The influence of large scanning eye movements on stereoscopic slant estimation of large surfaces

Raymond van Ee (1,2) & Casper J. Erkelens (2)

[1] *School of Optometry and Vision Science Program, University of California, Berkeley*

[2] *Helmholtz Institute, Princetonplein 5. 3584 CC, Utrecht, The Netherlands*

email: r.vanee@phys.uu.nl

Abstract. The results of several experiments demonstrate that the estimated magnitude of perceived slant of large stereoscopic surfaces increases with the duration of the presentation. In these experiments subjects were free to make eye movements. A possible explanation for the increase is that the visual system needs to scan the stimulus with eye movements (which take time) before it can make a reliable estimate of slant. We investigated the influence of large scanning eye movements on stereoscopic slant estimation of large surfaces. Six subjects estimated the magnitude of slant about the vertical or horizontal axis induced by large-field stereograms of which one half-image was transformed by horizontal scale, horizontal shear, vertical scale, vertical shear, divergence or rotation relative to the other half-image. The experiment was blocked in three sessions. Each session was devoted to one of the following fixation strategies: central fixation, peripheral (20 deg) fixation and active scanning of the stimulus. The presentation duration in each of the sessions was 0.5, 2 or 8 sec. Estimations were done with and without a visual reference. The magnitudes of estimated slant and the perceptual biases were not significantly influenced by the three fixation strategies. Thus, our results provide no support for the hypothesis that the time used for the execution of large scanning eye movements explains the build-up of estimated slant with the duration of the stimulus presentation.

1 Introduction

Magnitudes of perceived slant within a dichoptically presented stimulus increase over time (Gillam, Flagg & Finlay, 1984; Gillam, Chambers & Russo, 1988). For a presentation duration of 100 ms magnitudes of estimated slant are on average of the order of 20 %---in the absence of a visual reference--- and 55 %---in the presence of a visual reference---of the magnitudes of estimated slant for a presentation duration of 10 s (van Ee & Erkelens, 1996a). In the experiments of Gillam et al. and in our previous experiment, large stimuli were used and subjects were free to make eye movements. Enright (personal communication) suggested that sequential stereopsis (Enright, 1991, 1996)---which is based on a sequential comparison of near-foveal disparity using a number of back and forth eye saccades (see below)---is a possible explanation for the increase in the magnitude of estimated slant over time because the execution of saccades takes time. Sequential stereopsis would provide information about relative depth at various parts of the surface, and thereby about its slant.

As we will see, apart from sequential   stereopsis, there are other mechanisms mediated by eye movements that could help in stereoscopic estimation of slant. The influence of eye movements in general on slant perception has not been examined in the literature. Only if general eye movements do have an influence on the build-up time of slant estimation it is interesting to test the specific role of sequential stereopsis in this build-up. Thus, we made Enright's suggestion less specific by considering eye movements in general: We investigated the possibility that large scanning eye movements explain the build-up of perceived slant of large surfaces over time.

It may not seem very plausible that we should gain precise metrical slant information from scanning eye movements because oculomotor control is reputed both to provide poor cues for the estimation of absolute egocentric distance (Gogel, 1961; reviews in: Foley, 1980; Collewijn & Erkelens, 1990) and to be imprecise in many situations: Vergence position errors of up to 1-2 deg (Collewijn & Erkelens, 1990), vergence velocity errors of up to 1 deg/s (Steinman & Collewijn, 1980) and errors in cyclovergence of 10 minarc (Enright, 1990; van Rijn, van der Steen & Collewijn, 1994) are easily generated during natural behavior. One might expect that these errors degrade depth perception because stereopsis is sensitive to very small disparities. However, depth perception is largely unaffected by the retinal image movement caused by these eye movements (Westheimer & McKee, 1978; Steinman, Levinson, Collewijn & van der Steen, 1985; Patterson & Fox, 1984). These findings have led to the interpretation that depth perception depends upon relative disparities (Westheimer, 1979; Erkelens & Collewijn, 1985a,b) in which case

(errors in) eye movements are irrelevant. We will call stereopsis based on relative disparity, conventional stereopsis.

The facts that absolute distance information from oculomotor cues is poor, that oculomotor control is often imprecise, and that stereoscopic vision is regarded as based on relative disparities does not mean that eye movements do not contribute to depth perception. Under certain circumstances eye movements may enhance depth perception. This is an idea with a long and relatively controversial history (see a number of interesting references in Wright, 1951; Ogle, 1956; Enright, 1991). Indeed, it has been reported that changing fixation from one target to another target that is far to the side of the first target (in lateral, horizontal direction), improves the precision of relative *distance* judgments (Wright, 1951; Rady & Ishak, 1955[1] ; Enright, 1991). Analogously, scanning eye movements could help in stereoscopic *slant* perception. In principle, there are at least seven possible ways in which the visual system can take advantage of scanning eye movements for stereoscopic slant perception of large surfaces:

(1) by measuring the change in vergence,
(2) by sampling disparity during a saccade,
(3) by processing sequences of disparity,
(4) by processing sequences of disparity gradients,
(5) for the prevention of stereoscopic fatigue,
(6) for the prevention of disparity normalization and
(7) for the facilitation of fusion.

The first four can be regarded as potential methods that the subject can use for estimation of slant. The last three are advantageous effects of eye movements on binocular vision.

Measuring the change in vergence---Wright (1951) investigated the relative contributions of vergence (eye-movement condition) and disparity (steady-fixation condition) in a stereo acuity task in which he asked the subjects whether a target was in front of or behind another target. One of the targets was straight ahead. The other target was located eccentrically. In the eye-movement condition subjects were required to make only one eye movement from the straight-ahead target to the eccentric target. Fixation periods were unlimited. Wright suggested that the change in vergence state of the eyes from the first target location to the other target location makes a significant contribution to the stereoacuity result and becomes the predominant factor at eccentricities of 20 deg or larger. In order to isolate vergence-based stereopsis from conventional stereopsis, Wright (1951) made use of targets separated horizontally by an angular spacing that corresponded to the fovea-to-blind-spot distance: he showed that performing an eye movement permits stereoscopic vision in this situation where conventional stereopsis is impossible[2].

Sampling disparity during saccades---Ogle (1953) claimed that in Wright's experiment during an eye movement, midway between the two targets, the targets are relatively near to the fovea so that conventional stereopsis is possible during the saccade. Despite Wright's (1953) defense, Ogle (1956) concluded that there was no evidence that information was used about vergence provided by motor efference or by sensory feedback from the ocular muscles. While Wright's work was not considered in the literature after 1953, later it became clear that Ogle's explanation is unlikely (Enright, 1991; Howard & Rogers, 1995 p.177) because of saccadic suppression.

Processing sequences of disparity---Enright (1991) brought Wright's work back to the attention of researchers in stereopsis. Enright conducted experiments in which subjects had to do, what he called, a distance discrimination task: adjusting the distance of a target until it was perceived to be at the same egocentric distance as a reference target. He went a step further than Wright by claiming that distance discrimination was improved, relative to steady fixation on one target, by making back and forth saccades between the targets, also in cases where neither 1) Ogle's proposal (based on sampling of relative disparity during a saccade) nor 2) Wright's proposal (based on comparison of the vergence states at the target locations) would apply:

1) In order to exclude Ogle's explanation he presented the targets in alternation, with no temporal overlap;

2) He claimed that the use of the vergence state could be excluded (Enright, 1991, p. 1558) by using very brief (50 msec) target presentation durations. (The rationale here is based on the fact that completion

---

[1] Doubts about Rady and Ishak's results have been expressed (Ogle, 1956 and Enright, 1996) because they found reliable distance discrimination even with angular separations of as much as 50 deg.
[2] After finishing this manuscript, Brenner and van Damme (1998) claimed that people have access to reasonable accurate extra-retinal information on changes in ocular convergence.

of a vergence change is a relatively slow process.) He proposed that the underlying mechanism for this type of distance discrimination involves sequential comparison of near-foveal disparity. Enright (1991, 1994) emphasized that the visual system does not need to use ocular vergence in the discrimination task. In the case of a target that is nearly at the same distance as a reference, the subject could base distance discrimination on performing an iso-vergence saccade and evaluating the sign of the remaining disparity of the target (Enright, 1991, 1994)[3]. He called the underlying mechanism *sequential stereopsis*. Enright realized that sequential stereopsis makes stringent demands on the precision of the oculomotor system. In Enright's distance discrimination study most discriminations took as long as 15 sec which typically involved at least a dozen back and forth saccades between the targets (Enright, 1991, p. 1550). He stated that the saccade-to-saccade variability in adventitious vergence change (Enright, 1991, Table 1) in a sequential-stereopsis task was sufficiently small to permit the precise distance discrimination achieved by his subjects.

