

---

# Defaults in stereoscopic and kinetic depth perception

---

**Leonid L. Kontsevich**

*Smith-Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, CA 94115, USA (lenny@skivis.ski.org)*

This study presents three findings concerning the mechanisms of depth perception. First, the shape of the three-dimensional percept evoked by two-frame motion is defined solely by the rotation component around an axis in the frontoparallel plane; the visual system assigns a default value to this rotation component to arrive at a unique solution. Second, when the visual axes of two eyes are almost parallel, the visual system uses a default vergence value to reconstruct stereoscopic depth. Third, the default vergence and default rotation angles are highly correlated across subjects. This correlation implies that the two modalities share a common scaling default at an internal level.

**Keywords:** structure from motion; kinetic depth; stereopsis; interaction; shape perception

## 1. INTRODUCTION

The human visual system continuously reconstructs the structure of the three-dimensional world from the depth cues embedded in two-dimensional retinal projections. An important class of such cues combines different views of the same scene. These views may come from the two eyes, which look at the world from slightly different positions, or the views may change in time because objects move relative to us. Perceiving depth is called a stereoeffect in the first case and a kinetic depth effect in the second.

The stereoscopic and kinetic depth cues exhibit cross-adaptation (Rogers & Graham 1984; Nawrot & Blake 1989) and are indistinguishable under a wide range of conditions (Nawrot & Blake 1993). A predominant approach to depth processing (Marr 1982) postulates that the processing streams for these cues merge at a late stage after each depth cue has been processed in an independent mechanism into a depth signal. However, the available psychophysical data fail to substantiate this conjecture. Stereoscopic and kinetic depth effects are strikingly similar whichever property is analysed (Anstis 1970). In particular, their hyperacuity thresholds are similar over a wide range of separations (McKee *et al.* 1990), and the sensitivity curves for stereoscopic and kinetic depth modulations have similar shapes, both peaking between 0.2 and 0.5 cycles per degree (Rogers & Graham 1982). This similarity may be attributed to a similarity of independent processing mechanisms for these depth cues, but it also may be a result of a merge of two information streams and their processing by a single mechanism. A merge at any processing stage preceding the combined depth signals is consistent with the data described.

The details of the processing of the stereoscopic and kinetic depth cues can be revealed by studying their interaction: parallel processing implies little or no interaction, whereas a merge of the two streams would lead to a

strong interaction. There are two sets of data suggesting that an interaction may take place at very early stages. First, binocular neurons are sensitive to motion in the striate, prestriate and middle temporal visual areas (Pettigrew *et al.* 1968; Zeki 1979; Maunsell & Van Essen 1983) and, second, Bradshaw & Rogers (1992) reported subthreshold summation of the stereoscopic and kinetic depth cues. However, these data cannot be considered as definitive evidence for an early merge because kinetic depth may be produced in the monocular streams processing motion and the subthreshold depth signals may propagate up to the combination stage.

There are, however, convincing data showing that the stereoscopic and kinetic depth cues interact earlier than was postulated by Marr (1982). Johnston *et al.* (1994) found that the stereoscopic and kinetic depth cues after combination may produce more depth than each of these cues alone. To explain this finding, Landy *et al.* (1995) postulated that the stereoscopic and kinetic depth streams interact before they converge. The effect of this elaborate interaction is that the streams adjust their depth scales to provide consistent predictions. The merge of the two streams before the depth signals are generated can provide an alternative explanation that does not require a complicated interaction. On this basis the depth reconstruction mechanism may just combine different projections regardless of their origin, i.e. whether they come from another eye or from a different instant of time.

The present study investigates a deeper association between the stereoscopic and kinetic depth cues. It will be shown that, in the conditions where these cues do not provide sufficient information to reconstruct uniquely the three-dimensional structure, the visual system uses defaults to clarify three-dimensional interpretations. According to the analysis presented, if the defaults developed independently in separate processing streams, there should be a lack of correlation between these defaults. However, experimental analysis shows that the

defaults are highly correlated for a representative set of subjects. This correlation indicates that interaction between binocular and motion depth cues at the scaling stage is stronger than was previously thought. This result suggests that these cues may merge before the scaling stage.

## 2. RATIONALE

In this study, stereopsis will be compared with the simplest case of motion, known as two-frame motion, which is evoked by sequential presentation of two frames. Stereopsis and two-frame motion are similar in the sense that both produce a single disparity map based on the comparison of two views.

