23. Luther, G. W. III, Ferdelman, T. G., Kostka, J. E., Tsamakis, E. J. & Church, T. M. Temporal and spatial variability of reduced sulfur species (FeS$_2$, S$_2$O$_3^{2-}$) and porewater parameters in salt marsh sediments. *Biogeochemistry* **14**, 57–88 (1991).
24. Stookey, L. L. Ferrozine—a new spectrophotometric reagent for iron. *Anal. Chem.* **41**, 779–782 (1970).

....

Perceiving visual expansion without optic flow

Paul R. Schrater*†, David C. Knill†‡ & Eero P. Simoncelli§

* *Department of Neuroscience, University of Pennsylvania, 215 Stemmler Hall, Philadelphia, Pennsylvania 19104, USA*
‡ *Department of Psychology, University of Pennsylvania, 3815 Walnut Street, Philadelphia, Pennsylvania 19104, USA*
§ *Howard Hughes Medical Institute and Center for Neural Science and Mathematics, New York University, 4 Washington Place, Rm 809, New York, New York 10003, USA*

....

When an observer moves forward in the environment, the image on his or her retina expands. The rate of this expansion conveys information about the observer's speed[1] and the time to collision[2–4]. Psychophysical[5–7] and physiological[8,9] studies have provided abundant evidence that these expansionary motions are processed by specialized mechanisms in mammalian visual systems. It is commonly assumed that the rate of expansion is estimated from the divergence of the optic-flow field (the two-dimensional field of local translational velocities)[10–14]. But this rate might also be estimated from changes in the size (or scale) of image features[15]. To determine whether human vision uses such scale-change information, we have synthesized stochastic texture stimuli in which the scale of image elements increases gradually over time, while the optic-flow pattern is random. Here we show, using these stimuli, that observers can estimate expansion rates from scale-change information alone, and that pure scale changes can produce motion after-effects. These two findings suggest that the visual system contains mechanisms that are explicitly sensitive to changes in scale.

To test the hypothesis that the rate of image expansion on the retina might be estimated from changes in scale (Fig. 1), we designed three types of stochastic expansion stimuli (Fig. 2a). We created a sequence of uncorrelated white-noise images, and convolved each one with a band-pass filter whose centre frequency decreases exponentially with time. The result is a sequence of random textures whose average spatial scale grows according to a fixed rate of expansion. Because the individual stimulus frames are uncorrelated, both the optic-flow field and its divergence are (on average) zero, and thus cannot be used to compute expansion rate. Nevertheless, most people who have viewed these stochastic stimuli report a strong perception of expansion, emanating from the point of fixation. This suggests that humans have mechanisms that are
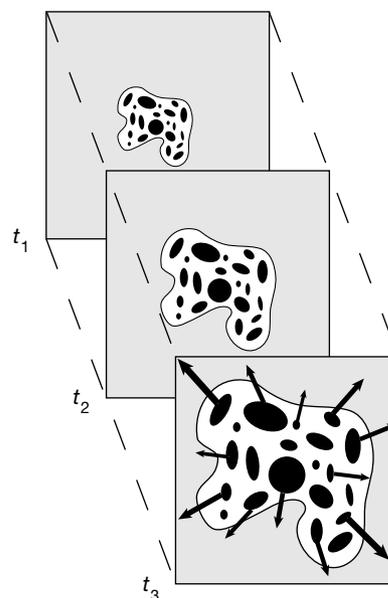
Figure 1 Illustration of visual input when an observer moves forward in the environment. Shown are three frames taken from a sequence of visual images (a movie). Superimposed on the last frame are example optic-flow vectors, indicating the local translational motion of selected points in the scene. The radiating pattern of these vectors indicates that the observer is moving forward. The divergence of this vector field provides one source of information about the rate of expansion. The change in scale of the objects or texture elements provides another source of information.
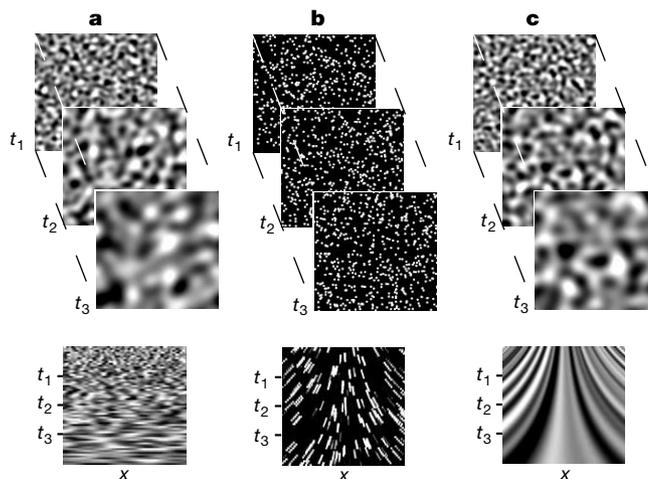
Figure 2 Illustration of the three types of stimuli used in our experiments. Top, example *x–y* frames of the stimulus movie; bottom, *x–t* cross-sections. **a**, Stochastic texture stimulus. Each frame contains a random texture synthesized by convolving a gaussian white-noise image with a band-pass filter. The peak spatial frequency of the band-pass filter decreases over time, producing a gradual change in scale of texture elements. The *x–t* slice illustrates the lack of consistent displacement information. Because each frame is generated from independent white noise, the resulting optic-flow field is random, and thus cannot be used to estimate expansion rate. **b**, Limited-lifetime random dots stimulus. Dots appear asynchronously in random locations, move according to the global expansion pattern for five frames, and then disappear. For this stimulus, the optic-flow field may be used to estimate expansion rate, but there is minimal scale-change information. **c**, Deterministic texture stimulus. Each frame is a re-scaled replica of a single image of band-pass-filtered gaussian white noise. Both optic flow and scale changes provide information about the expansion rate.

