

Estimating distance during self-motion: A role for visual–vestibular interactions

Kalpana Dokka

Department of Anatomy and Neurobiology,
Washington University in St. Louis, USA



Paul R. MacNeilage

Vertigo, Balance, and Oculomotor Research Center,
University Hospital of Munich, Germany



Gregory C. DeAngelis

Department of Brain and Cognitive Sciences,
University of Rochester, USA



Dora E. Angelaki

Department of Neuroscience,
Baylor College of Medicine, USA



A fundamental challenge for the visual system is to extract the 3D spatial structure of the environment. When an observer translates without moving the eyes, the retinal speed of a stationary object is related to its distance by a scale factor that depends on the velocity of the observer's self-motion. Here, we aim to test whether the brain uses vestibular cues to self-motion to estimate distance to stationary surfaces in the environment. This relationship was systematically probed using a two-alternative forced-choice task in which distance perceived from monocular image motion during passive body translation was compared to distance perceived from binocular disparity while subjects were stationary. We show that perceived distance from motion depended on both observer velocity and retinal speed. For a given head speed, slower retinal speeds led to the perception of farther distances. Likewise, for a given retinal speed, slower head speeds led to the perception of nearer distances. However, these relationships were weak in some subjects and absent in others, and distance estimated from self-motion and retinal image motion was substantially compressed relative to distance estimated from binocular disparity. Overall, our findings suggest that the combination of retinal image motion and vestibular signals related to head velocity can provide a rudimentary capacity for distance estimation.

Keywords: vestibular, visual motion, motion parallax, binocular disparity, depth perception, absolute distance, distance scaling, optic flow

Citation: Dokka, K., MacNeilage, P. R., DeAngelis, G.C., & Angelaki, D. E. (2011). Estimating distance during self-motion: A role for visual–vestibular interactions. *Journal of Vision*, 11(13):2, 1–16, <http://www.journalofvision.org/content/11/13/2>, doi:10.1167/11.13.2.

Introduction

Many visual cues are combined to construct an internal representation of the 3D structure of the environment (Backus & Banks, 1999; Backus, Banks, van Ee, & Crowell, 1999; Landy, Maloney, Johnston, & Young, 1995; Marr, 1982). Among the most salient cues are binocular disparity and motion parallax. These cues carry information about depth relative to the point of fixation, but additional scaling information is needed to reconstruct absolute distance. In the case of binocular disparity, the scale factor is fixation plane distance, which can be estimated from various cues including the vergence angle of the eyes and the vertical disparity field (Backus et al., 1999; Gogel & Tietz, 1973). In the case of motion parallax induced by self-motion relative to a stationary environment, the scale factor is head velocity. In the

absence of eye movements, absolute distance (Z) to a stationary scene element is equal to the ratio of its retinal velocity (R) to head velocity (V): $Z = V/R$ (Longuet-Higgins & Prazdny, 1980; MacNeilage, Banks, Berger, & Bulthoff, 2007; Nakayama & Loomis, 1974).

Despite this straightforward relationship, we are aware of no previous experiments that have tested specifically how distance perception depends on head velocity when the observer is moved passively. Note that passive movement allows us to assess the contribution of vestibular signals in the absence of efference copy generated by active self-movement. Numerous studies have investigated the role of extraretinal information in visual processing during perception of shape, depth, and distance, but it is not clear from most of these studies whether passive movements (e.g., sensory vestibular inputs) are sufficient to account for the extraretinal interactions. Furthermore, most previous studies have employed a fixation point that is world-fixed

rather than observer-fixed, such that eye movements covary with head movements, thereby adding additional extraretinal cues.

In particular, studies of *structure from motion* typically measure perception of surface curvature, slant, or shape and compare performance during stationary viewing of visual motion displays with performance during visual motion yoked to actively generated head movements. Such studies have concluded, for example, that actively generated head motion can disambiguate the sign of curvature (Cornilleau-Peres & Droulez, 1994), can facilitate curvature discrimination (van Damme & van de Grind, 1996), and can even qualitatively alter perceived shape (Wexler, Panerai, Lamouret, & Droulez, 2001). However, it is unclear how vestibular input may contribute to these effects.

Prior research has also investigated perception of *relative depth* using variations of the standard motion parallax protocol in which an observer moves his/her head laterally while simultaneously rotating the eyes to maintain fixation on a world-fixed target (Rogers & Graham, 1979). Consequently, it is known that extraretinal cues can be used to determine depth sign in visually ambiguous displays. In most experiments, both head and eye movements were actively generated and associated efference copy signals were available, such that it was not clear which extraretinal signals were most relevant. However, recent studies have demonstrated that oculomotor signals associated with smooth eye movements provide a sufficient extraretinal cue to perceive relative depth from motion parallax (Nadler, Angelaki, & DeAngelis, 2008; Nadler, Nawrot, Angelaki, & DeAngelis, 2009; Nawrot, 2003a, 2003b; Nawrot & Joyce, 2006; Nawrot & Stroyan, 2009). Moreover, some of these studies (Nadler et al., 2009; Nawrot, 2003b) have suggested that vestibular signals are not employed to estimate relative depth.

Most relevant to the present study are prior investigations of *absolute distance* perception during head motion. Early work on this question was conducted by Gogel (1965) and Gogel and Tietz (1973, 1979). They concluded that head motion cues can influence absolute distance perception but also uncovered large systematic biases. More recent studies have revisited this question and concluded that extraretinal head motion signals can be used to scale distance accurately during active movements in both lateral (Panerai, Cornilleau-Peres, & Droulez, 2002) and fore-aft (Peh, Panerai, Droulez, Cornilleau-Peres, & Cheong, 2002) directions, with little evidence of the biases reported by Gogel et al. (but see [Discussion](#) section). In all of these studies, head movements were actively generated such that the specific contribution of vestibular signals could not be assessed. Another variable sometimes overlooked in prior research is the amplitude of self-motion; in most studies, a variable visual signal was presented during a fixed head motion. In these cases, subjects may simply learn a calibration between retinal speed and distance (Panerai et al., 2002; Peh et al., 2002).

Some studies have addressed this concern by varying speed of active head movements. This manipulation was generally found to result in appropriate scaling of perceived depth (Ono & Ujike, 1994, 2005; Yajima, Ujike, & Uchikawa, 1998) and distance (Peh et al., 2002).

The present study builds on prior research in a few respects. First, passive head movements are used, such that signals accompanying active movements (efference copy) are not available. Second, an observer-fixed fixation point is used such that eye movements are not elicited. Finally, a range of head speeds is used to investigate the influence of head velocity on distance perception. The perceptual equivalence between distance from motion and distance from stereopsis is measured using a matching task that pits distance from motion against distance specified by binocular disparity and convergence (cf. Domini & Caudek, 2010; Nawrot, 2003a). The constellation of design features employed here is novel, and our findings suggest that vestibular cues to head velocity can be used to achieve rudimentary perception of distance.

Methods

Subjects

Twelve healthy young adults (age: 25–39 years) with normal or corrected-to-normal vision and no history of neurological or visual disorders participated in the experiment. Subjects were informed about the experimental procedures and written consent was obtained as per the guidelines of the Institutional Review Board at Washington University in St. Louis.

Equipment

Experiments were conducted using a human motion system at Washington University in St. Louis. Physical motion was generated by a hexapod motion platform (Moog 6DOF2000E). Subjects were seated on the platform in a padded seat and secured with a 5-point harness. Subjects' head was held in place by a deformable plastic mesh mask. Subjects viewed the optic flow stimulus through Crystal Eyes shutter glasses (30-Hz refresh rate per eye) that were used to generate stereoscopic images. Visual motion stimuli were front-projected onto a display screen (149 × 127 cm) that was fixed to the motion platform and was located ~64 cm in front of the eyes (Fetsch, Turner, DeAngelis, & Angelaki, 2009; Gu, Fetsch, Adeyemo, DeAngelis, & Angelaki, 2010). Thus, any visual object, such as the fixation target, that remained stationary on the display screen also remained head-fixed during motion of the platform. The field of view for each eye through the glasses was ~90° × 70°. A plastic film was placed in front of the glasses to blur the image and

discourage use of accommodative cues to depth (Watt, Akeley, Ernst, & Banks, 2005). Visual displays were programmed using OpenGL libraries in Visual C++. Subjects wore earplugs and headphones, which played white noise to mask sounds associated with platform motion. Responses were collected via a button box. In some subjects, left and right eye positions were recorded at 600 Hz via a video-based eye-tracking system (ISCAN) attached to the stereo glasses. In this case, the plastic blurring film was removed.

