Pergamon

0042-6989(95)00285-5

Vision Res., Vol. 36, No. 10, pp. 1479–1492, 1996
Copyright © 1996 Elsevier Science Ltd. All rights reserved
Printed in Great Britain
0042-6989/96 $15.00 + 0.00

Binocular Eye Movements Caused by the Perception of Three-Dimensional Structure from Motion

DARIO L. RINGACH,* MICHAEL J. HAWKEN,† ROBERT SHAPLEY†

Received 29 June 1995; in revised form 19 October 1995

We report that the perception of three-dimensional structure from monocular two-dimensional images changing over time—the kinetic depth effect (KDE)—can evoke binocular eye movements consistent with a three-dimensional percept. We used a monocular KDE stimulus that induced a vivid perception of a rigid three-dimensional sphere rotating in space. The gaze directions of both eyes were measured while observers pursued the motion of a patch on the surface of the perceived sphere as it went through a complete revolution. We found that the eyes converged when the patch was perceived on the front surface of the KDE sphere and diverged when the patch was perceived in the back. The pattern, magnitude and dynamics of binocular eye movements observed in the KDE experiment resembled those obtained when subjects viewed binocularly a light-emitting diode (LED) rotating in space and to the responses obtained with a dynamic stereogram simulating a rotating random dot sphere. Thus, the perception of three-dimensional structure from motion, stereopsis, or motion *and* stereopsis combined, were effective in guiding binocular eye movements. Copyright ©1996 Elsevier Science Ltd.

Kinetic depth effect   Structure-from-motion   Eye movements   Vergence   Sensorimotor system

INTRODUCTION

The kinetic depth effect (KDE)—or structure-from-motion (SFM)—is the perception of three-dimensional structure and motion from two-dimensional images that change over time (Wallach & O'Connell, 1953; Gibson & Gibson, 1957; Braunstein, 1976; Rogers & Graham, 1979; Ullman, 1979; Longuet-Higgins & Prazdny, 1980; Prazdny, 1980; Koenderink & van Doorn, 1975). The phenomenon is readily observed using two-dimensional image sequences in which any single frame is perceived as "flat", while a strong three-dimensional percept is evoked if the images are viewed in rapid succession. The existence of such stimuli, by itself, demonstrates that the human visual system is capable of integrating information across time and organizing it into a three-dimensional percept.

One hurdle in the study of the KDE has been devising reliable techniques to measure directly the *perceived depth* caused by a visual stimulus (Landy *et al.*, 1995). Subjective ratings of "*coherency*", "*rigidity*" and "*depth*" have been used in the past (Braunstein, 1962; Green, 1961; Dosher *et al.*, 1989b; Loomis & Eby, 1988; Petersik, 1979, 1980; Todd *et al.*, 1988). However, it is not clear how these measures relate to perceived depth. Attempts also have been made to study SFM based on objective psychophysical performance in a variety of tasks involving the detection, discrimination, or classification of a three-dimensional shape defined by two-dimensional motion information (Lappin *et al.*, 1980; Braunstein & Andersen, 1981; Andersen & Braunstein, 1983; Todd, 1984, 1985; Braunstein *et al.*, 1987; Sperling *et al.*, 1989; Siegel & Andersen, 1988; Hildreth *et al.*, 1990; Treue *et al.*, 1995). A major methodological problem arises in these studies, namely, deciding if a subject is performing the assigned psychophysical task on the basis of a *perceived three-dimensional shape* or is using some *alternative computation* involving only two-dimensional cues (Sperling *et al.*, 1989; Braunstein, 1994).

A different approach, that may avoid some of the above complications, is the use of motor activity measures to study three-dimensional visual perception. The basic idea is that the perception of absolute distances in space and of the three-dimensional shape of objects is

---

*To whom all correspondence should be addressed at: Center for Neural Science, New York University, 4 Washington Place, New York, NY 10003, U.S.A. [*Tel* 01-212-998-7614; *Fax* 01-212-995-4011; *Email* dario:@cns.nyu.edu].

†Center for Neural Science, New York University, New York, NY 10003, U.S.A.

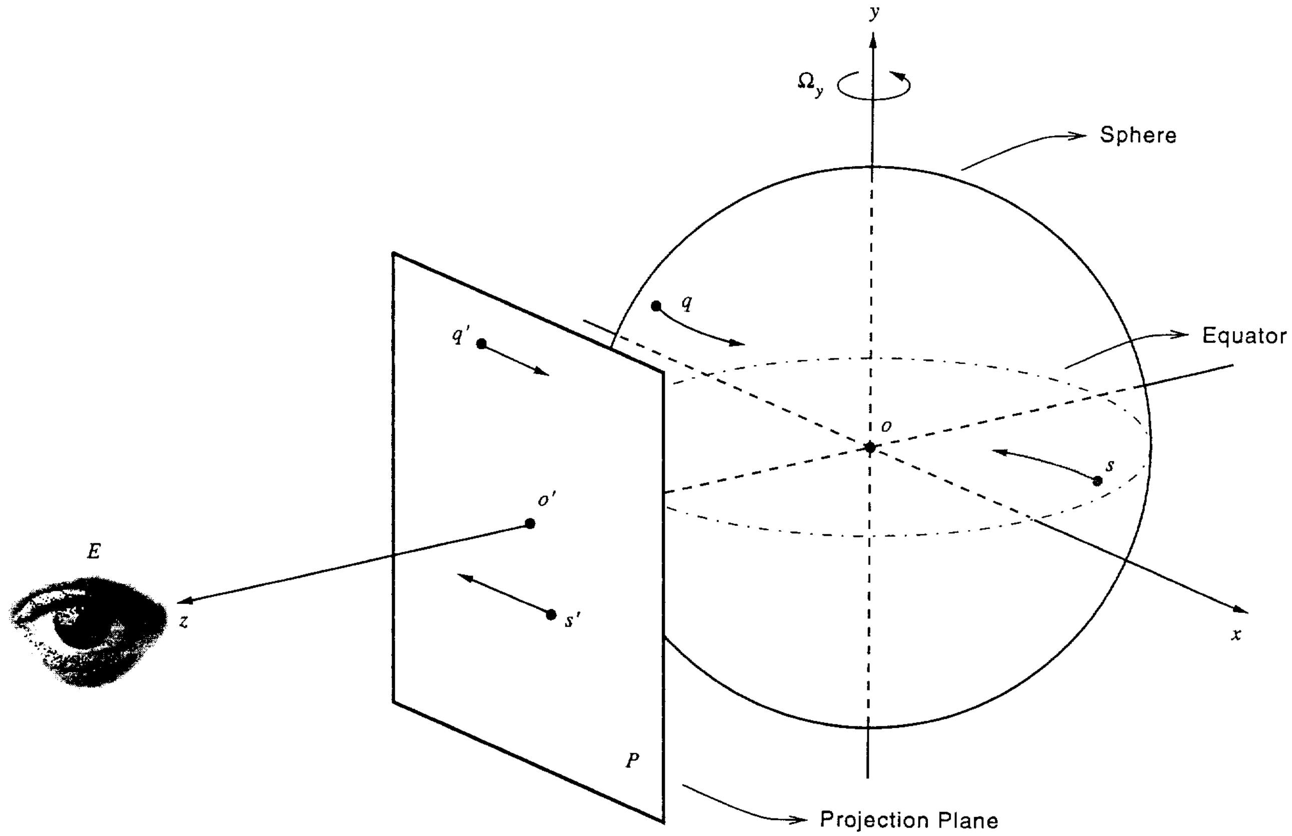

FIGURE 1. The viewing geometry used in the generation of the KDE stimulus. Uniformly distributed points were selected on the surface of a sphere. Only two dots are illustrated here. The simulated angular frequency of the sphere was $\Omega_y = 30$ deg/sec. The points were orthographically projected onto the plane P (the computer screen) which was parallel to the (x,y)-plane (see Methods section).

important in the planning and control of motor activities, such as reaching, grasping, manipulating objects and guiding binocular eye movements. As a consequence, it would be natural to expect some motor activities to reflect the perceived three-dimensional structure of the environment. We explored the possibility that the perception of three-dimensional structure-from-motion could influence binocular eye movements.

In this work we measured the gaze directions of both eyes while observers pursued the motion of a patch on the surface of a KDE sphere presented monocularly. To anticipate the results, we find that the perception of structure from motion causes binocular eye movements consistent with a three-dimensional percept. In other words, during the inspection of a KDE stimulus, subjects moved their eyes *as if* they were binocularly viewing a real three-dimensional object rotating in depth. Then we studied the dependence of the oculomotor response as a function of the number and lifetime of the dots in the KDE stimulus. The dynamical properties of the responses were investigated as well. Our findings suggest the use of binocular oculomotor responses as a possible measure of perceived depth.

EXPERIMENTS

Experiment 1. Binocular Eye Movements Caused by the Perception of Three-Dimensional Structure from Two-Dimensional Motion

An observer converges or diverges the eyes to fixate a visual target at different depths. Such vergence eye movements have been considered traditionally to be driven mainly by binocular disparity and image blur. The aim of this experiment was find out if changes in vergence also occur when a subject views a *monocular* KDE stimulus that evokes a strong three-dimensional percept. As shown below, the answer is positive. A natural question is how the oculomotor responses to the KDE compare to those recorded under more "natural" viewing conditions. To approach this issue, two additional experimental conditions were devised. In the rotating LED condition, subjects *binocularly* tracked the three-dimensional trajectory of a light-emitting diode (LED) which moved on a horizontal circular path in space. This stimulus contained binocular disparity, retinal image blur and changing-size (looming) information, but no relative motion cues. In the motion and stereo condition, subjects viewed a dynamic stereogram simulating a rotating random dot sphere; a stimulus that combined *both* relative image motion and binocular disparity information.

