Universiteit Utrecht

Bachelor Thesis

Physics Department
Universiteit Utrecht

Slant perception of asymmetric stimuli

Author: Paul Hop Universiteit Utrecht
Supervisor: Prof. dr. C.J. Erkelens Helmholtz Institute

Abstract. In this research slant perception of asymmetric stimuli was investigated. It is presumed that observers make assumptions like orthogonality and planarity to extract 3-D information from a 2-D image. When the retinal image is asymmetric this assumption leads to different slants for the different parts of the image. Interestingly, symmetric retinal images often become asymmetric when viewed obliquely, raising the question in which way this affects slant perception. Perceived slants of asymmetric retinal images were compared to computed slants. Computations were based on a vector model in which orthogonality and planarity constraints were implemented. Stimuli were projected rectangles slanted about the vertical/horizontal axis. Stimuli were either made asymmetrical on the screen (experiment 1), or made asymmetrical by slanting the screen (experiment 2). Results show that indicated slants can be predicted by local orthogonality/planarity assumptions, which leads to accurate predictions for perceived slants of obliquely viewed stimuli.

Keywords: Slant perception, depth perception, linear perspective, oblique viewing, asymmetry.

1. Introduction

In modern society we are presented with 2-D images all the time. Whether it is in art and pictures, or on our television and computer screen, we are able to perceive 3-D structures with ease. Creating a 3-D scene on a flat screen can be quite difficult, but our brain is confronted with an even more difficult task: it has to reconstruct the real world from the flat image. A task that humans (most of the time) perform automatically. This is remarkable, because in theory there is an infinite number of real world scenes that would create the same 2-D image. Practically, the understanding of picture perception enables us to create more realistic scenes on flat surfaces. In this study the perception of asymmetric retinal images is investigated, giving an insight in the perception of obliquely viewed images.

In order to extract 3-D information from a 2-D image, the brain uses different cues. Pictures often provide us with several cues like occlusion, shadowing and texture gradient, which are called pictorial cues (see Cutting & Vishton, 1995 for an overview). Even when minimal depth cues are available, humans have a sense of depth which is difficult to inhibit (Farran, Whitaker, & Patel, 2009). Pictorial cues can even dominate when opposing 3-D information is present. Fascinating examples are the 'reverspectives' made by artist Patrick Hughes. These paintings consist of a perspective-rich image painted on truncated 3-D pyramids protruding out of the picture plane. The 2-D depth cues on the faces of the pyramid suggest depth opposite to that of the physical depth of the pyramid it is painted on. Even when viewed binocularly, subjects perceive the depth implied by the 2-D cues under most viewing conditions (Papathomas, 2002).

1.1 Linear perspective

The most powerful cue contributing to this reverspective illusion is linear perspective (Papathomas, 2008; B. Rogers & Gyani, 2010). Aside of this illusion, several studies showed that perspective information is a strong depth cue (Andersen, Braunstein, & Saidpour, 1998; Todd, Thaler, & Dijkstra, 2005). A common example of linear perspective is a road receding into the picture. The edges of the road converge in the image, but still we perceive the road as having parallel edges that recede into the picture. Painters have utilized perspective for ages to strongly represent 3-D scenes in their art. To correctly interpret the image as the scene depicted, certain