sensitive to changes in scale per se.

To test this hypothesis more rigorously, we measured the ability of subjects to match the expansion rate of our stochastic stimuli to the expansion rate of a field of dots undergoing constant expansionary motion. The dots in the latter stimuli have a limited lifetime (Fig. 2b; and Methods), and thus contain optic-flow information about the rate of expansion but minimal scale-change information. On each trial, observers fixated on a point lying between a stochastic expansion stimulus and an expanding dots stimulus. The stimulus movies were repeated while the observer adjusted the rate of expansion of the dots until it appeared to match the expansion rate of the stochastic stimulus. We measured the matching expansion rates for four different stimulus conditions: two durations and two average texture scales. To provide a standard against which to compare these matches, observers were also asked to perform the same task using deterministic texture expansion movies with identical parameters (see Fig. 2c).

Figure 3 shows the results of the matching experiment for both stochastic and deterministic texture expansion stimuli. All obser-

vers show consistent increases in matched dot expansion rates as a function of the expansion rate simulated by the texture stimuli. Although matches are sometimes far from veridical, they are no more so for the stochastic stimuli than for the deterministic stimuli. Linear regression on the matching data shows some variation in slope and intercept across the four different stimulus conditions, but no consistent pattern emerges across subjects.

Although individual subjects show differences in the slopes of matching functions between stochastic and deterministic stimuli, group averages show no significant difference in the average slope for the two types of stimuli (analysis of variance (ANOVA): $F_{1,38} = 0.49$, $P = 0.49$, mean slope = 0.91 and 0.95, respectively). In addition, group