Experimental protocol

Subjects performed a two-interval and two-alternative forced-choice task. In the first interval (Figure 1), subjects experienced a 1-s rightward translation with a Gaussian velocity profile (Figure 2). One direction of movement (rightward) was used throughout the experiment because asymmetric sensitivity to leftward versus rightward motion (B.T. Crane, personal communication) would introduce noise if both directions of movement were used. The subject's displacement in this self-motion stimulus

can be represented as: $x(t) = x_{\max} \int_0^t e^{-\sqrt{2} \left(\frac{t-d}{\delta} \right)^2} dt$. Here, d , duration of the interval, is equal to 1 s; δ , number of standard deviations of the Gaussian per interval duration, is equal to 3; and x_{\max} , peak displacement, is equal to 5, 10, or 20 cm. The Gaussian velocity profile can thus be calculated as: $v(t) = \frac{d}{dt}(x(t))$. Thus, subjects experienced a maximum displacement of 5, 10, or 20 cm corresponding to a maximum velocity of 10, 20, or 40 cm/s, respectively. The corresponding acceleration profiles were biphasic with maximum acceleration values of 0.31, 0.62, and 1.24 m/s^2 , respectively.

The visual scene depicted a world-fixed frontoparallel plane of random dots at one of the three simulated distances (see Table 1). The right lens of the shutter glasses remained opaque in the first interval of the task, so binocular disparity cues were not available (“monocular” condition). After the movement was complete, the visual scene disappeared and the screen remained blank for 0.2 s. In the second task interval (Figure 1), the subject remained stationary, but both lenses of the shutter glasses were open. Another frontoparallel plane at a different distance was presented binocularly for 1 s. Subjects then indicated if the stereoscopic plane seen during the second interval was nearer or farther than the monocular, speed-specified plane viewed in the first interval. The distance of the stereoscopic plane in the second interval was varied according to a 1-up–1-down staircase procedure to find the matching stereoscopic distance equal to the distance perceived during the first interval. Note that the stereoscopic plane was rendered accurately to appear frontoparallel at the specified distance, including the correct gradients of horizontal and vertical disparities associated with each distance.

During both stimulus intervals of each trial, subjects fixated an observer-fixed monocular fixation cross visible only to the left eye. Dot density of the visual images was constant (0.009 dot/deg^2) during both intervals, so that subjects could not use dot density as a cue to distance. In addition, the diameter of the dots was randomized (Gaussian distribution: mean = 0.5° , $SD = 0.2^\circ$) in both intervals and was not an informative cue to distance.

In addition to the monocular condition, subjects performed the distance matching task in a “disparity control” condition. In this condition, the first stimulus interval included a rightward translation of the subject similar to the monocular condition. However, both the right and left shutter glasses operated normally in both the first and

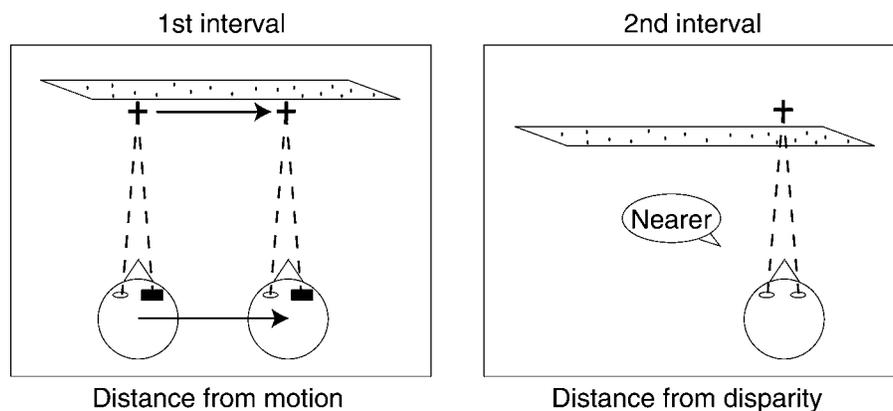


Figure 1. Illustration of the experimental protocol. In the first interval (left panel, view from above), the observer is physically translated to the right by a motion platform while viewing a world-fixed frontoparallel plane that is presented monocularly. In the second interval (right panel), the observer remains stationary and views a frontoparallel plane rendered stereoscopically at a different distance. The observer's task is to indicate if the second plane is perceived to be nearer or farther than the plane seen in the first interval. Note that the fixation cross was observer-fixed throughout the trial to minimize eye movements and was presented only to the left eye in both intervals.

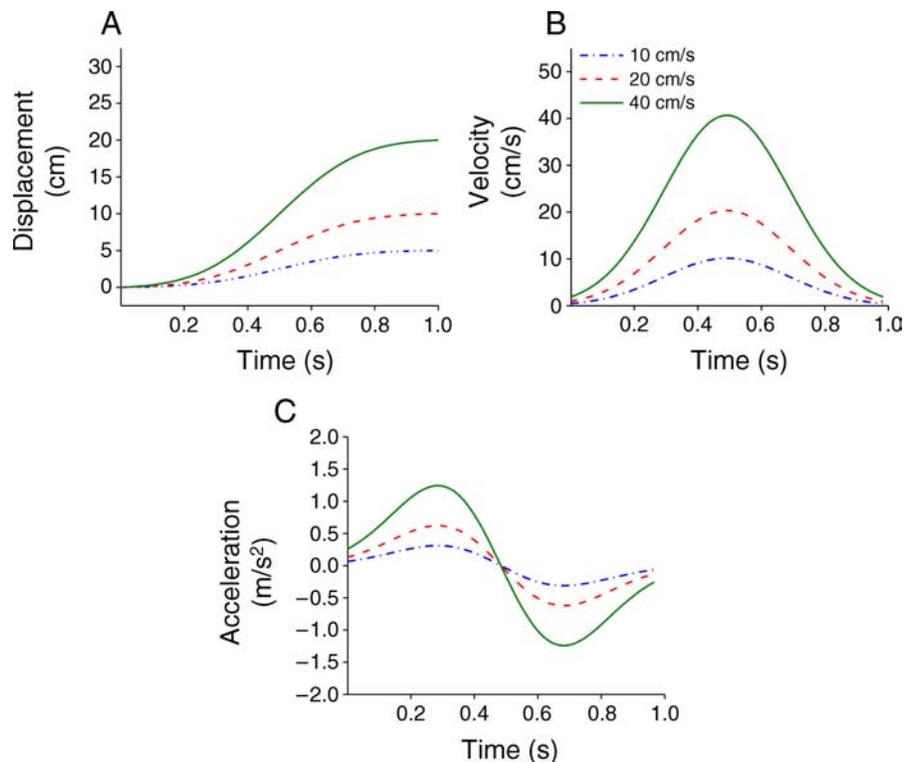


Figure 2. Time course of the self-motion stimuli. (A) Head displacement, (B) velocity, and (C) acceleration profiles used in the main experimental protocol. Subjects experienced one of three movement amplitudes during the first interval of each trial, with peak head velocities of 10, 20, and 40 cm/s, as denoted by the blue, red, and green curves, respectively.

second intervals, so responses provide a baseline measure of each subject's ability to match stereoscopic distance.

Nine combinations of head speed and simulated distance were tested 3 times total in each subject (Table 1). In each session, the three maximum head speeds (10, 20, and 40 cm/s) were combined with appropriate maximum retinal speeds (4, 9, 18, 36, or 72°/s) to yield three simulated distances (32, 64, and 128 cm) for both the monocular condition and disparity control. Trials for each of these six staircases (3 monocular, 3 disparity control) were interleaved in a given session. There were 40 trials per staircase, for a total of 240 trials per session. Across all sessions, each subject completed 120 trials (3 staircases) for each combination of head speed and simulated distance, for both the monocular condition and the disparity control condition (2160 trials in total).

The motion trajectories used here (Figure 2) are known from previous work to elicit robust self-motion perception in both monkeys and humans (Fetsch et al., 2009; Gu, Angelaki, & DeAngelis, 2008; Gu, DeAngelis, & Angelaki, 2007; MacNeilage, Banks, DeAngelis, & Angelaki, 2010; MacNeilage, Turner, & Angelaki, 2010), along with robust vestibular responses in neurons (Gu et al., 2010; Gu, Watkins, Angelaki, & DeAngelis, 2006) but very little if any translational vestibulo-ocular reflex (TVOR; Chowdhury, Takahashi, DeAngelis, & Angelaki, 2009). Sensory vestibular responses are likely to comprise the primary source of information about head speed in our

stimulus conditions; somatosensory and proprioceptive cues are unlikely to contribute strongly because self-motion perception in response to similar non-visual stimuli is greatly impaired when vestibular signals are eliminated by labyrinthectomy (Gu et al., 2007).

In addition to the main experiment described above, several control experiments were conducted: (1) the eye movement control, (2) the binocular fixation control, and (3) the default distance control.

The *eye movement control* was identical to the main experiment except that binocular eye movements were recorded. In this case, the plastic blurring film that was placed in front of the shutter glasses to reduce accommodative cues had to be removed. Five subjects (S1, S5, S6, S8, and S9) completed two staircases for each of the nine combinations of head speed and simulated distance (Table 1); one subject (S11) completed only one staircase.

		Head speed (cm/s)		
		10	20	40
Distance (cm)	32	18°/s	36°/s	72°/s
	64	9°/s	18°/s	36°/s
	128	4°/s	9°/s	18°/s

Table 1. Combinations of simulated distance (Z , rows) and maximum head speed (V , columns) presented. Each cell indicates the retinal velocity ($R = V/Z$) that was presented.