Methods

Stimuli and Procedure. The motion stimulus resembled the displays used in psychophysical studies of KDE (Green, 1961; Braunstein, 1962). Subjects monocularly viewed a computer-generated movie of translating dots obtained by the orthographic projection of points distributed on the surface of a sphere rotating about its vertical axis (Fig. 1). Uniformly distributed points ($N = 960$) were selected on the surface of a sphere, $x^2 + y^2 + z^2 = R^2$. The radius of the sphere was $R = 9.5$ cm and its angular velocity $\Omega_y = 30$ deg/sec, unless otherwise stated. The points were orthographically projected onto the plane P (the computer screen) which was parallel to the (x,y)-plane. The points q and s on the sphere, for example, project onto q' and s', respectively. The projection of the center of the sphere O onto the plane P is denoted by O'. The motion of the points on the projection plane had only a horizontal component. The simulated sphere was transparent; i.e., points on both the front and back surfaces were visible. The viewing direction corresponded to the positive z-axis which was centered between the observer's eyes. The viewing distance, $\overline{EO'}$, was 93 cm. Accommodative distance was approximately constant across the display, as the viewing distance to the sphere's boundary was only 93.48 cm. If one assumes a depth-of-focus of ±0.25 D (Davson, 1950; Campbell, 1957; Adamson & Finchman, 1939), it can be verified that a real sphere having a radius of $R = 9.5$ cm would be entirely in focus. Thus, blur is *not* in conflict with relative motion information. Interocular distance was 6.2 cm for subject MH, 6.0 cm for subject FM, and 6.8 cm for MY. Subjects MH and FM were aware of the goal of the experiment, and MY was naive. The screen background was dark and the surrounding room was dimly illuminated. A Silicon Graphics Elan R4000 computer generated the stimuli in real time. A Barco CCD 7651B monitor was used to display the stimulus. The screen measured 34.3 cm wide by 27.4 cm high. The radius of each white point (one pixel) on the screen subtended 30 sec arc. The mean luminance of the display, at the center of the screen, was 0.2 cd/m^2. The refresh rate of the monitor, and the frequency at which new image frames were presented, was 60 Hz.

The generated KDE stimulus contained only relative image motion information and induced a strong perception of a three-dimensional rotating sphere. Eye movements were restricted to the horizontal plane by instructing the subjects to select, and smoothly pursue, a surface patch lying on the sphere's equator. In most of the cases a red dot indicated a particular surface patch to be pursued. The directions of gaze of both eyes were measured simultaneously with a two-channel eyetracker. Both the opened and occluded eyes moved during the execution of the experiment.

Although the percept induced by such a KDE stimulus is usually bi-stable (the sphere can be perceived as rotating either clockwise or counterclockwise), *all* subjects in this study reported seeing a single direction of rotation during the execution of the task (counterclockwise when viewed from the top). In addition, they always started by looking at the surface patch perceived at the center of the *front* surface. We do not know the reason for the absence of reversals in the KDE condition. In separate experiments, we used dots with infinite and finite lifetimes, to rule out the possibility that subjects were tracking individual dots. In the finite lifetime condition, dots disappeared after 266 msec and were repositioned at random on the surface. In some experiments we included a red dot (3.7 min arc radius) on the sphere's equator, signaling a particular surface patch to be pursued. No significant differences were found between these conditions. The data presented here are collapsed across the above conditions.

In the rotating LED condition, subjects binocularly tracked the three-dimensional trajectory of a red LED which moved on a horizontal circular path in space. The minimal angular displacement of a 0.9 deg/step stepping motor (Alpha Products, Fairfield, CT, U.S.A.) was reduced 7.7 times to obtain an effective step size of 7 min arc. A black vertical pole, with an LED mounted at the top, was moved by the motor under computer control. Only part of the pole and the LED were visible during the experiments; the apparatus was occluded by black cardboard. The viewing distance to the LED's center of rotation was 93 cm. The distance from the LED to the center of rotation was $R = 9.5$ cm and its angular velocity was $\Omega_y = 30$ deg/sec. These parameters are same as in the KDE condition. The mean radius subtended by the LED during a revolution was 7.4 min arc. The room was dimly illuminated.

In the motion and stereo condition, a Tektronix SGS610 polarization modulator (average light transmission = 12%) was mounted in front of the computer screen to deliver different images to the two eyes at a frequency of 120 Hz (60 Hz for each eye). A pair of left and right circular polarizer filters was placed at approximately 2.5 cm in front of each eye and did not interfere with the optics of the eyetracker. A dynamic stereogram simulating a rotating random dot sphere was generated in real time by the computer. The angular difference between the viewing directions of the left and right eyes was equal to the vergence angle required by the subject to fixate at the center of the screen (under the assumption of perfect fixation).

Finally, in a fourth, control condition, observers were instructed to follow the motion of a single red dot on the computer screen. The presentation was monocular. The horizontal position of the dot was a sinusoidal function of time, such that the resulting motion was equivalent to that of the projection of a point on the sphere's equator in the KDE condition. Motion in depth was not perceived in this case. This control condition allowed us to estimate variation in vergence of the eyes due to changes in smooth pursuit eye velocity, to asymmetries between

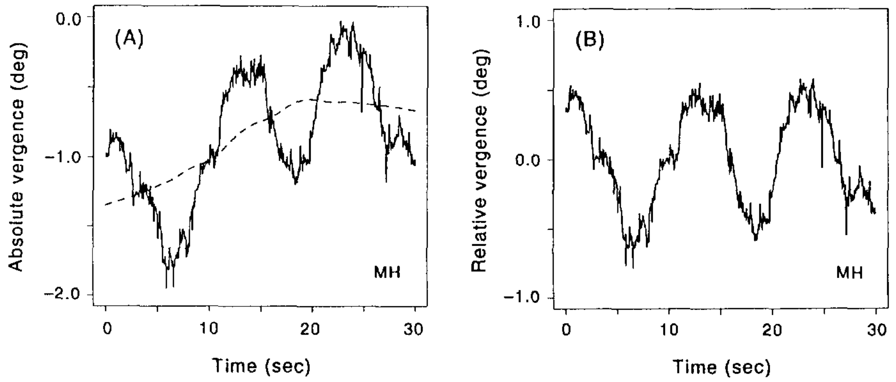

FIGURE 2. Raw vergence responses in the KDE condition. (A) An extreme example of a raw vergence response "riding" on top of a slow drift. The dashed curve shows the estimated drift. (B) The signal obtained after correcting for the drift by subtracting the trend from the raw vergence signal.

rightward and leftward pursuit (or between the occluded and unoccluded eyes) in a particular subject and to intrinsic noise of vergence eye movements.

Data Acquisition. A two-channel dual Purkinje image eyetracker (Crane & Steele, 1978) (Fourward Technologies, San Diego, CA, U.S.A.) was used to measure simultaneously the vertical and horizontal rotational components of gaze direction for both eyes. In the KDE and control conditions, monocular viewing was imposed using a black occluder positioned approximately 2.5 cm in front of one of the eyes. The occluder did not interfere with the optics of the apparatus. The eyetracker provided an accuracy of 1 min of arc rms. An infrared optometer, based on the Scheiner principle (Randle, 1970), was used to measure the accommodation of the left (unoccluded) eye with an accuracy of 0.1 D. A bite bar was used to minimize head movements. The signals were low-pass filtered at 40 Hz, digitally sampled at 100 Hz (12 bits A/D), and stored on disk for further analysis. The instrument was calibrated by having subjects fixate binocularly at targets on the horizontal and vertical axes separated at 1.7 deg intervals. We used least-squares regression to fit a line to the calibration data to estimate the gain and DC offset of each channel. No trends were obvious in the residuals, meaning that the system was linear within the range of the measurements. Calibrations were performed at the beginning and at the end of an experiment to have an estimate of the variability in the gains of each channel. For sessions approximately 20 min long, we found that the typical change in gains would be ≈ 2–4%.

Data Analysis. The horizontal components of gaze direction as a function of time can be described by $l(t)$ and $r(t)$; the horizontal directions of gaze of the left and right eyes, respectively. Our coordinates were defined so that the components $l = r = 0$ when the observer looked at infinity, and $l,r > 0$ for rightward eye rotations. From $l(t)$ and $r(t)$ we computed the version component (or mean gaze direction) $\mu(t) = (l(t) + r(t))/2$, and the vergence component $v(t) = l(t) - r(t)$ as a function of time (Hering, 1868).

In the KDE and control conditions, we frequently observed that the mean position of the occluded eye drifted slowly with time (Enright, 1987b; Erkelens & Regan, 1986). This caused a corresponding drift in the vergence signal. Figure 2(A) shows an extreme example of a raw vergence record in the KDE condition for which the drift was particularly large. A vergence response at the stimulus frequency is seen "riding" on top of a slow drift. The slow drift was estimated, and factored out, by fitting a linear local regression model to the raw vergence record (Cleveland *et al.*, 1992). The method is similar to computing a running window average of the data using a window width equal to the period of the stimulus. The dashed curve in Fig. 2(A) shows the estimated drift, which is subtracted from the original data to obtain the corrected vergence signal shown in Fig. 2(B). The corrected signal, denoted by $\Delta v(t)$, represents *relative changes* in vergence. The same analysis procedure was applied in all the experimental conditions.