averages of the r^2 values of the fits are not significantly different (t-test: $t_{38} = 1.6$, $P = 0.11$, mean $r^2 = 0.44$ and 0.52, respectively). Note, however, that subjects' matches for the deterministic stimuli do have a lower variance than those for the stochastic stimuli (mean s.d. = 0.53 and 0.65; $F_{1,158} = 12.7$, $P < 0.001$). This implies that observers can use the optic-flow information available in the deterministic expansion stimuli to
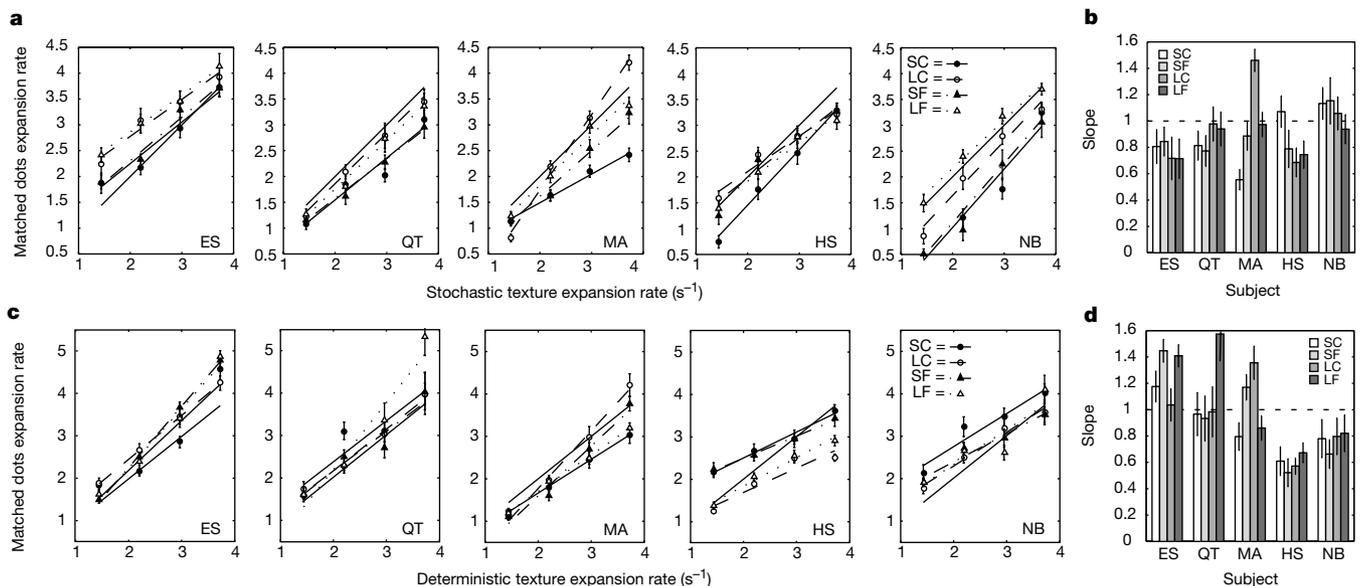
Figure 3 Perceptual matches of expansion rate between dot and texture stimuli.
a, Perceptually matched expansion rate of dot stimuli as a function of the stochastic texture expansion rate for five observers. Symbols indicate mean of 20–30 trials; bars indicate standard error. Stimulus conditions are denoted by different plot symbols: black, short duration (0.43 s); white, long duration (0.73 s); circles, course scale (geometric mean spatial frequency = 0.9 cycles per degree); triangles, fine scale (2.1 cycles per

degree). Thin lines indicate least-squares linear fits for each condition; thick line indicates a veridical match (unit slope). **b**, Slope estimates from **a**. Error bars indicate 95% confidence intervals. Dashed horizontal line denotes a veridical match. **c**, Perceptual matches between expansion rates of deterministic texture stimuli and dot stimuli (details as in **a**). **d**, Slope estimates from **c**.
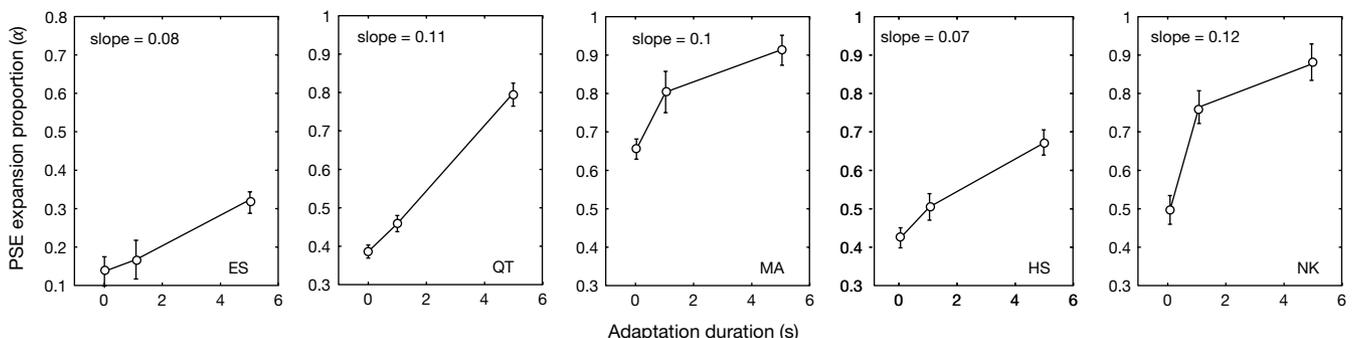
Figure 4 Effect of adapting to stochastic expansion stimuli on the perception of an ambiguous test stimulus for five subjects. The value of α yielding 50% expansion/contraction judgments is plotted against adaptation duration (zero corresponds to no

adaptation). Error bars indicate standard errors. Also indicated is the slope of the psychometric function that was fit to the raw data, which was constant across different adaptation conditions.

improve the reliability of their matches.