The *binocular fixation control* was identical to the main experiment except that the fixation cross was rendered binocularly in both the first and second intervals, always at the distance of the display screen (64 cm). Thus, the binocular fixation target was head-fixed and vergence eye movements were not required to maintain fixation in this control condition. In the main experiment, a monocular fixation cross was used for two reasons. First, we wanted to avoid having subjects use binocular fixation as a reference point to distance, so that judgments could not be made relative to the fixation point. Second, for disparity stimuli, there is a limited range of distances around fixation that subjects can fuse; with monocular fixation, subjects could change vergence to achieve fusion without breaking fixation. In contrast, the binocular fixation control was run in order to quantify the influence of stimulus-specified fixation distance on distance estimation. Each of 4 subjects (S8, S9, S10, and S11) completed one staircase for each of the nine combinations of head speed and simulated distance (Table 1).

The *default distance control* was run to measure the internal prior distance estimate of each subject. This experiment was identical to the main experiment except that the motion platform never moved. In the first interval, *stationary* subjects monocularly viewed a frontoparallel wall of random dots rendered at one of the three distances; however, no reliable cue to distance was available in this condition (by design). We intended for the stimulus to be identical for all 3 rendered distances in this control, such that subjects would have to rely on a default distance estimate. In the second interval, stationary subjects binocularly viewed another frontoparallel wall of dots at a different distance and then indicated if the second wall was nearer or farther than the first wall. Four subjects (S1, S6, S8, and S9) completed two staircases for each of the nine combinations of head speed and simulated distance (Table 1).

Data analysis

For each subject, perceived matching distance of the stereoscopic comparison stimulus was calculated for each combination of head speed and simulated distance by averaging the reversals for each staircase, excluding the first four reversals and averaging an even number of the remaining reversals. For the disparity control and monocular conditions, 18 ± 7 (mean \pm SD) and 16 ± 5 reversals were averaged for each staircase, respectively. In the experiments where subjects completed multiple staircases for each head speed/simulated distance combination, perceived matching distances were averaged across all staircases. To examine the influence of binocular fixation on distance matching behavior, mean matching distances for each head speed/simulated distance combination in the main experiment (averaged across three sessions) were compared with matching distances measured in the

binocular fixation control (single session). Similarly, to analyze the influence of a default distance estimate on matching behavior, mean matching distances from the main experiment (averaged across three sessions and three head speeds) were compared with those measured in the default distance control (averaged across two sessions). For these comparisons, data were taken only from subjects who participated in the main experiment as well as the corresponding control experiments.

For the analysis of eye movements, mean vergence angle (difference between the left and right eye positions) was calculated as the average vergence angle during the first 90 ms of each interval. Eye velocity was obtained by differentiating the left eye position and smoothing it with a boxcar filter. Mean eye velocity for each trial was calculated as the average left eye velocity during the 1-s duration of the first interval in which subjects experienced physical motion. Statistical analyses were performed using Statistica. We used both parametric and non-parametric regressions, as well as analysis of covariance (ANCOVA).

Results

We used a stereoscopically presented distance probe to measure the perceived distance of a monocular visual–vestibular stimulus. In the first interval of a trial (Figure 1), subjects experienced synchronized visual–vestibular rightward translation in the presence of a frontoparallel plane of random dots viewed monocularly. The second interval presented a frontoparallel plane having a different distance under normal binocular viewing conditions. Subjects then indicated if the second, disparity-defined plane appeared nearer or farther than the monocular speed-specified plane of the first interval. This procedure allowed us to find the disparity-defined distance that observers judged equivalent to the distance perceived during the monocular stimulus, for each combination of head and retinal speeds.

Main experiment

Results from a representative subject showing the effect of retinal and head speeds are shown in Figure 3. The top panels illustrate the staircase histories for a single session when the stimulus viewed during the first interval was presented with (Figure 3A, disparity control) and without (Figure 3B, monocular condition) binocular disparity. In these particular conditions, the subject experienced physical motion with a peak head velocity of 20 cm/s at three retinal speeds (36, 18, and 9°/s), thus simulating distances of 32, 64, and 128 cm (Figures 3A and 3B). The average of the staircase reversal points (excluding the first four reversals and considering an even number of the remaining reversals) was taken as the stereoscopic (3D) matching

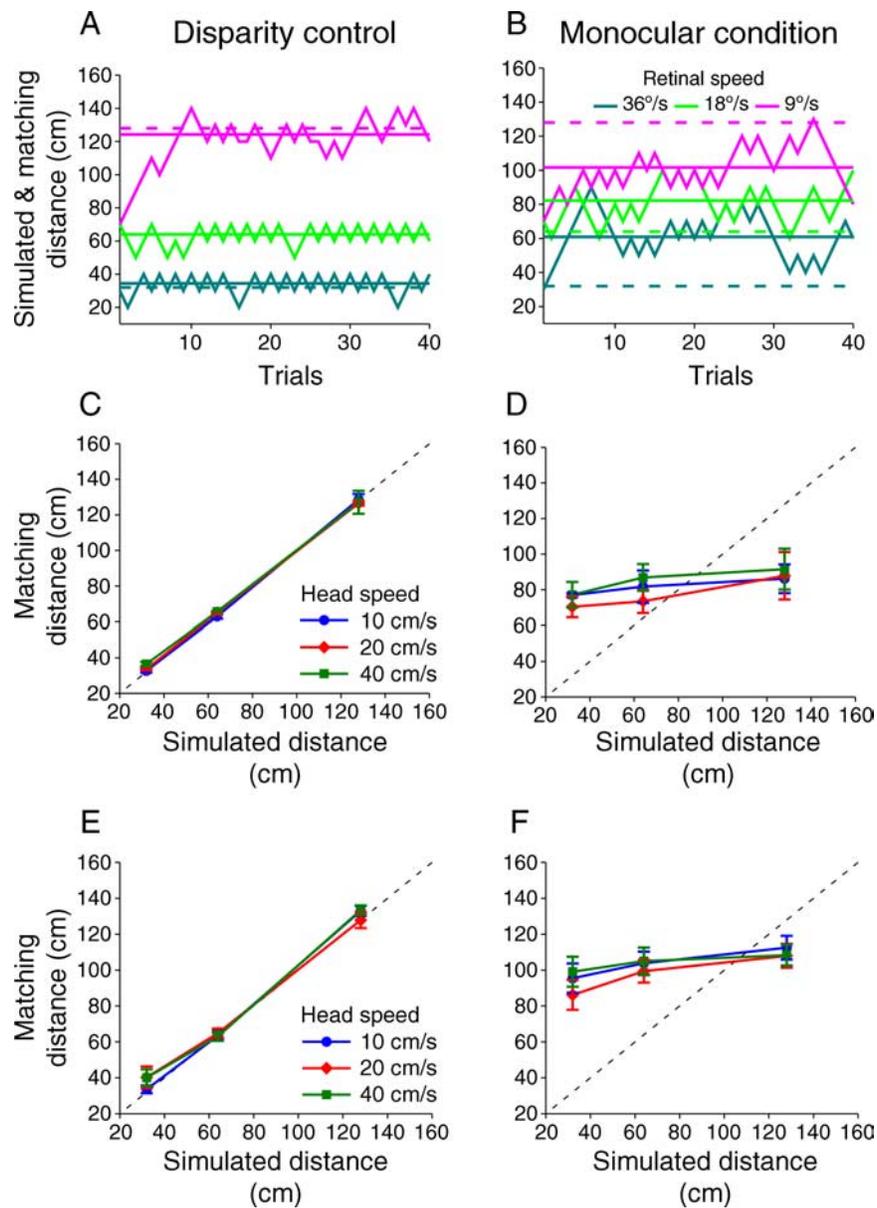


Figure 3. Sample subject and group data. The staircase histories for the (A) disparity control and (B) monocular conditions, respectively, at three different retinal speeds (same head speed: 20 cm/s). Horizontal dashed lines show the simulated distance (veridical matching); horizontal solid lines show the average of the staircase reversals, which approximates the point of subjective equality. Matching distances of Subject S1 in the (C) disparity control and (D) monocular conditions, respectively (see Table 2 for statistics). Matching distances averaged across all subjects in the (E) disparity control and (F) monocular conditions, respectively. Error bars indicate the standard error of the mean (SEM). The black dashed line represents the unity slope diagonal.

distance that observers estimated to be equivalent to the distance perceived in the monocular condition. Thus, the solid horizontal lines indicate the matching distance, whereas the dashed horizontal lines represent the actual simulated distance. Unlike in the disparity condition, the observer's matching behavior in the monocular condition was not good (Figure 3B, compare horizontal solid and dashed lines), with large offsets between the simulated distance and the matching distance. However, even in the monocular condition, the ordering of the three matching distances still agreed with the order expected from

simulated distance; that is, for the same head speed, perceived distance was farther for slow retinal speeds and nearer for fast retinal speeds.

Figures 3C and 3D show the subject's matching distance as a function of simulated distance for each of the three head speeds tested in the disparity control and monocular conditions, respectively. Across all conditions, the subject accurately matched the simulated distance of the visual scene during self-motion in the disparity control condition, indicating that movement of the platform did not influence the subject's ability to match disparity across

intervals. In the monocular condition, where distance was specified by the ratio of head speed to retinal speed, matching distance was not similar to the simulated distance. However, there was a significant correlation between simulated and matching distances (slope = 0.14, 95% confidence interval = [0.05, 0.23], $R = 0.82$, and $p = 0.008$).