We visualized the data in the plane $(\mu(t), \Delta v(t))$. Smooth eye movements describe a continuous curve (parametrized by t) $\Gamma(t) = (\mu(t), \Delta v(t))$ in the $(\mu, \Delta v)$-plane. Saccades are represented by a discontinuity (jump) in $\Gamma(t)$. These gaps were automatically generated by thresholding eye velocities at 15 deg/sec and traces clipped from 50 msec before to 50 msec after the occurrence of a saccade. Sometimes we observed that the two eyes executed saccades of different amplitudes. These events were considered legitimate contributions to the vergence signal (Enright, 1984). Therefore, *no* corrections were made to the raw vergence signal across saccadic gaps. We call the set of points $\{\Gamma(t)\}_{t=0}^{T}$ the *trace* of the eyes during the period of stimulation (from 0 to T sec). A trial consisted of $T = 30$ sec of stimulation, during which the sphere completed two-and-a-half revolutions about the center of rotation. The traces corresponding to the last two revolutions (from $t = 6$ to $t = 18$ sec and from $t = 18$ to $t = 30$ sec) were analyzed independently and constituted two observations in the data set. We assume that after the initial 6 sec the subject reached a "steady-state" representation of the surface.

A theoretical prediction for the eye movement trace was computed based on the viewing geometry in Fig. 1 as follows. We assumed perfect binocular fixation and that blur information, together with previous knowledge

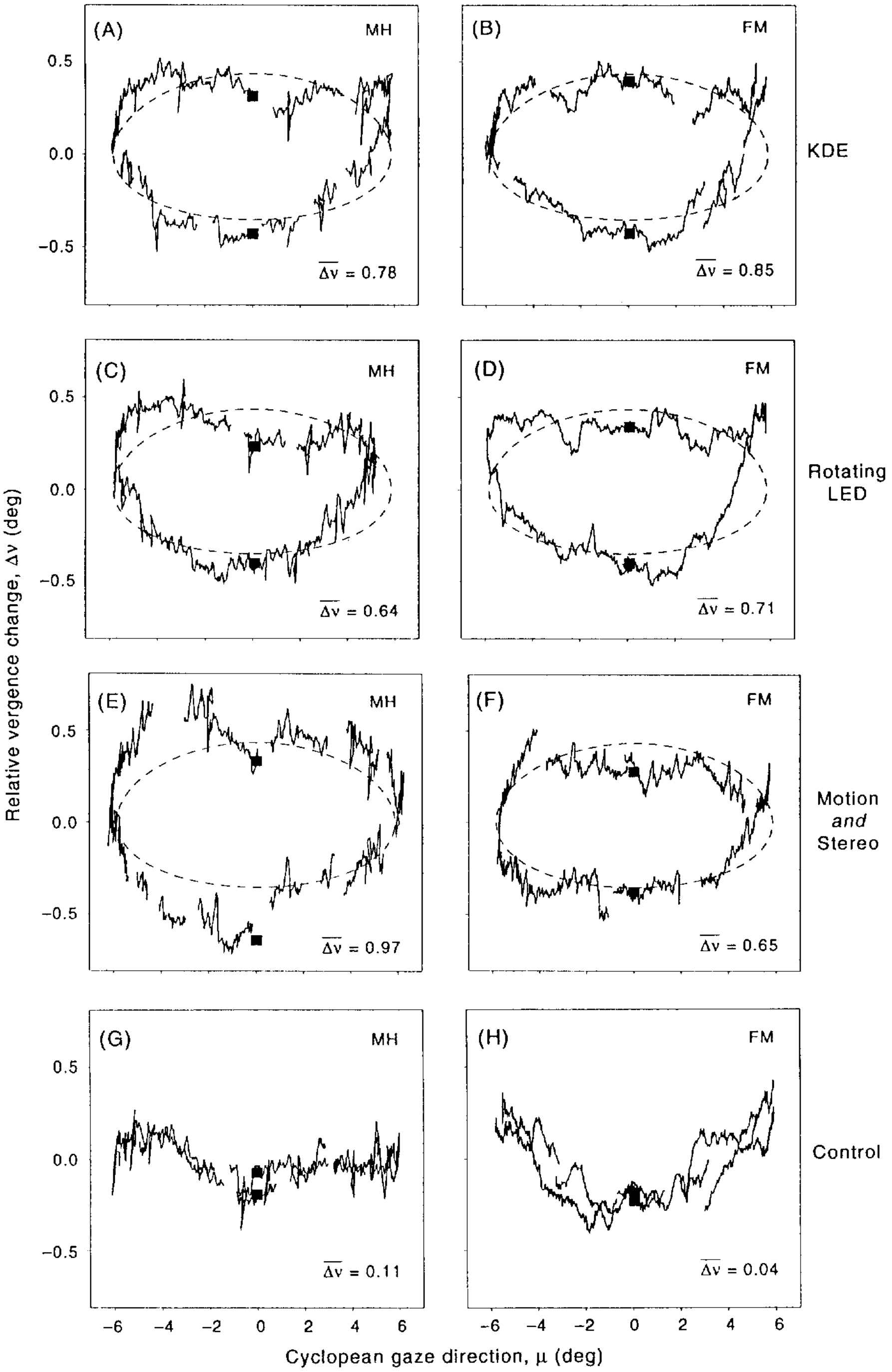

FIGURE 3. Basic findings. (A) and (B) Typical traces obtained when subjects monocularly viewed a rotating sphere produced by the KDE stimulus. The dashed curves show the shape of the traces expected from the viewing geometry illustrated in Fig. 1. Positive values of Δv represent converging eye movements. The solid squares represent the median values of the vergence component within the central 2 deg for the rightward ($d\mu(t)/dt > 0$) and leftward ($d\mu(t)/dt < 0$) segments of the trace. The magnitude of the estimated modulation in vergence, $\overline{\Delta v}$, defined as the absolute difference between the medians, is written as an inset. (C) and (D) Traces obtained under binocular smooth pursuit of a LED moving in a horizontal circular trajectory in space. (E) and (F) Traces measured in the motion and stereo condition, in which subjects viewed a dynamic stereogram simulating a random dot sphere rotating in space. (G) and (H) Traces obtained in the control condition, when the subjects followed the motion of a single red dot on the screen. The dot's motion was equivalent to that of the projection of a point on the sphere's equator in the KDE condition.

about the spatial layout of the experimental setup, provided the subjects with an estimated distance to the center of the perceived sphere equal to the distance between the observer and the screen. For each point on the sphere's equator we computed the length of the sides of the triangle formed by the eyes and the target as its vertices. We then solved for the vergence angle.

Results

The perception of three-dimensional shape caused a

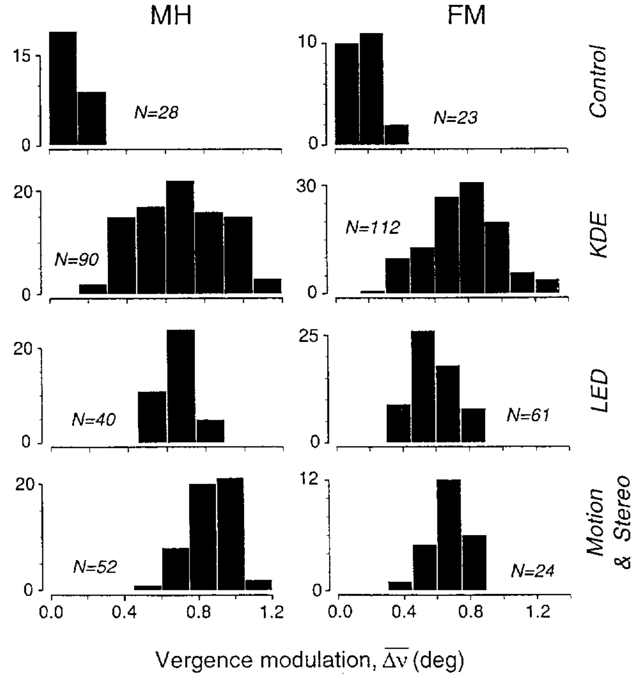

FIGURE 4. Distribution of vergence modulation values. The histograms show the distribution of $\overline{\Delta v}$ for each experimental condition.

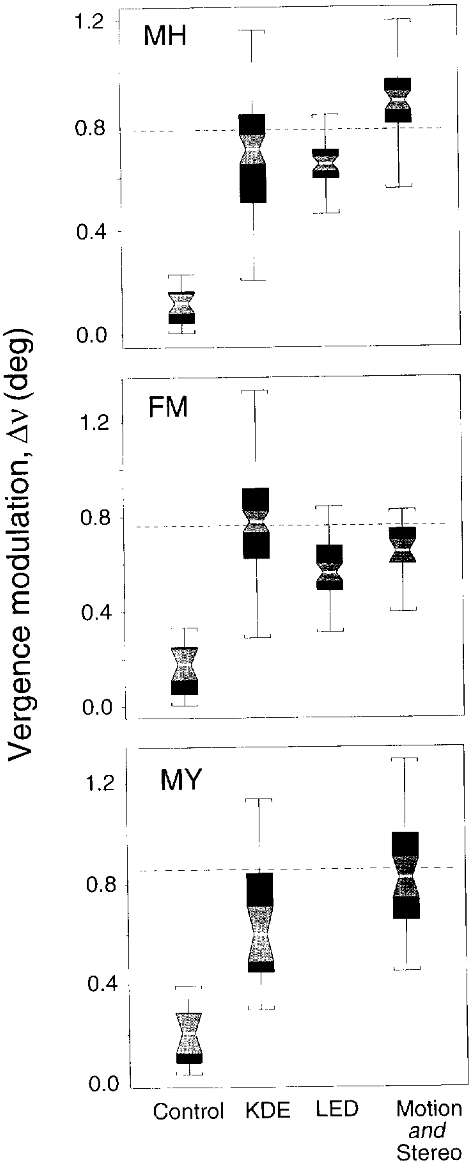

FIGURE 5. Summary of the data. The data obtained for three subjects are summarized by box plots (Hoaglin *et al.*, 1983). White lines indicate the location of the medians. Solid regions above and below the medians represent the upper and lower quartiles, respectively. Notched regions, shaded in gray, indicate 95% confidence intervals for the location of the medians. The "whiskers" of each box extend to the nearest value not beyond the standard range from the quartiles. The standard range used was 1.5 × the inter-quartile range. Points beyond these limits were considered outliers. The dashed horizontal lines represent the expected modulation in vergence calculated from the viewing geometry. The median of the responses in the KDE, rotating LED, and motion and stereo conditions were significantly larger than for the control condition.

modulation in vergence eye movements. Basic findings are illustrated in Fig. 3. Representative traces of one revolution, obtained in the KDE condition, are depicted in Fig. 3(A) and (B). The superimposed dashed curves show the shape of the traces expected from the viewing geometry illustrated in Fig. 1. In these graphs, *relative changes* in the vergence component, Δv, are plotted as a function of cyclopean gaze direction, μ. Positive values of Δv represent converging eye movements. The traces in Fig. 3(A)–(H) are arbitrarily displaced along the Δv axis.