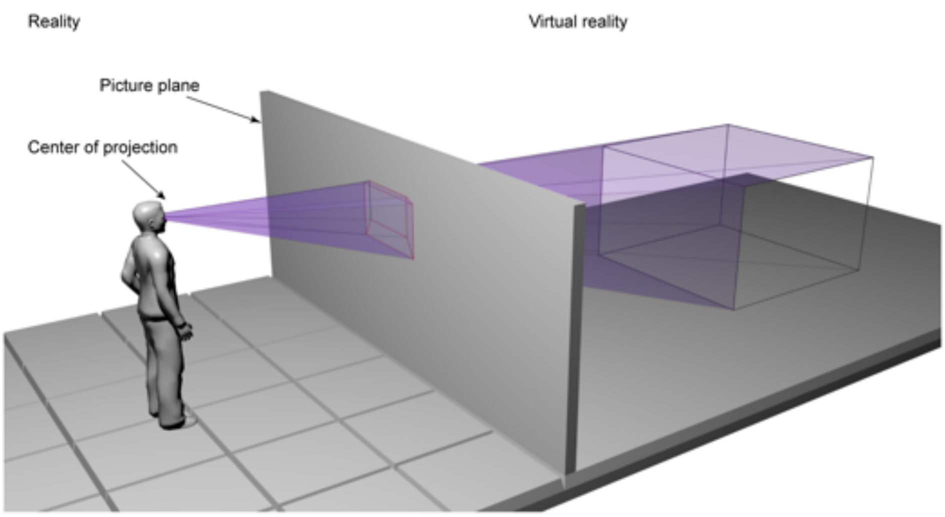


Figure 1. The construction of perspective. The lines emerging from the 3-D object cut through the picture plane. Source: (Ling, Nefs, Brinkman, Qu, & Heynderickx, 2013)

assumptions have to be made like 90° angles and parallel edges (Saunders & Backus, 2007).

Linear perspective is based on the array of light rays that emerge from a scene and reach the retina of the eye. When a surface is placed between the scene and the viewer (the surface being a painting or a screen for example), the array of light rays 'cut' through the surface (figure 1). The lines they create on the surface represent the scene on the 2-D surface. It is similar to a painter looking through a window and directly painting the scene exactly as seen on the window, the so-called 'Alberti window' (figure 2) (Palmer, 1999). Ideally the image is geometrically equivalent to the corresponding scene and the image closely resembles the real scene (Wijntjes, 2014). This is only true in the unique case when the viewer is positioned (his/her viewing point) in the centre of projection (CoP)[1]. The CoP is the geometric position from which the image is constructed; this place may exist physically like the eyes of the painter (figure 2), or might be virtual when constructing perspective on a computer. The consequence is that geometrical distortions are present in the array of lines that are projected onto the viewer's eye when the viewpoint and CoP do not coincide (Rosinski, Mulholland, Degelman, & Farber, 1980). This raises a problem in picture perception: what are the effects of variation of the viewing point and can these effects be predicted?

1 In reality a viewer has two viewing points, one for each eye. In this research the viewing point is defined in between the two eyes, the so-called 'cyclopean eye'.

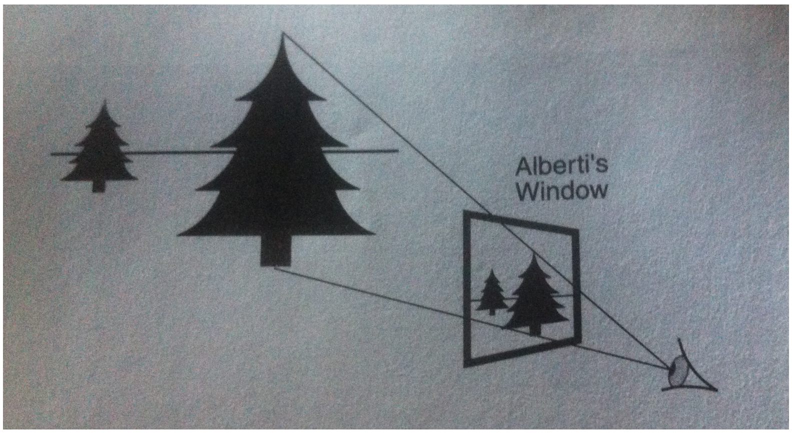


Figure 2. Alberti's window. The 3-D scene is painted directly on the window pane. When the position if the painter is fixed, the eyes of the painter correspond to the unique centre of projection. Source: (Palmer, 1999)

1.2 Previous studies

A considerable amount of research has already been conducted on the perception of pictures when the viewpoint is incorrect. The two main theories are the 'compensation theory' and the 'transformation theory' (Todorovic, 2008).

The first states that our visual system compensates for the incorrect viewpoint. Following this theory picture perception is robust: variation of the viewing point is not associated with significant perceptual effects. It has been proposed that the visual system recovers the correct CoP and reconstructs the correct image as if its seen from the centre of projection. Researchers have found evidence that observers exploit information about the picture surface to compensate for oblique viewing (Goldstein, 1987; Perkins, 1973; Vishwanath, Girshick, & Banks, 2005).