The matching experiment leaves open the possibility that subjects use a cognitive strategy to estimate the expansion rates of stochastic expansion stimuli (for example, by estimating the static texture size at two or more moments in time, and dividing by the time interval between them). To rule this out, we performed an additional experiment to measure motion after-effects induced by the stochastic stimuli, as these effects are generally believed to be insensitive to cognitive intrusion[16,17]. We used an experimental design analogous to one used previously to measure adaptation to fields of coherent translating dots[18]. In these experiments, adaptation stimuli are followed by ambiguously moving test stimuli containing mixtures of leftward and rightward moving dots. Adaptation to rightward motion causes a balanced mixture to appear to move leftward. The strength of this after-effect is measured as the percentage of test stimulus rightward dots required to produce a perception of no net motion.

Our experiment used stochastic expansion textures as adaptation stimuli. Test stimuli were created by mixing frames from expanding and contracting stochastic texture movies. In particular, we first created a set of stochastic expansion movies with frequency and duration parameters identical to those of the adaptation stimuli. We converted half of these into contraction movies by reversing the frame order. Ambiguous test stimuli were created by selecting each frame at random from either an expansion or contraction movie.

A mixture parameter, α, controlled the proportion of frames drawn from expansion movies (see Methods). Thus, $\alpha = 1$ corresponds to a stimulus whose frames are all drawn from expansion movies. On each trial, subjects viewed the adaptation stimulus for 0 (no adaptation), 1 or 5 s, followed by a test stimulus, after which they indicated whether the test stimulus appeared to be expanding or contracting. The α-value used to create the test stimuli was adjusted using a psychometric procedure to produce expanding responses 20, 50 and 80% of the time. The resulting data were fitted with a cumulative Normal function with a constant slope across adaptation conditions (nested model comparison tests did not support the use of multiple slopes for any of the subjects: $\chi^2(2) < 0.2$ for all subjects).

Figure 4 shows the null (50%) points estimated for each of five subjects as a function of adaptation time. With no adaptation, three out of five observers perceived the ambiguous stimuli as expanding, consistent with previously reported biases for forward/backward motion perception[19]. As the adaptation time increases, subjects' expansion null points increase, indicating that adaptation to stochastic expansion stimuli produces a bias toward perceiving the test stimuli as contracting. After 5 s of adaptation, the bias is substantial: stimuli that would have been judged ordinarily to be expanding 99% of the time became completely ambiguous. In addition, as the effect increases with adaptation duration, the shifts must be caused by a real after-effect rather than an induction or priming effect.

The results of our matching and adaptation experiments provide compelling evidence for the hypothesis that the human visual system uses scale-change information to compute expansion rate[15]. The strength of the evidence is due to the stochastic texture stimuli, which neatly separate the optic-flow information from the scale-change information. Related stereoscopic stimuli have been used to show that the perception of surface slant from stereopsis does not require point-wise disparity estimates[20], and related sound signals have been used to show neural sensitivity to frequency modulations in the mammalian auditory system[21].

The mechanism by which scale-change information is processed, and the method by which it is combined with optic-flow information remain unknown. Two basic possibilities exist. The visual system might use a single mechanism that is sensitive to both scale-change and optic-flow information. Alternatively, the visual system might have separate mechanisms for computing expansion from scale-change and optic-flow information, combining the

outputs of the two to arrive at a unitary estimate. For example, expansion can be computed separately from scale-change information by estimating the 'motion' of the signal's Fourier energy pattern along the spatial frequency axis. Given a set of neurons whose responses approximate local Fourier energy in a frequency band (like V1 complex cells), this computation could be neurally implemented as time-varying pooling of these neurons' responses adjusted at each moment to track the expected scale of the expanding image. These issues might be addressed best by physiological and psychophysical experiments using the stimuli introduced in this paper. □