Processing sequences in disparity gradients---In our experiment (van Ee & Erkelens, 1996a) saccades between widely separated parts on the surface provided sequences of disparity patches: one could regard the foveal area as a spotlight that moves over the stimulus, measuring local disparity gradients. Sequential comparison of disparity gradients could well be a factor in the enhancement of slant perception.

Prevent stereoscopic fatigue---The continued fixation of a stereoscopically presented pattern might lead to a general stereoscopic fatigue of the whole slanted field. This affects slant perception such that slanted patterns are perceived to be unslanted. If the eyes move back and forth between parts of a surface, disparity changes in both magnitude and position. The execution of eye movements prevents stereoscopic fatigue (Howard & Templeton, 1964).

Prevent disparity normalization---After fixating a three-dimensional pattern there might be a shift of the norm of fronto-parallelism such that slanted patterns come to look unslanted [the equidistance tendency as formulated by Gogel (1956)]. Howard & Rogers (1995, p.177) stated that scanning eye movements might prevent depth normalization.

Facilitation of fusion---Random dot stereograms portraying very complex surfaces can be fused with very short latencies, provided each corresponding feature is limited to Panum's fusional area (Julesz, 1978, 1986)[4]. However, even a very simple random dot stereogram consisting of a single square hovering over the background requires a long initial fusion time if the disparities of the square exceed Panum's fusional limit (Julesz, 1978, 1986). The subject has to learn the proper vergence strategies by aligning first one of the corresponding features, and then trying to reconverge on the other area slowly without breaking the fusion of the first area. Thus, the eye movements Julesz referred to were primarily to facilitate the fusion process; they bring the corresponding features of the retinal images into register. Gillam et al. (1988) and our previous experiment (van Ee & Erkelens, 1996a) investigated perceived slant evoked by simple stereograms consisting of relatively small disparities. Gillam et al. (1988) explicitly stated that the latencies of slant perception are post-fusional, which means that the period of time needed to accomplish fusion is not responsible for the increase of perceived slant over time. Therefore, performing eye movements in order to facilitate fusion cannot be an explanation for slant perception latencies in our previous experiment.

The mentioned possibilities are not exclusive: they might, for example, all work at once in slant perception. Investigating the possibility that eye movements explain the increase in estimated slant over the observation period is not only interesting in itself but may result in important implications for hypotheses about models of depth perception in general. If eye movements are not involved in the increase in slant over time then this increase is not an intrinsic property of disparity acquisition but is caused by disparity processing.

Existing studies concerning eye movements and shape perception in stereograms are only qualitative. Although one is able to fuse and recognize a complex stereogram with very short latencies it is still

---

[3] In Enright's experiments, distance discrimination thresholds involved nearly-equidistant targets. In the case of targets which are not nearly-equidistant, Enright (1994) proposed that the subject could base discrimination on the assumption that saccades between such targets almost always involve an undershoot in vergence change, meaning that the remaining disparity corresponds in sign to the new target's original depth difference.

[4] In earlier work, Julesz (1971) stated that the time needed to learn to fuse a random dot stereogram was related to the complexity of the surfaces that were portrayed. Later he explicitly stated that this statement is incorrect (Julesz, 1986). [We mention that his 1971-statement is in error because this statement is frequently referred to, in relation to the present study.]

unknown how good quantitative shape perception is. Therefore it is interesting to investigate the influence of eye movements on a metrical depth task (e.g. slant estimation); this has not been previously done.

The slant experiment in this paper is an extension of the one we did in van Ee and Erkelens (1996a). In that experiment we investigated temporal aspects of stereoscopic estimates---with free eye movements---of surface slant evoked by whole-field stereograms. This time the experiment was blocked in three sessions. Each session was devoted to one of the following fixation strategies: central fixation, peripheral fixation and active scanning of the stimulus. We expected poorer performance in the fixation conditions than in the free-scanning condition if eye movements help to facilitate slant perception. We included the peripheral fixation condition because from subjective slant estimation during pilot experiments we found this condition most difficult and we expected poor slant estimation in this condition (see also next paragraph). The stimuli were presented for either 0.5 sec, 2 sec or 8 sec. The presentation duration is an important parameter because if scanning eye movements are involved in the increase in estimated slant over time one would expect a larger difference in the results between the three fixation strategies for longer presentation durations; for longer presentation durations there is more time available to make scanning eye movements.

We also measured perceptual slant biases. Subjects generally show a small bias in their head-centric slant estimations when there is no frontoparallel (zero slant) reference plane. A plausible possibility is that subjects are biased towards determining the perceived slant relative to their (cyclopean) line of sight (see figure 1). This bias might be influenced by the fixation strategies. Regarding slant about the vertical axis, in case of fixation on the slanted plane eccentrically in the right visual field, this bias would increase the perceived slant; an objective frontoparallel plane (zero slant relative to the head) would be perceived as right side away (figure 1, top panel). Regarding slant about the horizontal axis, in case of fixation on the slanted plane in the lower visual field a slant estimation relative to the cyclopean eye increases slant estimates and would lead to a positive bias; again an objectively frontoparallel plane would be perceived with a positive slant angle (figure 1, lower panel). On the other hand, the recent model of Erkelens and van Ee (1998) in which disparity is processed in headcentric coordinates predicts the absence of a bias due to eccentric fixation.
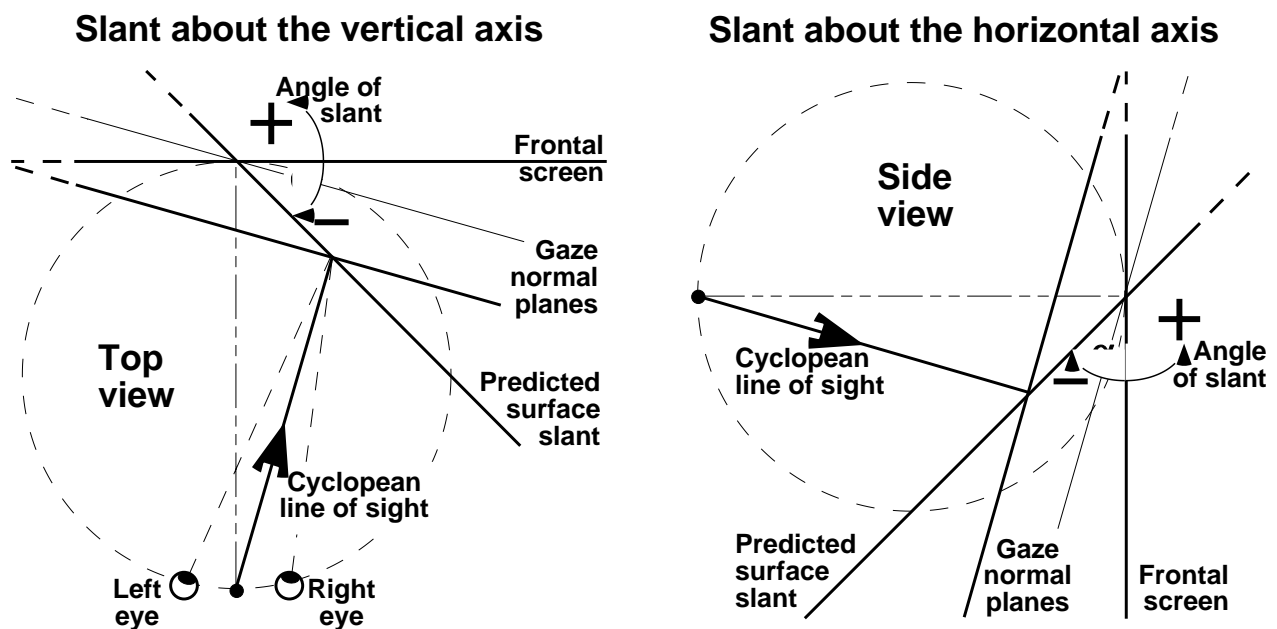

