

Binocular Eye Movements Evoked by Self-Induced Motion Parallax

Jared Frey¹ and Dario L. Ringach^{1,2,3}

¹Biomedical Engineering Interdepartmental Program and Departments of ²Psychology and ³Neurobiology, University of California, Los Angeles, Los Angeles, California 90095

Perception often triggers actions, but actions may sometimes be necessary to evoke percepts. This is most evident in the recovery of depth by self-induced motion parallax. Here we show that depth information derived from one's movement through a stationary environment evokes binocular eye movements consistent with the perception of three-dimensional shape. Human subjects stood in front of a display and viewed a simulated random-dot sphere presented monocularly or binocularly. Eye movements were recorded by a head-mounted eye tracker, while head movements were monitored by a motion capture system. The display was continuously updated to simulate the perspective projection of a stationary, transparent random dot sphere viewed from the subject's vantage point. Observers were asked to keep their gaze on a red target dot on the surface of the sphere as they moved relative to the display. The movement of the target dot simulated jumps in depth between the front and back surfaces of the sphere along the line of sight. We found the subjects' eyes converged and diverged concomitantly with changes in the perceived depth of the target. Surprisingly, even under binocular viewing conditions, when binocular disparity signals conflict with depth information from motion parallax, transient vergence responses were observed. These results provide the first demonstration that self-induced motion parallax is sufficient to drive vergence eye movements under both monocular and binocular viewing conditions.

Introduction

In humans, the lines of sight of the eyes converge onto a point of interest in space. The location of this point with respect to the head determines the appropriate angle of vergence of the eyes. Vergence eye movements reduce binocular disparities, allowing for fusion of the retinal images (Leigh and Zee, 2006). However, we know that monocular cues to depth are capable of evoking changes in vergence as well. These include retinal blur (Müller, 1843), changing size (looming) (Erkelens and Regan, 1986; McLin et al., 1988; Wismeijer and Erkelens, 2009), perspective (Enright, 1987; Wagner et al., 2009), shape from shading (Hoffmann and Sebald, 2007), and shape from motion in the kinetic depth effect (KDE) (Ringach et al., 1996). It is fairly well established, however, that when binocular disparity signals are available, they provide the main drive to vergence, with monocular cues to 3D depth rarely being able to compete if their signals are in conflict (Leigh and Zee, 2006).

Human ability to perceive depth from motion parallax has been studied before (Rogers and Graham, 1979; Durgin et al., 1995; Wexler et al., 2001; Medendorp et al., 2003; Braunstein, 2009; Rogers, 2009), but whether such percepts can trigger vergence eye movements has never been assessed in detail. One of

the goals of our study was to close this gap in the literature. Based on our prior study using the kinetic depth effect (Ringach et al., 1996), we hypothesized that motion parallax would be able to evoke vergence eye movements as well. Indeed, we report that during monocular viewing of a motion-parallax display the eyes change vergence in a way that correlates with perceived depth. Remarkably, even under binocular viewing conditions, when disparity information is unambiguous and indicates the stimulus lies flat on a frontoparallel plane, motion parallax was able to evoke vergence responses (albeit with a smaller amplitude and transient in time). This result contrasts with our prior study of shape from motion, where binocular viewing completely abolished vergence responses (Ringach et al., 1996). These findings indicate that information about depth from self-induced motion parallax can be used by the brain to control binocular eye movements, and it is strong enough to do so even when in conflict with binocular disparity signals.

Materials and Methods

A computer-simulated, transparent random-dot sphere was displayed on a video monitor (Panasonic High Definition Plasma Display, TH-50PF10UK, 1920 × 1080 pixels, 100 Hz refresh rate) mounted at eye level, 1 m away from the subject, in a dimly lit room (Fig. 1*a*). The locations of the dots were determined by perspective projection of a simulated, stationary sphere. Subjects stood unrestrained and viewed the display either binocularly or monocularly. As the subject moved, information obtained from a motion capture system was used to continuously render the sphere from the subject's vantage point (the point midway between the eyes) with one video frame delay. The display subtended 67° in the horizontal direction and 37° in the vertical direction. In the first experiment, the simulated sphere was 20° in

Received May 2, 2011; revised Sept. 23, 2011; accepted Sept. 29, 2011.

