Raymond van Ee and Casper J. Erkelens

Perception 24, 299-314, 1995.

Please send correspondence to the firs author:
Utrecht University, Vakgroep Fysica v/d Mens,
Princetonplein 5, NL–3584CC, Utrecht, The Netherlands;
Email: r.vanee@phys.uu.nl;
Fax: +31 30 522664

Abstract

From the literature it is known that the processing of disparity for slant is differe t in the presence and in the absence of a visual frame of reference. We elaborate the experimental findin that vertical disparity is not processed for slant perception in the presence of a visual reference. This theoretical analysis results in a reduction of the three basic first–orde transformations between the retinal half–images (Koenderink's divergence, rotation and deformation) to only two basic orthogonal transformations. The firs of these, horisontal scale, results in slant perception about the vertical axis, whereas the second, horisontal shear, results in slant perception about the horisontal axis. These transformations are based primarily on horisontal disparity. We show experimentally that in the presence of a frame of reference the amount of vertical transformation that is added to the two basic transformations (horisontal scale and shear) of a random–dot stimulus is indeed irrelevant for slant perception. We suggest that in the presence of a visual reference slant perception about oblique axes is based solely on linear combinations of the horisontal scale and horisontal shear transformations. Subjects are able to reproduce slants about oblique axes experimentally merely by combining horisontal scale and shear.

Introduction

In stereoscopic vision both eyes view slightly differe t aspects of visual space. Generally, the image of a stimulus on one retina can be regarded as a mathematical transformation of the image on the other retina. The difference between the two retinal projections of objects, which are determined by the geometry of binocular vision, are called binocular disparities or, in short, disparities. These disparities are sources of depth information.

The theoretical classificatio of the possible mathematical transformations between the retinal images (to first–orde approximation in spatial difference was fir worked out by Koenderink and Van Doorn (1976). They examined what kind of disparity information is in principle available for the computation of slant. They were able to decompose the first–orde disparity fiel into divergence, rotation and deformation and found that the deformation component specifie the gradient of the reciprocal distance. Their theory is attractive because it permits any transformation to be described solely in terms of divergence, rotation and deformation. In essence, their theory is a computational theory which can be used to derive depth from disparity field obtained from matching pictures taken by two cameras. Nevertheless, the theory has been applied frequently to human binocular vision and has been used to interpret experimental results.

It is important to verify whether Koenderink and Van Doorn's theory is in fact applicable to human vision. Gillam and Rogers (1991) recently investigated induced slant about the horizontal axis for stimuli relative to a visual frame of reference. They concluded that, contrary to Koenderink and Van Doorn's theory, perceived slant was not related to the deformation present but was predicted by the orientation disparity at the vertical meridian per se. From the study by Gillam and Rogers it is clear that the theory of Koenderink and Van Doorn does not always predict perceived slant correctly. However, it is still not certain whether Koenderink and Van Doorn's theory is applicable to human slant perception in the case of stimuli in which orientation disparity is of minor importance. Another problem with their theory is that it does not incorporate the asymmetry between horizontal and vertical disparity in the human visual system.

In the theory of Koenderink and Van Doorn (1976) horizontal and vertical disparity are equally important. However, both neuro–physiological and psychophysical studies show a strong anisotropy in human vision with respect to disparity in horizontal and vertical dimension. Stereopsis is based primarily on horizontal disparity but the role of vertical disparity in stereopsis is not entirely clear. In the literature there are clear reports about the ability of the human visual system to use vertical disparity for three–dimensional (3D) perception (e.g. Ogle 1950). According to a number of authors the role of vertical disparity could be to scale horizontal disparities for viewing distance. Recently, Rogers and Bradshaw (1993) have shown experimentally that in the case of large–fiel stimuli manipulations of vertical disparity do indeed in uence the perceived distance. However, there are also reports that vertical disparity is not used for 3D perception (Westheimer

1978; Cumming et al 1991; Sobel and Collett 1991). It may well be that vertical disparity is used for 3D perception, but only when very large retinal images are involved. It is possible that vertical disparity is not used for slant perception in the case of relatively small stimuli in the presence of a visual reference.

The influence of a visual reference

From eye movement studies it is also known that stereopsis is different with and without a visual frame of reference. A shift between the two parts of a stereogram relative to each other, without a reference, gives rise to vergence eye movements but not to perception of motion in depth (Erkelens and Collewijn 1985a, 1985b). By contrast, the same shift, in the presence of a visual frame of reference, gives rise to vivid perception of motion in depth. From these studies it can be concluded that absolute disparity is not a sufficient cue for stereopsis (for a review see Collewijn et al 1991a, 1991b).

Recently, we presented evidence that perception of slant is also affected by the presence or absence of a visual reference (Erkelens and van Ee 1993; van Ee and Erkelens 1993, 1994). We showed that vertical scaling or shearing of a half-image of a random-dot stereogram induced slant perception clearly in the absence of a visual reference but poorly in the presence of a visual reference (van Ee and Erkelens 1994). We also presented evidence that the time required for making reliable slant judgments due to horizontally scaled or sheared stimuli is of the order of hundreds of milliseconds with visual reference but is of the order of seconds (ten times as long) without visual reference.

The present paper is restricted to slant perception in the presence of a visual reference. We first examine how a strong asymmetry in the use of horizontal and vertical disparity affects the classification of the basic transformations (divergence, rotation, deformation). We elaborate the assumption that vertical disparity is irrelevant for slant perception relative to a visual frame of reference. We show that our assumption implies that slant can be described by a combination of two orthogonal transformations only: horizontal scale and horizontal shear. In two experiments we show that horizontal scale and horizontal shear are indeed sufficient to describe perception of slant about any axis relative to a visual frame of reference.

Theory

The disparity field (the angular relations between binocular visual directions) can be described theoretically by a vector field defined on the manifold of visual directions. The disparity function can be decomposed mathematically into elementary components up to any order by a Taylor expansion with respect to the position. The zero-order component of this expansion is the disparity value itself and represents translation of the retinal images relative to each other. The first-order component gives the rate of change in disparity (gradient). The second-order component gives the curvature of the disparity function.

Figure 1 about here.

The first-orde Taylor approximation of the disparity fiel can be decomposed into three elements (Koenderink and Van Doorn 1976). The elementary components are: divergence, rotation and deformation (see appendix A for the derivation). Divergence is identical to uniform scaling, i.e. uniform expansion or contraction. Deformation is a linear combination of expansion and contraction in orthogonal directions with conservation of area. Figure 1 shows the elementary first-orde transformations.

Figure 2 about here.

Two elementary first-orde transformations, which are commonly referred to in the literature (and which we will use too), are non-uniform scaling and shear in vertical or horizontal directions. Non-uniform scaling is a linear combination of deformation and divergence. Shear is a linear combination of deformation and rotation and is generally not a pure deformation. Examples of horizontal and vertical scale and shear are shown in figur 2.

Figure 3 about here.

If vertical disparity is irrelevant for slant perception when a visual frame of reference is present, then this has major consequences for the classificatio of the elementary transformations that can be used to describe slant perception. According to figur 3 this irrelevance suggests that 1) rotation and horizontal shear effecti ely induce similar slants (provided they contain the same amount of horizontal disparity), 2) divergence effecti ely induces the same slant as horizontal scale and 3) vertical scale and vertical shear will not induce slant. The proposition (2) that in the presence of a visual reference divergence induces the same slant as does horizontal scale is demonstrated in figur 4.

Figure 4 about here.

More specificall , we suggest that perceived slant relative to a visual frame of reference is related to a set of transformations of which only horizontal scale and horizontal shear are the basic orthogonal elements. Horizontal scale is associated with slant about the vertical axis of the stimulus. Horizontal shear is associated with slant about the horizontal axis. The transformations horizontal scale and horizontal shear are orthogonal in a mathematical sense (see appendix B for a proof of the orthogonality). This orthogonality means that these two transformations form a complete set. A complete set implies that the horizontal component of any transformation relative to an arbitrary axis (for instance an expansion in the 45 deg direction) can be described solely in terms of horizontal scale and horizontal shear.