Fig. 1: We define positive slant (α) about the vertical axis as right side away relative to the screen; about the horizontal axis as bottom side away. This figure illustrates how a slant estimate relative to the cyclopean line of sight might bias slant estimation. The location of the origin of the cyclopean line of sight is chosen midway between the nodal points of the eyes. If the subject fixates a peripheral mark on the plane either in the right side---in the case of slant about the vertical axis---or in the lower side of the visual field---in the case of slant about the horizontal axis---a slant estimate relative to the cyclopean eye increases slant estimates and might lead to a positive bias; an objectively frontoparallel plane would be perceived with a positive slant angle.
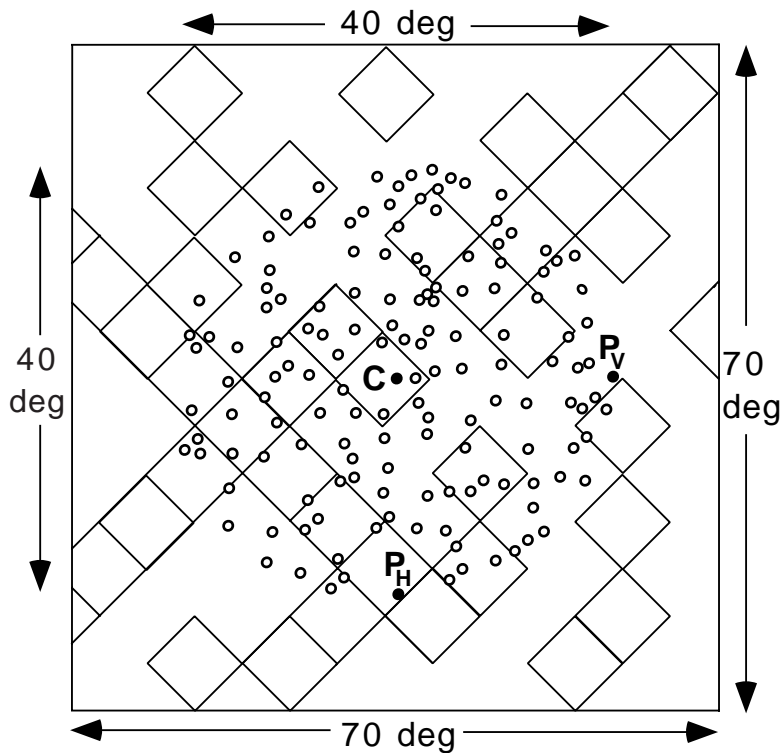
Fig. 2: A schematic illustration of the stimulus. The pattern of circles (diameter 40 deg) underwent the transformations (see text) between its half-images. These transformations evoked perceived slant. The distribution of the small circles was such that they covered about 10 % of the stereogram. Each of the circles had a diameter of 1.5 deg. The black dots mark the location where the fixation dot was shown for the central fixation condition (C) and the peripheral fixation condition (P_V, P_H). In the case of predicted slant about the vertical (horizontal) axis only P_V (P_H) was shown. In the with-reference condition the cross-hatched pattern was shown which served as a visual reference. Its half-images were untransformed relative to each other and presented in the plane of the screen. The visual reference subtended 70 by 70 deg. The diagonals of the individual squares were 15 deg. To prevent matching of false depth planes, not every possible square was shown (approximately six out of ten were shown). A different, randomly chosen configuration of circles and squares was presented every time a new stimulus was presented.

2   Methods

Each of the two half-images of the stereogram was generated at a frequency of 70 Hz. One image was projected in green light and was observed by the right eye through a green filter. A red filter was used to make the other image visible exclusively to the left eye. The transmission spectra of the filters were chosen such that they corresponded as closely as possible to the emission spectra of the projection TV. No crosstalk between the right and left eye views was observed when contrast and brightness of the projection TV were correctly adjusted. The subject was seated in front of the screen at a distance of 1.5 m. Head movements were restricted by a chinrest and a forehead support. Care was taken to ensure that the interocular axis was parallel to the frontal screen. The stereogram was circular (40 degrees diameter) and contained sparsely[5] distributed small circles (see figure 2). The distribution of the small circles was such that they covered about 10 % of the stereogram. The circles had a diameter of 1.5 deg each. A different, randomly chosen configuration of circles was presented on each trial.The task of the subject was to

---

[5] Among the cues for perceived slant that are available to the visual system are the horizontal disparity gradient, the texture gradient and foreshortening. The latter two are counter cues in our experiment because they indicate zero slant. A circular pattern with sparsely distributed circles minimized the effectiveness of these two cues. Note that slant of a pattern is defined by the disparities of just two circles on either side of the pattern. In order to increase the slant information that is available to the visual system we presented more circles.

estimate the magnitude of the perceived slant (figure 1) induced by the stereograms. Subjects were instructed to estimate the perceived slant relative to an (invisible) zero-slant plane. They were told that the screen had zero slant. This means that slant is defined exo-centrically (but because head movements were restricted, slant in exo-centric coordinates is, in practice, similar to slant in head-centric coordinates).

After each trial two binocularly visible lines (one fixed and one rotatable) appeared on the screen (see figure 4 in van Ee and Erkelens, 1996a). By manipulating the computer-mouse position, subjects set the angle between the rotatable line and the fixed line; this angle represented the estimated slant. The two lines were displayed as a flat 2D pattern on the screen and thus also served as a zero-slant reference between successive stimuli. Before starting the experiment we checked whether the subjects were able to estimate slant consistently in our procedure. We did this by means of a series of trials involving real and dichoptically projected slanted planes. The six subjects selected were not informed about the purpose of the experiment. The subjects were tested for good stereo-vision with Julesz random dot patterns. They showed average results in stereoscopic test tasks. They were all either faculty members or students of the Helmholtz Institute for Autonomous Systems Research. Three of the subjects had not participated before in stereoscopic slant estimation experiments but were familiar with performing experimental procedures for vision research. The other three subjects had participated in several stereoscopic slant judgment experiments.

As mentioned above, there were three fixation conditions: central, peripheral and free. In the central and the peripheral fixation condition the subject was instructed to fixate a mark that was located either in the center or in the periphery of the stimulus, respectively (see figure 2). This was a relatively easy task because the fixation mark was a disk with a diameter as large as 1.5 deg. Subjects were allowed to move their gaze over the fixation disc. Generally in vision research when strict fixation is required stimuli are presented with a short duration, like 75 msec, so that execution of a vergence eye movement is impossible. In pilot experiments, however, we found that conducting the task with such a short display duration is not feasible. We did not measure eye posture during the experiment. On the basis of subjective impressions during participation in preliminary experiments we consider it not very likely that occasional unintended large scanning eye movements act as significant contributors to our results. In addition, in our study it is not essential to have knowledge about the exact fixation location. For example it would not be a problem if a subject were eso- or exophoric. Fixation errors caused by eso- or exophoria would be nearly constant over the experiment. The only parameter we changed in the experiment is the fixation strategy and all other factors (even artifacts) stay presumably constant. In the free fixation condition the subject was instructed to scan the stimulus continuously.

Each of the experimental sessions was devoted to one fixation position; the subject participated in three sessions. In the sessions in which peripheral fixation was required the peripheral fixation dot was presented with such disparity that it was perceived in the slanted plane; in other words, the fixation dot had the same disparity as an element of the circular stimulus would have had at the same location. Prior to each trial the fixation dot had already been shown for 1 sec with the correct disparity. This was done to prevent (uncontrolled and involuntary) vergence changes in depth during the onset of the stereogram. Analogously, to prevent uncontrolled version and saccades the three fixation conditions were blocked. Thus, as far as possible, identical trials were presented under identical fixation conditions concerning the location and disparity of the fixation dot.