Author contributions: J.F. and D.L.R. designed research; J.F. and D.L.R. performed research; J.F. analyzed data; J.F. and D.L.R. wrote the paper.

The authors declare no competing financial interests.

Correspondence should be addressed to Dario L. Ringach at the above addresses. E-mail: dario@ucla.edu.

DOI:10.1523/JNEUROSCI.2192-11.2011

Copyright © 2011 the authors 0270-6474/11/3117069-05\$15.00/0

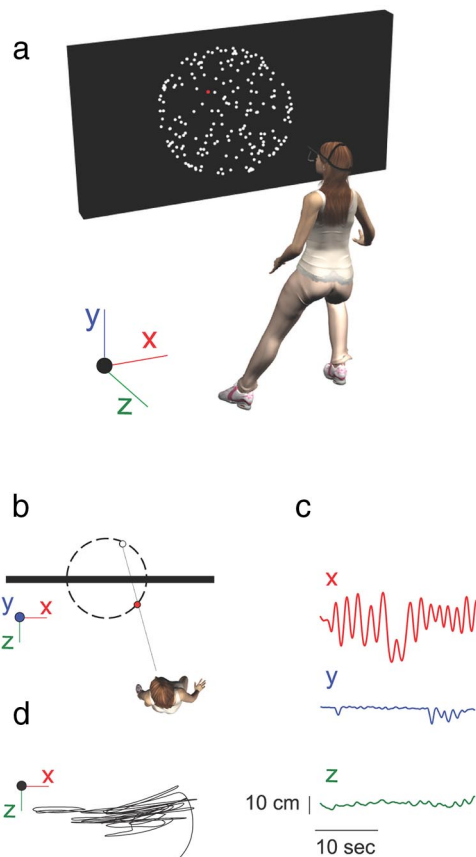


Figure 1. Experimental setup. *a*, Subjects wore a head-mounted eye tracker and stood about 1 m from a computer display depicting a random dot sphere. Head movements were tracked by a motion capture system. The display was updated continuously to render the projection of a static, transparent sphere from the observer's vantage point. *b*, The task consisting of visually tracking a red target dot that intermittently jumped between the front and back of the sphere along the subject's line of sight. *c*, *d*, Sample traces of self-induced subject movement during one trial in the experiment.

diameter (35.2 cm in space) and covered by 800 identical white dots (0.13° radius), uniformly distributed on the surface, which were replaced on each trial. One of the dots was designated as a tracking target and indicated in red. The initial position of the target dot was in the center of the display and located randomly on the front or back surface of the simulated sphere.

Subjects wore a head-mounted eye tracker (Eyelink II, SR Research) that sampled the position of both eyes at 250 Hz. Three infrared LED markers were affixed to the head-mounted eye tracker and four additional markers were affixed to the corners of the screen. The marker locations were tracked by an Optotrak Certus Motion Capture System (NDI), which allowed us to determine head position with submillimeter resolution. These data were used by a dedicated machine running MatLab (MathWorks) with PsychToolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) to update the display by rendering the sphere from the last recorded viewing point.

Each trial consisted of 24 s of voluntary, self-motion while subjects tracked the target dot on the surface of the sphere. The only motion on the display was due to the self-motion of the observer relative to the simulated sphere. A sample record of the head trajectory in one trial is shown in Figure 1, *c* and *d*. At random intervals (uniformly distributed) between 1 and 4 s, the target dot jumped to the opposite side of simulated sphere along the line of sight (Fig. 1*b*). This experimental design prevented displacements of the projected position of the red dot on the screen during a jump. Such a strategy was adopted to prevent saccadic eye movements accompanying the changes in vergence we wanted to measure. The occurrence of the jump was tagged by a TTL pulse that was time stamped by the eye tracker.