Experiments

The experiments reported here are designed for two purposes. First of all, we want to examine our suggestion that the three transformations (divergence, rotation and deformation) reduce to only two elementary transformations (horizontal scale and horizontal shear) in the case of disparity processing relative to a visual frame of reference. In the real world slants are usually not about the horizontal or vertical axis but are about oblique axes. Our theoretical results concerning the orthogonality of horizontal scale and horizontal shear form the basis of a model describing slant perception about oblique axes which is based solely on linear combinations of the two basic transformations. Therefore, we also want to investigate whether linear combinations of the latter transformations are sufficie t to describe slant about oblique axes.

Figure 5 about here.

The stimuli are generated at a frequency of 70 Hz by an HP 750 graphics computer. Subsequently, the stimuli are back–projected on a fronto–parallel translucent screen by a projection TV (Barco Data 800). The subject is seated about 1.5 m from the screen. One image is projected on the screen in green light and is observed by the right eye through a green filte: A red filte is used to make the other image visible exclusively to the left eye. The transmission spectra of the filte1 (anaglyph glasses, Schott Tiel, the Netherlands) are chosen such that they correspond as far as possible to the emission spectra of the projection TV. No crosstalk between the right and left eye views is observed when contrast and brightness of the projection TV are correctly adjusted. Figure 5 shows the experimental set–up. The stimuli are viewed in a completely dark room. Neither the screen (or its boundaries) nor other objects in the room are visible.

Figure 6 about here.

The stimulus contains two random–dot patterns and a reference pattern. Schematic drawings of the stimulus are shown in figur 6. The subjects are asked to match the slant of pattern 2 (by manipulating the computer–mouse position) to the pre-set slant of pattern 1. The whole–fiel visual reference (width 69 deg and height 56 deg) consists of a cross–hatched pattern. The sizes of the two random–dot patterns are slightly differe t to prevent subjects from using the size of the patterns in their judgments. The subjects are explicitly asked to match parallelism (slant) and not to match maximum disparities, which are larger for the larger pattern. The size of pattern 1 is: width 9.6 deg and height 7.5 deg. The size of pattern 2, which contains a random-dot pattern similar to that of pattern 1, is: width 8.9 deg and height 6.8 deg. The dot diameter (0.2 deg) and the dot density (1 dot/deg) are the same in both patterns. The reference pattern contains a window (width 10.1 deg and height 21.6 deg) to minimize the in uence of a possible depth contrast effec (Werner 1938)[1]. The checkered whole–fiel reference pattern consists of a fiel of adjacent squares with diagonals of 7.2 deg. Since the test and match stimuli are small relative to the fixe visual frame of reference, the experimental set–up discourages torsional eye movements (see Kertesz (1991) for an analysis of the minimum dimensions of stimuli to drive torsion).

Experiment 1

Methods Ten subjects (8 males and 2 females, ages 23–52 years) take part in the experiment. None of them shows any visual or oculomotor pathology except for one subject who shows refraction anomalies which are corrected by his own glasses. Four of the subjects are experienced in stereoscopic experiments and no subjects (except for the authors) have been informed about the purposes of the experiment.

The half–image of pattern 1 viewed by the left eye is transformed relative to the half–image viewed by the right eye. The following transformations are presented: horizontal scale (-6 % – 6 %, step–size 2 %), vertical scale (-6 % – 6 %, step–size 4 %), horizontal shear (-3 deg – 3 deg, step–size 1 deg)[2], vertical shear (-3 deg – 3 deg, step–size 2 deg), divergence (-6 % – 6 %, step–size 2 %) and rotation (-3 deg – 3 deg, step–size 1 deg). These

[1] The depth contrast eect could result in an undesired slant of the reference pattern at the place of the random{dot stimulus due to the disparity of the random{dot stimulus. An undesired depth contrast eect is also to be expected between pattern 1 and 2 during the matching procedure. However, the better the matching, the less the depth contrast eect, because nally both patterns are parallel at equal depth.

[2] This means that the red and the green vertical contours (or horizontal contours in the case of vertical shear) of the stimulus were rotated over -3 deg to 3 deg relative to each other by the shear operation; see also g. 2. By the amount of shear of a stimulus we do not mean the perceptually induced slant, which can be tens of deg.

amounts of transformation comprise more or less the entire range of fusible disparities. Fusion problems are therefore prevented.

When pattern 1 contains horizontal scale, vertical scale or divergence, which are transformations that are normally associated with slant about the vertical axis, pattern 2 is presented below pattern 1 (as shown in fig. 6a). In such cases the subject operates the mouse in order to control the horizontal scale (slant about the vertical axis) of the half-images of pattern 2. When pattern 1 contains horizontal shear, vertical shear or rotation, pattern 2 is presented to the left of pattern 1 (fig. 6b) because these transformations are normally associated with slant about the horizontal axis. This time the mouse is used to control the horizontal shear (slant about the horizontal axis) of the half-images of pattern 2. The fact that patterns 1 and 2 are presented adjacently along the direction of rotation makes it easier to match the parallelism of both patterns, which in turn prevents undesired influences of the depth contrast effect (see also footnote 1).

The subjects should make their decisions within 5 seconds. The pre-set slants are presented in random order. A series of trials consists of 36 presentations. Each subject views five series without feedback. Between stimuli the screen is blanked for two seconds. The subjects are not restricted with ragard to their head or eye movements.

Results In our experiments the subject always obtains a stable percept of slant after the presentations of the stimulus without latencies. Furthermore, depth contrast effects are successfully prevented. No significant differences are found between subjects. The results of experiment 1, averaged over the ten subjects, are presented in figure 7. Each pre-set vertical scale (fig. 7a) and vertical shear (fig. 7b) is matched by horizontal scale of about 0 % (which means no horizontal scale) and horizontal shear of about 0 deg, respectively. This means that under our experimental conditions neither vertical scale nor vertical shear induces perception of slant. Pre-set divergences are matched by similar percentages of horizontal scale. This means that divergence induces effectively the same slant about the vertical axis as does horizontal scale, with the same amount of horizontal transformation (fig. 7a). When the horizontal disparity is similar, rotation induces effectively the same slant about a horizontal axis as does horizontal shear (fig. 7b).

Figure 7 about here.

Three subjects repeated the experiment several times at intervals of a week. There were no significant differences in the results. To check whether we were successful in discouraging torsional eye movements we additionally recorded these movements for two of our subjects while they performed the entire matching experiment. For this purpose we ran another series of stimuli with an exposure time of 20 seconds. Torsional eye movements are measured by the three-dimensional scleral-coil technique as described by Ferman et al (1987). We did not fin any correlation between cyclovergence responses and transformations of the stimuli.

Experiment 2

Methods In the former experiment we investigated the consequences of the irrelevance of vertical disparity (when a frame of reference is present) for the classificatio of transformations. This irrelevance leads theoretically to a set of two basic transformations, horizontal scale and shear, which in principle can serve to describe slant about oblique axes. We proceed by investigating whether linear combinations of horizontal scale (-6 % – 6 %, step-size 2 %) and horizontal shear (-3 deg – 3 deg, step-size 1 deg) are perceived as slant about oblique axes. The matching procedure of experiment 1 is repeated with three subjects. Two of them are experienced in stereoscopic experiments. The experimental set-up of figur 6a is used. The subjects are asked to match the slant of pattern 2 to the the pre-set slant of pattern 1 by operating the mouse. This time the horizontal position of the computer mouse represents slant about the horizontal axis of pattern 2. The vertical position represents slant about the vertical axis of pattern 2. A series of trials consists of 49 presentations. Each subject views three series without feedback.