From the eye movement records shown in Fig. 3(A) and (B) it is clear that there are two different values of the vergence angle v on the trace for each mean gaze direction μ. The maximum difference between these v values occurs at the center of the sphere, where $\mu = 0$. This means that the observer was fixating at different depth planes, by converging and diverging the eyes, when looking at the front and the back surface of the perceived sphere, respectively. We were able to determine this correlation because subjects always started to pursue the central surface patch perceived at the front surface of the sphere (see Methods section). The subjects who participated in the study never reported reversals in the direction of rotation of the sphere during the experiments.

Typical traces measured when subjects binocularly viewed the LED rotating in space are shown in Fig. 3(C) and (D). There were no obvious differences between the shape of the traces obtained in the KDE and the rotating LED conditions. Similar records were also obtained in the motion and stereo condition, as shown in Fig. 3(E), (F). In the control condition, however, eye movement traces were very different from all conditions in which depth was perceived. There was only a single vergence angle (within noise levels) for each mean gaze direction, as shown in Fig. 3(G) and (H).

The amplitude of the modulation in vergence was estimated from the trace, $\Gamma(t)$, as follows. We segmented the trace into two components, one corresponding to rightward gaze movement ($d\mu(t)/dt > 0$) and the other to leftward gaze movement ($d\mu(t)/dt < 0$). For each of these segments, we selected those points lying within the central two degrees of visual angle, -1 deg $< \mu(t) < 1$ deg. The medians of the vergence angles were computed

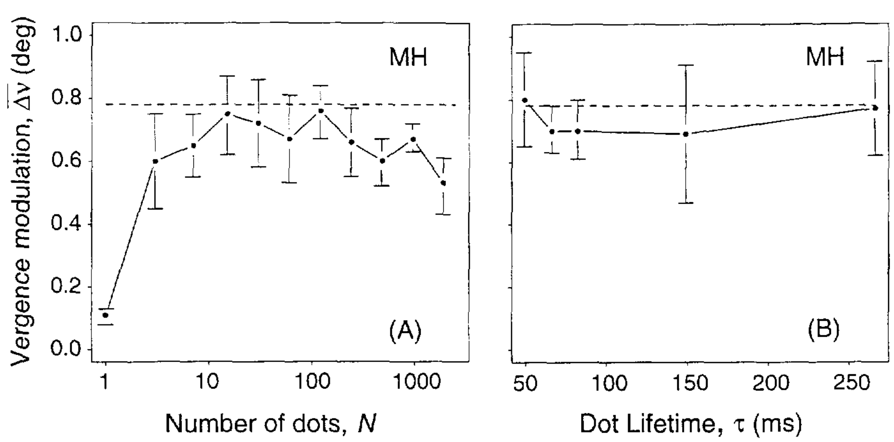

FIGURE 6. (A) Dependence of vergence modulation on the number of dots, N, distributed on the surface of the sphere. The responses initially increase to saturate at $N \approx 15$. Also, there seems to be a small tendency of the responses to diminish for large N. (B) Dependence of vergence modulation on dot lifetime, τ. The modulation in vergence is approximately constant for 50 msec $< \tau <$ 266 msec. The bars represent 95% confidence intervals for the location of the medians.

for the points within each of the groups and are marked by solid squares in the traces depicted in Fig. 3(A)–(H). The estimated modulation in vergence, denoted by $\overline{\Delta v}$, is defined as the absolute difference between the medians. The measure is centered around $\mu = 0$ because this is where we expect the signal to be maximal.

The distributions of the vergence modulations $\overline{\Delta v}$ for the four different experimental conditions, shown in Fig. 4 for two subjects, illustrate the large effect of monocular KDE stimulation on vergence modulation. The median of the responses in the KDE, rotating LED and motion and stereo conditions were close to each other and similar to the predictions based on the viewing geometry. In comparison, the modulation of vergence in the control condition was significantly smaller (Wilcoxon rank-sum test, $P < 0.001$ in all tests). The response variance in the KDE condition was larger than those in the rotating LED and the motion and stereo conditions (F-test, $P < 0.001$ in all tests). The origin of this difference is unknown.

A summary of the data, for all three subjects, is provided by the box plots (Hoaglin *et al.*, 1983) shown in Fig. 5. The modulation in vergence expected from the viewing geometry (=0.79 deg for MH, 0.76 deg for FM and 0.86 deg for MY) is shown by the dashed lines. For subject MH, the median response in the motion and stereo condition was significantly larger than in the KDE and LED conditions (Wilcoxon rank-sum test, $Z = -5.5662$, $P < 0.001$), while for FM the response was intermediate. To corroborate the initial results, an additional subject, MY, ran only a subset of the experiments with a smaller number of trials (≈ 35) per condition (this is the reason for the larger confidence intervals). Subject MY's responses to the KDE were significantly larger than in the control condition, but smaller than those in the motion and stereo condition.

We found no detectable accommodation changes within the precision of our measurements (±0.1 D) in any of the experimental conditions. This is not very surprising. For a physical sphere of the size simulated in our experiments, shifting fixation between the front and the back surfaces of the sphere requires only a change of 0.2 D in accommodation. As all the dots are on the plane of the screen, no blur information is available to drive accommodation. In addition, the expected modulation in accommodation caused by the CA/C crosslink—the convergence-linked accommodation—is of the order of 0.12 D, if we assume a typical CA/C gain of 0.1 D/Δ (Finchman & Walton, 1957).

The Effect of Dot Numerosity and Lifetime

We have done preliminary experiments to gain some insight into the importance of (a) the number of dots on the surface of the sphere; and (b) the lifetime of the dots, on the magnitude of vergence modulation. Classic studies of the KDE have used subjective ratings of "rigidity", "coherence" and "depth" in multidot displays to explore these questions. We investigated whether the magnitude of vergence modulation behaves in a similar way to the subjective judgments used in the past.

First, the KDE condition was repeated while varying the number of dots present on the sphere. All the dots had infinite lifetimes. A red fixation dot always appeared on the equator. For each experimental trial all the other dots were uniformly distributed on the surface of the sphere to generate a new stimulus. The graph in Fig. 6(A) illustrates the modulation in vergence as a function of the number of dots in the display. The bars indicate 95% confidence intervals. It can be seen that the responses saturate at $N \approx 15$ points. In addition, there seems to be a small decrease at high densities.

To study the effect of the dot lifetimes we fixed the number of dots at $N = 960$ and each dot was presented for τ msec on the display. After the elapsed time the dot was repositioned on the surface of the sphere at random. The dots were asynchronously updated. A fixation point was not used in this case. Instead, the subject was instructed to select, and to smoothly pursue, a "surface patch" on the equator of the sphere. The plot in Fig. 6(B) shows that vergence modulation was essentially constant for $\tau > 50$ msec (three frames). For $\tau = 33$ msec (two frames), the display looked like flicker noise and the subject was not able to do the task.

Discussion

Traditionally, binocular disparity and retinal image blur are considered the primary physical stimuli to the vergence and accommodation eye movement systems (Maddox, 1893; Phillips & Stark, 1977; Stark *et al.*, 1980; Judge, 1991). However, our results demonstrate that under monocular viewing conditions binocular eye movements still occur, and that they reveal aspects of the perceived three-dimensional structure of the environment. This opens the possibility of using oculomotor responses to study three-dimensional visual perception in human subjects and non-human primates. An experiment similar to the one described above could be performed on monkeys to determine if they perceive a three-dimensional shape when presented with a KDE stimulus.

The findings indicate that the recovery of three-dimensional shape, based solely on monocular relative motion information, is able to drive vergence eye movements with a "gain" similar to that produced by disparity, blur and looming combined, and to the "gain" obtained when relative disparity and motion are combined. The comparable responses produced by the KDE stimulus, the LED condition and the stereoscopic viewing of a rotating three-dimensional sphere support the hypothesis that three-dimensional structure derived from two-dimensional motion and from binocular disparity are similarly represented in the visual system (Nawrot & Blake, 1993). It is possible that there is a single three-dimensional representation of surfaces where depth information from multiple visual cues is combined (Marr, 1982) and that the estimated three-dimensional structure drives vergence eye movements. This hypothesis is consistent with our results and previous work (Enright, 1987a,b; Erkelens & Regan, 1986; McLin *et al.*, 1988), which suggest that different visual cues, through their influence on perceived three-dimensional shape and motion in depth, may serve as effective stimuli to vergence eye movements. It would be of interest to determine if vergence responses can also be elicited using other monocular visual cues to three-dimensional shape, such as shading and texture gradients (Gibson, 1950).