This study will deal with weak perspective projection, which combines the orthographic projection and scaling in the projection plane (the scaling component is included to account for the change of the projection size with distance). Two weak perspective views uniquely specify the shape of the projected object, if the angle between the projection axes in the object-centred coordinate system is known (Bennett *et al.* 1989; Koenderink & Van Doorn 1991; Kontsevich 1993). For stereopsis, this angle should be available as long as the visual system senses the visual directions of the two eyes. For motion, however, the visual system typically has no access to the motion parameters and at least three views are required to guarantee uniqueness of the interpretation (Ullman 1979). The fact that two-frame motion evokes a three-dimensional percept implies that the visual system makes assumptions to choose between possible three-dimensional interpretations.

### (a) *Motion default*

Any motion in space can be decomposed into the components of translation and rotation around some origin. Any translation in space can be further decomposed into translation along the projection axis and translation in the frontoparallel plane. In the weak perspective projection, scaling of the image represents the first component. This change creates a percept of general approach/recession and does not reveal the three-dimensional shape of the projected object. The second component of translation merely shifts the image in the projection plane; it does not evoke a depth percept. Therefore, the kinetic depth effect is entirely a result of the rotation component.

Any rotation around some origin can be decomposed uniquely into two simpler rotations around the same origin (Ullman 1979; Todd & Bressan 1990; Koenderink & Van Doorn 1991; Kontsevich 1993): rotation in the frontoparallel plane and rotation around some axis in the frontoparallel plane. Again, the frontoparallel component rotates the image in the image plane and does not produce depth percept. Therefore, only the rotation 'out' of the frontoparallel plane carries information about the three-dimensional structure in the two-dimensional projection; the experimental results of Loomis & Eby (1988) corroborate this conclusion. This depth-modulating rotation is specified by two parameters: orientation of the axis in the frontoparallel plane and rotation angle. These two parameters are

appropriate to define the perceived three-dimensional structure.

Nature, however, may not follow this ideal scheme and the motion components that are irrelevant to three-dimensional structure may affect, and perhaps diminish, the depth percept in comparison to that evoked by the depth-modulating rotation only. This possibility will be tested in the experiment and rejected. The experiments will also reveal that orientation of the depth-modulating rotation axis has little or no effect on the perceived structure of the studied test object. It will be shown that across all conditions tested, two two-frame motions of the same object, which evoke similar three-dimensional percepts, have the same depth-modulating rotation angle. That is, the three-dimensional structure from two views is defined solely by the depth-modulating angle.

To arrive at a three-dimensional interpretation based on two-frame motion, the visual system has to guess the depth-modulating angle. There are two possible strategies for solving this task. The first is to choose the depth-modulating angle depending on the object, so as to make the object look familiar or symmetrical in space (Kontsevich 1996) or commensurate the depth with the width and height (Caudek & Profitt 1993). The second strategy is to use an object-independent default value for the depth-modulating rotation. For only the first strategy the perceived three-dimensional shape should not depend on the rotation magnitude; any correlation between shape and magnitude would indicate the involvement of the default strategy. Such dependence does indeed exist, as observers are apparently able to make metric judgements (Liter *et al.* 1993; Johnston *et al.* 1994); see Todd & Bressan (1990) and Eagle & Blake (1995) for alternative points of view on this issue.

To measure the default rotation angle, the effect of the object-dependent strategy was diminished by choosing a test object whose depth was commensurate with its dimensions in the frontoparallel plane.

### (b) *Stereoscopic default*

When two eyes fixate a point in space, the relationship between binocular images can be characterized as the rotation of the viewed object by the vergence angle around the vertical axis. Because the vertical axis belongs to the frontoparallel plane, the vergence is the same as the depth-modulating rotation and, therefore, knowledge of the vergence guarantees uniqueness of the three-dimensional interpretation from two images. Stereopsis, although it has access to vergence information, systematically distorts the reconstruction of the three-dimensional structure of simple objects (Gogel 1960; Foley 1980; Johnston 1991). An experiment done in this study shows that when the visual axes are close to parallel (either converging or diverging), the stereoscopic system disregards the information about the vergence and uses, instead, a default vergence angle. This distortion, in particular, eliminates the intrinsic instability of the three-dimensional reconstruction at small vergence angles.