Results The results of experiment 2, averaged over the three subjects, are presented in table 1. Again no significa t difference are found between subjects. Each pre-set combination of horizontal scale (firs column of table 1) and horizontal shear (firs row of table 1) is matched by a combination of horizontal scale and horizontal shear set by the subjects. The mouse positions selected by the subject are converted to measured horizontal scale and horizontal shear.

Table 1 about here.

The results indicate that perceived slants about arbitrary axes are uniquely related to linear combinations of horizontal scale and horizontal shear.

Discussion

Taken together, our results show that human perception of slant about oblique axes relative to a visual frame of reference depends on the combination of only two orthogonal mathematical transformations between the half-images of a stereogram. The firs one is *horizontal scale* which is associated with slant about the *vertical axis*. The second transformation is *horizontal shear* which is associated with slant about the *horizontal axis* of the stimulus. Combinations of horizontal scale and horizontal shear represent horizontal disparity gradients in oblique directions and are therefore associated with slant about oblique axes.

The theory of Koenderink and Van Doorn

The theory about the relationship between induced slant and the geometry of binocular vision, which is developed by Koenderink and Van Doorn (1976), is in essence a computational theory. The theory can be applied to artificial vision but has also been used as a basis for developing experiments concerning human vision. It is common practice to compare experimental results with the theory of Koenderink and Van Doorn. However, the theory does not hold[3] for human perception of slant relative to a visual frame reference because vertical disparity is irrelevant for perception of slant when a frame of reference is present. Vertical disparity is intrinsically present in transformations like divergence and rotation. The irrelevance of a vertical disparity gradient in the presence of a visual reference means that for perception of slant: 1) divergence is effectively identical to horizontal scale, 2) rotation is effectively identical to horizontal shear and 3) vertical scale and vertical shear do not induce slant.

A visual frame of reference

Our results are different from several reported results including the results of Ogle (1950) and more recently the results of Rogers (1992) and Howard and Kaneko (1993). Before discussing the experimental results of other authors we will distinguish between disparity processing with and without a visual frame of reference.

One reason why we have to take into account the role of a visual reference (like for instance a stimulus background, a dimly lit room or the boundaries of a projection screen) is that disparity processing for depth is different with and without a frame of reference (Erkelens and Collewijn 1985a, 1985b; Howard and Zacher 1991; Erkelens and van Ee 1993; van Ee and Erkelens 1993, 1994). The study of perception of depth has been dominated by a psychophysical approach, whereas oculomotor behaviour has been more often inferred than adequately measured. As a result several authors have confused absolute and relative disparity and have failed to recognize the influence of a visual frame of reference (for a review see Collewijn et al 1991a, 1991b). Erkelens and Collewijn (1985a, 1985b) found that disparity without a visual frame of reference (that is, absolute disparity or the vergence angle of the eyes) is not a cue for perception of motion in depth, whereas it is a cue with a visual frame. Howard and Zacher (1991) found that

[3]We fully support the remark of a reviewer `that there is nothing wrong with the fact that some good theories do not work under certain conditions'.

cyclodisparity relative to a visual reference, not absolute cyclodisparity, is a cue for slant perception. There have been indications (Erkelens and van Ee 1993; see also Gillam et al 1988b; Stevens and Brooks 1987) that linear transformations between the entire half-images of a stereogram without a visual reference elicit perception of slant less succesfully than these transformations with a visual reference. Very recently, we have shown that reliable judgments of slant require observation periods (latencies) of the order of hundreds of milliseconds in the presence of a visual reference but about ten times as long in the absence of a reference (van Ee and Erkelens 1994).

The presence of a visual reference means that there are disparity relations between differe t stimuli. On the other hand, the absence of a visual reference means that the whole retinal image is subjected to the transformation (and that eye–movements like cyclovergence are dependent on the stimulus orientation). The reason for the differenc in stereopsis with and without a visual reference is not entirely clear. A possible reason is that disparity without visual reference is less reliable because the disparity could be caused by eye movements or head movements. This implies that the processing of disparity without visual reference requires compensation for eye and head movement–induced disparity. In the presence of a visual reference, on the other hand, disparity relations between the stimulus and the visual reference are independent of eye or head movements and thus invariant. We restricted our study to perception of slant of a stimulus relative to a visual frame of reference. The presence of invariant disparity relations between stimulus and reference may be the reason why in our experiment subjects obtain stable depth perception without latencies (see also Gillam et al 1988b). Slant perception without a visual reference takes a few seconds (van Ee and Erkelens 1994). This latency could be caused by a recalibration of stereopsis due to extra–retinal signals about the eye and head position. The distinction into conditions with and without a visual frame of reference helps us to compare our experimental results with other reports.

Slant perception without visual reference

Ogle (1950) found that uniform divergence of retinal images relative to each other (uniform aniseikonia) does not lead to slant perception. When we remove the visual reference in our experiment slant perception does indeed vanish in the case of divergence. Thus, our results do not contradict the results of Ogle (1950). Rogers (1992) and Howard and Kaneko (1993) investigated horizontal shear, vertical shear and rotation of the right retinal image relative to the left retinal image. Unlike us these authors investigated slant perception with large (75 75 deg or more) stimuli and in the absence of a visual frame of reference (cyclovergence was therefore possible). Rogers (1992) concluded that vertical disparities of corresponding elements close to the horizontal meridian are used to drive cyclovergent eye movements, whilst horizontal disparities close to the vertical meridian are used as a source of information about the 3D shape of surfaces. Howard and Kaneko

(1993) suggest that the differenc between the horizontal shear and the vertical shear of the retinal images is the primitive for perception of slant about the horizontal axis. In the case of disparity processing without a frame of reference, conclusive claims about the validity of Koenderink and Van Doorn's theory are premature because cyclovergence can contribute to the perceived slant about the horizontal axis. Rogers (1992) and Howard and Kaneko (1993) did indeed report cyclotorsion. Cyclotorsion is important because it can contribute to the perceived slant about the horizontal axis. Our experimental design differ from the design of Howard, Kaneko and Rogers in that it discourages cyclotorsion. Control measurements showed that none of our stimuli in fact induced cyclotorsion. Howard and Kaneko repeated their experiment in the presence of a visual reference and confirme (Howard, personal communication, Aug. 1993) our results (van Ee and Erkelens 1993), which are the same as those described in this report[4].

Slant perception with visual reference

Gillam and Rogers (1991) investigated induced slant about the horizontal axis caused by stimuli (10 deg diameter patterns) relative to a fixe visual frame of reference (in their case a dimly visible room). They observed that rotation induces slant about the horizontal axis (as does horizontal shear) but that vertical shear does not induce slant. Gillam and Rogers concluded that contrary to Koenderink and Van Doorn's theory, perceived slant was not related to the deformation present but was predicted by the orientation disparity at the vertical meridian per se. We have been able to corroborate these results quantitatively on the basis of horizontal disparity. Westheimer's (1978) results are also in agreement with ours. He reported that divergence of a stimulus induces slant about the vertical axis (as does horizontal scale). He reported also that vertical disparity alone without horizontal disparity does not induce slant. Westheimer (1978) and Gillam and Rogers (1991) did not mention explicitly the presence of a visual reference in their experimental set-up. However from the description of their methods it can be inferred that a visual frame of reference was present. In our view the presence of a visual reference is the reason why Westheimer (1978) found no slant due to vertical disparity whereas Ogle (1950) did, and that Gillam and Rogers (1991) found slant due to rotation whereas Howard and Zacher (1991) did not.

[4] After completing this paper Howard and Kaneko (1994), studying shear transformations, and Kaneko and Howard (1994), studying scale transformations, reported on the fact that vertical shear and vertical scale do clearly induce slant in the absence but not in the presence of a frontal dot pattern which is untransformed for both eyes. These results also conrm our results. (They call their visual reference `zero{disparity surround'. We would suggest that visual reference is a preferable term. A zero{disparity surround may be confused with the horopter which is not what they intended.)