The transformation theory states that the visual system retrieves depth from the new transformed geometry. By changing the viewpoint observers see a related, but geometrically transformed 3-D scene. It has also been suggested that the visual system combines these two systems (Yang & Kubovy, 1999).

So far there is no general agreement whether the visual system compensates, transforms, or uses both mechanisms. Several studies have reported agreement between perceived depth and the depth specified by geometric transformation in perspective images. However, a systematic underestimation of slant is reported in most studies (Andersen et al., 1998; Erkelens, 2013; Saunders & Backus, 2006; Todd et al., 2005; Todorovic, 2009).

Other studies suggest that the geometric transformation validity is relatively small or does not even exist (Perkins, 1973; Rosinski et al., 1980; Vishwanath et al., 2005). Following this interpretation observers are aware of their different viewpoint and compensate for the oblique viewing angle. There are however obvious cases in which picture perception does change when altering the viewpoint. Apparent rotations have been reported in several studies (Goldstein, 1979; Gombrich, 1972). A striking example is the 'your country needs you' poster used in the First World War (figure 4). Lord Kitchener is pointing and staring at you, and almost magically follows you when moving along the poster. At the same time the surroundings (the physical wall it hangs on) do not rotate along with you. Another example is the 'egocentric road', in which road receding into the picture rotates along with the observer (Todorovic, 2005). Changing the viewing angle seems to have a small effect on the perception of the layout of the display, but has a large effect on the orientation of the objects depicted on the display (Goldstein, 1987; Koenderink, Van Doorn, Kappers, & Todd, 2004). Another problem with the compensation theory is the 'double' perception involved. The viewer has to look at a scene, determine the viewing angle, and then create the new compensated view. This seems more difficult and computationally complicated than normal perception.

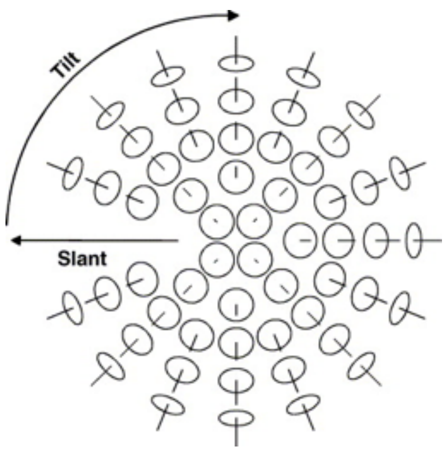


Figure 3. Definition of slant and tilt. Source: (Norman, Todd, Norman, Clayton, & McBride, 2006)

Figure 4. Lord Kitchener, note that he 'follows' you.

1.3 Terminology

In the following section a few terms will be used, which will be briefly described here. Generally the orientation of a surface is defined by slant and tilt (Stevens, 1983). The slant is then quantified by the angle between the surface normal and the line of sight (0° ≤ 90°). The direction of the slant can be specified by tilt (0° ≤ 360°) (figure 3). Because the usage of slant and tilt would become needlessly confusing in the experiments in this study, only the slant definition is used. The direction of slant is specified by the axis the object is slanted about (e.g. the vertical axis or the horizontal axis).

Furthermore Saunders & Knill defined the spin of an object as the angle between the axis of symmetry and the direction of tilt (Saunders & Knill, 2001). Thus this defines the rotation in the plane of the object. These definitions together completely describe the orientation of the surface.

Three classes of stimuli can be defined for a 2-D perspective image: the virtual stimulus, the proximal stimulus and the physical stimulus. The physical stimulus is the stimulus as it is depicted on the screen. The proximal stimulus is the retinal stimulus the image produces (Erkelens, 2013b). The virtual stimulus is the 3-D object that produces the same proximal image as the physical stimulus. For example: a 2-D trapezoid image produces the same retinal stimulus as a certain slanted rectangle. Important in this study are two types of slants: the depicted slant (Φ) is the slant of the rectangle as it is projected on the frontoparallel screen. The screen slant (σ) is the physical slant of the screen. The virtual slant (v) is the slant of the computed virtual object for a given depicted slant (Φ) and screen slant (σ). Note that both Φ and v are slants in virtual space. Φ will be exclusively used for the computed slants of stimuli as depicted on the frontoparallel screen, while v will be exclusively used for computed slants of stimuli on a slanted screen.

