not different from zero for object-movement judgements \( (0.11\pm0.35, t(7) = 0.32, p = 0.76, \) Cohen’s \( d = 0.11 \) ). This result indicates that discrepant form cues affected self-movement but not object-movement judgments and suggests that the identification of object
movement does not rely on a prior estimate of self-movement direction (see also Warren, Rushton and Foulkes, 2013).

The contribution of local motion cues

It has been shown that both local and global flow motion processing contribute to the identification of scene-relative object movement (Warren and Rushton, 2009a). For the object movement judgment task, the probe dot is surrounded by scene dots so it is possible that the perceived object movement is not due to the processing of global flow but rather local motion contrast. We ran a control experiment in which we removed dots within a 10° aperture around the probe to rule out this possibility. We ran 8 participants and found again that there was a significant difference between the between the two conditions ($t(7)=16.4, p<0.001$, Cohen’s $d = 5.8$) and the mean of the normalized difference scores was significantly larger than zero for self-movement judgments ($2.61\pm0.35, t(7) = 7.39, p < 0.001$, Cohen’s $d = 2.6$), but not different from zero for object-movement judgments ($-0.24\pm0.35, t(7) = -0.68, p = 0.52$, Cohen’s $d = 0.24$). The results and conclusions were thus not changed by the removal of local dot motion.

Experiment 2

Figure 4 shows the percentage of trials in which the judgements was consistent with the flow-defined (left axis) and the position-defined (right axis) direction of movement. Results are shown for the self-movement and object-movement judgement tasks and for both display conditions.

[ Approximate Figure 4 location ]

For the self-movement judgement task, the percentage of trials in which the judgements were consistent with the optic flow in both the phi (99%) and the reverse-phi (97%) motion displays was high and significantly different from the chance level of 50% ($t(7) = 105.7, p < 0.001$, Cohen’s $d = 37$ and $t(7) =51.6, p < 0.001$, Cohen’s $d = 18$, respectively). This result demonstrates that the
stroboscopic nature of the display did not present a problem, and the range of disparities used was appropriate for our participants who understood the task and performed to a high standard.

For the object movement judgement task, in the phi motion display, the percentage of trials in which the judgements were consistent with the optic flow was again high (95%) and significantly different from chance ($t(7) = 27.25, p < 0.001$, Cohen’s $d = 9.6$). This demonstrates the standard flow-parsing effect (e.g. Rushton & Warren, 2005). In the reverse-phi motion display, in line with other work using reverse-phi, e.g. Bours, Kroes & Lankheet (2009), there was a slight drop performance (in 87% of trials, the judgements were consistent with the optic flow) but the percentage was still high and well above chance ($t(7) = 12.7, p < 0.001$, Cohen’s $d = 4.5$).

Overall, these data suggest that optic flow is the primary cue in both self-movement and scene-relative object movement judgements.

Discussion [979/1500]

With two very different types of visual displays, a minimal random dot display, and a stereo 3D display containing discrete objects, we found the same result: Optic flow has the sole influence on the perceived scene-relative object movement during self-movement.

Why do other relevant retinal cues (such as form and position cues) not contribute to the identification of scene-relative object movement? The most accurate and precise perceptual estimates are derived from the weighted combination of information from all available cues (Landy, Maloney, Johnston and Young, 1995). Consequently, it seems counter-intuitive that only a sole retinal cue, optic flow, would be used. However, if we assume that it takes longer to process, weight and combine cues then the demands of evolutionary history (catching a prey or avoiding a predator) might have favoured a faster but less precise solution based on a single source of retinal information.

A process reliant solely on optic flow for the identification of scene-relative object motion could be very fast. The entire process could be carried out within hMT+/V5 (Pelah, Barbur, Thurrell and Hock, 2015): MSTd is sensitive to flow components so activity in MSTd could modulate the activity of neurons in hMSTI or hMT, sites that respond to local motion (see Layton and Fajen, 2016, for a
worked example of a flow-parsing model based on hMT and hMST). Another possibility is that all
the computations are conducted within hMT (see Royden, Sannicandro and Webber, 2015, for a
worked example of a flow-parsing model based on MT-like speed and direction tuned neurons), or all
the computations are conducted in hMST (see Krekelberg, Paolini, Bremmer, Lappe, Hoffmann, 2001
for a proposal that MST neurons could identify scene-relative object movement).

A question that arises is how the relevant information might reach hMT+. We have shown that
flow parsing does not make use of form cues provided by Glass patterns. Single-cell (Krekelberg,
Dannenberg, Hoffmann, Bremmer and Ross, 2003) and human imaging work (Krekelberg, Vatakis
and Kourtzi, 2005) suggests that visual areas along the dorsal pathway to hMT+ are sensitive to Glass
patterns. Not all neurons along the dorsal motion pathway are sensitive to Glass patterns but it seems
unlikely that there would be two parallel sub-pathways within the dorsal stream, one sensitive to
Glass patterns and one not. Another possibility is that a parallel pathway to hMT+ is involved.

A number of strands of evidence (e.g., Guy and Zeki, 1995; Azzopardi and Hock, 2011)
suggest a direct pathway to hMT+ that bypasses V1. Using EEG and MEG, ffytche, Guy and Zeki
(1995) found that for fast motion, MT+ neurons showed an earlier response than V1 neurons
supporting the existence of a direct link. Based on a Conditional Grainger Causality analysis of fMRI
data, Gaglianese, Costagli, Bernardi, Ricciardi and Pietrini (2015) concluded that fast motion
information is routed from LGN to MST, bypassing V1, slower motion is routed from LGN to MT.
Recent behavioural work (using reverse-phi) suggests that “objectless” (Azzopardi and Hock, 2011)
motion information is carried along the direct pathway. Further, behavioural evidence specifically
suggests that the non-occipital pathway can transmit optic flow information (Pelah, Barbur, Thurrell
and Hock, 2015). The reported direct pathway to hMT+ has the required characteristics to fit the data
we report here, i.e., the change in object position and form cues available in our stimuli would not be
carried by this pathway but optic flow (pure-motion) would.