Orientation disparity

Gillam and Rogers (1991) explained their results in terms of orientation disparity and suggested that perceived slant is predicted from the orientation disparity at the vertical meridian per se. They concluded that orientation disparity at the horizontal meridian does not induce perception of slant. The notion that orientation disparity can in uence slant perception is very interesting if for instance one is trying to understand the anisotropy reported in the detection of slant thresholds about the horizontal and vertical axis (e.g. Cagenello and Rogers 1993, but see also Mitchison and McKee 1990; Gillam and Ryan 1992). Gillam and Rogers (1991), who did not study thresholds for slant detection, explained their results in terms of orientation disparity. However, they did not vary the contents of orientation disparity in their stimuli; they merely investigated random–dot stimuli. Cagenello and Rogers (1993) already suggested the co-variance of both positional disparity and orientation disparity in the stimuli of Gillam and Rogers. We suggest another explanation for the results of Gillam and Rogers, namely an explanation that is based on positional disparity: their results may be due to the irrelevance of vertical disparity in slant perception in the presence of a visual reference. If the stimuli used by Gillam and Rogers (1991) are expressed in terms of mathematical transformations between the retinal half–images (instead of in terms of orientation disparity), their stimuli form a special class (the shear transformations) of the stimuli used in our study. Transformations which contain orientation disparity at the vertical meridian are in fact horizontal shear transformations and thus induce perception of slant about the horizontal axis. Transformations which contain orientation disparity at the horizontal meridian are in fact vertical shear transformations and thus do not induce perception of slant. In our view, the perception of slant of a particular stimulus may be due to the underlying transformation of horizontal disparity rather than to orientation disparity itself.

Vertical disparity

In the literature there are clear reports about the ability of the human visual system to use vertical disparity for depth perception. First, vertical scaling of a single retinal image, optically by means of an aniseikonic lens in front of one eye leads to perception of slant about the vertical axis if the observer is presented with the vertical scaling for a considerable period of time (Ogle 1950; Gillam et al 1988a). Secondly, Ogle (1950) did not observe slant effect for overall aniseikonia (divergence which contains similar amounts of vertical and horizontal disparity), as mentioned above. Finally, Rogers and Bradshaw (1993) recently showed that vertical size ratios in the medial plane can in principle be used to derive stimulus distance and that experimental manipulations of vertical disparity by means of a 80 80 deg stimulus do indeed in uence the perceived distance. In these reports large field of disparity without a visual reference were used, (see also section 12 of the discussion of the recent paper by Bishop (1994) about the globality of vertical disparity processing). The significanc of a vertical disparity gradient for perception of slant relative to a frame of reference can be questioned. There have been indications that vertical disparity is not used for slant perception in the case of stimuli relative to a reference (Westheimer 1978; Cumming et al 1991; Sobel and Collett 1991). Very recently we have shown that either a vertically scaled or sheared half-image of a stereogram leads to reliable slant perception in the absence of a visual reference but leads to only poor perception of slant in the presence of a visual frame of reference (van Ee and Erkelens 1994).

In conclusion

Thus, for disparity processing in binocular depth experiments it is important to distinguish between conditions with and without a visual reference. Our assumption about the irrelevance of vertical disparity for slant perception in the presence of a visual frame of reference is based on this distinction. The modificatio of Koenderink and Van Doorn's theory on the basis of this assumption has resulted in a model for perception of slant about oblique axes in the presence of a visual frame of reference.

References

Bishop P O, 1994 "Size constancy, depth constancy and vertical disparities: a further quantitative interpretation" *Biological Cybernetics* 71 37-47

Cagenello R, Rogers B J, 1993 "Anisotropies in the perception of stereoscopic surfaces: the role of orientation disparity" *Vision Research* **33** 2189-2201

Collewijn H, Steinman R M, Erkelens C J, Regan D, 1991a "Binocular fusion, stereopsis and stereoacuity with a moving head" In *Binocular vision* Ed D Regan (London: Macmillan Press) pp 121–136

Collewijn H, Van der Steen J, Van Rijn L J, 1991b "Binocular eye movements and depth perception" in *Representations of vision, trends and tacit assumptions in vision research* Ed A Gorea (Cambridge: University Press) pp 165-183

Cumming B G, Johnston E B, Parker A J, 1991 "Vertical disparities and the perception of three–dimensional shape" *Nature* **349** 411-413

van Ee R, Erkelens C J, 1993 "Slant perception in relation to the geometry of binocular vision: Effect of local and global disparity fields Poster presentation at the International Conference and Nato Advanced Workshop on Binocular Stereopsis and Optic Flow, Toronto, Canada.

van Ee R, Erkelens C J, 1994 "Time–dependent perception of large–fiel disparity gradients: validation of a model for slant perception during dynamic viewing" *Perception* **23** 35

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, 1993 "Disparity processing for depth perception under dynamic viewing conditions" Oral presentation at the International Conference and Nato Advanced Workshop on Binocular Stereopsis and Optic Flow, Toronto, Canada.

Ferman F, Collewijn H, Jansen T C, Van den Berg A V, 1987 "Human gaze stability in the horizontal, vertical and torsional direction during voluntary head movements, evaluated with a three-dimensional scleral induction coil technique" *Vision Research* 27 811-828

Gillam B, Chambers D, Lawergren B, 1988a "The role of vertical disparity in the scaling of stereoscopic depth perception: An empirical and theoretical study" *Perception & Psychophysics* 44 473-483

Gillam B, Chambers D, Russo T, 1988b "Postfusional latency in stereoscopic slant perception and the primitives of stereopsis" *Perception & Psychophysics* 14 163-175

Gillam B, Rogers B, 1991 "Orientation disparity, deformation, and stereoscopic slant perception" *Perception* 20 441-448

Gillam B, Ryan C, 1992 "Perspective, orientation disparity, and anisotropy in stereoscopic slant perception" *Perception* 21 427-439

Howard I P, Kaneko H, 1993 "Shear disparity, cyclovergence and the perception of surface inclination" Oral presentation at the International Conference and Nato Advanced Workshop on Binocular Stereopsis and Optic Flow, Toronto, Canada.

Howard I P, Kaneko H, 1994 "Relative shear disparity and the perception of surface inclination" *Vision Research* 34 2505-2517

Howard I P, Zacher J E, 1991 "Human cyclovergence as a function of stimulus frequency and amplitude" *Experimental Brain Research* 85 445-450

Kaneko H, Howard I P, 1994 "Compression—magnificatio disparities and the perception of surface slant" *Perception* 23 33

Kertesz A E, 1991 "Cyclofusion" In *Binocular vision* Ed D Regan (London: Macmillan Press) pp 111–120

Koenderink J J, Van Doorn A J, 1976 "Geometry of binocular vision and a model for stereopsis" *Biological Cybernetics* 21 29-35

Ogle K N, 1950 *Researches in binocular vision* (Philadelphia: Saunders) pp 173–199

Rogers B J, 1992 "The perception and representation of depth and slant in stereoscopic surfaces" in *Artifici and Biological Vision Systems* Eds G Orban, H Nagel (Hamburg: Springer) pp 241-266

Rogers B J, Bradshaw M F, 1993 "Vertical disparity, differe tial perspective and binocular stereopsis" *Nature* 361 253-255

Sobel E C, Collett T S, 1991 "Does vertical disparity scale the perception of stereoscopic depth?" *Proceedings of the Royal Society (London B)* 244 87-90

Stevens K A, Brooks A, 1987 "Depth reconstruction in stereopsis" *Proceedings of the fir. IEEE international conference on computer vision London* 682-686

Werner H, 1938 "Binocular depth contrast and the conditions of the binocular field *American Journal of Psychology* 51 489-497

Westheimer G, 1978 "Vertical disparity detection: Is there an induced size effect? *Investigative Ophthalmology and Visual Science* 17 545-551

Acknowledgments—We are grateful to J.J. Koenderink for valuable discussions and for being one of the subjects. We thank A.J. Van Doorn and A. Kappers for carefully reading this paper and P. Schiphorst for developing the stimuli. This work was supported by the Netherlands Organisation for Scientifi Research (NWO), grant no. 810-404-006/1.