Methods

Stimuli

All stimuli were displayed at 60 Hz in 8 bits of precision on a calibrated gray-scale monitor. Stimuli were viewed monocularly from a distance of 55.7 cm, and subtended 16.5 degrees of visual angle. Stochastic textures were constructed by convolving independent frames of gaussian white noise with band-pass filters that are described in the spatial frequency domain by: $F(\omega_r) = 0.5 + 0.5 \cos_r \log_2[\omega_r/2f(t)]$, where $\cos_r$ is a cosine function restricted to one cycle, $f(t)$ is the peak frequency of the filter at time t, and $\omega_r = \sqrt{(\omega_x^2 + \omega_y^2)}$.

For a constant expansion rate a, $f(t) = \exp(\log[f_{initial}] - at)$. We used four different expansion rates: 0.96, 1.44, 2.0 or 2.75 s^{-1}. For each rate, we used two different durations (0.43 and 0.73 s) and two geometric-mean frequencies (0.9 and 2.1 cycles per degree). For a given geometric mean frequency f_{mean}, the initial and final frequencies ($f_{initial}$ and f_{final}) were defined by $\log[f_{initial,final}] = aT/2 \pm \log[f_{mean}]$, where T is the duration. We constructed 125 different versions of each stochastic stimulus type by interleaving every third frame of 5 different movies. The deterministic stimuli were made from the stochastic texture stimuli by taking the first frame and performing band-limited sub-sampling and interpolation to digitally expand the initial image. For the dots stimuli, 200 constant size (4 pixels), randomly placed dots were moved according to: $x(t+1) = x(t) + ax(t)$, which approximates $dx/dt = ax$. Dots had limited lifetimes of five frames (83 ms), and expired asynchronously. When the lifetime was exceeded, or when dots advanced past the edge of the image, they were randomly re-placed in the image. This procedure produces a constant expected density of dots across the image, and prevents the dots from carrying significant size-change information.

Matching experiment

Observers were asked to match the perceived expansion rate of the stochastic textures with that of the dots. Observers were given no feedback, and were queried at the end of the experiment regarding the phenomenal appearance of the stimuli and their matching strategy. Each trial consisted of the following events. Observers fixated on a spot between a stochastic texture movie centred 9.1 degrees left, and a dots movie centred 9.1 degrees right of fixation. Both movies were repeated with a 0.33-s delay between cycles. At any time observers could press keys to increase or decrease the expansion rate of the dots movie, or indicate that the expansion rates appeared to match. Between 24 and 34 matches were collected for each of the 16 different stochastic texture movies, the first 4 of which were discarded to remove early learning effects. The observer's matches for each condition were fit by least-squares linear regression.

Adaptation experiment

Ambiguous test stimuli for the adaptation experiment were constructed by sampling frames from stochastic texture expansion movies, such that the nth frame of the test movie is copied from the kth frame of an expanding stochastic texture movie according to the distribution: $p(k|n, \alpha) = \alpha\phi(n) + (1 - \alpha)\phi(N - n + 1)$, where $\phi(n)$ denotes a Normal density function with mean n and variance 1.5, rounded to the nearest integer, and N is the total number of frames in the movie. Observers viewed these stimuli centrally and after each presentation chose whether the stimuli appeared to be expanding or contracting. The QUEST adaptive psychophysical procedure[22] was used to estimate the α producing 20, 50 and 80% expansion responses. The test stimuli were preceded by adaptation stimuli of three different durations (0, 1 and 5 s). The 5-s adaptation condition also included an initial 20-s adaptation period. For each adaptation condition, 400 trials were collected and all response data were fit by a cumulative Normal psychometric function through maximum likelihood. Fifty per cent null points were determined from these psychometric function fits, and standard errors of these estimates were estimated using 1,000 fits to bootstrap re-samples of the response data. Differences in slope across adaptation duration were assessed using a nested-hypothesis test procedure[23]. The test determines whether the goodness-of-fit of a model with independent slopes for each adaptation duration is significantly better than a single slope model by computing $x = 2 \ln[L_i/L_r]$, where L_i and L_r are the likelihoods of the fit for the independent and reduced models, respectively. The statistic x is χ^2 distributed with 2 degrees of freedom, computed from the difference in the number of free parameters in the two models.