The dependence of vergence modulation as a function of numerosity, N, and lifetime, τ, of dots in the stimulus is qualitatively similar to the functions obtained in previous psychophysical studies of KDE using subjective ratings. Braunstein (1962) found that depth judgments increased linearly with the logarithm of the number of points in the range $2 < N < 6$. Similar results have been obtained by Green (1959). After this initial increase, depth judgments saturate for $16 < N < 32$ (Green, 1961; Dosher *et al.*, 1989b). The approximately constant value of Δv in the range $15 < N < 120$ (exact Wilcoxon signed-rank test, alternative hypothesis $\mu = 0.73$ deg, $P > 0.2$) indicates that texture density cues were not a significant factor in determining the magnitude of the responses in the KDE condition. Salient SFM is perceived in displays with short dot lifetimes (Dosher *et al.*, 1989a; Husain *et al.*, 1989; Treue *et al.*, 1991). Consistent with this observation, it has been found that performance in a shape classification task (Sperling *et al.*, 1989) saturates at $\tau \approx 50$ msec. The modulation in vergence is approximately constant for $50 \text{ msec} < \tau < 266$ msec (exact Wilcoxon signed-rank test, alternative hypothesis $\mu = 0.73$ deg, $P > 0.12$). The similarities between the behavior of $\overline{\Delta v}$ and various subjective judgments as a function of dot numerosity and lifetime provides additional support to the idea that vergence modulation is related to perceived depth.

Experiment 2. Dynamic Properties of the Responses

An alternative way to explore a hypothetical correlation between the "perceived depth" evoked by the KDE and the magnitude of vergence modulation is by studying the dynamics of the phenomenon. On the perceptual side one can measure how "perceived depth" depends on the angular speed of the KDE sphere. Similarly, on the motor side, it is possible to measure how the modulation in vergence depends on angular speed. Once these data are available it should be possible to determine if there is a correlation between perceived depth in the KDE and the oculomotor responses. In this experiment, we provide data about the dynamics of oculomotor responses to the KDE.

The dynamical properties of the responses in the different experimental conditions that evoke motion in depth may also shed light on the underlying mechanisms. For example, if the responses to the KDE and the rotating LED were to have dramatically different temporal frequency tuning curves, one could argue safely that disparity information is processed in a different way from relative motion information in the generation of oculomotor responses.

The aim of this experiment was to measure the dynamics of eye movements in the KDE condition and compare it to those obtained in the motion and stereo and rotating LED conditions. This was done by computing, in each case, the amplitude and phase of the first harmonic component for both the version and vergence signals, as the angular speed of the simulated sphere was varied. Except for the difference in angular velocity, the experimental conditions were identical to those described in Experiment 1. At all rotation speeds, a trial consisted of 30 sec of stimulation.

Results

Segments of typical raw vergence records obtained at various temporal frequencies of stimulation are depicted in Fig. 7. A similar dependence of responses is observed in all three experimental conditions. A significant oscillation of the vergence signal at the driving frequency is observed up to frequencies of 1 Hz. In the KDE condition, the responses are seen "riding" on top of a slow drift of the signal, as previously described. The drift

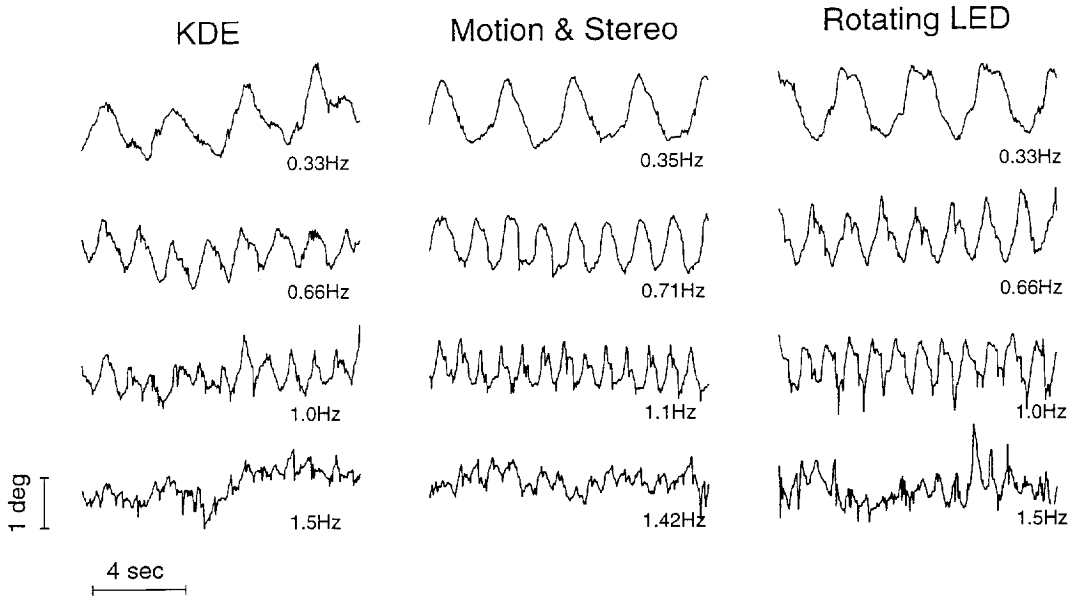

FIGURE 7. Examples of raw vergence responses at different temporal frequencies and experimental conditions (Subject FM). Qualitatively similar responses are observed in all three experimental conditions. The responses cutoff at approximately 1 Hz.

is absent in the motion and stereo and rotating LED condition, in which both eyes are open.

Bode plots, showing the dependence of the first harmonic amplitudes and phases of the version and vergence components in the KDE and motion and stereo conditions for two subjects, are presented in Fig. 8. Open squares represent the responses in the KDE condition and solid triangles represent those in the motion and stereo condition. The results are qualitatively similar in both conditions. The responses are low-pass with a high frequency cut-off of ≈ 1 Hz. The phase of the vergence component changes more rapidly than the phase of the version component. The rate of change of the phases at high frequencies can be accounted for by a pure delay of ≈ 90 msec in the version component and of 215 msec in the vergence component for subject MH, and of ≈ 80 msec in version and ≈ 130 msec in vergence for subject FM. Similar results (not shown) were also obtained in the rotating LED condition for both subjects.

Discussion

We were unable to dissociate the mechanisms generating the oculomotor responses in the KDE, stereo and motion and rotating LED conditions. Instead, we found the dynamics of version and vergence to be similar in all experimental conditions. There are at least two non-exclusive possibilities that may explain this result. First, it could be that the system is limited on the perceptual front-end: subjects reported that all types of stimuli were perceptually weakened for frequencies above 1 Hz. This suggests a possible correspondence between perceived depth and the amplitude of vergence responses. Additional data on the perceived extent of motion in depth under similar experimental conditions would be necessary to establish this conclusion firmly.* Another possibility is that the motor system is limiting the performance. Smooth pursuit of an oscillating target in the frontoparallel plane is known to cut-off at about 1 Hz (Stark, Vossius & Young, 1962; Dallos & Jones, 1963). Similarly, vergence responses to small targets oscillating in depth fall off at about 1 Hz (Jones & Kerr, 1972; Rashbass & Westheimer, 1961; Riggs & Niehl, 1960; Westheimer & Mitchell, 1969). These earlier studies investigated pursuit and vergence in isolation. Our results show that similar characteristics are observed during binocular tracking of the periodic motion of a target in three-dimensional space, when version and vergence responses occur simultaneously.

Experiment 3. Voluntary Control and Perceptual Bias

Subjects are known to have voluntary control of their eye movements (Westheimer, 1957; Finchman, 1962; Malmstrom & Randle, 1976; McLin & Schor, 1988). We investigated whether the vergence eye movements in response to the KDE stimulus could have been generated by voluntary eye movements.

As shown in Fig. 5, similar results were obtained in Experiment 1 for the naive (MY) subject and those aware of the goal of the experiment (FM and MH). In addition, subjects failed to replicate the eye movement traces obtained in the KDE condition without visual stimulation (in the dark). To assess further the extent of voluntary eye movements, we instructed subjects to bias (either increase or decrease) their responses intentionally when the visual stimulus was present. These results, described

---

*Some temporal properties of the KDE have been studied in the past (Hildreth *et al.*, 1990; Loomis & Eby, 1988; Todd & Bressan, 1990; Todd *et al.*, 1988; Eby, 1992). However, the methods and type of stimuli used differ too much from those in our study, precluding a direct comparison of the results.