Each of the sessions was divided into two series: one series without and one series with a visual reference plane[6] (see also Howard & Kaneko, 1994; van Ee & Erkelens, 1995, 1996a; Kaneko & Howard, 1996). The series without a visual reference were presented in a completely dark room, nothing being visible except the circular stimulus[7]. These series were preceded by a dark-adaptation period of six

---

[6] At present, research into stereoscopic depth perception is concentrating more and more on developing models that are valid for ecological conditions and/or during movements of the eyes and the head. Usually, in daily circumstances objects have a disparity relative to a static visual frame of reference. During movements of the eyes and the head, the binocular disparity field of the object and its surrounding change continuously. Characteristic for these disparity changes is that they are essentially whole-field changes (e.g. van Ee and Erkelens, 1996b) without a static visual frame of reference. To cover both situations we investigated the role of the fixation strategies both with and without a reference.

[7] Note that the distinction between "with" and "without" reference refers to the disparity gradient. Without reference means that there is only one disparity gradient present in the visual field. With reference means that there is at least one other disparity gradient present in the visual field. Thus, this distinction does not refer to

minutes. Because the subjects were dark adapted, experiments could be done with low contrast and low brightness settings of the projection TV, without loss of visibility of either half-image of the stereogram. This means that the screen did not (as far as possible) serve as an illuminated plane (a visual reference). After dark adaptation the relative brightness of the red and green half-images was adjusted such that the two half-images were experienced as equally bright when viewed through the anaglyph glasses. During the series of trials in which there was a visual reference, a whole-field reference consisting of a fronto-parallel cross-hatched pattern (70 by 70 deg) was projected onto the screen (see figure 2). The cross-hatched pattern consisted of a field of adjacent diagonal squares with diagonals of 15 deg. The reference pattern was changed randomly each time a new stimulus appeared. To prevent wallpaper (aliasing) effects (matching of false depth planes), not every possible square was shown; instead approximately six out of every ten were shown. In the condition in which a reference was present, the room was dimly lit, which effectively prevented depth contrast effects from causing slant of the reference.

The transformations of the left eye's half-image relative to the right eye's half-image of the stereogram were either horizontal scale, horizontal shear, vertical scale, vertical shear, divergence or rotation. All of these transformations are interesting because Howard and Kaneko (1994), van Ee and Erkelens (1995) and Kaneko and Howard (1996) showed that the transformations horizontal and vertical scale and shear can be regarded as basic transformations for slant judgments[8]. Horizontal scale, vertical scale and divergence evoke slant about the vertical axis; horizontal shear, vertical shear and rotation evoke slant about the horizontal axis. We define slants about the vertical axis as positive if the right side is away from the observer. A slant about the horizontal axis is defined as positive if the bottom side is away. A positive magnitude of horizontal scale or horizontal shear of the right eye's half-image relative to the left eye's half-image leads to a positive angle of perceived slant. A positive magnitude of vertical scale or vertical shear of the right eye's half-image relative to the left eye's half-image leads to a negative angle of perceived slant. A positive magnitude of divergence or rotation of the right eye's half-image relative to the left eye's half-image leads to a positive angle of perceived slant (van Ee and Erkelens, 1998). The magnitudes of horizontal scale, vertical scale and divergence were either –6 % or 6 %. The magnitudes of horizontal shear, vertical shear and rotation were either -3.3 deg or 3.3 deg. The magnitudes were chosen such that the amount of theoretically predicted slant were identical. For example, 6 % scale evokes, theoretically, the same slant as 3.3 deg shear (see van Ee & Erkelens, 1996a). In the periphery, at 20 deg, disparities were 1.2 deg at their maximum. Subjects reported seeing the entire field slanted. Subjects did not base their slant estimations on just the central area of texture. (See figure 7 for a demonstration.) The transformations not only changed the large-field circular pattern into an oval pattern, they also transformed each of the individual circles into ovals according to the local transformation.

As mentioned before, each of the sessions was divided into a series without and a series with a visual reference. We already know from previous studies that local vertical scale (van Ee & Erkelens, 1995; Kaneko & Howard, 1996) and vertical shear (Howard & Kaneko,1994; van Ee & Erkelens, 1995) in the presence of a whole field reference do not evoke a percept of a slanted surface. Therefore, we did not present the vertical scale and vertical shear transformations in the series containing a visual reference.

In summary, subjects had to perform 12 series of slant estimations: the three fixation conditions both in presence and in the absence of a visual reference and for perceived slant about the vertical and about the horizontal axis. With 7 replicates of each condition, this means that they had to estimate a total of 1260 slants because there were 3 fixation conditions (central, peripheral and free), 2 magnitudes of transformations (-6 %, 6 % or -3.3 deg, 3.3 deg), and 3 presentation durations (0.5, 2, 8 sec). For each of those variables, there were 10 transformation conditions: 4 transformations (horizontal scale, horizontal shear, divergence and rotation) both with and without a reference (which makes 8), 2 transformations (vertical scale, vertical shear) without a reference. The order of the experimental sessions with regard to the fixation conditions was: central fixation, peripheral fixation, free fixation. In each of the sessions the subject started with a series of trials without a visual reference. We had a particular reason for this choice because from other research (van Ee, Backus and Erkelens, 1996) we know that there might be a learning effect in stereoscopic slant estimations. In the used order we finished with the condition which we

---

relative disparity compared with absolute disparity. Even in the without-reference condition, relative disparities are present throughout the visual field. Each pattern element has disparity relative to every other pattern element.

[8] Stereoscopically perceived orientations of planar surfaces about the vertical axis are related to the difference between horizontal scale and vertical scale disparities. Perceived slants about the horizontal axis are related to the difference between horizontal shear and vertical shear disparities. Rotation is a combination of identical magnitudes of horizontal and vertical shears of the same polarity. Divergence is a combination of identical magnitudes of horizontal and vertical scales of the same polarity.

assumed to be the least difficult one (assumed from personal experience during pilot experiments). We anticipated that otherwise the learned slant in the easy `free-fixation' condition with reference, could have improved slant estimations in the more difficult conditions (i.e. it could have introduced a response bias without a real change in the percept). The order of the experimental series with regard to the slant direction was: slant about the vertical axis, then, slant about the horizontal axis. Trials were presented in random order within one series.
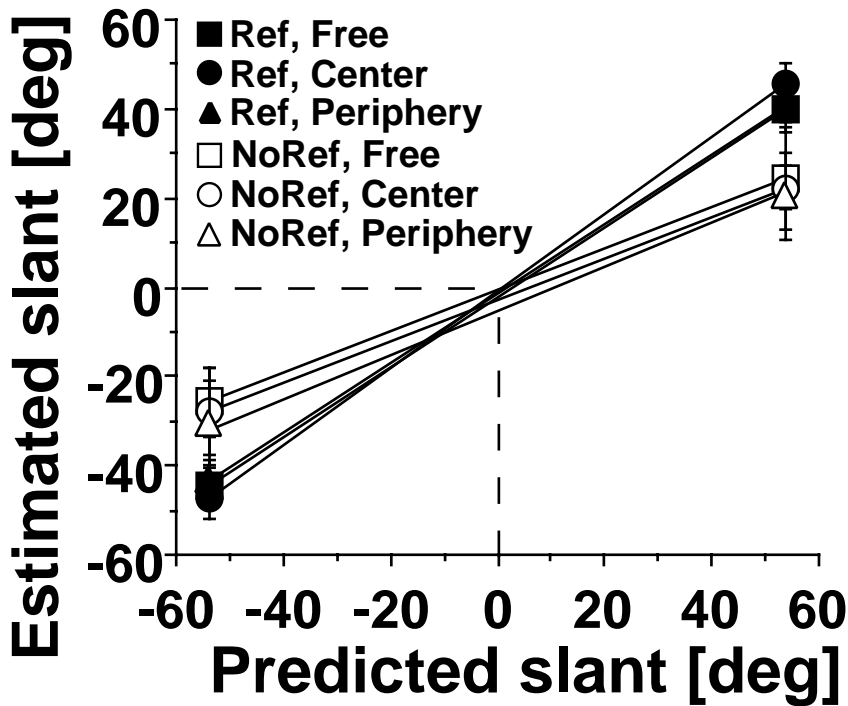
Fig. 3: This figure illustrates the method of data processing and shows the estimated slant as a function of predicted slant of a typical subject (JE), for the three fixation conditions, with (Ref) and without (NoRef) a visual reference. The data are obtained for a horizontally scaled stereogram and a presentation duration of 2 sec. `Free', `Center', `Periphery' denote free, central and peripheral fixation, respectively. The error bars represent standard deviations based on seven slant judgments. This subject showed no significant bias when a visual reference was present. For peripheral fixation he showed a significant bias when a visual reference was absent.