The default vergence strategy does not imply that the stereoscopic system cannot use smaller angles than the default in reconstructing depth; for example, motion of one object in the scene can promote recalibration of the

vergence angle for the whole scene (Econopouly & Landy 1995).

(c) **The external factors affecting defaults**

According to the data to be presented, the effective vergence angle employed by the visual system is veridical at only one viewing distance, the vergence angle of which is close to the default value (see experiment 2 below). This veridical vergence has been extensively studied in the literature (actually, these studies were concerned with veridical depth perception, which can be converted easily into veridical vergence given the interocular distance of the observer). There is a well-documented link between the value of default vergence and two factors external to the visual system: (i) a typical default vergence angle corresponds to a distance of about 80 cm (Gogel 1972; Foley 1980; Johnston 1991), which is close to the arm-length distance at which the objects held in a hand are explored; (ii) the default vergence is correlated with the physiological resting state of convergence that eyes assume in complete darkness (Owens & Leibowitz 1980). These two correlations do not necessarily imply that the external factors are causal, although this is a feasible possibility.

The default rotation angle for the kinetic depth effect is potentially related to a completely independent set of external factors, such as the rotational and translational speed of the observer and the observed objects. 'Effortless' vergence has no relation to the motion default; arm-length distance rather corresponds to static binocular viewing conditions and, therefore, also is not related to the motion default. This argument leads us to the following prediction.

(d) **Possible outcomes**

If scaling is learned by the visual system independently for stereoscopic and kinetic depth cues, the experiment should show no correlation between the default vergence and default rotation. Correlation between the defaults would indicate that the processing streams for both cues interact strongly at the scaling stage or that these cues merge before the scaling.

### 3. METHODS

The paradigm employed in the experiments was based on comparison of a perceived three-dimensional shape with a mental model known from a verbal description. Observers controlled the depth modulation in a three-dimensional percept by varying the depth-modulating angle. The observers' task was to adjust the perceived three-dimensional shape to the mental model.

To make a judgement about the amount of depth in the three-dimensional percepts, a version of the 'apparently circular cylinder' paradigm (Johnston 1991) was employed. Because the motion three-dimensional percept may flip relative to the frontoparallel plane, the convex cylinder stimulus used in the original study was modified to balance the convex and concave parts of the stimulus. The new stimulus had two attached cylindrical surfaces of opposite curvature, as shown in figure 1. The test cylinders had circular profiles, judgement of the circularity of which was the core in the experiments to be described. The cylinder axes were always orientated vertically. The transparent surface of the cylinders was covered with 1000

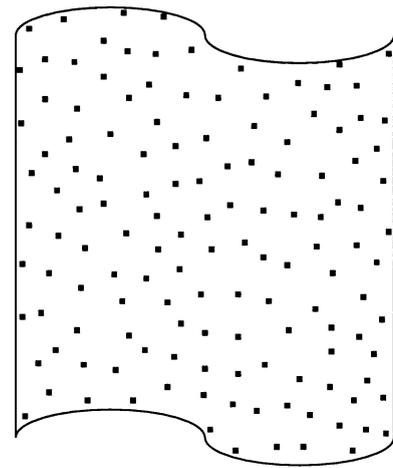


Figure 1. The test object was a surface with 1000 randomly distributed dots.

randomly placed black dots. When the object was orientated along the frontoparallel plane, the black dots were distributed uniformly on a white background.

For each trial, the computer generated two views of the test object. The transformation of the object between the two views was a superposition of translation, scaling and rotation in the screen plane, and (depth-modulating) rotation about an axis in the screen plane. The  $X$  and  $Y$  components of translation, scaling factor, rotation angle in the screen plane and orientation of the axis had to be set before each experiment. The depth-modulating rotation angle (i.e. the angle between the projection axes) was the only parameter controlled by the observer in the course of the experiment.

The two views were presented either side-by-side on the monitor screen in the binocular task or alternating at the same position in the motion task. The view aperture was  $3.2^\circ \times 3.2^\circ$  in both tasks, the distance between the view centres in the binocular task was  $5^\circ$  (10 cm), and the texture dot was  $2 \text{ arc min} \times 2 \text{ arc min}$ . In the experiments with two-frame motion, the aperture of one (frontoparallel) view was 10% larger than the aperture of the other, so that the edges of the stimulus did not provide any shape cues. The frame duration was 300 ms with no gap between the frames, unless otherwise stated.