We ran experiments under monocular and binocular viewing conditions. In the monocular condition, each subject performed one block of 30 trials, yielding ~ 270 depth jumps per subject. Since we anticipated lower signal-to-noise ratios for the binocular condition, two blocks of 30 trials were run in this condition. In both conditions, trials beginning with the target dot in front surface were randomly interleaved with trials starting with the target dot in back. The eye tracker was calibrated with a nine-point calibration display at the beginning of each block and after the 15th trial. Calibration was always binocular. Fixed to the eye tracker was an occluder that covered one eye during the monocular viewing condition without interfering with the operation of the eye tracker, so the movements of both eyes could be recorded.

In a second experiment, subjects viewed monocularly six different sphere sizes (diameters subtending 27.2° , 22.16° , 18.3° , 13.14° , 7.91° , and 2.64° or, in space, 48.4, 39.2, 32.2, 23.03, 13.83, and 4.6 cm, respectively) to test whether changes in vergence scaled with the physical size of the simulated jumps in depth. The number of dots for each sphere (11,025, 7225, 4900, 2500, 900, and 100, respectively) was chosen to maintain constant dot density on the surface of the sphere. The dot density was higher in this experiment because at the smallest diameter, a minimum number of dots were needed to create a compelling impression of a rigid sphere. The dots were smaller in size (0.066° radius) to reduce visual clutter at higher dot densities. Additionally, the brightness of the white dots decreased with sphere size so as to maintain constant luminous flux across the display. In this experiment subjects ran one block of 30 trials (five trials for each of the six sphere sizes, randomly interleaved).

Subjects. Experiments were approved by the University of California, Los Angeles Institutional Review Board and subjects provided their informed consent for participation. Subjects had normal or corrected to normal vision. Some subjects participated in more than one experiment. In the first experiment, the monocular viewing condition was performed by the two authors (1 and 2, both male) and four subjects with little or no previous experience with oculomotor experiments and naive as to the objectives of the study (females: 4, 5; males: 3, 6). The binocular condition was performed by three subjects that had previously participated in the monocular condition (1, 3, and 5) and three new subjects (males: 7, 9; female: 8) with little or no previous experience with oculomotor experiments and naive as to the objectives of the study. In the second experiment, which examined vergence responses as a function of sphere size, one of the authors (1) participated along with two new naive, inexperienced subjects, 10 and 11 (both female). Naive subjects were given the following instructions: "At the beginning of each trial, you will first see a cross in the center of the screen. When you are ready to begin the trial, fixate on the cross and press the button to begin the trial. Many white dots and one red dot will then appear on the screen. These dots will move on the display as you move side-to-side. Your task is to follow the red dot as you move continuously until the dots disappear." Further clarifications and a sample trial were provided to naive subjects as necessary, but there was no suggestion that the dots represented depth or three-dimensional shape. All of our subjects reported a vivid 3D percept when debriefed after the experiment.

During the analysis of the data, we noticed that the vergence responses of one of the authors, 2, were consistent with perceived depth but approximately four times larger in magnitude than the other subjects. This subject was aware of the goals of the study in advance and his responses, while large, were also substantially slower than the other subjects, suggesting the possibility of top-down control (although he reported not being consciously aware of any intention to modulate his vergence state.) Based on these considerations, we excluded this subject's data from subsequent analyses.

Data analysis. The vergence angle was calculated from head-referenced eye position data and aligned to the occurrence of the target jump in depth for averaging across trials. Data points with a velocity above three SDs were clipped to eliminate saccades. Since intervals between the jumps were as short as 1 s, segments were truncated at the occurrence of the next jump to prevent contamination from adjacent segments.