1. Clifford, C. W., Beardsley, S. A. & Vaina, L. M. The perception and discrimination of speed in complex motion. *Vision Res.* **39,** 2213–2227 (1999).
2. Lee, D. N. A theory of visual control of braking based on information about time-to-collision. *Perception* **5,** 437–459 (1976).
3. Schiff, W. & Detwiler, M. L. Information used in judging impending collision. *Perception* **8,** 647–658 (1979).

4. Kaiser, M. K. & Hecht, H. Time-to-passage judgments in non-constant optical flow fields. *Percept. Psychophys.* **57**, 817–825 (1995).

5. Regan, D. & Beverley, K. I. Illusory motion in depth: aftereffect of adaptation to changing size. *Vision Res.* **18**, 209–212 (1978).

6. Regan, D. & Beverley, K. I. Binocular and monocular stimuli for motion in depth: changing-disparity and changing-size feed the same motion-in-depth stage. *Vision Res.* **19**, 1331–1340 (1979).

7. Morrone, M. C., Burr, D. C., Di Pietro, S. & Stefanelli, M. Cardinal directions for visual optic flow. *Curr. Biol.* **9**, 763–766 (1999).

8. Orban, G. A. *et al.* First-order analysis of optical flow in monkey brain. *Proc. Natl Acad. Sci. USA* **89**, 2595–2599 (1992).

9. Duffy, C. J. & Wurtz, R. H. Sensitivity of MST neurons to optic flow stimuli: I. a continuum of response selectivity to large-field stimuli. *J. Neurophysiol.* **65**, 1329–1345 (1991).

10. Gibson, J. J. *The Perception of the Visual World* (Houghton Mifflin, Boston, 1950).

11. Warren, W. H., Morris, M. W. & Kalish, M. L. Perception of translational heading from optical flow. *J. Exp. Psychol. Hum. Percept. Perform.* **14**, 646–660 (1988).

12. Warren, W. H., Blackwell, A. W., Kurtz, K. J., Hatsopoulos, N. G. & Kalish, M. L. On the sufficiency of the velocity field for perception of heading. *Biol. Cybern.* **65**, 311–320 (1991).

13. Perrone, J. A. Model for the computation of self-motion in biological systems. *J. Opt. Soc. Am. A* **9**, 177–194 (1992).

14. Crowell, J. A. & Banks, M. S. Perceiving heading with different retinal regions and types of optic flow. *Percept. Psychophys.* **53**, 325–337 (1993).

15. Beverley, K. I. & Regan, D. Texture changes versus size changes as stimuli for motion in depth. *Vision Res.* **23**, 1387–1400 (1983).

16. Anstis, S. M. The perception of apparent movement. *Phil. Trans. R. Soc. Lond. B* **290**, 153–168 (1980).

17. Cavanagh, P. & Mather, G. Motion: the long and short of it. *Spatial Vis.* **4**, 103–129 (1989).

18. Blake, R. & Hiris, E. Another means for measuring the motion aftereffect. *Vision Res.* **33**, 1589–1592 (1993).

19. Perrone, J. A. Anisotropic responses to motion toward and away from the eye. *Percept. Psychophys.* **39**, 1–8 (1986).

20. Tyler, C. & Sutter, E. Depth from spatial frequency difference: An old kind of stereopsis? *Vision Res.* **19**, 859–865 (1979).

21. Tian, R. & Rauschecker, J. P. Processing of frequency-modulated sounds in the cat's posterior auditory field. *J. Neurophysiol.* **79**, 2629–2642 (1998).

22. Watson, A. B. & Pelli, D. G. Quest: A Bayesian adaptive psychophysical method. *Percept. Psychophys.* **33**, 113–120 (1983).

23. Kendall, M. K. & Stuart, A. *The Advanced Theory of Statistics: Vol. 3. Design, Analysis, and Time Series* (Hafner, New York, 1966).