1.4 Geometric transformation theory

A recent study presented new evidence for the geometric transformation theory (Erkelens, 2013a, 2013b). In this study observers were presented rectangular as well as rhombic grids that were virtually slanted about the vertical axis. Screen cues were defined as the cues that depend on the rotation of the screen (e.g. knowledge of screen slant and blur). Sets of physical stimuli for different screen slants were computed such that the virtual stimuli were identical (and thus independent of screen slant). This way it could be tested whether screen cues were significant in slant perception. The results showed that the screen cues were insignificant and that linear perspective dominated.

Furthermore the study showed that linear perspective can predict perceived slant when the screen is slanted. Stimuli were created that have a certain slant on a frontoparallel screen (Φ). The screen on which these stimuli were presented was subsequently slanted (σ). The study showed that for each combination of Φ and σ, there exists a new rectangular solution with a virtual slant (v). This virtual slant is shown to be predictable for perceived slants.

1.5 Focus of this research

In this research I further explore the predictive quality of virtual slant in oblique viewing as proposed by Erkelens. In particular this research focuses on the perception of asymmetric stimuli. I present a computational model to reconstruct the virtual object corresponding to a 2-D image. The computations were made for perspective-rich stimuli that consists of rectangular grids.

The computational model is based on vector calculus, in which vectors represent the rays joining the retina and the 2-D image (see Appendix A for a description of the model). Elongating these vectors and applying constraints to the relations between the vectors allows me to reconstruct the virtual object. The mathematical reconstruction of a real world scene faces the problem of the ambiguity of the 3-D world lying behind the 2-D image. Although computationally there is an infinite number of possible objects a 2-D image can represent, the human observer does not seem to encounter this problem in daily life. This is presumably because of certain assumptions humans make. For rectangular surfaces there are two constraints that provide the viewer with cues about depth: orthogonality and parallelism. This means that the opposite edges are parallel and the vertices are 90°. In this study projections of rectangles are constructed that are slanted about one of their symmetry axes (vertical or horizontal). In this case orthogonality and parallelism assumptions provide the same image when viewed from the centre of projection. This is different for rectangles that are slanted about another axis, in that case the orthogonality and parallelism assumptions provide different solutions (Saunders & Backus, 2007). In this study I focus on rectangles that are <u>virtually</u> (Φ) slanted about one of their symmetry axes, the corresponding image was <u>physically</u> (σ) slanted about another axis than the symmetry axis.

When the centre of projection coincides with the viewpoint, the computational model provides an unambiguous solution for a trapezoid image, namely a slanted rectangle. Several researchers have investigated the effect of changing the viewpoint along the depth axis (or scaling the image, which is equivalent) (Juricevic & Kennedy, 2006; Saunders & Backus, 2006). Erkelens studied rectangular and rhombic objects that were slanted about the vertical axis, presented on a screen that was slanted about the vertical axis. For both cases (translation along the depth axis, rotation about the vertical axis), the geometry predicts a new rectangular virtual object. This can be geometrically explained by looking at the symmetry plane of the image. In these studies the rotations and translations ensured that the viewer was still positioned in the symmetry plane of the image. The symmetry plane for the rectangular stimuli is the plane through the centre of the image (figure 5). When moving in this symmetry plane compression and/or scaling takes place, but the proximal stimulus will still be a symmetric trapezoid, and thus has an unique corresponding virtual object having a rectangular shape.[2]