Average data ($\pm SEM$) from all subjects are summarized in Figures 3E and 3F for the disparity and monocular conditions, respectively. As expected, matching and simulated distances were in agreement for the disparity control condition (slope = 0.98, 95% confidence interval = [0.94, 1.03], $R = 0.92$, and $p < 0.0001$), demonstrating that subjects could accurately match and compare disparity-defined stimuli during translation by the motion platform. By contrast, in the monocular condition, matching distance was not similar to the simulated distance. For simulated distances of 32, 64, and 128 cm, matching distances were (mean $\pm SEM$): 93.6 ± 4.7 cm, 102.7 ± 3.8 cm, and 109.6 ± 3.6 cm, respectively. Nevertheless, there was a significant positive correlation between simulated and matching distances (ANCOVA: $F(1, 300) = 28.02$, slope = 0.16, 95% confidence interval = [0.09, 0.22], $R = 0.2$, and $p < 0.0001$). The strength of this correlation varied

significantly across subjects (interaction between subject and simulated distance: $F(11, 300) = 3.80$, $p < 0.001$); 4 out of 12 individual subjects exhibited a significant correlation between simulated and matching distances (Table 2). Thus, while subjects did not correctly match distances in the monocular condition, the interaction between head speed and retinal speed influenced how they perceived distance in the absence of disparity cues.

To further explore these effects, two additional analyses were performed. First, we examined whether, for a given retinal speed, there was a significant effect of head speed on matching distance and vice versa (Figure 4). As illustrated in Figure 4A and Table 1, there were 4 successive pairs of stimuli for which the retinal speed was identical (9, 18, or 36°/s) but head speed differed. A scatter plot of the average matching distance for each of these 4 pairs is shown in Figure 4C (data from 12 subjects, $12 \times 4 = 48$ data points). For each stimulus pair with a common retinal speed, the ordinate displays the matching distance for the higher head speed (simulated far distance), whereas the abscissa shows matching distance for the lower head speed (simulated near distance). For example, subjects viewed the retinal speed of 9°/s at two head speeds: 20 cm/s and 10 cm/s, corresponding to simulated

Subject	Condition	Regression			
		Slope	CI	R	p
S1	Disparity	0.97	[0.94, 1.00]	0.99	0.0001*
	Monocular	0.14	[0.05, 0.23]	0.82	0.008*
S2	Disparity	1.08	[1.03, 1.13]	0.99	0.0001*
	Monocular	0.06	[-0.09, 0.21]	0.32	0.39
S3	Disparity	1.14	[1.00, 1.29]	0.99	0.0001*
	Monocular	0.14	[-0.05, 0.33]	0.54	0.13
S4	Disparity	1.15	[1.09, 1.21]	0.99	0.0001*
	Monocular	0.01	[-0.17, 0.18]	0.03	0.93
S5	Disparity	1.09	[0.91, 1.27]	0.98	0.0001*
	Monocular	0.10	[-0.02, 0.22]	0.59	0.08
S6	Disparity	0.93	[0.85, 1.01]	0.99	0.0001*
	Monocular	0.23	[0.02, 0.44]	0.7	0.03*
S7	Disparity	0.63	[0.35, 0.91]	0.90	0.001*
	Monocular	0.07	[-0.24, 0.37]	0.20	0.61
S8	Disparity	1.09	[1.04, 1.14]	0.99	0.0001*
	Monocular	0.03	[-0.09, 0.15]	0.22	0.56
S9	Disparity	0.91	[0.75, 1.07]	0.98	0.0001*
	Monocular	0.05	[-0.13, 0.23]	0.24	0.54
S10	Disparity	1.22	[1.12, 1.31]	0.99	0.0001*
	Monocular	0.67	[0.44, 0.89]	0.93	0.0002*
S11	Disparity	1.04	[0.89, 1.17]	0.98	0.0001*
	Monocular	0.44	[0.14, 0.73]	0.79	0.01*
S12	Disparity	0.57	[0.20, 0.95]	0.81	0.008*
	Monocular	-0.03	[-0.26, 0.20]	-0.12	0.79

Table 2. Linear regression between simulated and matching distances in the disparity control and monocular conditions, shown separately for each subject. CI indicates 95% confidence interval, and R indicates the correlation coefficient, shown along with the corresponding p -value. *Indicates a significant regression slope ($p < 0.05$). Comparable p -values were obtained using non-parametric regression analysis.

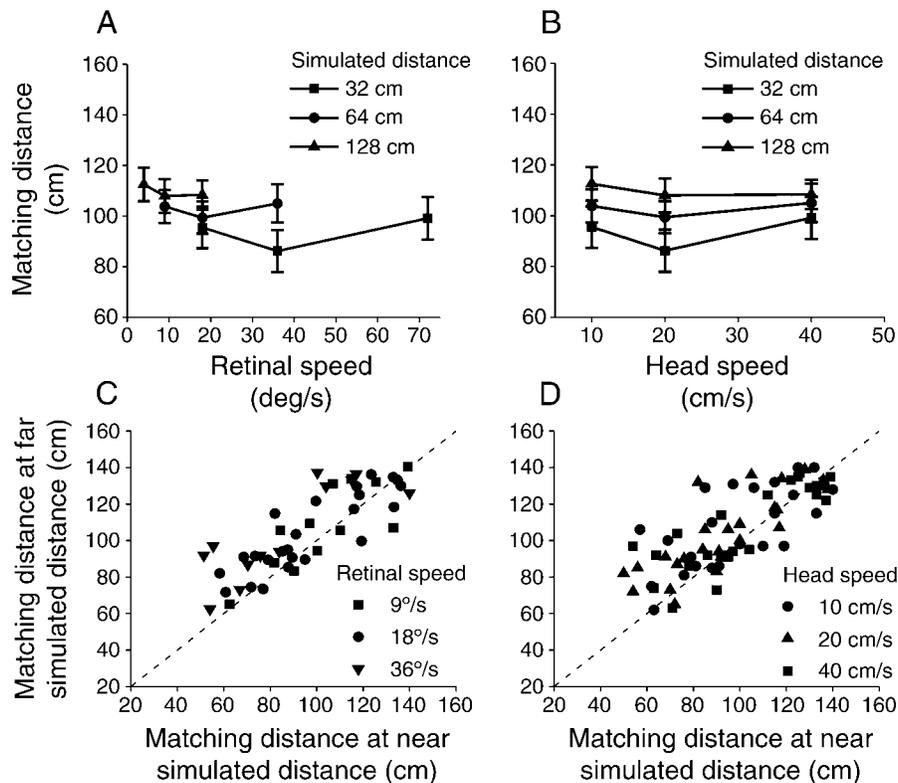


Figure 4. Summary of interactions between retinal speed and head speed in the monocular condition. (A) Distance matches as a function of retinal speed, sorted according to simulated distance. (B) Distance matches as a function of head speed, sorted by simulated distance. Error bars indicate *SEM*. (C) For each retinal speed, the distance match at a faster head speed (simulating far distance; ordinate) is plotted versus the distance match at a slower head speed (simulating near distance; abscissa). (D) For each head speed, the distance match at a slower retinal speed (simulating far distance; ordinate) is plotted versus the distance match at a faster retinal speed (simulating near distance; abscissa). The black dashed line is the unity slope diagonal.

distances of 128 and 64 cm, respectively. Thus, for the retinal speed of $9^\circ/\text{s}$, the ordinate shows the matching distance of each subject at the head speed of 20 cm/s (128-cm simulated distance) and the abscissa shows matching distance for 10 cm/s (64-cm simulated distance).

In Figure 4C, the majority of the data points lie above the unity slope diagonal, indicating that, for the same retinal speed, higher head speeds lead to farther matching distances (signed rank test, $p < 0.001$). Figures 4B and 4D present a similar comparison, with data now grouped according to head speed; here, there were 6 successive pairings in which head speed was identical but retinal speed differed (for a total of $12 \times 6 = 72$ data points). As shown in Figure 4D, the majority of the data points again lie above the unity slope diagonal, indicating that, for a given head speed, lower retinal speeds lead to farther matching distances and higher retinal speeds lead to nearer matching distances (signed rank test, $p < 0.001$).

A second analysis approach was based on the theoretical relationship between distance (Z), head speed (V), and retinal speed (R): $R = V/Z$. By a simple mathematical transformation, the logarithm of distance can be expressed

as the difference in the logarithms of the head speed and retinal speed:

$$\log(Z) = \log(V) - \log(R). \quad (1)$$

Considering the logarithm of matching distance as the dependent variable, we evaluated whether perceived distance, $\log(Z)$, can be modeled as a weighted linear combination of $\log(V)$ and $\log(R)$. In other words, we examined if $\log(Z)$ could be described by the function: $\log(Z) = a * \log(V) + b * \log(R)$, where a and b represent the slopes associated with head speed and retinal speed, respectively. There was indeed a significant influence of both head speed ($\log(V)$; ANCOVA: $F(1, 310) = 14.99$, $p < 0.001$) and retinal speed ($\log(R)$; ANCOVA: $F(1, 310) = 28.13$, $p < 0.0001$) on perceived matching distance ($\log(Z)$) when data were pooled across subjects. Importantly, the dependence of perceived matching distance on head speed had a positive slope (or weight) of 0.14 (95% confidence interval = [0.05, 0.24]), whereas the dependence of perceived matching distance on retinal speed had a negative slope of -0.13 (95% confidence interval = $[-0.20, -0.07]$).