Appendix A

We show the properties of the transformations, non–uniform scale and shear. We fir show the first–orde Taylor approximation of a vector fiel in the neighbourhood $(\mathbf{r} + d\mathbf{r})$ of $\mathbf{r}$, decomposed into its characteristic three components: divergence, rotation and deformation. This decomposition is valid for first–orde approximations of any vector fiel is the disparity vector fiel r is a visual direction. is define on the manifold of visual directions. The notation of the mathematics is the same as in Koenderink and Van Doorn (1976).

The first–orde approximation of the differenc between $(\mathbf{r} + d\mathbf{r})$ and $(\mathbf{r})$ can be expressed as:

$$(\mathbf{r} + d\mathbf{r}) \qquad (\mathbf{r}) = (\frac{\partial}{\partial \mathbf{r}})d\mathbf{r} \ .$$

Consider the matrix form of $\frac{\partial}{\partial \mathbf{r}}$ and its symmetric and antisymmetric parts:

$$\frac{\partial}{\partial \mathbf{r}} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & \frac{1}{2}(a_{12} + a_{21}) \\ \frac{1}{2}(a_{12} + a_{21}) & a_{22} \end{pmatrix} + \begin{pmatrix} 0 & \frac{1}{2}(a_{12} \quad a_{21}) \\ \frac{1}{2}(a_{21} \quad a_{12}) & 0 \end{pmatrix} .$$

These matrices are define with respect to a Cartesian coordinate system. In the case of the disparity fiel we defin the x–axis of the Cartesian coordinate system to be co–linear with the interocular axis. The antisymmetric part of the matrix $\frac{\partial}{\partial \mathbf{r}}$ contains the curl of the vector fiel $(curl \quad = a_{21} \quad a_{12})$:

$$\frac{\partial}{\partial \mathbf{r}}_{antisymm} = \frac{1}{2} curl \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} .$$

It is always possible to fin a coordinate transformation R such that the symmetric part of the matrix $\frac{\partial}{\partial \mathbf{r}}$ can be represented in diagonal form:

$$\frac{\partial}{\partial \mathbf{r}}_{symm} = R^{1} \begin{pmatrix} {}_{max} & \\ 0 & {}_{min} \end{pmatrix} R \ ; \ ({}_{max}, \ {}_{min} \ real; \ {}_{max} \quad {}_{min}) \ .$$

R is a rotation that specifie the angle ( ) of the axis of expansion or contraction:

$$R = \begin{pmatrix} \cos & \sin \\ \sin & \cos \end{pmatrix} .$$

The matrix $\frac{\partial}{\partial \mathbf{r}}_{symm}$ can be decomposed into two parts so that one part is traceless:

$$\frac{\partial}{\partial \mathbf{r}}_{symm} = \frac{1}{2}({}_{max} \quad {}_{min})R^{1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} R + \frac{1}{2}({}_{max} + {}_{min})R^{1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} R \ .$$

If we defin $div = {}_{max} + {}_{min}$, $curl = a_{21} \; a_{12}$, and $def = {}_{max} \; {}_{min}$, the matrix $\frac{\partial}{\partial r}$ can be decomposed into divergence, rotation and deformation:

$$\frac{\partial}{\partial r} = \frac{1}{2} div \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2} curl \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \frac{1}{2} def \;\; R^{\;} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} R \; .$$

We present the elementary transformations in their canonical matrix forms:

$$divergence : \begin{pmatrix} & 0 \\ 0 & \end{pmatrix},$$

with real;

$$rotation : \begin{pmatrix} & 0 \\ 0 & \end{pmatrix},$$

with imaginary;

$$deformation : \begin{pmatrix} & 0 \\ 0 & \end{pmatrix},$$

with real.

The eigenvalues of an arbitrary transformation are of interest because they reveal information about the kind of transformation under consideration. The canonical forms of horizontal scale, vertical scale, horizontal shear and vertical shear are:

horizontal scale : $\begin{pmatrix} & 0 \\ 0 & 0 \end{pmatrix},$

one eigenvalue zero, real;

vertical scale : $\begin{pmatrix} 0 & 0 \\ 0 & \end{pmatrix},$

one eigenvalue zero, real;

horizontal shear : $\begin{pmatrix} 0 & c \\ 0 & 0 \end{pmatrix},$

eigenvalues zero, c is a real quantity;

vertical shear : $\begin{pmatrix} 0 & 0 \\ c & 0 \end{pmatrix},$

eigenvalues zero.

Horizontal scale and horizontal shear are chosen as the elementary orthogonal (see appendix B) transformations of the first-orde decomposition. The decomposition is:

$$(r + dr) \qquad (r) = [\begin{pmatrix} c_1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & c_2 \\ 0 & 0 \end{pmatrix}]dr \ .$$

This decomposition provides the basis for the description of perceived local slant about any axis for human binocular vision (see also figur 3). The numbers c_1 ans c_2 are real quantities, not necessarily equal. They describe the magnitudes of the horizontal scale and horizontal shear, respectively.

Appendix B

We show that horizontal scale and horizontal shear are orthogonal transformations. Consider the general transformation A:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} .$$

Consider also the linear vector fiel r which depends on the position x:

$$r(x) = Ax = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} .$$

We are interested in changes of the horizontal component of the vector fiel. Consider the partial derivatives of the horizontal component of the vector fiel r(x) in horizontal and vertical direction:

$$\frac{\partial r_1(x)}{\partial x_1} = a_{11} \ , \ \frac{\partial r_1(x)}{\partial x_2} = a_{12} \ .$$

The matrix form of horizontal scale (hscale) and horizontal shear (hshear) is:

$$A_{hscale} = \begin{pmatrix} c_1 & 0 \\ 0 & 0 \end{pmatrix} \ , \ A_{hshear} = \begin{pmatrix} 0 & c_2 \\ 0 & 0 \end{pmatrix} \ ,$$

respectively. For horizontal scale, the partial derivatives of the horizontal component of r(x) in horizontal and vertical direction are:

$$\frac{\partial r_{1(hscale)}(x)}{\partial x_1} = c_1 \ , \ \frac{\partial r_{1(hscale)}(x)}{\partial x_2} = 0 \ ;$$

and for horizontal shear:

$$\frac{\partial r_{1(hshear)}(x)}{\partial x_1} = 0 \ , \ \frac{\partial r_{1(hshear)}(x)}{\partial x_2} = c_2 \ .$$

These are components of two vectors which form an orthogonal basis because they are perpendicular:

$$\begin{pmatrix} c_1 \\ 0 \end{pmatrix} \perp \begin{pmatrix} 0 \\ c_2 \end{pmatrix} .$$

Therefore, horizontal disparity gradients in oblique directions (in the two-dimensional plane) are composed of these two vectors.

Figure captions

Fig. 1:
The first-orde approximation of the disparity fiel can be described mathematically as a superposition of elementary geometrical transformations (divergence, rotation, and deformation).

Fig. 2:
Scale and shear transformations. M define the magnificatio factor as a percentage %, the angle define the magnitude of the shear transformation in degrees.

Fig. 3:
The irrelevance of a vertical disparity gradient for slant perception has major consequences for the elementary transformations of figure 1 and 2. The small arrows in the figure on the left indicate the nulling of the vertical disparity component. The irrelevance suggests that 1) divergence and horizontal scale effecti ely induce similar slants, 2) rotation effec tively induces the same slant as horizontal shear and 3) vertical scale and vertical shear are not expected to induce slant.