The experiments were conducted in a dark room where all external references were eliminated. The distance between the monitor screen and the observer was 2.4 m for the motion task and 1.2 m for the stereoscopic task. The experiments with two-frame motion used a monocular viewing condition: the non-dominant eye was covered by an eye-patch. In the stereoscopic experiments the views were cross-fused. Observers were instructed to keep two Nonius lines aligned in the centre of the view. Vergence angle was set by means of prisms in front of the observers' eyes.

The task for the observer was to adjust the depth of the cylindrical surface to make circular depth profiles. The observer could vary only the depth-modulating angle; all other motion parameters were set at the beginning of and kept constant during a block of the adjustment trial. The initial value of the depth-modulating angle in each adjustment trial was set randomly by the computer. The adjustment step was initially set at  $1^\circ$ ; it could be changed by a factor of four, up or down, to make coarser or finer adjustments. The time given for each

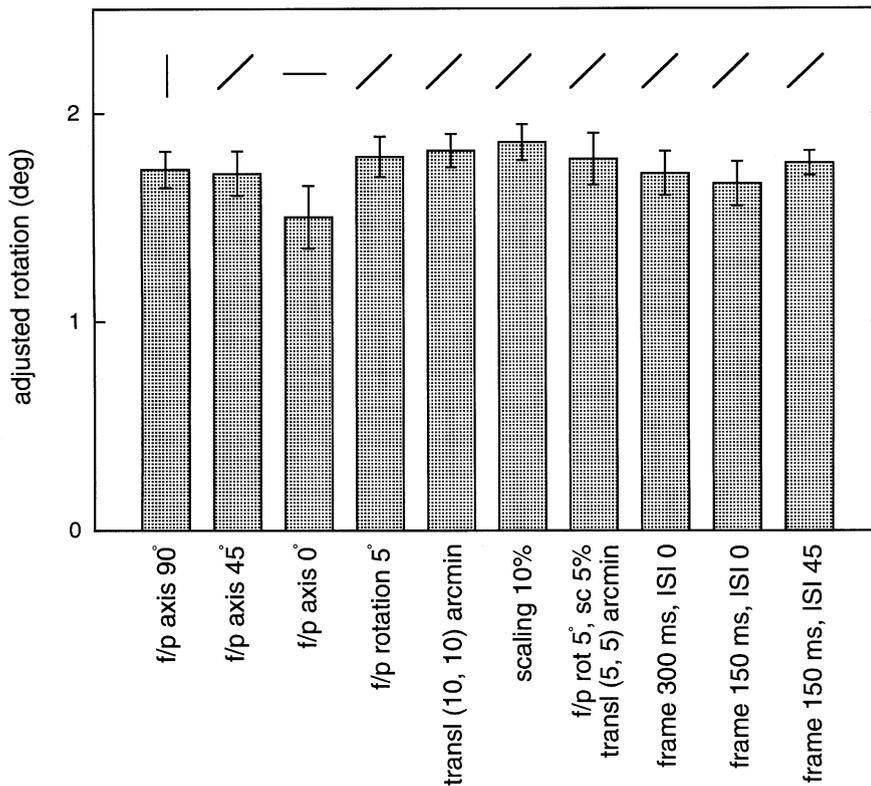


Figure 2. Bar graph of the adjustment experiment results for one observer. The value of the adjusted depth-modulating rotation does not depend on the orientation of the axis (first three bars), additional motion components in the frontoparallel plane (next four bars) and the temporal parameters of the stimulus (the last three bars). The segments on top of the bars depict the axis orientation for the depth-modulating rotation. Error bars show the standard errors of the estimates.

adjustment step was unlimited; in motion experiments the observers were instructed to allow the three-dimensional percept to flip to avoid adaptation (Uomori & Nishida 1994). After an adjustment step was made, the views were presented with a new random texture to avoid a pattern adaptation and local references. The observers were instructed to avoid reducing the adjusted angle to zero, which might provide a reference point between the trials. Each datum point in this study represents the average of five adjustments.

In the first two experiments, the subjects were the author (36 years old) and a paid volunteer (16 years old), both with normal vision. The volunteer subject was unaware of the goals of the experiments. In the third experiment, ten subjects (aged 16 to 68 years, with normal and corrected-to-normal vision) were used, three of whom knew the goals of the experiments.