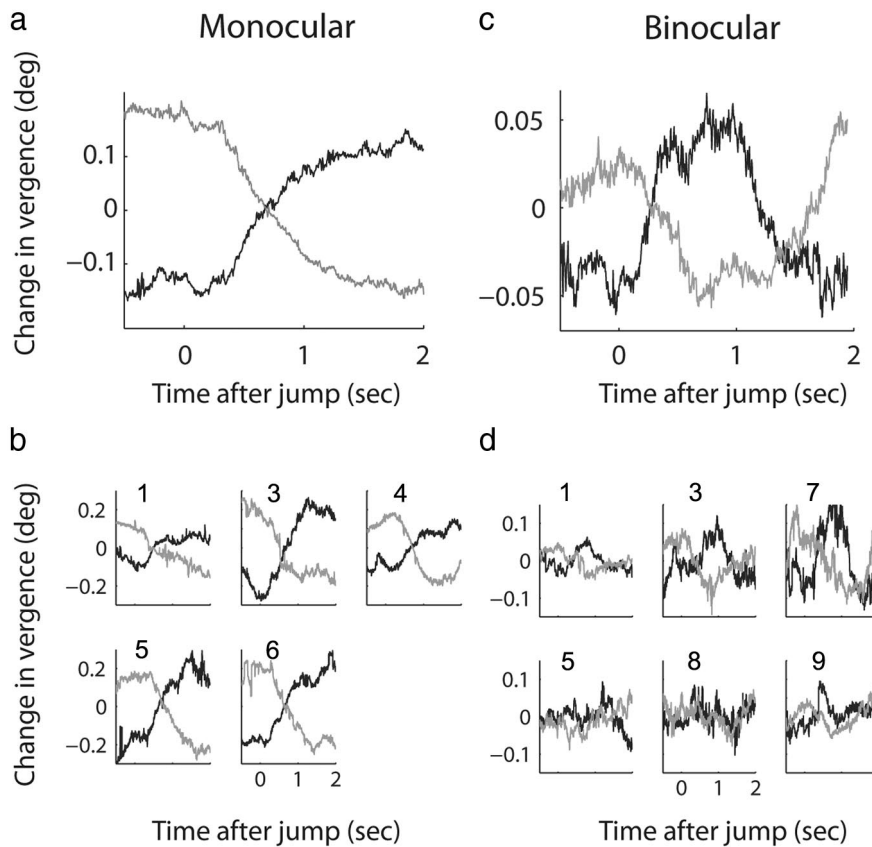


Figure 2. Self-induced motion parallax evokes vergence eye movements. Vergence signals were aligned to the occurrence of target jumps and the mean was subtracted. The segments were then separated by direction of jump (front-to-back and back-to-front) and subsequently averaged. Black traces are changes in vergence when the target jumped from the back of the sphere to the front (when we expect the eyes to converge), while the gray traces show the vergence change when the target jumps from front to back (when we expect the eyes to diverge). Average responses were different in both monocular and binocular conditions (*a, c*). The effect was clear in all individuals during monocular viewing (*b*), but more variable and weaker to nonsignificant in the binocular viewing condition (*d*). Subjects' identification numbers appear at the inset.

The mean vergence angle of each jump segment was subtracted from that segment. Finally, the segments were averaged across simulated jumps having the same direction (front-to-back or back-to-front) to yield a mean change in vergence over time according to sign of depth change.

To estimate the latencies of the responses, we first subtracted the mean back-to-front response from the mean front-to-back response. From this differential signal, we then subtracted its baseline during the 500 ms prior to the onset of the jump and fitted the initial segment of the signal (from 500 ms before to 360 ms after the target jump) response with the following empirical function:

$$\alpha(t) = \begin{cases} 0 & t < d \\ c(t - d) & t \geq d \end{cases}$$

Here α is the vergence angle and t is time. The fitting parameters are the slope, c , and the latency, d . To find the optimal parameters, we used the MatLab functions *nlinfit* and *nlparci*, which perform nonlinear least-squares optimization using the Levenberg–Marquardt algorithm and calculates confidence intervals for the parameters, respectively. A similar approach was adopted earlier by Busetini et al. (2001).