Although the absolute values of these weights are much smaller than the ideal unity values expected from the above equation, they are nevertheless equal in magnitude and opposite in polarity as expected from the mathematical relationship between distance, head speed, and retinal velocity. Figure 5 shows the estimated slopes (i.e., weights) for each subject. In each case, the 95% confidence intervals on the slopes overlap with the negative diagonal line (see Table 3) consistent with the idea that each individual subject weighted head speed and retinal speed about equally, as expected from the above equation. Note also that the magnitudes of the slopes in Figure 5 vary considerably among subjects. As shown in Table 3, 2 out of 12 subjects exhibited a significant slope associated with head speed, whereas 4 out of 12 subjects exhibited a significant slope associated with retinal speed. These data suggest that some subjects appear to make good use of vestibular signals during the distance matching task, whereas other subjects place little weight on these cues to distance. Thus, collectively, the data provide support for the idea that at least some subjects can use a combination of head speed and retinal speed to achieve rudimentary distance perception in the monocular condition.

Binocular fixation control

Given the compression of matching distances observed in the monocular condition (Figures 3B, 3D, and 3F) and the corresponding weak effect of the visual and head speed cues, we wondered to what extent matching distance behavior was driven by the vergence angle of the eyes. To investigate this question, we repeated the experiment in a subset of four subjects. These subjects performed the same distance matching task while fixating

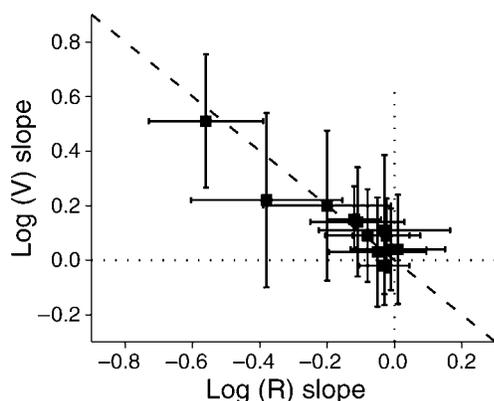


Figure 5. Testing the head speed to retinal speed ratio hypothesis, according to which $\log(Z) = \log(V) - \log(R)$. Data shown represent slopes of distance matches ($\log(Z)$) as a function of head speed ($\log(V)$) and retinal speed ($\log(R)$) for each subject (monocular condition). Error bars indicate 95% confidence intervals. Black dashed line represents the negative unity slope diagonal.

a binocular fixation cross. As observed in Figure 6, subjects' perceived matching distance in this condition was compressed over a narrow range around the distance of the binocular fixation target. Thus, presentation of a binocular fixation point led to compression of perceived distance in the monocular interval. In the absence of valid cues to distance, monocular stimuli seem to be compressed onto the fixation plane, and matching behavior may be mediated by perceived distance of the fixation cross and/or vergence angle of the eyes. This supports the hypothesis that eye vergence and perhaps the Equidistance Tendency (Gogel, 1965; see Discussion section) play a role in the compression observed in the monocular condition of the main experiment as well. Although binocular fixation acted to reduce the correlation between simulated and matching distances, this correlation was still significant (slope = 0.06, 95% confidence interval = [0.002, 0.11], $R = 0.34$, and $p = 0.04$), as compared to results for the same subset of four subjects in the main experiment (slope = 0.30, 95% confidence interval = [0.17, 0.42], $R = 0.62$, and $p < 0.001$). It should be noted that subjects performed fewer trials in the binocular fixation control (40 trials for each head speed/simulated distance combination) than in the main experiment (120 trials for each speed/distance combination; see Methods section for details). Despite fewer trials in this control experiment, the data still reveal a significant correlation between simulated and matching distances. Thus, while it is certainly influential, binocular localization of the fixation point (and associated vergence behavior) was not the only cue subjects used to estimate distance in the monocular condition; the distance estimate from retinal speed and head speed also played a role.

Eye movement control

The above results suggest that perceived distance of the fixation point and/or vergence angle of the eyes may play substantial roles in determining perceived distance in the monocular condition. We therefore repeated the experiment in six subjects while recording binocular eye movements. There was no significant influence of vergence angle in the first interval (ANCOVA: $F(1, 90) = 0.12$, $p = 0.73$) on perceived matching distance (Figure 7A) nor was there a systematic correlation between eye velocity ($F(1, 90) = 2.82$, $p = 0.1$) and perceived matching distance (Figure 7B). Importantly, the significant positive correlation between matching and simulated distances persisted (ANCOVA: $F(1, 90) = 5.78$, $p = 0.018$) after accounting for the effects of eye velocity and vergence angle in the first interval.

We then investigated if a change in vergence from the first interval to the second interval influenced matching behavior in the monocular condition. One would expect such an effect if, for example, distance estimates in both intervals are strongly driven by the vergence angle of the

Subject	Regressor	Slope	Regression		
			CI	<i>R</i>	<i>p</i>
S1	log(<i>V</i>)	0.15	[0.03, 0.27]	0.72	0.02*
	log(<i>R</i>)	−0.12	[−0.20, −0.04]	−0.81	0.01*
S2	log(<i>V</i>)	−0.02	[−0.13, 0.08]	−0.17	0.60
	log(<i>R</i>)	−0.03	[−0.11, 0.04]	−0.33	0.32
S3	log(<i>V</i>)	0.09	[−0.08, 0.26]	0.43	0.24
	log(<i>R</i>)	−0.08	[−0.21, 0.04]	−0.57	0.14
S4	log(<i>V</i>)	0.04	[−0.11, 0.19]	0.25	0.53
	log(<i>R</i>)	−0.01	[−0.11, 0.10]	−0.01	0.98
S5	log(<i>V</i>)	0.14	[−0.06, 0.34]	0.55	0.13
	log(<i>R</i>)	−0.11	[−0.25, 0.03]	−0.59	0.11
S6	log(<i>V</i>)	0.20	[−0.07, 0.48]	0.51	0.12
	log(<i>R</i>)	−0.20	[−0.40, −0.01]	−0.72	0.04*
S7	log(<i>V</i>)	0.11	[−0.16, 0.39]	0.37	0.35
	log(<i>R</i>)	−0.03	[−0.23, 0.16]	−0.16	0.69
S8	log(<i>V</i>)	0.09	[−0.04, 0.23]	0.57	0.14
	log(<i>R</i>)	−0.04	[−0.14, 0.06]	−0.34	0.34
S9	log(<i>V</i>)	0.03	[−0.17, 0.23]	0.14	0.74
	log(<i>R</i>)	−0.05	[−0.19, 0.09]	−0.32	0.44
S10	log(<i>V</i>)	0.51	[0.27, 0.76]	0.63	0.002*
	log(<i>R</i>)	−0.56	[−0.73, −0.39]	−0.95	0.0002*
S11	log(<i>V</i>)	0.22	[−0.10, 0.54]	0.33	0.14
	log(<i>R</i>)	−0.38	[−0.61, −0.16]	−0.81	0.006*
S12	log(<i>V</i>)	0.04	[−0.16, 0.24]	0.18	0.66
	log(<i>R</i>)	0.01	[−0.13, 0.15]	0.08	0.84

Table 3. Multiple linear regression model (applied separately to data from each subject) characterizing the dependence of the logarithm of matching distance in the monocular condition on head speed and retinal speed (also logarithmically transformed). CI indicates 95% confidence interval, and *R* indicates the partial correlation coefficient, shown along with the corresponding *p*-value. *Indicates significant partial correlation ($p < 0.05$).

eyes. A change from near to far convergence would be associated with far judgments, while a change from far to near vergence would be associated with near judgments. To assess this, we performed an ROC analysis in which we compared the distribution of vergence change in trials associated with “near” responses with the distribution in trials associated with “far” responses (see Gu et al., 2007 for details on this type of analysis). If the two distributions being compared are identical, one expects an ROC value near 0.5 and a $p > 0.05$, suggesting that change in vergence across stimulus intervals does not play a role. An ROC value different from 0.5 and $p < 0.05$ suggests divergent distributions, and therefore a significant association between vergence change and distance matching behavior.

As observed in Table 4, the results of the ROC analysis varied considerably from subject to subject. Overall, change in vergence angle from the first to the second interval was significantly associated with distance matching behavior in 4 out of 6 subjects (S5, S8, and S11, $p < 0.05$; S9, $p = 0.05$). Of these subjects, only one (S11) showed a significant correlation between simulated and

perceived matching distance (Table 2). In this case, it could be that head speed contributed to both the vergence response and the associated distance percept. In general, however, the correlation between vergence change and perceptual decisions (Table 4) does not predict the effectiveness of head speed and retinal speed cues to distance in the monocular condition. Thus, it seems very unlikely that our main results were a consequence of vergence eye movements.