Data were analyzed as described in van Ee and Erkelens (1996a): We determined mean estimated slant as a function of geometrically predicted slant separately for each combination of subject, condition, transformation, and presentation duration. The advantage of using predicted slants is that slant about the vertical axis and slant about the horizontal axis, which are caused by different transformations, can be treated in an identical way. Estimated slant as a function of geometrically predicted slant was fit by a linear relationship. Previous work has shown that, to a good approximation, there exists a linear relationship between estimated and predicted slant (van Ee and Erkelens, 1996a). In this study we are more interested in the role of sequential fixation in slant estimation than in slant estimation per se. Therefore in this study we measured perceived slant for only two predicted slants. Figure 3 illustrates the method of data analysis. The slope of the fitted line is the coefficient **s** in the equation:

estimated slant = **s** × geometrically predicted slant

and, therefore, s represents estimated slant as a fraction of geometrically predicted slant. These s-values (each one based on 14 trials derived from 7 repetitions times 2 magnitudes of transformation) characterize subject's behavior and are plotted in the figures (4 and 5). A subject who would perform the

task veridically (based on the geometrically present stereo information) would consistently exhibit s-values equal to unity. The intersection of the fitted line with the axis where predicted slant is zero reflects the subject's bias in the estimated slant. Measuring the bias in eccentric gaze is potentially interesting because subjects might be biased to a slant setting relative to gaze normal instead of relative to fronto-parallel.

The transformations vertical scale, vertical shear, divergence and rotation as used in the experiment do not mimic objects in the real world (for the given eye posture). This means that there is no geometrical relationship between the magnitudes of these transformations and slant predicted from these magnitudes. Therefore, our figures (4 and 5) show the *normalized* estimated slant as a fraction of the predicted slant. Normalized means that the estimated slant is divided by the predicted slant of horizontal scale or horizontal shear. As an example: We present 3 % vertical scale for a presentation duration of 8 seconds. Say the estimated slant is -15 deg. The geometrically predicted slant of 3 % horizontal scale would be 34 deg (for an observation distance of 150 cm and an inter-ocular distance of 6.5 cm). The estimated slant divided by the predicted slant is -15/34 = -0.44 . Furthermore, in order to be able to compare in one figure the results of vertical shear and vertical scale with the results of horizontal shear, horizontal scale, rotation and divergence we determined the absolute value of this fraction for vertical shear and scale. Thus, in our example the normalized estimated slant is 0.44.

3   Results

The trends of the results were similar for all subjects although quantitative differences in the slopes between subjects were large. Therefore it was pointless to present the mean slopes across the six subjects. There appeared to be a clear difference between the practiced and the unpracticed subjects. Figures 4 and 5 show the effects of the three different fixation positions on the slopes (normalized slant estimation) for the six transformations and the three observation periods. The mean results of the three practiced subjects are shown in figure 4. The error bars in figure 4 (and 5) represent the cross-subject standard error for the mean values of the subjects. Figure 5 shows the mean results of the three unpracticed subjects. The main result is that there is no significant difference between the mean estimates of slant for the three fixation positions. While we cannot absolutely exclude occasional unintended large scanning eye movements as a contributor to our results, our results provide no support for the hypothesis that large scanning eye movements are important for stereoscopic slant perception of large surfaces. Inspection of the raw data shows that this result also holds for the individual subject data (not shown). Apparently, neither pronounced stereoscopic fatigue nor pronounced depth normalization was present in the experiments.

In previous slant estimation experiments subjects were allowed to have free fixation. The present results for the free fixation condition confirm previous reports on slant perception. Slant estimations increase over time to a greater extent in the presence than in the absence of the reference. Slant estimations in the presence of a reference are also larger than in the absence of a reference (Gillam et al., 1984, 1988; van Ee and Erkelens, 1996a). Estimates of slant by practiced subjects are larger than estimates of slant by unpracticed subjects especially for small presentation durations (0.5 sec and 2 sec) and in the absence of a reference (see also van Ee et al., 1996). Unpracticed subjects require longer presentation times to allow the build-up of slant. Consequently, practiced subjects show less increase in their slant estimates over time than inexperienced subjects. The magnitude of slant due to divergence is about equal to the magnitude of slant due to horizontal scale minus slant due to vertical scale (Kaneko and Howard, 1996; van Ee and Erkelens, 1998). Similarly, the magnitude of slant due to rotation is equal to the difference between the magnitude of slant due to horizontal shear and the magnitude of slant due to vertical shear (Howard and Kaneko, 1994; van Ee and Erkelens, 1998).
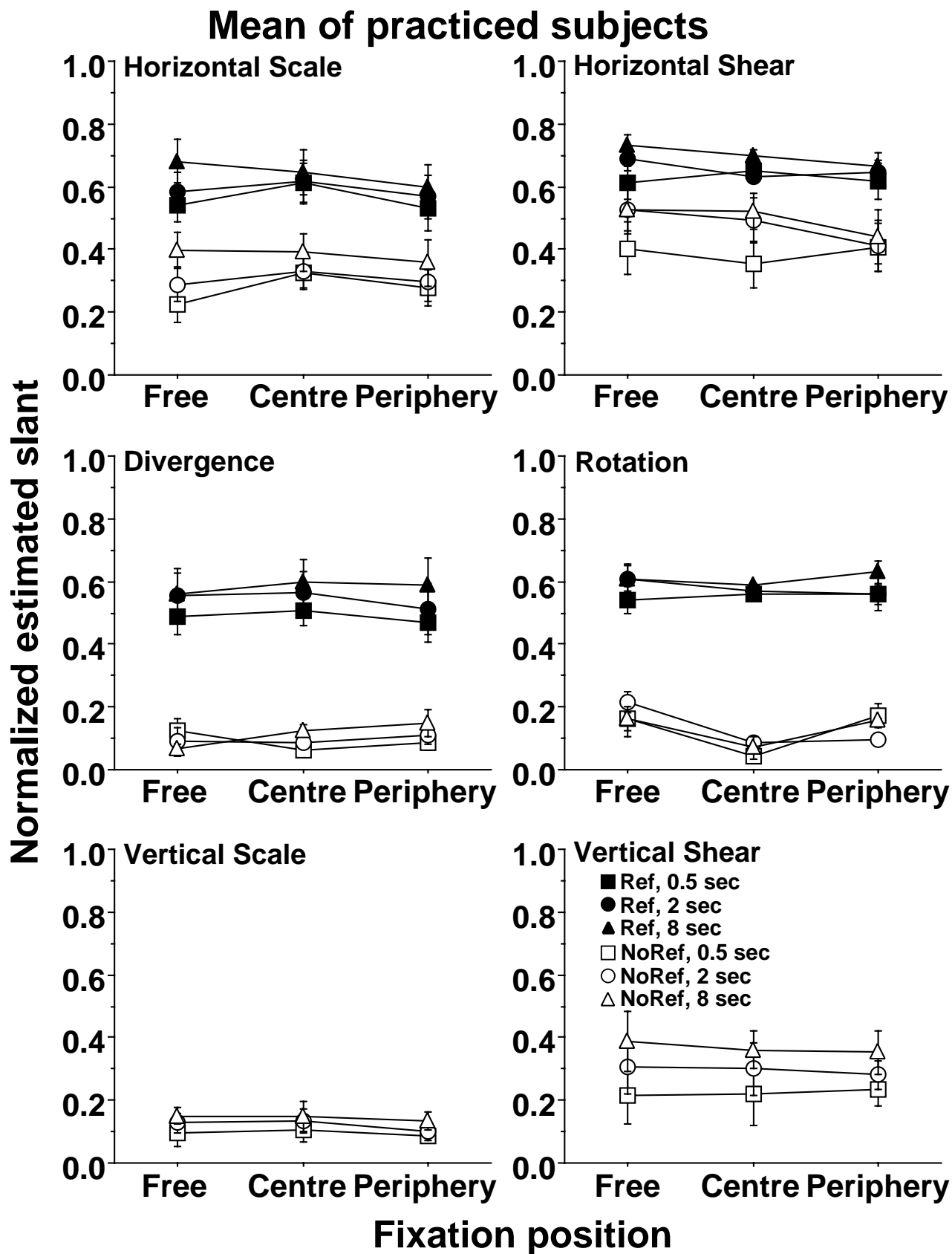
Mean of practiced subjects