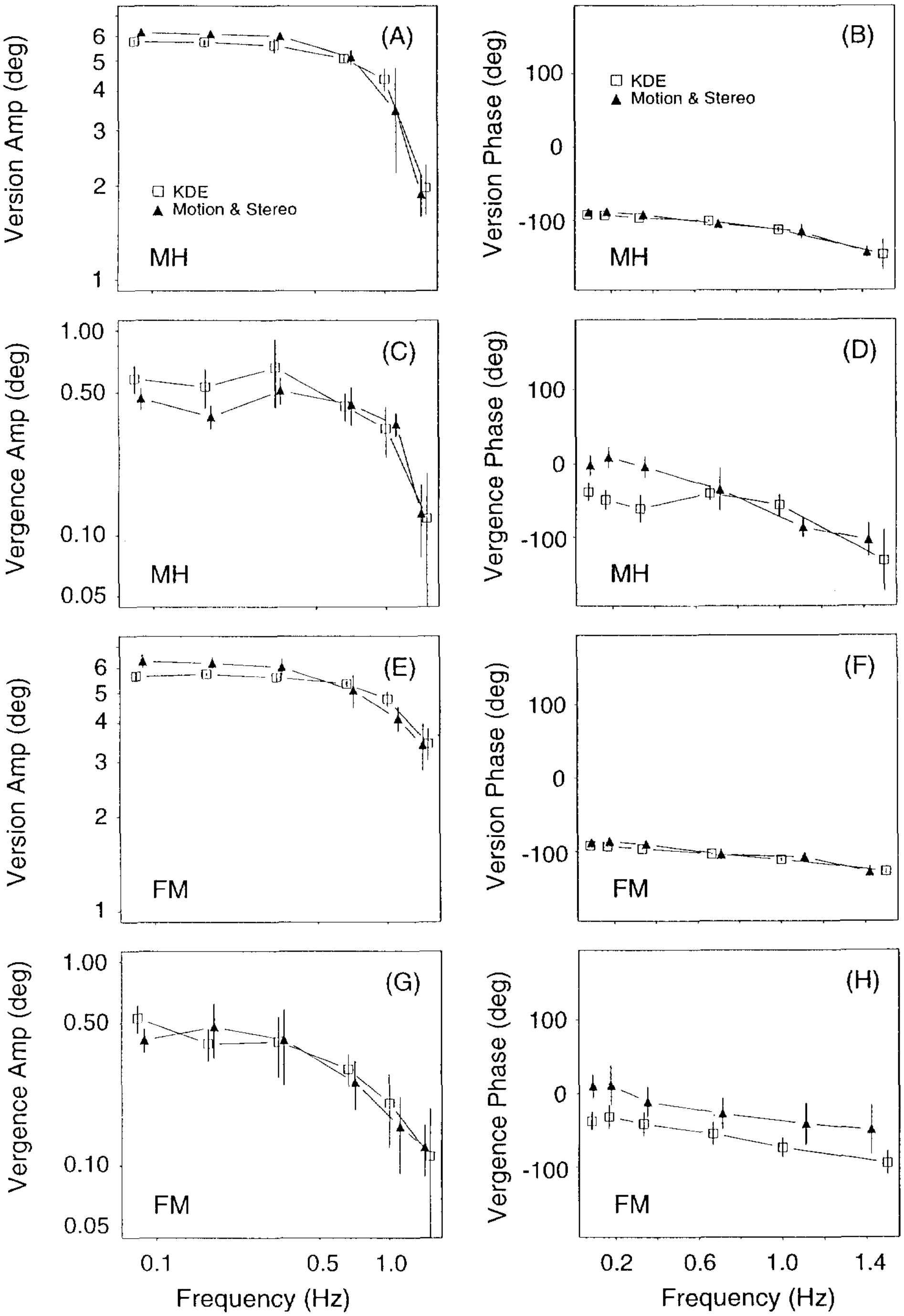

FIGURE 8. Bode plot summary of the dynamic responses. The graphs show the first harmonic amplitudes and phases of the version and vergence components in the KDE (open squares) and motion and stereo (solid triangles) conditions. (A)–(D) Subject MH; (E)–(H) subject FM.

next, also indicate that voluntary eye movements did not play a major role in the generation of the responses.

In one experiment, we used the same stimulus as in the control condition—a single dot oscillating back and forth. This time, however, we asked the subjects to *imagine that the red fixation dot was actually moving in depth*, describing a horizontal circular path in space. We will refer to this experiment as the three-dimensional bias condition. The idea is to try to force the subject to generate a large modulation in vergence, even though the KDE stimulus is very weak. In this condition, subject FM showed a very small increase in the responses, relative to the original control experiment (0.3 ± 0.2 deg). In contrast, MH's responses were large and statistically indistinguishable from the KDE condition (0.8 ± 0.3 deg).

In an additional experiment, we used the same visual stimulation as in the KDE condition (described in Experiment 1), but instructed the subjects to follow the motion of the fixation dot on the *plane of the screen, ignoring any perceived motion in depth*. We will refer to this condition as the two-dimensional bias condition.

Here, we attempt to determine if subjects can abolish the modulation in vergence voluntarily when they view a stimulus that induces a strong three-dimensional percept. In this case, MH showed similar responses to the original experiment 0.69 ± 0.30 deg; FM's responses were bimodally distributed, with modes at 0.3 and 0.75 deg. Subject FM's binocular eye traces in the two-dimensional bias condition, however, were more discontinuous than in the KDE condition; i.e., they showed an increased frequency of saccades. In spite of FM's frequent use of saccades in the two-dimensional bias condition many traces still exhibited large modulations in vergence, represented by the mode at $\overline{\Delta v} = 0.75$ deg.

These results can be summarized as follows. When the motion cues were weak, but still consistent with a simple three-dimensional interpretation, as in the case of a single oscillating fixation dot, only MH showed the capability to respond with large and smooth modulations in vergence, while FM showed only a small increase. However, when the motion stimuli gave a strong indication of a three-dimensional shape, neither subject was able to abolish their vergence responses without an increase in the frequency of saccades.

A possible explanation for these results is that the effect of the different instructions was simply to bias the subject's interpretation of the display as either two-dimensional or three-dimensional when the stimulation was weak. In fact, MH reported that he could "see" the single oscillating dot as rotating in depth in the three-dimensional bias condition, in accord with the large magnitude of his responses. Johansson (1958) has reported also that a sinusoidally oscillating dot is sometimes perceived as a circular motion in depth. In the two-dimensional bias condition, in which a strong KDE stimulus is present, subjects were not able to perceive the stimulus as two-dimensional, and their responses were similar to those for the KDE condition unless they switched to other strategies (as evidenced by FM's increasing frequency of saccades).

In order to test the idea that perceptual bias was responsible for the changes in the responses—and *not* a direct voluntary control of eye movements—we repeated the experiments using a KDE stimulus whose associated control condition could not be easily perceived as two- or three-dimensional simultaneously. This was done by imposing a more complicated motion on the sphere. For each time, t, we generated the image as the orthographic projection of the random dot sphere which was first rotated about the vertical y-axis by $\Omega_y t$ deg and then rotated about the x-axis by $10\sin(5\Omega_y t)$ deg. In the resulting display the sphere appears as rotating about the vertical axis while "rocking" back and forth at the same time. The percept of a three-dimensional sphere is still very strong. In fact, subjects reported that the perceived three-dimensional structure of the sphere was enhanced by the "rocking" motion, compared to constant rotation about the vertical axis. In the associated control condition, we presented a single fixation dot on the sphere's equator. The motion of the fixation dot no longer described a simple sinusoidal movement on the screen. Instead, it followed a wiggly path which is difficult to interpret as a simple motion in three-dimensional space. The plot in Fig. 9(A) illustrates the two-dimensional trajectory described by the fixation dot in this case. The dot oscillates back and forth on the same curve.

Only MH ran this experiment, as he was the one who produced large responses in the three-dimensional bias condition. Figure 9(B) shows the results for all four

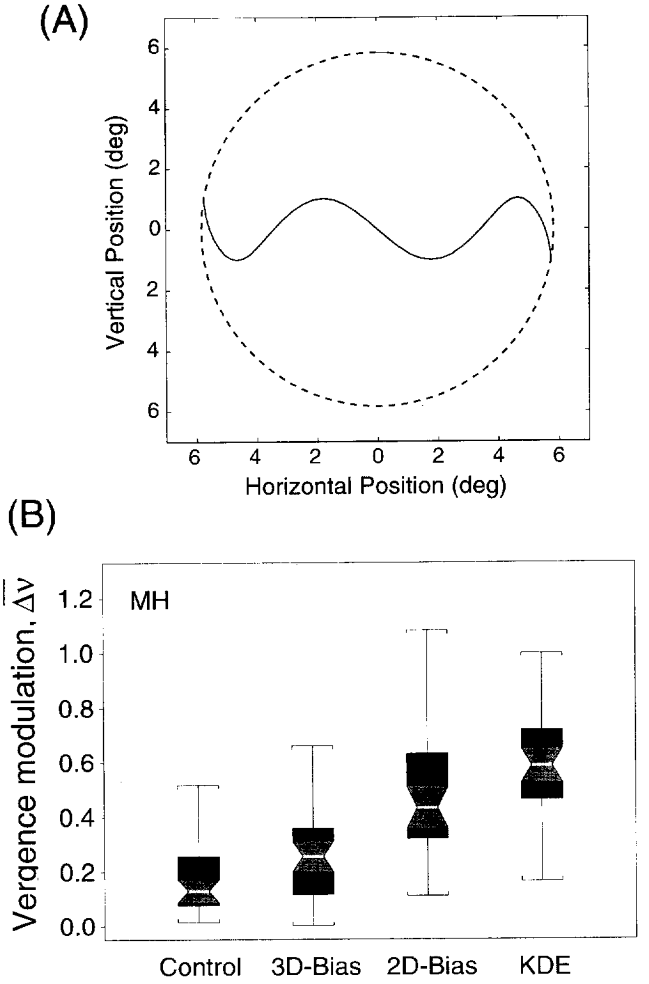

FIGURE 9. (A) The projected motion of a point on the equator of the sphere for a complex motion in three dimensions. The fixation point oscillates back and forth on the same curve. The dashed line represents the occluding boundary of the sphere. (B) Box plot summary of vergence modulation when a complex motion of the sphere was used. In the control and three-dimensional bias conditions subjects viewed a single red dot oscillating back and forth on the curved trajectory shown in (A). In the control condition, subjects were asked to follow the motion of the dot on the plane of the screen; in the three-dimensional bias condition, subjects were instructed to imagine the dot was moving in depth. In the two-dimensional bias and KDE conditions, $N = 960$ dots were distributed on the surface, and subjects perceived a rotating KDE sphere "rocking" back and forth. In the two-dimensional bias condition, subjects were instructed to ignore any perceived motion in depth and pursue the fixation dot on the plane of the screen; in the KDE condition, subjects were asked to follow the fixation dot on its perceived three-dimensional trajectory. There was a significant difference between all the conditions. The magnitude of the responses can be listed in ascending order: control < three-dimensional bias < two-dimensional bias < KDE.

conditions using the new stimulus display. There is a slight decrease in the two-dimensional bias condition relative to the KDE condition. There is also a small increase in the three-dimensional bias condition relative to the control condition. However, the pooled responses of the three-dimensional bias and control conditions were significantly smaller than those for the two-dimensional bias and KDE conditions. In particular, the responses in the three-dimensional bias condition were now much reduced compared to the case in which a simple rotational motion of the sphere was used. These results seem to indicate that the modulations of the responses in the previous experiment were a consequence of perceptual biases induced by the instructions, and that direct voluntary control played a minor role.