Default distance control

We also considered the possibility that the combination of retinal and head speed cues to distance has a modest effect because subjects give high weight to an internal default estimate of distance when more robust cues to distance are not available (i.e., when the stimulus is monocular). To measure this default distance estimate, we ran the default distance control experiment in a subset of four subjects. In this control, subjects were asked to perform the distance matching task in the absence of

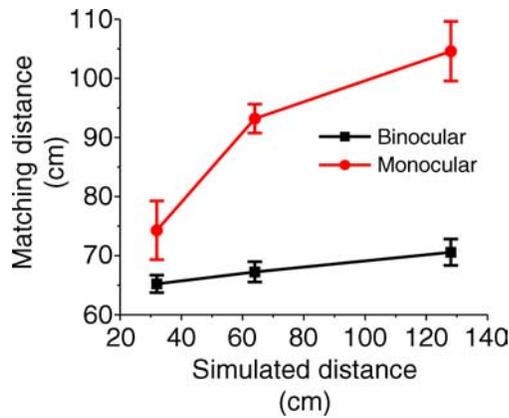


Figure 6. Distance matching behavior in four subjects who viewed the visual stimuli with either a monocular or binocular fixation cross. With monocular fixation: slope = 0.30, 95% confidence interval = [0.17, 0.42], $R = 0.62$, and $p < 0.0001$. With binocular fixation: slope = 0.06, 95% confidence interval = [0.002, 0.11], $R = 0.34$, and $p = 0.04$.

self-motion. In this case, there are no cues to distance in the monocular first interval. Note that subjects performed fewer trials in this control (80 trials for each simulated distance) than in the main experiment (see [Methods](#) section).

As expected, matching distances in this control condition do not correlate with rendered distance (Figure 8, black squares; ANCOVA: $F(1, 7) = 2.23$, slope = -0.12 , 95% confidence interval = $[-0.32, 0.07]$, $R = -0.18$, and $p = 0.18$). This simply confirms that there were no effective cues to distance in this control condition, whereas the interaction between retinal and physical motion was the primary cue in the main experiment. Indeed, the same four subjects showed a significant correlation between simulated and matching distances when self-motion was present (Figure 8, red circles; ANCOVA: $F(1, 7) = 8.65$, slope = 0.11, 95% confidence

	ROC analysis	
	ROC statistic	p -value
S1	0.52	0.17
S5	0.56	0.01*
S6	0.5	0.47
S8	0.56	0.003*
S9	0.54	0.052
S11	0.57	0.03*

Table 4. Results of the ROC analysis looking at the relationship between change in vergence angle from 1st to 2nd interval and subjects' responses of "near" and "far". $p > 0.05$ indicates that the distribution of vergence change associated with "near" and "far" responses are similar. $p < 0.05$, marked by *, indicates that vergence change distributions associated with the two responses are significantly different.

interval = [0.02, 0.2], $R = 0.56$, and $p = 0.02$). Across subjects, average default distance was between 100 and 120 cm. This may reflect the combined action of the Specific Distance Tendency (Gogel & Tietz, 1973), which would draw perceived distance toward ~ 2 m, and the Equidistance Tendency (Gogel, 1965), which would draw perceived distance toward the distance of the fixation target, presumably somewhere near the actual screen distance of 64 cm. As discussed in more detail below, these mechanisms could also explain the compression and bias in matching distances observed in the main experiment.

Discussion

As we move through the world, sensory consequences of our self-motion are registered continuously by the

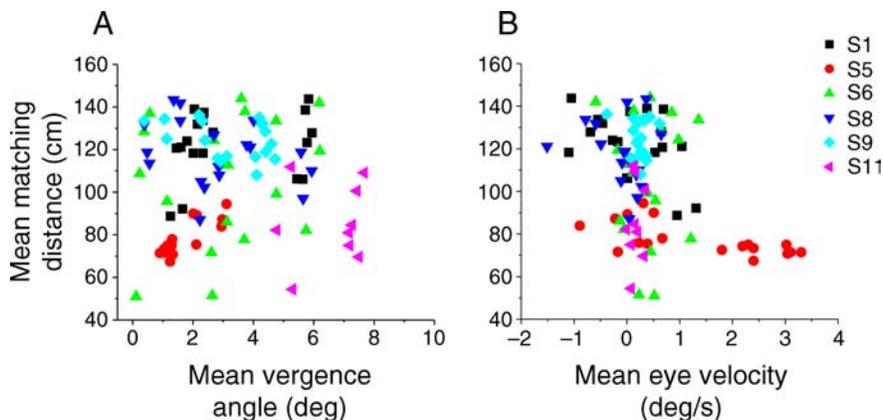


Figure 7. Relationship between distance matching behavior and eye movements for each subject across all 9 combinations of head speed and simulated distance. (A) Mean matching distance is plotted as a function of mean vergence angle (computed during the initial 90 ms of each trial). (B) Mean matching distance is plotted as a function of mean eye velocity (average velocity of the left eye during monocular stimulus presentation).

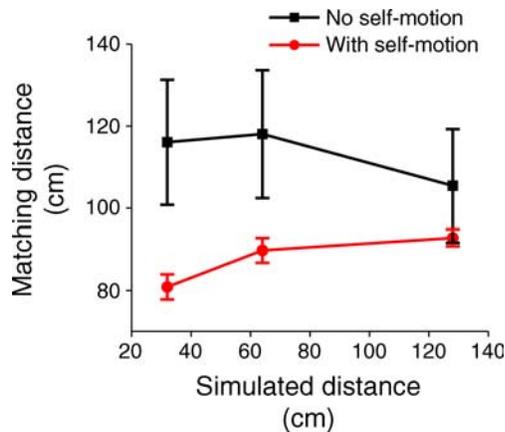


Figure 8. Default matching distance measured in four subjects who viewed the visual stimuli in the absence of physical motion. Black squares show average matching distances from 4 subjects in the default distance control experiment in which the stimulus was monocular and stationary in the first interval. As expected, matching distances are not correlated ($F(1, 7) = 2.23$, slope = -0.12 , 95% confidence interval = $[-0.32, 0.07]$, $R = -0.18$, and $p = 0.18$) with simulated distances, instead reflecting some default distance estimate. By comparison, red circles show average matching distances of the same four subjects on trials in which the platform moved in the first interval. Average distance matches are correlated with simulated distance ($F(1, 7) = 8.65$, slope = 0.11 , 95% confidence interval = $[0.02, 0.2]$, $R = 0.56$, and $p = 0.02$).

visual and vestibular systems. Because these senses respond in diverse fashion to the same stimulus, the information they convey is complimentary, and there is potential for beneficial multisensory interactions. Here, we demonstrate such an interaction in which a vestibular estimate of head velocity may be used to scale distance during self-motion. We show that for a fixed head speed, slower retinal speeds are perceived to be farther. Similarly, for a fixed retinal speed, faster head speeds lead to perception of farther distances. In other words, perceived distance is influenced by the ratio of retinal and head speeds. We suggest that a vestibular signal is the most likely source of information about head velocity because we used passive head motion and an observer-fixed fixation target, such that efference copy signals related to active head and eye movements are minimized.

Absolute distance perception during self-motion

The relationship between head motion and perception of absolute distance (i.e., absolute motion parallax) has been studied in detail before, both theoretically (Gibson, 1950; Koenderink, 1986; Longuet-Higgins & Prazdny, 1980; Nakayama & Loomis, 1974) and experimentally (Beall, Loomis, Philbeck, & Fikes, 1995; Dees, 1966; Eriksson,

1974; Ferris, 1972; Gogel & Tietz, 1973, 1979; Johansson, 1973; Panerai et al., 2002; Peh et al., 2002). Most of these studies differ from ours in that head movements were active, and corresponding eye movements were elicited to maintain fixation on a world-fixed target. Therefore, a contribution of vestibular signals cannot be concluded. Similar to our findings, these studies showed that the combination of visual motion and extraretinal signals related to head motion acts as a cue to absolute distance.

In one study (Beall et al., 1995), subjects moved the head from side to side and judged the distance of a glowing sphere viewed monocularly in darkness. Sphere size and distance were manipulated such that visual angle remained constant and the only cue to distance was absolute motion parallax. Distance estimates were heavily biased but revealed correct depth ordering, similar to the present study. Another relevant study (Gogel & Tietz, 1979) compared the relative effectiveness of oculomotor cues, such as convergence and accommodation, with absolute motion parallax cues. Subjects moved their head from side to side while viewing an object at one of three physical distances, which provided the stimulus for convergence and accommodation. At each distance, 3 gains consistent with 3 absolute motion parallax distances were applied to the visual object speed concurrent with head motion. The dependent measure was perceived absolute distance of the object, probed with verbal reports but also with another indirect procedure (see Gogel & Tietz, 1979 for details). Results showed that convergence was the strongest cue to absolute distance. Accommodation and absolute motion parallax were about equally effective; once again, all distance estimates were biased. Such biases were also reported in a study in which subjects were asked to reach to targets that were presented monocularly during active head movements (Bingham & Pagano, 1998).