#### 4. EXPERIMENT 1: INVARIANCE OF THE MOTION DEFAULT ANGLE

In this experiment, the depth-modulating rotation angle was compared for three sets of diverse conditions. The first set was designed to test for anisotropy of the orientation of the axis of depth-modulating rotation. Measurements were taken for the axis tilted at 0°, 45° or 90° to the horizon.

The second set of conditions was designed to test the interaction of the depth-modulating rotation with other motion components. Rotation in the frontoparallel plane, translation in the frontoparallel plane and scaling were added to the depth-modulating rotation separately and in combination. The magnitudes for all the components studied were set close to the limits (within 30%) where the motion correspondence breaks down (these limits varied between the subjects). In all conditions, the axis of

the depth-modulating rotation was orientated at 45° to the horizontal.

The third set of conditions evaluated the effect of temporal parameters on depth modulation. The comparison was made between three temporal sequences: 300 ms frame exposure with no delay between the frames, 150 ms frame exposure with no delay, and 150 ms frame exposure with a 45 ms blank field between the frames.

The values of the temporal parameter studied delimit a narrow range where observers perceived motion of a solid object. At 75 ms frame rate, the object was perceived as blurry and jittery; at 600 ms frame rate, the percept was severely flattened during the static intervals; the increased delay to 90 ms caused degradation of motion.

For both observers, the adjusted depth-modulating rotation was constant across all the conditions of the experiment; the data for the naive observer are shown in figure 2. Thus, for the two observers studied, the three-dimensional percept evoked by two-frame motion was independent of all spatial and temporal parameters (except the depth-modulating rotation angle).

The consequences of this result should be considered in more detail. The adjusted depth-modulating angle does not depend on orientation of the axis. This result indicates that the effect of the depth-modulating rotation was isotropic for the studied subjects and test stimulus. The generality of this conclusion requires further study with more observers and objects.

Moreover, translation, rotation in the frontoparallel plane and scaling had no effect on the adjusted depth-modulating rotation. This result indicates that the visual system implements the ideal scheme that reconstructs three-dimensional structure solely on the basis of the depth-modulating rotation. Implementation of this

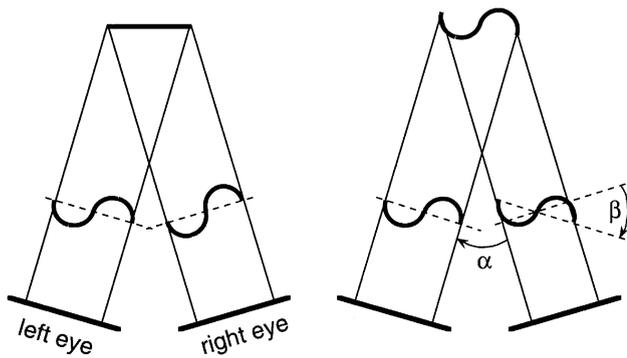


Figure 3. The stereoscopic stimulus, for the purpose of illustration, can be described as though each eye looks at its own physical replica of the object. The viewing conditions can be specified by the vergence angle ( $\alpha$ ) and the depth-modulating rotation angle ( $\beta$ ) of the object.

scheme requires the visual system to compute the depth-irrelevant motion components and nullify their effect. There is a simple linear algorithm for this task, which may have a biologically plausible implementation (Kontsevich 1993).

## 5. EXPERIMENT 2: VERGENCE DEFAULT

In the stereoscopic experiments, the views for two eyes corresponded to a test object slightly rotated around the vertical axis. The stereoscopic stimulus in this experiment can be described as though each eye looks at its own physical replica of the object. When both eyes see identical views, there is no binocular disparity between the images and the stereoscopically perceived three-dimensional structure is a flat plane. When the views are different owing to the depth-modulating rotation of one of the replicas by the vertical axis (angle  $\beta$ ), a three-dimensional structure emerges (figure 3). The task, for a given vergence angle  $\alpha$ , was to adjust the depth-modulating rotation angle  $\beta$  until the surface appeared to be comprised of circular cylindrical surfaces.