Fig. 4: Normalized estimated slant versus the fixation position for the six transformations and the three presentation durations, both with (Ref) and without (NoRef) the visual reference. The data are the means of the three practiced subjects. The error bars represent the cross-subject standard error for the mean values of the subjects.
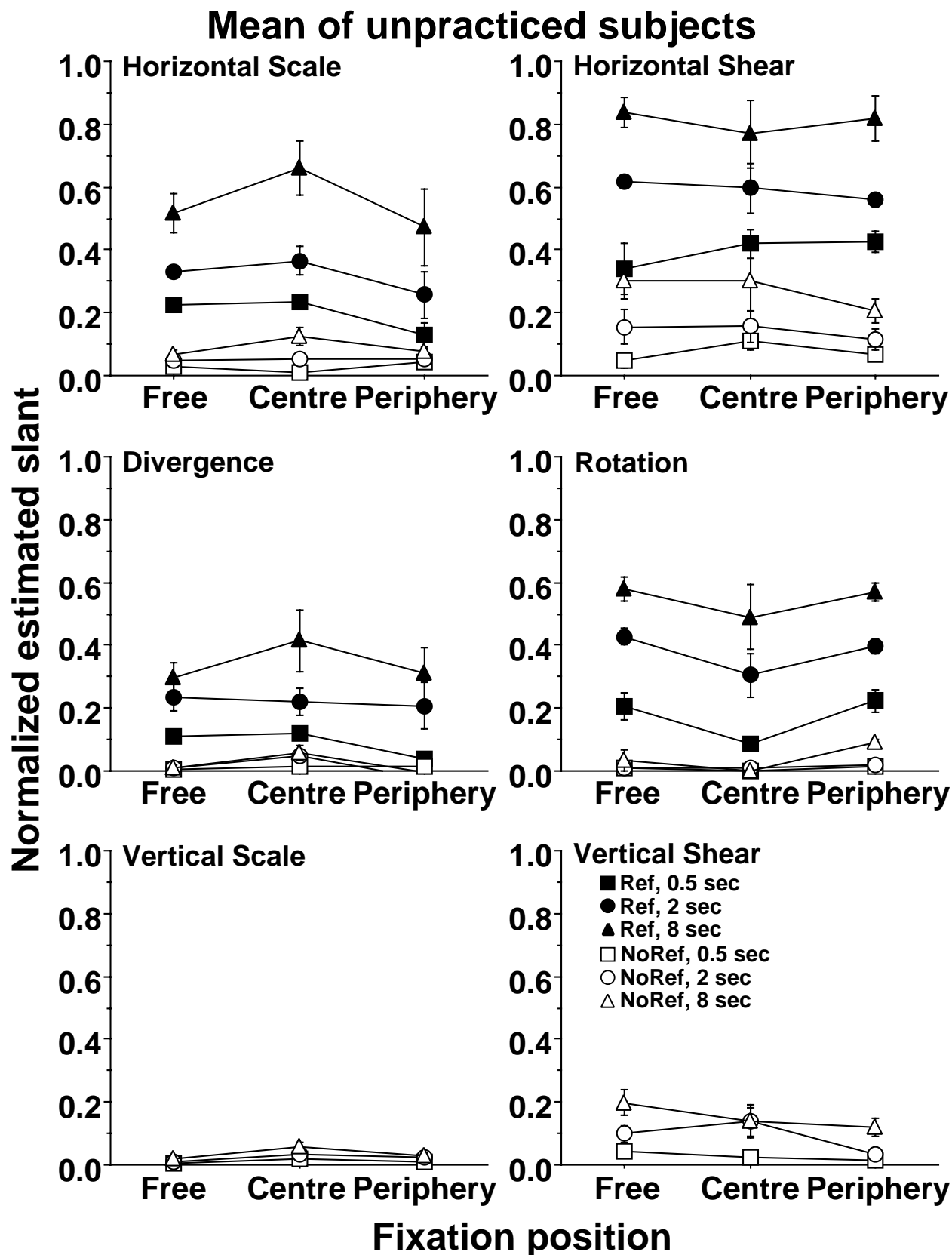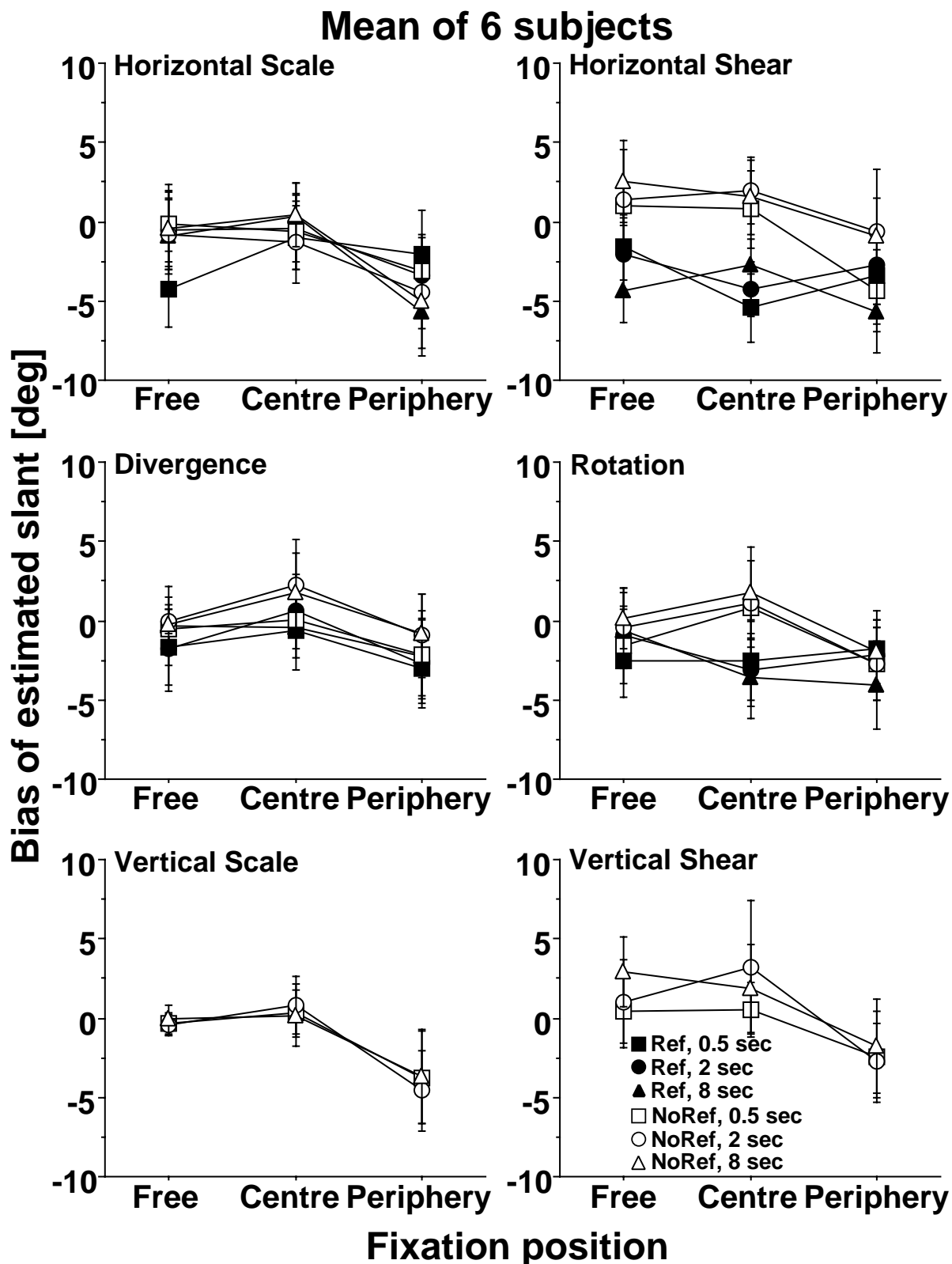
Mean of unpracticed subjects

Fig. 5: Same as figure 4, but for the three unpracticed subjects.