More recently, experiments were conducted (Panerai et al., 2002; Peh et al., 2002) that compared perception of absolute distance between two conditions, one in which the observer moved and viewed a stationary sphere composed of random dots (self-motion condition) and another in which the sphere moved and was viewed by a stationary observer (object motion condition). These studies reported near-veridical (unbiased) perception of absolute distance during the self-motion condition, contrary to our results and previous findings that report significant bias and only a weak effect of absolute motion parallax cues (Beall et al., 1995; Gogel & Tietz, 1973, 1979). Although the conclusions of Panerai et al. (2002) differ sharply from our findings, the difference in results is likely smaller than it appears. Panerai et al. found accurate distance judgments when head velocity was constant in the self-motion condition, but subjects also showed accurate distance estimates in the object motion condition. This suggests that subjects simply learned a calibration between retinal velocity and distance. Indeed, when head velocity was randomized across trials (as in the present study), distance perception in the object motion condition

was largely abolished. In the self-motion condition, extraretinal cues clearly allowed some distance perception, but the authors did not measure the point of subjective equality in this condition, so it is unclear to what extent the subjects accurately judged distance when head velocity was varied. Two other factors may also have contributed to the near-veridical distance judgments in the studies of Panerai et al. and Peh et al. (2002). First, subjects were asked to categorize distance into one of four groups that spanned the range of true distances, and this may have helped subjects learn to calibrate retinal velocity to distance. Second, a screening procedure was employed in which subjects were given feedback about the accuracy of their distance judgments before participating in the experiment. Thus, our findings with passive translation of subjects may not be inconsistent with the results of Panerai et al. using active self-motion.

In summary, while distance estimates appear to be influenced by absolute motion parallax, the effects are often weak and biases are well documented. The Specific Distance Tendency (Gogel & Tietz, 1973) describes the observation that objects are perceived to be at one particular distance (~ 2 m) when presented in the absence of reliable distance cues. The Equidistance Tendency (Gogel, 1965) describes the observation that objects of ambiguous distance are perceived to be at the same distance as adjacent objects. These tendencies have recently been proposed to result from prior expectations based on statistical regularities in the environment (Yang & Purves, 2003). Thus, perception of absolute distance appears to depend not only on sensory signals but also on internal tendencies (i.e., priors) that can bias distance estimates based on sensory cues. Such tendencies may account for biases in our results as well. Unlike some of the experiments described above, our design does not directly assess how subjects estimate absolute distance. Instead, we obtain equivalence between distance from motion and distance from stereopsis. Nevertheless, to the extent that such tendencies act more strongly on estimates of distance from motion than they do on estimates of distance from stereopsis (Gogel & Tietz, 1973), they could still account for our results.

We attempted to characterize these internal tendencies in the binocular fixation and default distance control experiments (Figures 6 and 8). In the binocular fixation control, perceived matching distances were close to the simulated distance of the fixation cross (64 cm), consistent with the Equidistance Tendency. This tendency may have played a role even when the fixation cross was presented monocularly. It is likely that subjects assigned a distance to the fixation cross, and perceived distance of the random-dot field may have been drawn to this distance. For example, a presumed fixation cross distance of ~ 100 cm would be consistent with the compression that was often observed in the main experiment. Regarding presumed fixation cross distance, it should be pointed out that, while we rendered an *observer-fixed* fixation cross of ambiguous

distance, an alternative interpretation is possible. Specifically, an object that remains stationary on the display screen during observer motion (like our fixation cross) is also consistent with a *world-fixed* object at infinite distance. However, it is unlikely that the stimulus was interpreted this way because vergence angles measured in the eye movement control condition, which should provide a rough indication of presumed fixation distance, did not cluster near zero (Figure 7A).

Biases were also examined in the default distance control. Subjects remained stationary, so no visual-vestibular cues to distance were available. These measurements might therefore reflect the combined action of the Specific Distance and Equidistance Tendencies, which would draw estimates toward a distance of ~ 2 m and ~ 64 cm, respectively (assuming that the fixation cross was perceived to be at the screen distance). Mean matching distances were between 110 and 120 cm, about halfway between the ~ 2 m and ~ 64 cm distances, roughly consistent with the predictions of these tendencies.

In the monocular condition of the main experiment, simulated distances of 32, 64, and 128 cm were perceptually matched to stereoscopically defined distances of 94, 103, and 110 cm. These results reflect a bias consistent with the tendencies described above but also indicate that the ratio of head speed to retinal speed can act as a weak cue to distance, perhaps similar to accommodation in its effectiveness. The relative weakness of this cue may be partly due to the fact that it is based on a ratio computation and therefore incorporates variability from both visual and vestibular signals. More variable cues are typically weighted less during cue combination (Backus & Banks, 1999; Hillis, Watt, Landy, & Banks, 2004; Landy et al., 1995), thus prior expectations or tendencies would have more influence. Alternatively, the weak action of this speed ratio cue could be due to uncertainty about the stationarity of the visual scene. If retinal image motion is partly attributable to the independent movement of objects in the world, then distance estimated from monocular visual motion and head speed may be unreliable.

Oculomotor signals and distance perception

Oculomotor signals, such as convergence, have been shown to act as robust cues to distance. Even the Specific Distance Tendency is observed to correlate with the default convergence angle of the eyes in darkness (Foley, 1978; Owens & Leibowitz, 1976). The idea that convergence influences judgments of absolute distance, even in the absence of stereoscopic cues, is consistent with our ROC analysis (Table 4), which shows that convergence behavior was correlated with observer responses in 4 out of 6 subjects.

Despite the substantial influence of convergence, we suggest that vestibular signals related to head speed also

influenced distance judgments in our study. This might seem surprising because prior studies ruled out vestibular signals in favor of oculomotor signals in determining relative depth sign (not absolute distance) from motion parallax (Nawrot, 2003a, 2003b; Nawrot & Joyce, 2006). In particular, Nawrot (2003a) included a stimulus condition with a head-fixed fixation point and concluded that depth sign was determined by an oculomotor signal that was needed to cancel the reflexive TVOR, not by the vestibular signal itself. In our experiment, oculomotor TVOR cancellation signals may have been available. Although the TVOR is weak under the conditions of our experiments (Angelaki & Hess, 2005), we nevertheless cannot rule out the possibility that the head velocity-related signal used for scaling absolute distance is driven by oculomotor rather than purely sensory vestibular information. Even if oculomotor TVOR cancellation signals participate in distance scaling, this information still has its origin in the vestibular system, as there are no other cues present in our experiments that can provide information about head speed.

Neurophysiological implications

These findings have important implications for neurophysiological investigations of brain areas that signal depth or distance. In particular, MT neurons contribute to depth perception from disparity (Chowdhury & DeAngelis, 2008; DeAngelis & Uka, 2003; Uka & DeAngelis, 2006) and also carry information about depth from motion parallax (Nadler et al., 2008, 2009). Although MT neurons do not respond to vestibular stimuli when presented in isolation (Chowdhury et al., 2009), the present findings suggest that neuronal responses to retinal image motion might be modulated by head speed. Alternatively, the necessary interaction between retinal motion and head speed may not occur until downstream multisensory areas, like the dorsal medial superior temporal area (MSTd), where neurons are tuned not only to optic flow and binocular disparity (Duffy & Wurtz, 1991a, 1991b; Gu et al., 2008; Roy & Wurtz, 1990) but also to vestibular cues (Gu et al., 2007, 2006; Page & Duffy, 2003). At present, the brain areas that are involved in coding the ratio of head speed to retinal speed remain unknown.

Conclusions

In summary, we examined how vestibular cues influence absolute distance perception. Our results indicate that monocularly perceived distance during passive self-motion depends on both head speed and retinal image speed. However, large biases were observed in perceived distances, and some subjects appeared to be much more capable than others of using vestibular input to estimate

distance. Overall, these findings suggest that vestibular signals related to self-motion, along with retinal image motion, can provide a crude estimate of distance during self-motion.

Acknowledgments

This work was supported by NIH R01 DC007620 (to DEA), NIH EY016178 (to GCD), NIH Institutional National Research Service Award 5-T32-EY13360-07 (to PRM), and the German Federal Ministry of Education and Research under Grant Code 01 EO 0901 (to PRM). The authors bear full responsibility for the content of this publication.

Author contributions: Kalpana Dokka and Paul R. MacNeilage contributed equally to this work.

Commercial relationships: none.

Corresponding author: Dora E. Angelaki.

Email: angelaki@bcm.edu.

Address: Department of Neuroscience, MS: BCM 295, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030.