The motor side of depth vision

Kai Schreiber*†, **J. Douglas Crawford**†‡, **Michael Fetter**§
& Douglas Tweed*†‡

* *Departments of Physiology and Medicine, University of Toronto,*
1 King's College Circle, M5S 1A8 Toronto, Canada
† *Canadian Institutes of Health Research, Group for Action and Perception*
‡ *Centre for Vision Research, York University, 4700 Keele Street, M3J 1P3 Toronto, Canada*
§ *Department of Neurology, University Hospital Tübingen, Hoppe-Seyler-Straße 3, 72072 Tübingen, Germany*

To achieve stereoscopic vision, the brain must search for corresponding image features on the two retinas[1]. As long as the eyes stay still, corresponding features are confined to narrow bands called epipolar lines[2,3]. But when the eyes change position, the epipolar lines migrate on the retinas[4–6]. To find the matching features, the brain must either search different retinal bands depending on current eye position, or search retina-fixed zones that are large enough to cover all usual locations of the epipolar lines. Here we show, using a new type of stereogram in which the depth image vanishes at certain gaze elevations, that the search zones are retina-fixed. This being the case, motor control acquires a crucial function in depth vision: we show that the eyes twist

about their lines of sight in a way that reduces the motion of the epipolar lines, allowing stereopsis to get by with smaller search zones and thereby lightening its computational load.

Eye motion shifts the epipolar lines—the retinal bands where corresponding image features, or stereo matches, project in the two eyes. In the simulation shown in Fig. 1, if the small open circles represent point images cast onto the right retina, and the eyes are converged 30° while looking level, then the corresponding images on the left retina must lie somewhere along the thick lines. When the eyes look 30° down, still converged 30°, the images on the left retina must lie on the thin lines in Fig. 1. When the eyes look 30° up, the images lie on the dashed lines. If the brain searches the wrong lines for stereo matches, it will not find them and stereopsis will fail.

This is a fundamental problem facing any creature that wants both stereopsis and mobile eyes, and there are just two possible solutions. Either the matching algorithm searches different retinal bands depending on current eye position, or it searches eye-fixed zones large enough to encompass all possible locations of the epipolar lines in all usual eye configurations (outlined regions in Fig. 1). Most theories of stereopsis assume that the brain searches narrow bands, so as to reduce its computational load. We used visual stimuli like those in Fig. 2 to show that in fact the brain uses large, eye-fixed zones.

Discs a1 and a2 in Fig. 2 form a cyclorotated stereogram (b1 and b2 form another, interleaved to save space on the page). Disc a1, to be viewed by the right eye, is rotated 2° counterclockwise, and a2 is rotated 2° clockwise, for a total of 4° incyclorotation. When humans converge and look down, our eyes incycloverge, with the upper poles of both eyes rotating medially[5,7,8]. When we converge and look up, our eyes excycloverge. If stereo matching adjusts its search to correct for eye position, this cyclovergence should make no difference to vision; but if the matching algorithm does not correct for eye position, we should see incyclorotated stereograms better when we look down, and excyclorotated stereograms better when we look up. If you fuse the stereograms of Fig. 2 as instructed in the figure legend, you should see the stereoscopic images only on downgaze.
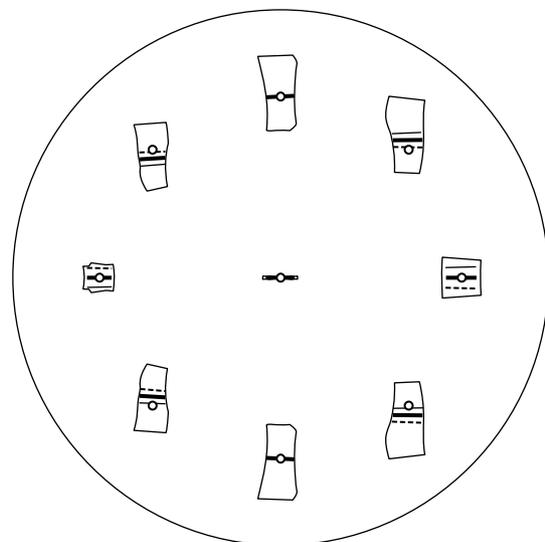
Figure 1 Epipolar lines vary with eye position. The nine open circles are images of nine objects projected onto the right retina, one foveal and the rest 15° eccentric (the perimeter circle marks 22.5° from the fovea). Projections of these same objects onto the left retina must lie somewhere on the thick line segments—pieces of epipolar lines—when the eyes converge 30° and look level. They lie on the dashed lines when the eyes look up 30° and on the thin lines when they look down 30°, always with 30° of convergence. Outlines mark the entire range of motion of the epipolar segments when the eyes also move ± 20° horizontally. Eye motion in this simulation is a realistic combination of Listing's law and L2 (see Methods).

819