For the second experiment, in which we measured the responses to spheres of different sizes, we compared the geometrically expected versus measured changes in vergence for each jump the target made in depth. For each case, we first computed the expected change in vergence for each simulated change in depth. This is calculated as the vergence change required to shift gaze between the two points on the simulated sphere intersected by the line joining the vantage point (middle of the eyes of

the observer) and the red target. We then computed the measured change in vergence as the difference between a “pre-jump” vergence angle obtained as the mean of the 125 samples (500 ms of data) immediately before the target jump, and a “post-jump” vergence angle obtained as the mean of the 125 samples from 2 s after the jump to 2.5 s. This window was selected from the analysis of the mean responses (Fig. 2), which saturate ~ 2 s after the jump. Finally, we calculated the correlation coefficient between the measured and expected vergence changes and its statistical significance.

Results

In the monocular viewing condition, motion parallax information to depth was sufficient to evoke convergence of the eyes when the target dot jumped from back to front and a divergent movement when the target jumped from front to back (Fig. 2*a*). The magnitude of the response was significant and robust in all subjects tested (Fig. 2*b*). The effect was also significant under binocular viewing of the same stimuli (Fig. 2*c*). Here, however, the responses were clearly smaller in magnitude and transient in time. In addition, we noted a larger individual variability in the binocular condition, with some subjects showing little or no effect (such as subjects 5 and 8 in Fig. 2*d*).

We then took a closer look at the early dynamics of the average responses in the monocular and binocular conditions. The average responses to jumps in opposite depth directions were subtracted to produce a differential response (Fig. 3*a*). The monocular response raises and saturates reaching a steady-state value at ~ 1.5 s after the target jump, while the binocular response peaks at a lower amplitude at ~ 1 s after the target jump and decays back to baseline shortly after.

To ensure that these results were not due to individual differences, the data were reanalyzed using only the three subjects who performed the task under both viewing conditions (1, 3, 5). While the pooled responses for the three subject group were more variable, they were not noticeably different from the average responses of the entire subject group (data not shown).

To look for differences in response delay between the two conditions, we analyzed the initial 360 ms (Fig. 3*b*). The estimated latencies from the fits (d in the first equation) were 226 ± 33 ms and 192 ± 37 ms ($\pm 95\%$ confidence intervals) for the monocular and binocular responses, respectively. The binocular responses were thus slightly faster than monocular ones.

To determine whether the vergence response scaled with changes in the simulated depth jumps of the target, we plotted the measured versus expected changes in three subjects (Fig. 4). We found a modest but statistically significant correlation between the expected and measured change in vergence ($r = 0.12$, $p < 0.005$, best-fitting line $y = 0.1x - 0.02$). Thus, while the evoked response was much smaller than expected (the slope of the line is 0.1 instead of unity), the magnitude of the response was nevertheless correlated with the magnitude of the simulated change in depth.