Fig. 4:
In the presence of a visual reference divergence induces the same slant as does horizontal scale. Upper stereogram: the right half-image has a horizontal scale of -8 % relative to the left half-image. Lower stereogram: the right half-image has a divergence of -8%

Fig. 5:
The experimental set-up.

Fig. 6:
Schematic drawing of the stimuli; the dimensions are not to scale. Figure a) shows pattern 1 and pattern 2, used for matching the perceived slant about the vertical axis (in the case of pre-set divergence, horizontal scale, and vertical scale). Pattern 2 consists of a random-dot pattern similar to that of pattern 1 but is smaller. Figure b) shows stimuli used for matching the perceived slant about the horizontal axis (in the case of pre-set rotation, horizontal shear, and vertical shear).

Fig. 7:
The results of experiment 1. Figure a) shows measured percentages (and standard deviations) of divergence and vertical scale relative to horizontal scale. Divergence is represented by squares, vertical scale by circles. Each pre-set vertical scale was matched by horizontal scale of about 0 %. Pre-set divergences were matched by similar percentages

of horizontal scale. Figure b) shows measured amounts (and standard deviations) of rotation and vertical shear relative to horizontal shear. Rotation is represented by squares, vertical shear by circles. The dimension along both axes is in degrees, as define in fig 2. Each pre-set vertical shear was matched by horizontal shear of about 0 deg. Pre-set rotations were matched by similar amounts of horizontal shear

Tabel 1:
Combinations (horizontal scale (%); horizontal shear (deg)) estimated by the subjects which are not the same as the presented combinations are denoted explicitly. Estimated combinations which are the same as the presented combinations are denoted by a dot. The deviations are of the order of 1 % (scale) and 1 deg (shear), which re ects the smallest steps of adjustment of the mouse in our device. In other words perceived slants about arbitrary axes can be matched within an accuracy of 1 % and 1 deg by combinations of the elementary transformations, horizontal scale and horizontal shear.

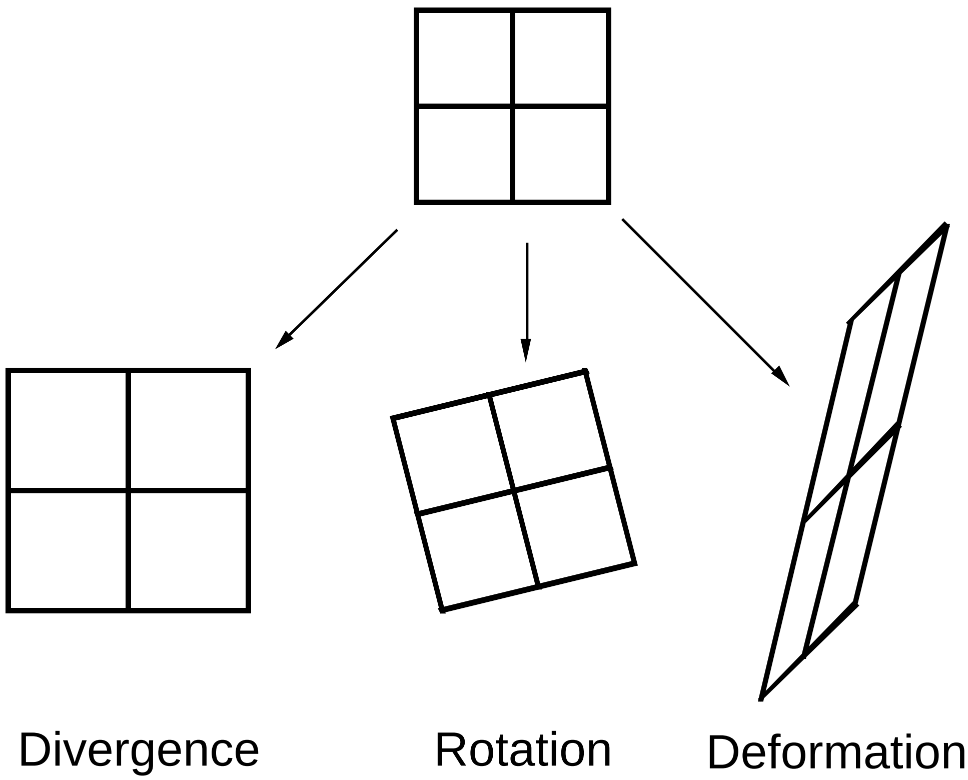
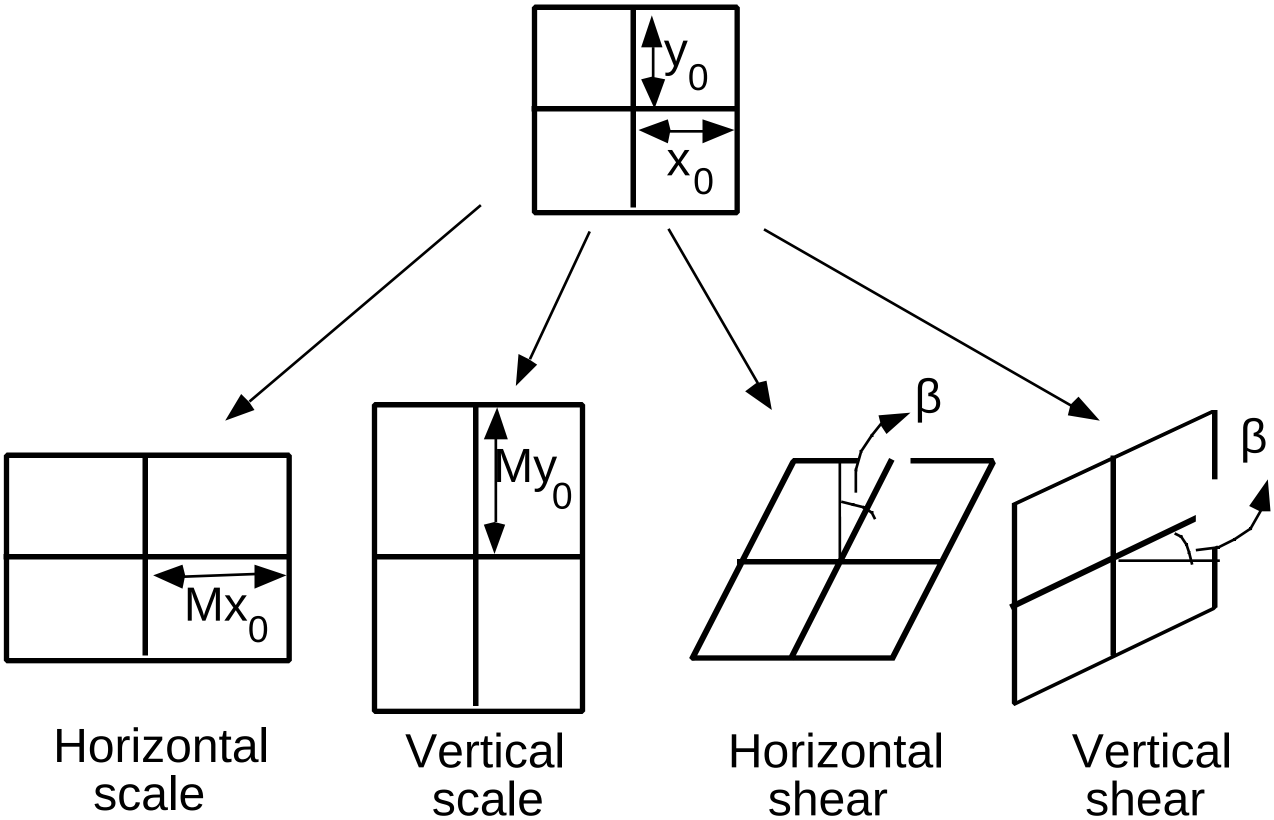