References

- Angelaki, D. E., & Hess, B. J. (2005). Self-motion-induced eye movements: Effects on visual acuity and navigation. *Nature Reviews Neuroscience*, *6*, 966–976.
- Backus, B. T., & Banks, M. S. (1999). Estimator reliability and distance scaling in stereoscopic slant perception. *Perception*, *28*, 217–242.
- Backus, B. T., Banks, M. S., van Ee, R., & Crowell, J. A. (1999). Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Research*, *39*, 1143–1170.
- Beall, A. C., Loomis, J. L., Philbeck, J. W., & Fikes, T. G. (1995). *Absolute motion parallax weakly determines visual scale in real and virtual environments*. Paper presented at the Proceedings of Society of Photo-Optical Instrumentation Engineers, Bellingham, WA.
- Bingham, G. P., & Pagano, C. C. (1998). The necessity of a perception-action approach to definite distance perception: Monocular distance perception to guide reaching. *Journal Experimental Psychology: Human Perception and Performance*, *24*, 145–168.
- Chowdhury, S. A., & DeAngelis, G. C. (2008). Fine discrimination training alters the causal contribution of macaque area MT to depth perception. *Neuron*, *60*, 367–377.
- Chowdhury, S. A., Takahashi, K., DeAngelis, G. C., & Angelaki, D. E. (2009). Does the middle temporal

- area carry vestibular signals related to self-motion? *Journal of Neuroscience*, *29*, 12020–12030.
- Cornilleau-Peres, V., & Droulez, J. (1994). The visual perception of three-dimensional shape from self-motion and object-motion. *Vision Research*, *34*, 2331–2336.
- DeAngelis, G. C., & Uka, T. (2003). Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *Journal of Neurophysiology*, *89*, 1094–1111.
- Dees, J. W. (1966). Accuracy of absolute visual distance and size estimation in space as a function of stereopsis and motion parallax. *Journal of Experimental Psychology*, *72*, 466–476.
- Domini, F., & Caudek, C. (2010). Matching perceived depth from disparity and from velocity: Modeling and psychophysics. *Acta Psychologica (Amsterdam)*, *133*, 81–89.
- Duffy, C. J., & Wurtz, R. H. (1991a). Sensitivity of MST neurons to optic flow stimuli: I. A continuum of response selectivity to large-field stimuli. *Journal of Neurophysiology*, *65*, 1329–1345.
- Duffy, C. J., & Wurtz, R. H. (1991b). Sensitivity of MST neurons to optic flow stimuli: II. Mechanisms of response selectivity revealed by small-field stimuli. *Journal of Neurophysiology*, *65*, 1346–1359.
- Eriksson, E. (1974). Movement parallax during locomotion. *Perception & Psychophysics*, *16*, 197–200.
- Ferris, S. H. (1972). Motion parallax and absolute distance. *Journal of Experimental Psychology*, *95*, 258–263.
- Fetsch, C. R., Turner, A. H., DeAngelis, G. C., & Angelaki, D. E. (2009). Dynamic reweighting of visual and vestibular cues during self-motion perception. *Journal of Neuroscience*, *29*, 15601–15612.
- Foley, J. M. (1978). Primary distance perception. In R. Held, H. W. Leibowitz, & H. L. Teuber (Eds.), *Handbook of sensory physiology* (vol. 8, pp. 181–213). Berlin, Germany: Springer-Verlag.
- Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.
- Gogel, W. C. (1965). Equidistance tendency and its consequences. *Psychological Bulletin*, *64*, 153–163.
- Gogel, W. C., & Tietz, J. D. (1973). Absolute motion parallax and specific distance tendency. *Perception & Psychophysics*, *13*, 284–292.
- Gogel, W. C., & Tietz, J. D. (1979). A comparison of oculomotor and motion parallax cues of egocentric distance. *Vision Research*, *19*, 1161–1170.
- Gu, Y., Angelaki, D. E., & DeAngelis, G. C. (2008). Neural correlates of multisensory cue integration in macaque MSTd. *Nature Neuroscience*, *11*, 1201–1210.
- Gu, Y., DeAngelis, G. C., & Angelaki, D. E. (2007). A functional link between area MSTd and heading perception based on vestibular signals. *Nature Neuroscience*, *10*, 1038–1047.
- Gu, Y., Fetsch, C. R., Adeyemo, B., DeAngelis, G. C., & Angelaki, D. E. (2010). Decoding of MSTd population activity accounts for variations in the precision of heading perception. *Neuron*, *66*, 596–609.
- Gu, Y., Watkins, P. V., Angelaki, D. E., & DeAngelis, G. C. (2006). Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. *Journal of Neuroscience*, *26*, 73–85.
- Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision*, *4*(12):1, 967–992, <http://www.journalofvision.org/content/4/12/1>, doi:10.1167/4.12.1. [PubMed] [Article]
- Johansson, G. (1973). Monocular movement parallax and near-space perception. *Perception*, *2*, 135–146.
- Koenderink, J. J. (1986). Optic flow. *Vision Research*, *26*, 161–179.
- Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research*, *35*, 389–412.
- Longuet-Higgins, H. C., & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B: Biological Sciences*, *208*, 385–397.
- MacNeilage, P. R., Banks, M. S., Berger, D. R., & Bulthoff, H. H. (2007). A Bayesian model of the disambiguation of gravito-inertial force by visual cues. *Experimental Brain Research*, *179*, 263–290.
- MacNeilage, P. R., Banks, M. S., DeAngelis, G. C., & Angelaki, D. E. (2010). Vestibular heading discrimination and sensitivity to linear acceleration in head and world coordinates. *Journal of Neuroscience*, *30*, 9084–9094.
- MacNeilage, P. R., Turner, A. H., & Angelaki, D. E. (2010). Canal-otolith interactions and detection thresholds of linear and angular components during curved-path self-motion. *Journal of Neurophysiology*, *104*, 765–773.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W. H. Freeman and Company.
- Nadler, J. W., Angelaki, D. E., & DeAngelis, G. C. (2008). A neural representation of depth from motion parallax in macaque visual cortex. *Nature*, *452*, 642–645.
- Nadler, J. W., Nawrot, M., Angelaki, D. E., & DeAngelis, G. C. (2009). MT neurons combine visual motion

- with a smooth eye movement signal to code depth-sign from motion parallax. *Neuron*, *63*, 523–532.
- Nakayama, K., & Loomis, J. M. (1974). Optical velocity patterns, velocity-sensitive neurons, and space perception: A hypothesis. *Perception*, *3*, 63–80.
- Nawrot, M. (2003a). Depth from motion parallax scales with eye movement gain. *Journal of Vision*, *3*(11):17, 841–851, <http://www.journalofvision.org/content/3/11/17>, doi:10.1167/3.11.17. [PubMed] [Article]
- Nawrot, M. (2003b). Eye movements provide the extra-retinal signal required for the perception of depth from motion parallax. *Vision Research*, *43*, 1553–1562.
- Nawrot, M., & Joyce, L. (2006). The pursuit theory of motion parallax. *Vision Research*, *46*, 4709–4725.
- Nawrot, M., & Stroyan, K. (2009). The motion/pursuit law for visual depth perception from motion parallax. *Vision Research*, *49*, 1969–1978.
- Ono, H., & Ujike, H. (1994). Apparent depth with motion aftereffect and head movement. *Perception*, *23*, 1241–1248.
- Ono, H., & Ujike, H. (2005). Motion parallax driven by head movements: Conditions for visual stability, perceived depth, and perceived concomitant motion. *Perception*, *34*, 477–490.
- Owens, D. A., & Leibowitz, H. W. (1976). Oculomotor adjustments in darkness and the specific distance tendency. *Perception & Psychophysics*, *20*, 2–9.
- Page, W. K., & Duffy, C. J. (2003). Heading representation in MST: Sensory interactions and population encoding. *Journal of Neurophysiology*, *89*, 1994–2013.
- Panerai, F., Cornilleau-Peres, V., & Droulez, J. (2002). Contribution of extraretinal signals to the scaling of object distance during self-motion. *Perception & Psychophysics*, *64*, 717–731.
- Peh, C. H., Panerai, F., Droulez, J., Cornilleau-Peres, V., & Cheong, L. F. (2002). Absolute distance perception during in-depth head movement: Calibrating optic flow with extra-retinal information. *Vision Research*, *42*, 1991–2003.
- Rogers, B., & Graham, M. (1979). Motion parallax as an independent cue for depth perception. *Perception*, *8*, 125–134.
- Roy, J. P., & Wurtz, R. H. (1990). The role of disparity-sensitive cortical neurons in signalling the direction of self-motion. *Nature*, *348*, 160–162.
- Uka, T., & DeAngelis, G. C. (2006). Linking neural representation to function in stereoscopic depth perception: Roles of the middle temporal area in coarse versus fine disparity discrimination. *Journal of Neuroscience*, *26*, 6791–6802.
- van Damme, W. J., & van de Grind, W. A. (1996). Non-visual information in structure-from-motion. *Vision Research*, *36*, 3119–3127.
- Watt, S. J., Akeley, K., Ernst, M. O., & Banks, M. S. (2005). Focus cues affect perceived depth. *Journal of Vision*, *5*(10):7, 834–862, <http://www.journalofvision.org/content/5/10/7>, doi:10.1167/5.10.7. [PubMed] [Article]
- Wexler, M., Panerai, F., Lamouret, I., & Droulez, J. (2001). Self-motion and the perception of stationary objects. *Nature*, *409*, 85–88.
- Yajima, T., Ujike, H., & Uchikawa, K. (1998). Apparent depth with retinal image motion of expansion and contraction yoked to head movement. *Perception*, *27*, 937–949.
- Yang, Z., & Purves, D. (2003). A statistical explanation of visual space. *Nature Neuroscience*, *6*, 632–640.