The results of this experiment for two observers are presented in figure 4. At large vergence angles, the adjusted depth-modulating angle depends almost linearly on the actual vergence angle, although well below the veridical slope of 1.0 shown by the tilted dashed line. In this range the linear slope is 0.44 (s.e.=0.03) for M.K. and 0.47 (s.e.=0.04) for L.K., which is close to the estimates of vergence gain provided in other studies (Foley 1985; Johnston 1991). At small vergence angles, the adjusted depth-modulating angle becomes independent of the vergence angle, which suggests that in this range vergence information is not taken into account by the visual system. Therefore, the default vergence can be evaluated by adjusting the depth-modulating rotation at parallel visual axes.

In natural viewing conditions the angles  $\alpha$  and  $\beta$  are always equal. The veridical vergence is defined as a point of equality between adjusted depth-modulating rotation (i.e. assumed vergence) and actual vergence. On the graphs, this point is the intersection of the measured curve and the tilted dashed line representing the equality relationship. For both observers, the default vergence, shown by the horizontal dashed line, is close to veridical

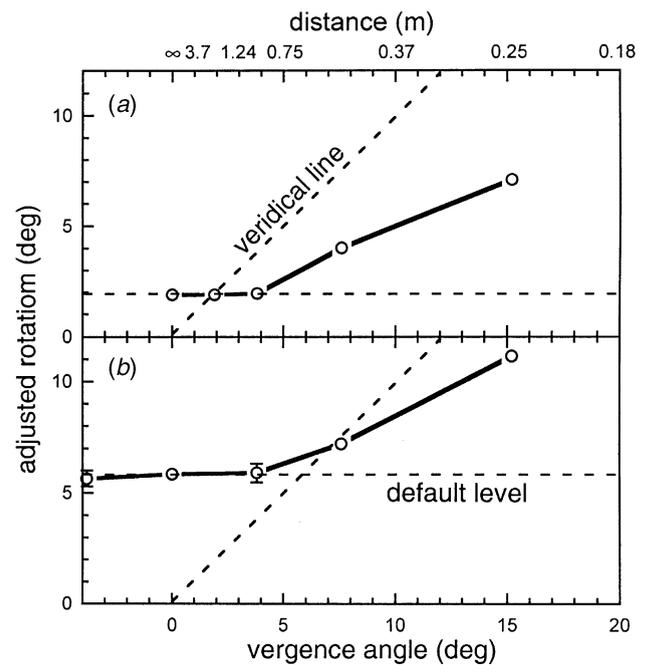


Figure 4. The adjusted depth-modulating rotation as a function of the vergence for two observers: (a) M.K. and (b) L.K. Error bars show the standard error of the estimates. The equivalent distance calculated for an inter-ocular distance of 65 mm is shown on the top.

vergence. This similarity is important for the arguments presented in § 2.

## 6. EXPERIMENT 3: CORRELATION BETWEEN THE MOTION AND VERGENCE DEFAULTS

In this experiment, the motion and vergence defaults were measured for ten subjects. The motion default, i.e. default depth-modulating rotation, was estimated for the condition where its axis was tilted by  $45^\circ$  to horizontal in the frontoparallel plane and no other motion component was present. This particular condition was chosen rather arbitrarily: experiment 1 demonstrated that all motion parameters other than the depth-modulating rotation angle have no effect on depth. The default vergence was estimated as an adjusted rotation for parallel visual axes ( $\alpha=0$ ). The results are shown in figure 5.

The estimates of the default vergence varied over a wide range from  $1.9^\circ$  to  $15.3^\circ$  (200–25 cm) with a geometric mean of about  $5^\circ$  (about 75 cm). Despite the large variance of the defaults across observers, the data points on the graph are clustered along a straight line. A correlation coefficient (computed in linear coordinates) of  $r=0.91$  ( $p<0.01$ ) shows the degree of interdependence of the motion and vergence defaults. This correlation indicates that the studied defaults are not independent, and the binocular and motion mechanisms interact more than has been commonly thought.