Fig. 6: Bias of estimated slant versus the fixation position for the six transformations and the three presentation durations, both with (Ref) and without (NoRef) the visual reference. The data are the means of the six subjects. The error bars represent the standard error across the six subjects.

Figure 6 shows the biases in the slant estimations across the six subjects. Contrary to the results of the slant estimates, the biases could not be clearly divided into a practiced and unpracticed group. From the large error bars it can be seen that the individual subject data differed considerably. Subjectively, the biases in slant tend to be larger when fixation is in the periphery. However, in our results this tendency is not significant. More precise studies (many more repetitions per trial) are necessary to check for the presence of this tendency. If there is a bias, it is in the opposite direction from that predicted in figure 1. If the bias was caused by an effect of slant estimation relative to the cyclopean line of sight we expect a positive bias in the peripheral fixation condition only. The absence of a bias would be consistent with the recently published model of Erkelens and van Ee (1998) in which disparity is processed in headcentric coordinates.

That a negative bias is present in the free-fixation condition is probably an artifact of the set up. On average the part of the surface that was perceived away from the observer (behind the screen) had a slightly smaller slant than the part of the surface that was nearer to the observer (in front of the screen). However, although the effect of the artifact is consistent, it is smaller than about 2.5 deg.

Uncrossed   Crossed
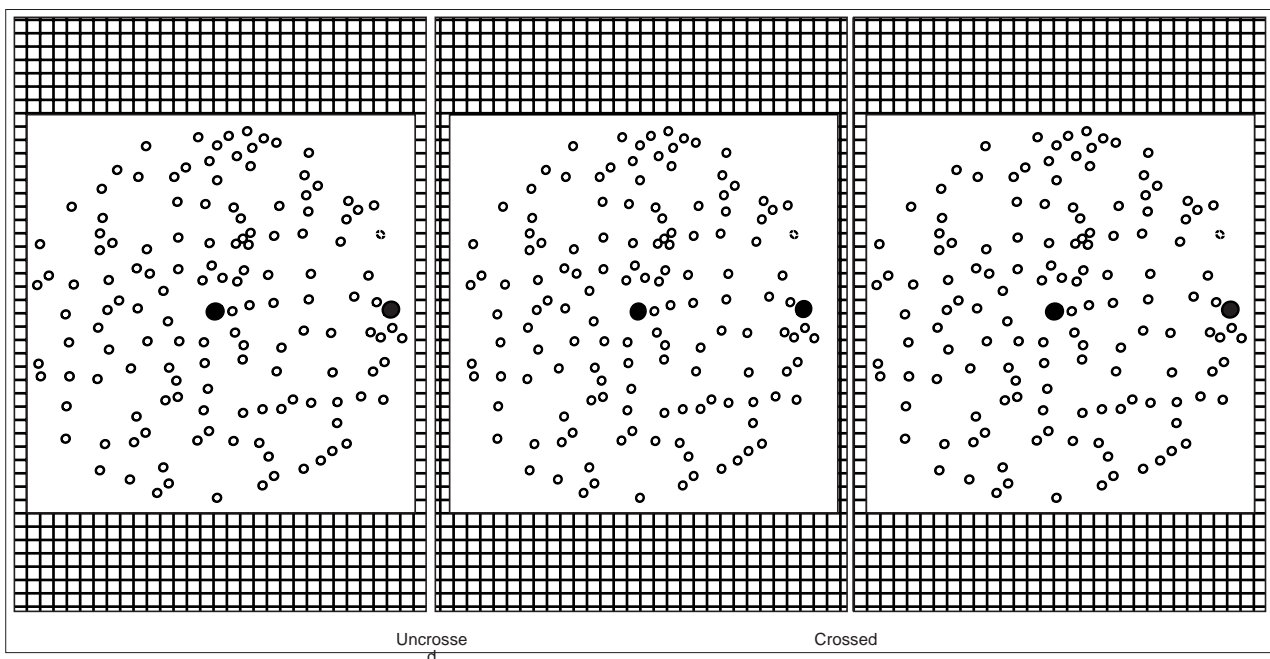
Fig. 7: Stereogram depicting an example of a slanted plane about the vertical axis. In this demonstration a horizontal scale of 6 % (the maximum transformation that was applied in our experiment) has been applied between the two half-images of the circle pattern. Observers who have best fusion when their eyes are crossed (which means that the half-image on the right side is seen by the left eye), should fuse the two half-images on the right side of this figure; uncrossed fusers should view the two images on the left. After fusion the right side of the circle pattern will be seen in front of the plane of the paper (negative angle of slant). The black dots mark the approximate location of fixation in the case of central fixation and peripheral fixation. This demonstration supports our statement that in our experiment the entire circle pattern was perceived to be slanted and that observers did not base their slant judgments on just the central texture area. Slant estimation does not require a sharp image of the pattern elements throughout the visual field (and consequently does not require saccades) because a sharp image is not a prerequisite for the perception of gross ordering effects of objects that are relatively near to each other in lateral direction. On the other hand, it can be seen that counting the number of circles in a cluster in the periphery is a difficult task which requires a sharp image and consequently saccades. [This demonstrating gives an acceptable idea of how the stimulus looked like in the experiment but is far from being accurate. The most accurate replication of the experimental situation is obtained if one views this stereogram from a very short distance. The viewing distance has to be as short as 6 cm to create a stimulus with a diameter of 40 deg. But, on the other hand, for this short distance the geometrically presented slant is very small in case of a 6 % horizontally scaled pattern because slant is geometrically proportional to the viewing distance. An unfortunate property of free-fusion stereograms is that they require eye movements in order to bring the two half-images into register. Maintaining fixation for near viewing in this free-fusion stereogram is also much more

13

difficult than in the real experiment where we utilized red/green images. In the real experiment the percept was stable and fusion was achieved quickly (see also Gillam et al., 1988).]

4   Discussion

We investigated whether large scanning eye movements explain the build-up of perceived slant of large surfaces over time. While we cannot absolutely exclude occasional unintended large scanning eye movements as a contributor to our results, on the basis of subjective impressions during participation in pilot experiments we consider it unlikely that these eye movements acted as significant contributors to our results. Thus, our results provide no support for the hypothesis that the time used for the execution of large scanning eye movements---and consequently sequential stereopsis as suggested by Enright---explains the build-up of estimated slant with the duration of the stimulus presentation. Apparently, the advantage mediated by fixation shifts on distance discrimination of discrete objects (Wright, 1951; Enright, 1991) does not extend to the estimation of the orientation of a slanted surface rendered with adjacent elements.

We emphasize that our results do not contradict the results of Wright and Enright. They studied a different aspect of the visual system. Their task required a sharp image of the targets of which the relative distance was discriminated. Saccades are essential for the accurate recognition of many types of visual targets because visual details are resolved best when imaged in the fovea. Sharp images are not a prerequisite in order to perceive gross ordering effects of objects that are relatively near to each other in lateral direction. Apparently our slant estimation task does not require a sharp image of the pattern elements throughout the visual field and consequently does not require saccades. Figure 7 provides a supporting demonstration for this statement. We stress that our results were obtained with plane surfaces and considerable slants. If the surface were to be non-planar, the shape based on the sharp image in the central visual field would not be generalizible to the whole field and shape perception would probably require saccades. Similarly, if the slant were to be very small, for instance below threshold, it may well be recoverable from information provided by scanning eye movements.