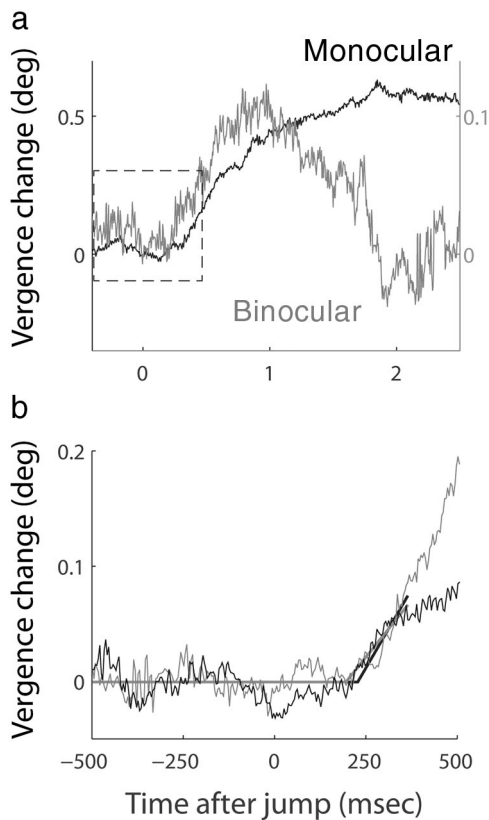


Figure 3. Dynamics of the vergence responses. Average responses from Figure 2, *a* and *c*, were subtracted to compare the dynamics of binocular and monocular conditions. *a*, The difference between back-to-front and front-to-back is shown in black for monocular viewing and gray for binocular viewing. Note the binocular response is plotted at one-fifth the scale of the monocular and shows the entire time course of the response. *b*, An expanded view of the early phase of the responses (dashed square in *a*) along with their fits in solid black and gray lines (which overlap before response onset). Note the initial responses up to 350 ms after the target jump are very similar in both cases.

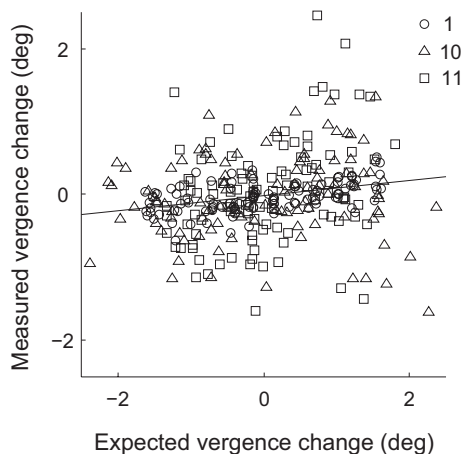


Figure 4. Magnitude of the mean vergence response as a function of object size. Data from three subjects show a correlation between the measured and expected change in depth. This implies that the size of the vergence changes correlate with the simulated size of the objects. Subjects' identification numbers appear at the inset.

Discussion

The goal of the present study was to find out whether depth information derived from self-motion through a static scene can be used by the brain to evoke binocular eye movements in accor-

dance with a 3D percept. Indeed, we found changes in vergence evoked in both directions of simulated motion parallax, consistent with perceived jumps in depth in both naive and experienced subjects. The size of the effect was larger and more robust across subjects in the monocular viewing condition. The magnitude of the response correlated with the size of the simulated object. When the stimulus was viewed binocularly, a weaker and transient response was nevertheless detected. This was a surprising result. Typically, when monocular cues to depth are in clear conflict with binocular disparity, the latter dominates. However, our results show that the perception of depth from motion parallax is so compelling that it can transiently evoke vergence movements that are in conflict binocular disparity.

The estimated response latencies to motion-parallax information are in general agreement with those reported earlier for vergence changes in response to other monocular cues without anticipation (Erkelens and Regan, 1986; Ringach et al., 1996; Leigh and Zee, 2006). The initial raising phase of the vergence response is similar under both monocular and binocular viewing. However, ~ 150 ms after the response begins (~ 350 ms after the target jump), the two conditions diverge. The monocular response continued to rise for several hundred milliseconds, while the binocular response reached an inflection point and soon began to decline back to baseline. This is likely the result of increased binocular disparity signaling the error incurred by the evoked movement. This interpretation is consistent with the observed 150 ms delay, which is similar to the one obtained by driving vergence with changes in binocular disparity (Rashbass and Westheimer, 1961; Erkelens and Collewijn, 1991; Leigh and Zee, 2006).