Because a special effort was made to diminish the role of the object-dependent three-dimensional interpretation strategy, which may distort the default estimates, equal defaults for each observer should be expected. Indeed, the data in figure 5 support this idea; the data points are clustered along the dashed line, which corresponds to the

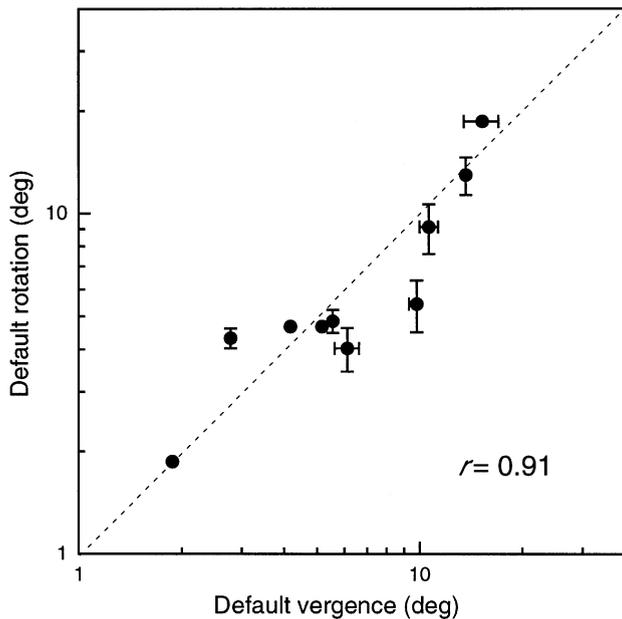


Figure 5. The relation between default vergence and default rotation for ten subjects. Each point represents one subject; the error bars represent the standard error of the estimate in the corresponding dimension. The points cluster along the diagonal (dashed line) corresponding to equality between the defaults.

equality relationship between the defaults. Linear regression analysis shows that the slope (1.01, s.e.=0.16) and  $y$ -intercept ( $-0.57$ , s.e.=1.38) are within one standard error from the values of 1 and 0 that would correspond to the identical defaults. The normalized mean standard deviation of the data point from the equality line is equal to 0.21 (the confidence interval of 0.39 for  $p < 0.05$ ), whereas the defaults vary by a factor of eight.

## 7. DISCUSSION

Following the logic proposed in §2, the result obtained is compatible with two options: there is either strong interaction between the parallel processing streams at the scaling stage or these streams merge before that stage. Is there a way to resolve between them?

The assumption of a merge before the scaling stage would imply that both streams must have the same scaling mechanism. Because the vergence signal affects scaling of the stereoscopic depth, it should correspondingly affect scaling of the three-dimensional structure produced by the kinetic depth effect. Although I am unaware of a study addressing this issue directly, there are some indications reported in fig. 7 of Johnston *et al.* (1994) that support such a possibility. Future studies may give a definitive answer to this intriguing question.

## 8. CONCLUSIONS

Four major experimental results were obtained.

1. When the visual axes approach parallel directions, stereopsis discards information about vergence and uses a default vergence angle to clarify three-dimensional structure.

2. The depth in the kinetic depth effect is defined by the angle of rotation around an axis in the frontoparallel plane. All other motion components have no effect on the reconstructed depth.
3. To elucidate two-frame motion, the visual system assumes a default value for the rotation angle.
4. The values for motion and vergence defaults are highly correlated across observers. The correlation found suggests two options: either (i) the stereoscopic and kinetic depth are processed by two independent mechanisms, which interact strongly at the scaling stage; or (ii) the stereoscopic and kinetic depth cues merge before the scaling stage during which they get scaled to provide a consistent solution.

The study was supported by the grant NEI 7890. I thank Christopher W. Tyler for comments on the manuscript.