Tasks in which eye movements are of help in vision include establishing fusion and segregation. Prior to fusion, eye movements are essential in order to bring the two disparate retinal images into register (Julesz, 1971). In order to segregate complex textures into target and background during visual search, saccades are important instead of attentional shifts (He & Kowler, 1992). Future research is necessary to investigate the role of attentional shifts in stereoscopic slant estimation.

The increase in perceived slant over time appears to be an intrinsic property of disparity processing (stereopsis). This is in accordance with the model recently proposed by van Ee and Erkelens (1996b) which provides a possible explanation for the observations in the literature that whole-field disparity gradients, such as horizontal scale and horizontal shear, do not evoke vivid perception of slant. They found that the disparity fields caused by whole-field horizontal scale and shear are similar to the disparity fields brought about by head rotations which means that these disparity fields are ambiguous to interpret if the visual system attempts to calculate the slant of a surface relative to the body: During navigation the visual system should determine whether the whole-field disparity gradient is caused by the slanted surface or by a head rotation. This determination introduces an extra source of noise which decreases the reliability of whole-field disparity gradients. According to van Ee and Erkelens' model, the weight given to disparity cues relative to non-stereo cues [such as perspective, texture (e.g. Johnston, Cumming & Parker, 1993; Frisby, Buckley, Wishart, Porrill, Gårding & Mayhew, 1995) etc.] is smaller for whole-field disparity gradients[9] than for disparity gradients in the presence of a visual reference. A plausible reason for the increase in perceived slant over time is that the visual system needs time to overrule conflicting non-stereo cues, which are usually present in stereoscopic computer displays (e.g. Ryan & Gillam, 1994).

---

[9] Strictly speaking, our model applies to disparity gradients on a computer screen and not on the retinas.

5   References

Brenner, E. & van Damme, W. J. M. (1998) Judging distance from ocular convergence. Vision Research, 38, 493-498.

Collewijn, H. & Erkelens, C. J. (1990). Binocular eye movements and perception of depth. In Kowler, E. (Ed.), Eye movements and their role in visual and cognitive processes (pp 213-261). Amsterdam: Elsevier Science Publishers B.V.

Enright, J. T. (1990). Stereopsis, cyclotorsional "noise" and the apparent vertical. Vision Research, 30, 1487-1497.

Enright, J. T. (1991). Exploring the third dimension with eye movements: better than stereopsis. Vision Research, 31, 1549-1562.

Enright, J. T. (1994). Saccade-Vergence interactions? An evolutionary perspective on unbalanced saccades. In Ygge, J., Lennerstrand, G. (Eds.), Eye movements in reading (pp 117--134). Pergamon Press.

Enright, J. T. (1996). Sequential stereopsis: a simple demonstration. Vision Research, 36, 307-312.

Erkelens, C. J. & Collewijn, H. (1985a). Motion perception during dichoptic viewing of moving random-dot stereograms. Vision Research, 25, 583-588.

Erkelens, C. J. & Collewijn, H. (1985b). Eye movements and stereopsis during dichoptic viewing of moving random-dot stereograms. Vision Research, 25, 1689-1700.

Erkelens, C. J. & van Ee, R. (1998). A computational model of depth perception based on headcentric disparity. Vision Research, 38, 2999-3018.

Foley, J. M. (1980). Binocular distance perception Psychological Review, 87, 411-434.

Frisby, J. P., Buckley, D., Wishart, K. A., Porrill, J., Gårding, J. & Mayhew, J. E. W. (1995). Interaction of stereo and texture cues in the perception of three-dimensional steps. Vision Research, 35, 1463-1472.

Gillam, B., Chambers, D. & Russo, B. (1988). Postfusional latency in stereoscopic slant perception and the primitives of stereopsis. Journal of Experimental Psychology: Human Perception and Performance, 14, 163-175.

Gillam, B., Flagg, T. & Finlay, D. (1984). Evidence for disparity change as the primary change as the primary stimulus for stereoscopic processing. Perception & Psychophysics, 36, 559-564.

Gogel, W. C. (1956). The tendency to see objects as equidistant and its inverse relation to lateral separation. Psychological Monographs, 70, Whole No. 411, 1-17.

Gogel, W. C. (1961). Convergence as a cue to absolute distance. Journal of Psychology, 52, 287-301.

He, P. & Kowler, E. (1992). The role of saccades in the perception of texture patterns. Vision Research, 32, 2155-2163.

Howard, I. P. & Kaneko, H. (1994). Relative shear disparities and the perception of surface inclination. Vision Research, 34, 2505-2517.

Howard, I. P. & Rogers, B. J. (1995). Binocular vision and stereopsis. Oxford: Oxford University Press.

Howard, I. P. & Templeton, W. B. (1964). The effect of steady fixation on the judgment of relative depth. The quarterly journal of experimental psychology, 16, 193-203.

Johnston, E. B., Cumming, B. G. & Parker, A. J. (1993). Integration of depth modules: Stereopsis and texture. Vision Research, 33, 813-826.

Julesz, B. (1971). Foundations of cyclopean perception. Chicago: University of Chicago Press.

Julesz, B. (1978). Global stereopsis: cooperative phenomena in stereoscopic depth perception. In Handbook of Sensory Physiology, Vol VIII, Perception, 215-256. Springer, Berlin.

Julesz, B. (1986). Stereoscopic vision. Vision Research, 26, 1601-1612.

Kaneko, H. & Howard, I. P. (1996). Relative size disparities and the perception of surface slant. Vision Research, 36, 1919-1930.

Ogle, K. N. (1953). The role of convergence in stereoscopic vision. Proceedings of the Physical Society of London, B, 66, 513-514.

Ogle, K. N. (1956). Stereoscopic acuity and the role of convergence. Journal of the Optical Society of America, 46, 269-273.

Patterson, R. & Fox, R. (1984). Stereopsis during continuous head motion. Vision Research, 24, 2001-2003.

Rady, A. A. & Ishak, I. G. H. (1955). Relative contributions of disparity and convergence to stereoscopic acuity. Journal of the Optical Society of America, 45, 530-534.

van Rijn, L. J., van der Steen, J. & Collewijn, H. (1994). Instability of ocular torsion during fixation: cyclovergence is more stable than cycloversion. Vision Research, 34, 1077-1087.

Ryan, C. & Gillam, B. (1994). Cue conflict and stereoscopic surface slant about horizontal and vertical axes. Perception, 23, 645-658.

Steinman, R. M. & Collewijn, H. (1980). Binocular retinal image motion during active head rotation. Vision Research, 20, 415-429.

Steinman, R. M., Levinson, J. Z., Collewijn, H. & van der Steen, J. (1985). Vision in the presence of known natural retinal image motion. Journal of the Optical Society of America A, 2, 226-233.

van Ee, R., Backus, B. T. & Erkelens, C. J. (1996). Perceptual learning in stereoscopic slant estimation. Internal Report of the Helmholtz Institute, UU-PA-pmi-088.

van Ee, R. & Erkelens, C. J. (1995). Binocular perception of slant about oblique axes relative to a visual frame of reference. Perception, 24, 299-314.

van Ee, R. & Erkelens, C. J. (1996a). Temporal aspects of binocular slant perception. Vision Research, 36, 43-51.

van Ee, R. & Erkelens, C. J. (1996b). The stability of stereoscopic depth perception with moving head and eyes. Vision Research, 36, 3827-3842.

van Ee, R. & Erkelens, C. J. (1998). Temporal aspects of stereoscopic slant estimation: An evaluation and extension of Howard and Kaneko's theory. Vision Research, 38, 3871-3882.

Westheimer, G. (1979). Cooperative neural processes involved in stereoscopic acuity. Experimental brain Research, 18, 893-912.

Westheimer, G. & McKee, S. P. (1978). Stereoscopic acuity for moving retinal images. Journal of the Optical Society of America, 68, 450-455.

Wright, W. D. (1951). The role of convergence in stereoscopic vision. Proceedings of the Physical Society of London, B, 64, 289-297.

Wright, W. D. (1953). A reply to Ogle (1953). Proceedings of the Physical Society of London B, 66, 514-515.