Fig 1, van Ee



y0
x0
Mx0
My0
β
β
Horizontal scale
Vertical scale
Horizontal shear
Vertical shear

Fig 2, van Ee

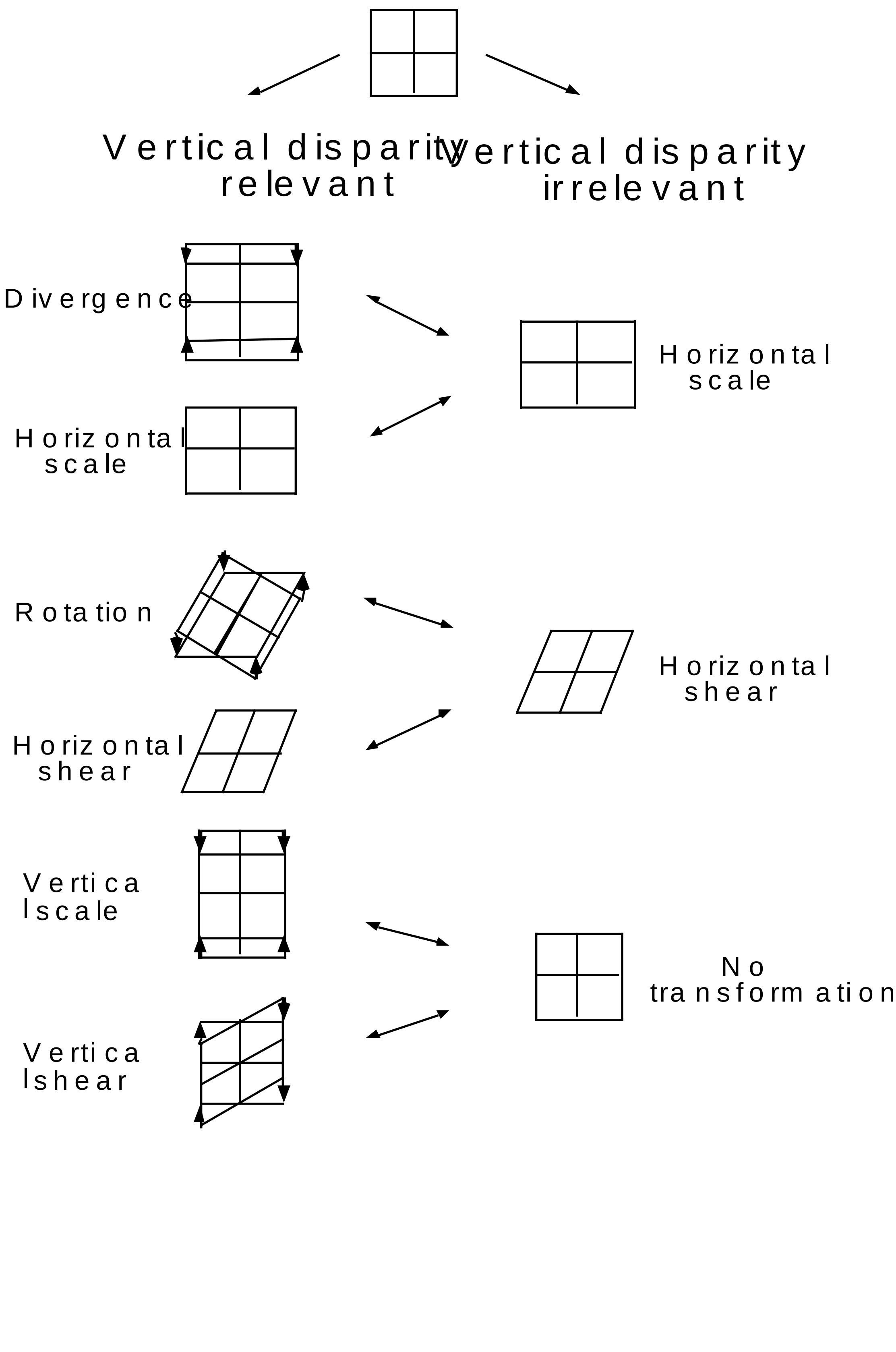

Vertical disparity relevant
Vertical disparity irrelevant
Divergence
Horizontal scale
Horizontal scale
Rotation
Horizontal shear
Horizontal shear
Vertical scale
No transformation
Vertical shear

Fig 3, van Ee

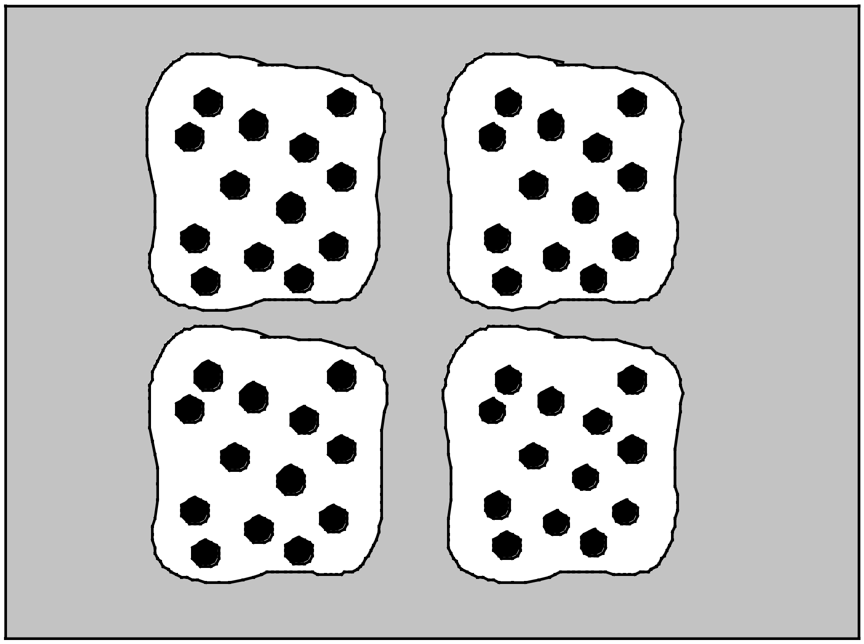
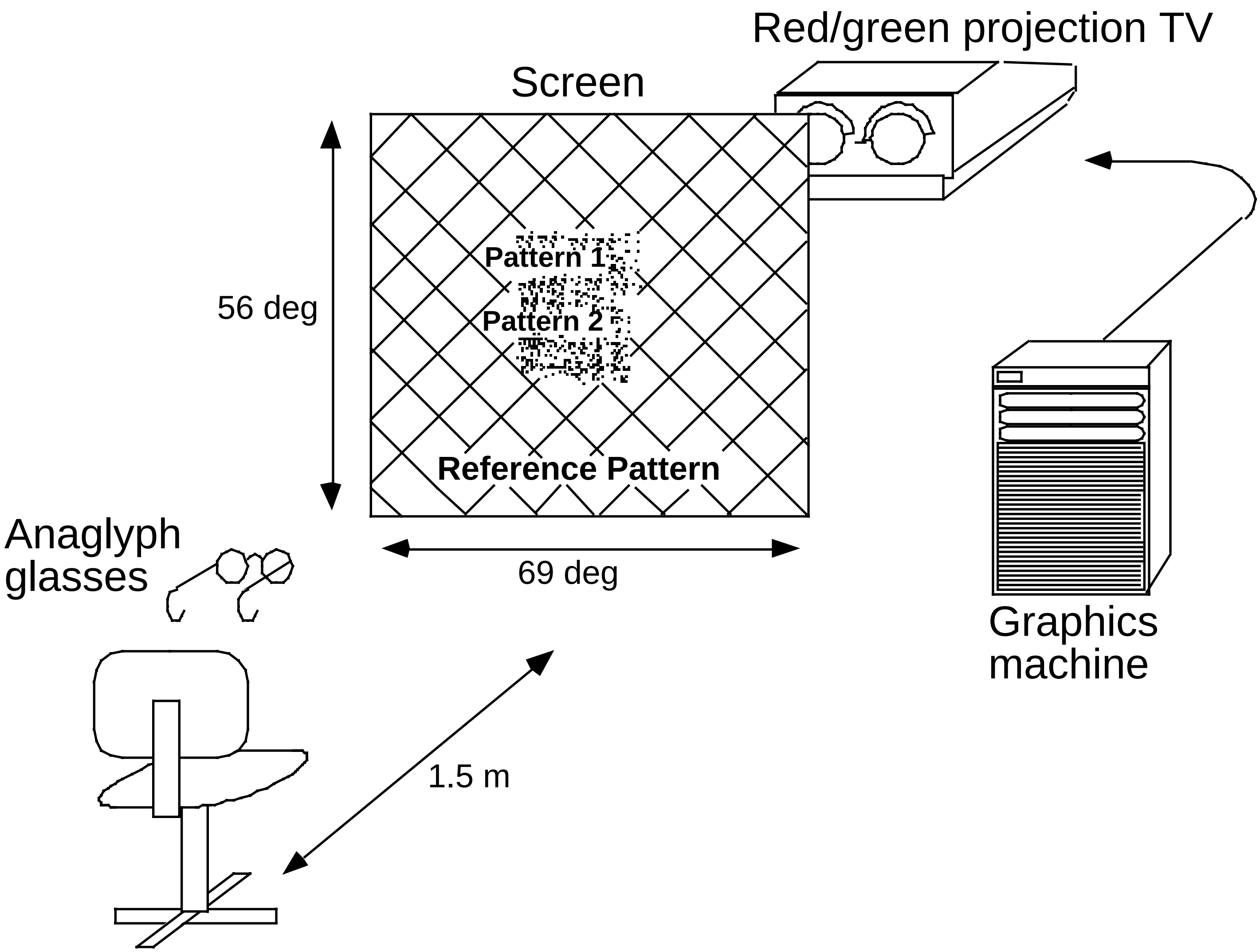