McLin & Schor (1988) showed that voluntary changes in accommodation and vergence occur in a ratio that is similar to the AC/A ratio and different from the CA/C ratio, suggesting that voluntary effort drives accommodation primarily, while vergence is a secondary response through the AC/A cross-link. As already mentioned, we were not able to measure any accommodation changes in any of our experiments (within ±0.1 D). This supports the notion that our responses were not generated by voluntary control. In addition, the magnitude of vergence changes generated under voluntary control are large (6–14 Δ D) (McLin *et al.*, 1988). It is not clear if subjects have the finer volitional control required to generate the smaller vergence modulations we observed. Furthermore, it is unlikely that voluntary responses would have a mean amplitude so close to the one predicted by the viewing geometry, as found in Experiment 1. It is interesting to note that some of the subjects in previous studies (McLin *et al.*, 1988; Westheimer, 1957; Malmstrom & Randle, 1976) reported the use of visual imagery as a strategy to generate voluntary changes in accommodation, consistent with an indirect use of a three-dimensional internal representation.

GENERAL DISCUSSION

We have shown that the recovery of three-dimensional shape, based solely on monocular relative motion information, was able to drive vergence eye movements with a "gain" similar to that produced by disparity, blur and looming combined, and to that obtained when relative stereo disparity and motion were combined. Similar binocular eye movement traces were produced by the KDE stimulus, by the real physical rotation of an LED, and by a dynamic stereogram of a rotating random dot sphere. The results are consistent with the hypothesis that three-dimensional structure derived from two-dimensional motion information and from binocular disparity are similarly represented in the visual system. This finding provides objective evidence for our introspection that a "true" sense of depth is elicited by a KDE stimulus.

The recorded eye movement responses could be mediated by a single three-dimensional representation of surfaces where depth information from multiple visual cues is combined (Marr, 1982). The estimated three-dimensional structure could be used by the brain to plan different motor activities, such as binocular eye movements. Previous work suggests that under monocular viewing the perception of three-dimensional depth from perspective (Enright, 1987a,b) and looming (changes in size) (Erkelens & Regan, 1986; McLin *et al.*, 1988), can induce changes in vergence. Thus, there has been some indication in the past that monocular three-dimensional perception can drive binocular eye movements.

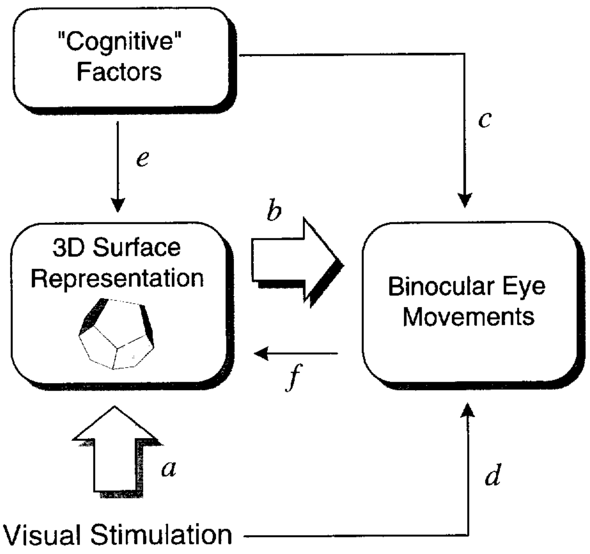

FIGURE 10. Multiple factors influencing binocular eye movements. (*a*) The visual system may construct an internal three-dimensional model of the environment based on the retinal stimulation. (*b*) The estimated three-dimensional model can be used to plan motor activities, such as binocular eye movements. (*c*) Direct voluntary control can influence eye movements. (*d*) In binocular viewing conditions examples can be found of stimuli that will induce a change in binocular eye movements *without* eliciting a three-dimensional percept. (*e*) Subjects can be instructed to bias their three-dimensional perception of the scene; this is most evident when the visual stimulus has multiple two-dimensional and three-dimensional interpretations. (*f*) Modulation in vergence may influence the perceived depth.

Under binocular viewing conditions, disparity-driven vergence is expected to dominate the vergence state (Enright, 1987a; Judge, 1991).* The results show that under monocular viewing, binocular eye movements still occur in a way which is consistent with our three-dimensional perception of space. This phenomenon might be useful in studying how *monocular* cues give rise to the perception of shape and depth and how they combine to yield a single percept. We doubt the methodology would be applicable when binocular disparity cues are available (see below).

Eye movements are known to be under the control of both visual stimulation and of "cognitive" factors such as attention, expectation, memory, and voluntary control

---

*In fact, when we repeat the KDE experiment with *binocular* vision, with all the dots having zero disparity, we see that the modulation in vergence is virtually abolished. This is not surprising, as in this case very small changes in vergence ($\approx$ 20 min arc) would cause diplopia (Ogle, 1952).

(Kowler, 1990). Figure 10 shows different factors affecting binocular eye movements. We hypothesize that visual stimulation is processed by the visual system to create a three-dimensional surface representation. In turn, the *perceived* three-dimensional structure aids in the planning of eye movements and other motor activities.

Relevant to the present discussion is the finding that under *binocular* viewing conditions there are situations in which a stimulus may elicit binocular eye movements without the observer reporting a change in perceived depth (Kertesz, *et al.*, 1983; Houtman *et al.*, 1977; Schor *et al.*, 1992; Erkelens & Collewijn, 1985; Regan *et al.*, 1986). This fact, indicated by the arrow labeled *d* in Fig. 10, seems to cast doubt on the applicability of oculomotor measurements as an indicator of perceived depth under binocular viewing conditions. In contrast, under *monocular* viewing conditions—when there is no disparity-driven vergence—and constant accommodation, we are not aware of any clear examples that were either (a) visual stimuli that while perceived as "flat" induce changes in vergence, nor (b) visual stimuli that while perceived as three-dimensional do *not* cause changes in vergence. We have shown that relative image motion can be quite effective in driving vergence. In our experiments, *monocular* viewing was used to minimize the possibility that binocular stimulation was driving vergence directly without eliciting a three-dimensional percept and to avoid the conflict of cues that would arise between binocular disparity and relative motion for a visual stimulus presented on a flat computer screen.

Top-down "cognitive" factors can affect eye movements *directly* via voluntary control (Westheimer, 1957; Finchman, 1962; Malmstrom & Randle, 1976; McLin *et al.*, 1988), or *indirectly* by influencing the current three-dimensional surface representation. These possibilities are indicated by arrows labeled *c* and *e* in Fig. 10. The findings in Experiment 3 suggest that one can bias the perceived interpretation of the scene as either two- or three-dimensional by giving appropriate instructions to the subjects. Our results suggest that the changes in the responses with different instructions are mediated *indirectly* through a change in the current three-dimensional representation of the scene (the path $e \rightarrow b$) and *not* by a direct voluntary control of the eyes (arrow *c*). Finally, it is possible that eye movements themselves influence the perception of three-dimensional surfaces, as shown by the arrow labeled *f* in the model.

REFERENCES

Adamson, J. & Finchman, E. F. (1939). The effect of lenses and convergence upon the state of accommodation of the eye. *Transcripts of the Ophthalmic Society of the UK*, *59*, 163–179.

Andersen, G. J. & Braunstein, M. L. (1983). Dynamic occlusion in the perception of rotation in depth. *Perception & Psychophysics*, *34*, 356–362.

Braunstein, M. L. (1962). Depth perception in rotating dot patterns: Effects of numerosity and perspective. *Journal of Experimental Psychology*, *64*, 415–420.

Braunstein, M. L. (1976). *Depth perception through motion.* New York: Academic Press.

Braunstein, M. L. (1994). Structure from motion. In Smith, A. T. & Snowden, R. J. (Eds) *Visual detection of motion*, Ch. 13, (pp. 367–393). London: Academic Press.

Braunstein, M. L. & Andersen, G. J. (1981). Velocity gradients and relative depth perception. *Perception & Psychophysics*, *29*, 145–155.

Braunstein, M. L., Hoffman, D. D., Shapiro, L. R., Andersen, G. J. & Bennett, B. M. (1987). Minimum points and views for the recovery of three-dimensional structure. *Journal of Experimental Psychology: Human Perception & Performance*, *13*, 335–343.

Campbell, F. W. (1957). The depth of field of the human eye. *Optica Acta*, *4*(4), 157–164.

Cleveland, W. S., Grosse, E. & Shyu, W. M. (1992). Local regression models. In Chambers, J. M. & Hastie, T. J. (Eds) *Statistical models in S.* UK: Wadsworth & Brooks.

Crane, H. D. & Steele, C. M. (1978). Accurate three-dimensional eyetracker. *Applied Optics*, *17*, 691–705.

Dallos, P. J. & Jones, R. W. (1963). Learning behavior of the eye fixation control system. *IRE Transcripts in Automatic Control*, *8*, 218–227.

Davson, H. (1950). *The physiology of the eye.* Philadelphia: The Blakiston Company.