These findings demonstrate, for the first time, that vergence control can be influenced by self-induced, motion parallax information. A previous report using a KDE stimulus (Ringach et al., 1996) did not evoke a vergence response during binocular viewing. The main difference is that in motion parallax the retinal motion is induced by one's voluntary movement through the environment, while in the KDE it is generated by the rotation of an object and a static observer. In the case of KDE, depth sign is ambiguous. A rigid, rotating object could be moving clockwise or counterclockwise, and nothing about the motion itself distinguishes between these two possibilities. As a result, the percept is bistable, sporadically switching between the clockwise and counterclockwise interpretations (Nawrot and Blake, 1989). Self-induced motion parallax, on the other hand, is unambiguous, as the observer also has information about his/her own motion (from both efference copy and vestibular signals). Further, if the relative velocity between object and observer is known, motion parallax can provide an absolute indicator of distance independent of other cues (Ferris, 1972; Ono et al., 1986). These qualities make self-induced motion parallax a potentially stronger source of depth information compared to KDE. In agreement with this idea, Wexler et al. (2001) compared depth perception from motion parallax and KDE directly and found that motion parallax yielded a stronger percept. Similarly, we show here that motion parallax is perhaps the strongest of the monocular cues to vergence, and capable of driving binocular eye movements even when in direct conflict with binocular disparity information.

References

- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436.
- Braunstein ML (2009) Motion parallax with and without active head movements. *Perception* 38:912–913; discussion 917–919.
- Busettoni C, Fitzgibbon EJ, Miles FA (2001) Short-latency disparity vergence in humans. *J Neurophysiol* 85:1129–1152.

- Durgin FH, Proffitt DR, Olson TJ, Reinke KS (1995) Comparing depth from motion with depth from binocular disparity. *J Exp Psychol Hum Percept Perform* 21:679–699.
- Enright JT (1987) Perspective vergence: oculomotor responses to line drawings. *Vis Res* 27:1513–1526.
- Erkelens CJ, Collewijn H (1991) Control of vergence: gating among disparity inputs by target selection. *Exp Brain Res* 87:671–678.
- Erkelens CJ, Regan D (1986) Human ocular vergence movements induced by changing size and disparity. *J Physiol* 379:145–169.
- Ferris SH (1972) Motion parallax and absolute distance. *J Exp Psychol* 95:258–263.
- Hoffmann J, Sebald A (2007) Eye vergence is susceptible to the hollow-face illusion. *Perception* 36:461–470.
- Kleiner M, Brainard D, Pelli D (2007) What's new in Psychtoolbox-3? *Perception* 36 ECVF Abstract Supplement.
- Leigh RJ, Zee DS (2006) *The neurology of eye movements*, Ed 4. New York: Oxford UP.
- McLin LN Jr, Schor CM, Kruger PB (1988) Changing size (looming) as a stimulus to accommodation and vergence. *Vis Res* 28:883–898.
- Medendorp WP, Tweed DB, Crawford JD (2003) Motion parallax is computed in the updating of human spatial memory. *J Neurosci* 23:8135–8142.
- Müller J (1843). *Elements of physiology*. London: Taylor and Walton.
- Nawrot M, Blake R (1989) Neural integration of information specifying structure from stereopsis and motion. *Science* 244:716–718.
- Ono ME, Rivest J, Ono H (1986) Depth perception as a function of motion parallax and absolute-distance information. *J Exp Psychol Hum Percept Perform* 12:331–337.
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442.
- Rashbass C, Westheimer G (1961) Disjunctive eye movements. *J Physiol* 159: 339–360.
- Ringach DL, Hawken MJ, Shapley R (1996) Binocular eye movements caused by the perception of three-dimensional structure from motion. *Vis Res* 36:1479–1492.
- Rogers B (2009) Motion parallax as an independent cue for depth perception: a retrospective. *Perception* 38:907–911.
- Rogers B, Graham M (1979) Motion parallax as an independent cue for depth perception. *Perception* 8:125–134.
- Wagner M, Ehrenstein WH, Pappathomas TV (2009) Vision in reverspective: percept driven versus data-driven eye control. *Neurosci Lett* 449:142–146.
- Wexler M, Panerai F, Lamouret I, Droulez J (2001) Self-motion and the perception of stationary objects. *Nature* 409:85–88.
- Wismeijer DA, Erkelens CJ (2009) The effect of changing size on vergence is mediated by changing disparity. *J Vis* 9:12.