The involvement of the non-occipital pathway to hMT+ would also potentially reduce processing
latencies. Optic flow could reach hMT+ and be processed in time to be combined with retinal image
object motion that had progressed up the occipital pathway. If optic flow is not fed along the non-
occipital pathway, slower “re-entrant processing” (information passing up the hierarchical processing
pathway before feeding back down; Lamme, Super and Spekreijse, 1998) would be required, as would
a storage-delay system to solve the problem of optic flow information being available later than
information about the movement of the object in the retinal image.

We noted that early work demonstrated a limited role for extra-retinal information in the
identification of scene-relative object movement (e.g., Gogel and Tietz, 1974; Wallach, Stanton and
Becker, 1974). Recent work suggests that when both retinal and extra-retinal information is available,
the two sources are used in conjunction (Wexler, 2003; Tcheang, Gilson and Glennerster, 2005;
Work on judgements of self-movement has identified MSTd as a site at which vestibular and retinal
signals to self-movement are combined (Gu, Angelaki and DeAngelis, 2008). With a combined
estimate of self-movement in MSTd, flow-parsing could either occur in MSTd (Krekelberg et al,
2001) or the output of MSTd could feed or modulate processing in other areas.

Reflecting on the two tasks we examined here (the identification of scene-relative object
movement and the perception of self-movement direction), we raise the interesting question of what
purposes optic flow processing serves. The idea of optic flow was first introduced in the context of
judging and controlling self-movement (Gibson, 1950, 1958; but see also Grindley, 1942, as discussed
by Mollon, 1997; and Calvert, 1950). However, a host of studies (Llewellyn, 1971; Vishton and
Cutting, 1995; Beall and Loomis, 1996; Beusmans, 1998; Rushton, Harris, Lloyd and Wann, 1998; Li
& Cheng, 2013; Li & Niehorster, 2014) have now challenged the hypothesis that optic flow is the sole
cue used in the perception of self-movement direction or visual guidance of walking and steering. In
contrast, the evidence for a central role of optic flow processing in the identification of scene-relative
object movement during self-movement has grown progressively stronger. This provides a new focus
and motivation for ongoing neurophysiological, psychophysical and modelling work on optic flow
processing.

In summary, the results of the current study indicate that optic flow is the primary, and potentially
sole retinal cue used for the identification of scene-relative object movement.
References


481-494.


moving observers. Exp Brain Res 221, 177-189.

Calvert ES (1950) Visual aids for landing in bad visibility, with particular reference to the transition
from instrumental to visual flight. Trans Illium Engng Soc Lond 15, 183-219.


Dupin L, Wexler, M (2013) Motion perception by a moving observer in a three-dimensional


Fajen BR, Matthis JS (2013) Visual and non-visual contributions to the perception of object motion
during self-motion. PLoS ONE 8: e5546

ffytche DH, Guy CN, Zeki, S (1995) The parallel visual motion inputs into areas V1 and V5 of human

not V1-mediated, functional influence between the thalamus and middle temporal complex in the human brain is modulated by the speed of visual motion. Neuroscience 284: 833-844.


Flow processing for the identification of object movement

Rushton et al

Neurophys 94: 4373-4386.


Flow processing for the identification of object movement  


Figure Legends

Figure 1: Schematic illustration of the animated Glass pattern display. Dot pairs are oriented to form a radial pattern with the form-defined focus 10° to the left or right of the middle of the screen indicated by the blue cross (not part of stimulus). Dot pairs move (indicated by red arrows) in a radial pattern with the motion-defined focus in the middle of the screen indicated by the red cross (not part of stimulus). Probe dot (yellow, 3° to the right of the motion-defined focus) moves upwards. Perceived trajectory (yellow arrow) is the vector sum of the actual movement (white arrow) and the induced lateral component (green arrow) of movement in the probe due to flow parsing.

Figure 2: Left panel illustrates the classic regular phi and reverse phi motion effects. In a two-frame display, a dot is shown first in an initial position and then in a position to the right. If the timing and displacement are within the appropriate range, then the dot is seen to move rightwards (regular phi motion). If the dot changes polarity between frame one and frame two, then the dot is seen to move leftward (reverse phi motion). Middle panel shows phi-motion principle applied to the displacement of a viewpoint relative to a scene. Between frames scene objects are moved and transformed in a way that is compatible with a lateral translation and counter-rotation of the head to keep the centre of the volume straight-ahead. Right panel shows the perceived self-movement with regular and reversed phi motion displays.

Figure 3: Effect of discrepant form cues on self-movement and object movement judgements. Participants made judgements on the animated Glass pattern displays with the motion- and form-defined foci separated laterally by 10° and on the flow-only displays with the form cues removed. Graph shows the difference in judgements between the animated Glass pattern and the flow only displays, expressed as normalized difference scores. Data shown for 8 participants with mean indicated by large disks.
Figure 4. Percentage of trials in which self-movement (left) and object movement (right) judgements are consistent with the motion cues. Data shown for 8 participants with mean indicated by large disks.