## REFERENCES

- Anstis, S. M. 1970 Phi movement as a subtraction process. *Vision Res.* **10**, 1411–1430.
- Bennett, B. M., Hoffman, D. D., Nicola, J. E. & Prakash, C. 1989 Structure from two orthographic views of rigid motion. *J. Opt. Soc. Am. A* **6**, 1052–1069.
- Bradshaw, M. F. & Rogers, B. J. 1992 Subthreshold interactions between binocular disparity and motion parallax. *Invest. Ophthalmol. Vis. Sci.* **33**, 1332.
- Caudek, C. & Proffitt, D. R. 1993 Depth perception in motion parallax and stereokinesis. *J. Exp. Psychol.* **19**, 32–47.
- Eagle, R. A. & Blake, A. 1995 Two-dimensional constraints on three-dimensional structure from motion. *Vision Res.* **35**, 2927–2941.
- Econopoulou, J. C. & Landy, M. S. 1995 Stereo and motion combined rescale stereo. *Invest. Ophthalmol. Vis. Sci.* **36**, S665.
- Foley, J. M. 1980 Binocular distance perception. *Psychol. Rev.* **87**, 411–434.
- Foley, J. M. 1985 Binocular distance perception: egocentric distance tasks. *J. Exp. Psychol.* **11**, 133–139.
- Gogel, W. C. 1960 The perception of shape from binocular disparity cues. *J. Psychol.* **50**, 179–192.
- Gogel, W. C. 1972 Scalar perceptions with binocular cues of distance. *Am. J. Psychol.* **85**, 477–497.
- Johnston, E. B. 1991 Systematic distortions of shape from stereopsis. *Vision Res.* **31**, 1351–1360.
- Johnston, E. B., Cumming, B. G. & Landy, M. S. 1994 Integration of stereopsis and motion shape cues. *Vision Res.* **34**, 2259–2275.
- Koenderink, J. J. & Van Doorn, A. J. 1991 Affine structure from motion. *J. Opt. Soc. Am. A* **8**, 377–385.
- Kontsevich, L. L. 1993 Pairwise comparison technique: a simple solution for depth reconstruction. *J. Opt. Soc. Am. A* **10**, 1129–1135.
- Kontsevich, L. L. 1996 Symmetry as a depth cue. In *Human symmetry perception and its computational analysis* (ed. C. W. Tyler), pp. 331–348. Utrecht, The Netherlands: VSP.
- Landy, M. S., Maloney, L. T., Johnston, E. B. & Young, M. 1995 Measurement and modeling of depth cue combination: in defence of weak fusion. *Vision Res.* **35**, 389–412.
- Liter, J. C., Braunstein, M. L. & Hoffman, D. D. 1993 Inferring structure from motion in two-view and multiview displays. *Perception* **22**, 1441–1465.
- Loomis, J. M. & Eby, D. W. 1988 Perceiving structure from motion: failure of shape constancy. In *Proc. Second Int. Conf. Comput. Vis.*, pp. 383–391. Washington, DC: IEEE.
- McKee, S. P., Welch, L., Taylor, D. G. & Bowne, S. F. 1990 Finding the common bond: stereoacuity and the other hyperacuties. *Vision Res.* **30**, 879–891.

- Marr, D. 1982 *Vision*. San Francisco, CA: Freeman.
- Maunsell, J. H. & Van Essen, D. C. 1983 Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *J. Neurophysiol.* **49**, 1148–1167.
- Nawrot, M. & Blake, R. 1989 Neural integration of information specifying structure from stereopsis and motion. *Science* **244**, 716–718.
- Nawrot, M. & Blake, R. 1993 On the perceptual identity of dynamic stereopsis and kinetic depth. *Vision Res.* **33**, 1561–1571.
- Owens, D. A. & Leibowitz, H. W. 1980 Accommodation, convergence, and distance perception in low illumination. *Am. J. Optomet. Physiol. Opt.* **57**, 540–550.
- Pettigrew, J. D., Nikara, T. & Bishop, P. O. 1968 Binocular interaction on single units in cat striate cortex: simultaneous stimulation by single moving slit with receptive fields in correspondence. *Exp. Brain Res.* **6**, 391–410.
- Rogers, B. J. & Graham, M. E. 1982 Similarities between motion parallax and stereopsis in human depth perception. *Vision Res.* **22**, 261–270.
- Rogers, B. J. & Graham, M. E. 1984 Aftereffects from motion parallax and stereoscopic depth. In *Sensory experience, adaptation and perception* (ed. L. Spillmann & B. R. Wooten), pp. 603–619. New York: Lawrence Erlbaum.
- Todd, J. T. & Bressan, P. 1990 The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Percept. Psychophys.* **48**, 419–430.
- Ullman, S. 1979 *The interpretation of visual motion*. Cambridge, MA: MIT Press.
- Uomori, K. & Nishida, S. 1994 The dynamics of the visual system in combining conflicting KDE and binocular stereopsis cues. *Percept. Psychophys.* **55**, 526–536.
- Zeki, S. M. 1979 Functional specialization and binocular interaction in the visual areas of rhesus monkey prestriate cortex. *Proc. R. Soc. Lond. B* **204**, 379–397.

As this paper exceeds the maximum length normally permitted, the author has agreed to contribute to production costs.