Fig 4, van Ee



Red/green projection TV
Screen
56 deg
Pattern 1
Pattern 2
Reference Pattern
69 deg
Anaglyph glasses
Graphics machine
1.5 m

Fig 6, van Ee

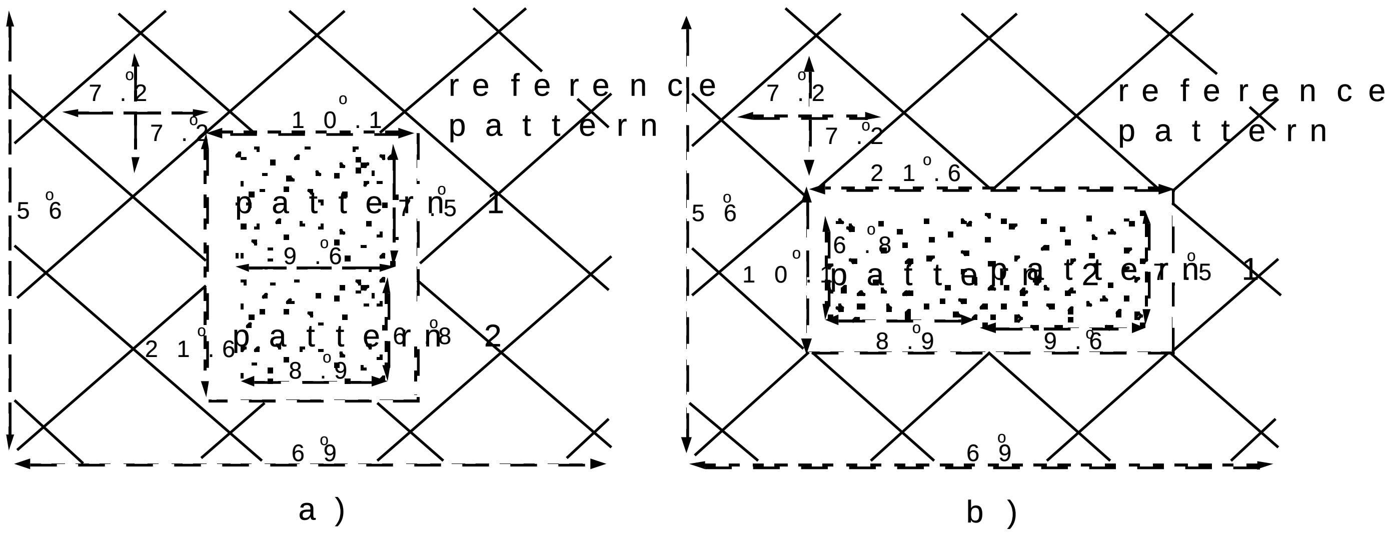
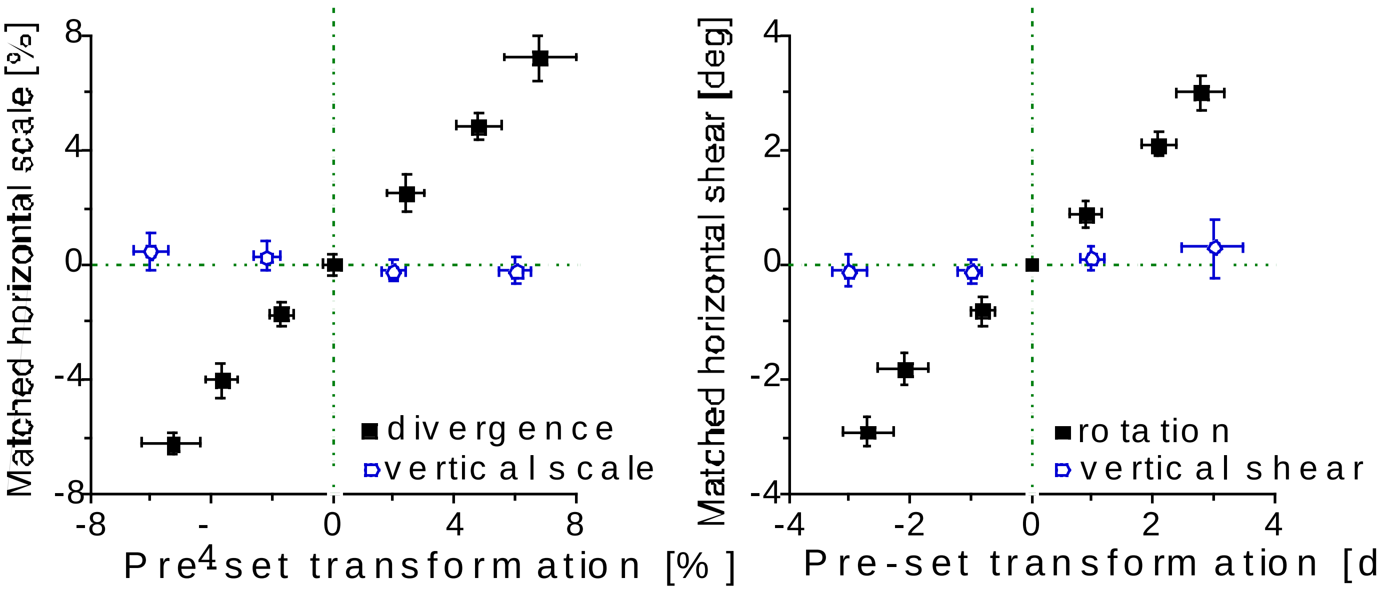

reference pattern
pattern 1
pattern 2
7.2°
10.1°
56°
21.6°
9.6°
8.9°
6.8°
69°
a)
b)

Fig 6, van Ee



a)

b)

Fig 7, van Ee