Dosher, B. A., Landy, M. S. & Sperling, G. (1989a). Kinetic depth effect and optic flow I. 3D shape from Fourier motion. *Vision Research*, *29*, 1789–1813.

Dosher, B. A., Landy, M. S. & Sperling, G. (1989b). Ratings of kinetic depth in multidot displays. *Journal of Experimental Psychology: Human Perception & Performance*, *15*(4), 816–825.

Eby, D. W. (1992). The spatial and temporal characteristics of perceiving 3-D structure from motion. *Perception & Psychophysics*, *51*(2), 163–178.

Enright, J. T. (1984). Changes in vergence mediated by saccades. *Journal of Physiology*, *350*, 9–31.

Enright, J. T. (1987a). Art and the oculomotor system: Perspective illustrations evoke vergence changes. *Perception*, *16*, 731–746.

Enright, J. T. (1987b). Perspective vergence: Oculomotor responses to line drawings. *Vision Research*, *27*, 1513–1526.

Erkelens, C. J. & Collewijn, H. (1985). Eye movements and stereopsis during dichoptic viewing of moving random-dot stereograms. *Vision Research*, *25*, 1698–1700.

Erkelens, C. J. & Regan, D. (1986). Human ocular vergence movements induced by changing size and disparity. *Journal of Physiology*, *379*, 145–169.

Finchman, E. F. (1962). Accommodation and convergence in the absence of retinal images. *Vision Research*, *1*, 425–440.

Finchman, E. F. & Walton, J. (1957). Reciprocal action of accommodation and convergence. *Journal of Physiology, London*, *201*, 21–47.

Gibson, E. J. & Gibson, J. J. (1957). Continuous perspective transformations and the perception of rigid motion. *Journal of Experimental Psychology*, *54*, 129–138.

Gibson, J. J. (1950). *The senses considered as perceptual systems.* Houghton-Mifflin.

Green, B. F., Jr (1959). Kinetic depth effect. Technical Report Psychology Group 58, Quarterly Progress Report, MIT, Lincoln Laboratory.

Green, B. J. Jr (1961). Figure coherence in the kinetic depth effect. *Journal of Experimental Psychology*, *62*, 272–282.

Hering, E. (1868). In Bridgeman, B. & Stark, L. (Eds) *The theory of binocular vision*, New York: Plenum Press.

Hildreth, E. C., Grzywacz, N. M., Adelson, E. H. & Inada, V. K. (1990). The perceptual buildup of three-dimensional structure from motion. *Perception & Psychophysics*, *48*, 19–36.

Hoaglin, D. C., Mosteller, F. & Tukey, J. W. (1983). *Understanding robust and exploratory data analysis*, New York: Wiley.

Houtman, W. A., Roze, J. H. & Scheper, W. (1977). Vertical motor fusion. *Documents in Ophthalmology*, *44*, 179–185.

Husain, M., Treue, S. & Andersen, R. A. (1989). Surface interpolation in 3-D structure-from-motion perception. *Neural Computation*, *1*, 324–333.

Johansson, G. (1958). Rigidity, stability and motion in perceptual space. *Acta Psychologica*, *52*, 359–370.

Jones, R. & Kerr, K. W. (1972). Vergence eye movements to pairs of disparity stimuli with shape selections cues. *Vision Research*, *12*, 1425–1430.

Judge, S. J. (1991). Vergence. In Carpenter, R. H. S. (Ed.), *Eye movements*, Ch. 7 (pp. 157–174). London: Macmillan Press.

Kertesz, A. E. (1983). Vertical and cyclofusional disparity vergence. In Schor, C. M. & Ciuffreda, K. (Eds), *Vergence eye movements: Basic and clinical aspects* (pp. 317–348). Boston: Butterworth.

Koenderink, J. J. & van Doorn, A. J. (1975). Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Optica Acta*, *22*,(9), 773–791.

Kowler, E. (1990). The role of visual and cognitive processes in the control of eye movement. In Kowler, E. (Ed.), *Eye movements and their role in visual and cognitive processes*, Volume 4 of *Reviews of oculomotor research*, Ch. 1 (pp. 1–70). Amsterdam: Elsevier.

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.

Lappin, J. S., Doner, J. F. & Kottas, B. L. (1980). Minimal conditions for the visual detection of structure and motion in three dimensions. *Science*, *209*, 717–719.

Longuet-Higgins, H. C. & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B*, *208*, 385–397.

Loomis, J. M. & Eby, D. W. (1988). Perceiving structure from motion: Failure of shape constancy. In *Proceedings of the 2nd International Conference on Computer Vision* (pp. 383–391). Tampa, FL: IEEE Computer Society Press.

Maddox, E. E. (1893). *The clinical use of prisms*. Bristol: John Wright and Sons.

Malmstrom, F. & Randle, R. J. (1976). Effects of visual imagery on the accommodation response. *Perception & Psychophysics*, *19*, 450.

Marr, D. (1982). *Vision*. New York: W. H. Freeman & Company.

McLin, L. N. & Schor, C. M. (1988). Voluntary effort as a stimulus to accommodation and vergence. *Investigations in Ophthalmology & Visual Science*, *29*, 1739–1746.

McLin, L. N., Schor, C. M. & Kruger, P. B. (1988). Changing size (looming) as a stimulus to accommodation and vergence. *Vision Research*, *28*, 883–898.

Nawrot, M. & Blake, R. (1993). On the perceptual identity of dynamic stereopsis and kinetic depth. *Vision Research*, *33*(11), 1561–1571.

Ogle, K. N. (1952). On the limits of stereoscopic vision. *Journal of Experimental Psychology*, *44*, 253–259.

Petersik, J. T. (1979). Three-dimensional object constancy: Coherence of a simulated rotating sphere in noise. *Perception & Psychophysics*, *25*, 328–335.

Petersik, J. T. (1980). The effects of spatial and temporal factors on the perception of stroboscopic rotation simulations. *Perception*, *9*, 271–283.

Phillips, S. & Stark, L. (1977). Blur: A sufficient accommodative stimulus. *Documenta Ophthalmologa*, *43*, 65–89.

Prazdny, K. (1980). Egomotion and relative depth map from optical flow. *Biological Cybernetics*, *36*, 87–102.

Randle, R. J. (1970). Volitional control of visual accommodation. Technical Report 82, Advisory Group for Aerospace Research and Development, Garmisch-Partenkirchen, West Germany.

Rashbass, C. & Westheimer, G. (1961). Disjunctive eye movements. *Journal of Physiology (London)*, *159*, 339–360.

Regan, D., Erkelens, C. J. & Collewijn, H. (1986). Necessary conditions for the perception of motion in depth. *Investigations in Ophthalmology & Visual Science*, *27*(4), 584–597.

Riggs, A. L. & Niehl, E. W. (1960). Eye movements recorded during convergence and divergence. *Journal of the Optical Society of America*, *50*, 913–920.

Rogers, B. J. & Graham, M. E. (1979). Motion parallax as an independent cue for depth perception. *Perception*, *8*(2), 125–134.

Schor, C. M., Alexander, J., Cormack, L. & Stevenson, S. (1992). Negative feedback control model of proximal convergence and accommodation. *Ophthalmic & Physiological Optics*, *12*, 307–318.

Siegel, R. M. & Andersen, R. A. (1988). Perception of three-dimensional structure from motion in monkey and man. *Nature*, *331*, 259–261.

Sperling, G., Landy, M. S., Dosher, C. A. & Perkins, M. E. (1989). Kinetic depth effect and identification of shape. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(4), 826–840.

Stark, L., Keynon, R. V., Krishnan, V. V. & Ciuffreda, K. J. (1980). Disparity vergence: A proposed name for a dominant component of binocular vergence eye movements. *American Journal of Optics & Physiological Optics*, *57*(9), 606–609.

Stark, L., Vossius, G. & Young, L. R. (1962). Predictive control of eye tracking movements. *IRE Transcripts of Human Factors & Electronics*, *3*, 52–57.

Todd, J. T. (1984). The perception of three-dimensional structure from rigid and nonrigid motion. *Perception & Psychophysics*, *36*, 97–103.

Todd, J. T. (1985). The analysis of three-dimensional structure from moving images. In Ingle, D., Jeannerod, M. & Lee, D. (Eds), *Brain mechanisms and spatial vision* (pp. 73–93). Netherlands: Martinus Nijhoff.

Todd, J. T., Akerstrom, R. A., Reichel, F. D. & Hayes, W. (1988). Apparent rotation in three-dimensional space: Effects of temporal, spatial, and structural factors. *Perception & Psychophysics*, *43*, 179–188.

Todd, J. T. & Bressan, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception & Psychophysics*, *48*, 419–430.

Treue, S., Andersen, R. A., Ando, H. & Hildreth, E. C. (1995). Structure-from-motion: Perceptual evidence for surface interpolation. *Vision Research*, *35*(1), 139–148.

Treue, S., Husain, M. & Andersen, R. (1991). Human perception of structure from motion. *Vision Research*, *31*, 59–75.

Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.

Wallach, H. & O'Connell, O. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, *45*, 205–217.

Westheimer, G. (1957). Accommodation measurements in empty fields. *Journal of the Optical Society of America*, *47*, 714.

Westheimer, G. & Mitchell, D. E. (1969). The sensory stimulus for disjunctive eye movements. *Vision Research*, *9*, 749–755.

Acknowledgements—Supported by NIH-EY01472 and a Goldsmith Foundation Fellowship. We thank M. Landy, L. Maloney, L. Kaufman, N. Rubin, J. Gordon, P. Glimcher, G. Sapiro, J. Enright and anonymous reviewers for useful comments on earlier versions of this manuscript.