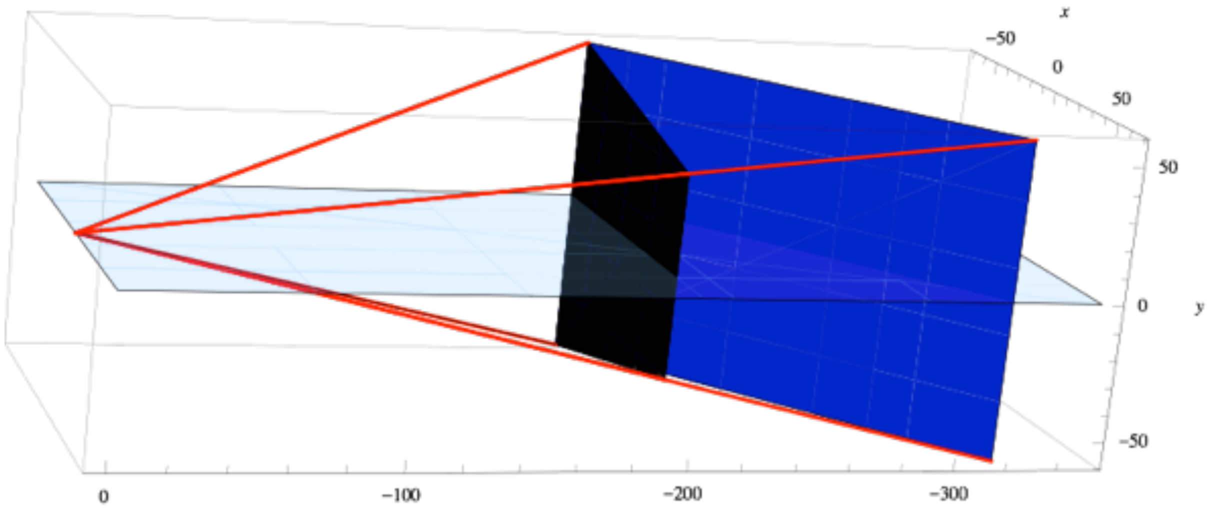


Figure 5. An example of a virtual rectangular surface corresponding to a trapezoid on the screen. The black image is the image as presented on the screen. The blue surface is the corresponding slanted rectangle. The transparent surface is the symmetry plane of the image.

[2] As is explained in Appendix A, the shape and orientation of the rectangle are unique. However, there is an infinite number of parallel solutions with increasing size. The absolute size of an image can not be determined by perspective geometry.

The aim of this research is to study the effects of a viewpoint outside of the symmetry plane, such a viewpoint creates an asymmetric proximal stimulus. When the proximal stimulus is asymmetric (by constructing a asymmetric stimulus on the frontoparallel scene, or by rotating a symmetric image), there is no such a rectangular planar solution anymore. The question rises whether human observers still perceive a slanted object and whether the perceived slant can be predicted. It is possible that the visual system ignores the fact that the solution is not rectangular. Several studies have shown that humans are likely to perceive symmetry, even when there is no symmetry in the image (Barlow & Reeves, 1979; King, Meyer, Tangney, & Biederman, 1976). It is also possible that observers do notice asymmetry, but knowledge of the original shape makes them accept the distortions.

Here I explore the possibility of 'near rectangles', in which one pair of vertices is 90° and the other pair deviates from 90°. By applying fewer constraints in the computational model, the possible solutions can be explored. The model is adjusted so that the constraints for the virtual object are 1) one pair of 90° angles on one side of the quadrangle and 2) the object has to be planar. In this case the corresponding virtual objects contain one pair of 90° angles and one pair of angles that is shifted from 90°. The shift depends on the amount of asymmetry of the physical stimulus (equivalent to amount of rotation of the screen in this case). Interestingly, there are different sets of solutions now, because there are four pairs of adjacent vertices. So there are two solutions for which the angles on one side of the quadrangle are 90° (b & c or a & d in figure 6) and thus the lower and upper

Figure 6. Example of a trapezoid, which corresponds to a slanted rectangle when 90° angles are assumed.

edges are parallel. And there are two solutions for which the lower or upper vertices are 90° (a & b or c & d in figure 6), and thus the left and right edges are parallel.

To my knowledge few studies have investigated how humans perceive such distortions. Research by VanGorp focused on IBR (image base rendering)[3]. The research focused on the extent in which distortions are accepted by observers (Vangorp et al., 2013). In this research pictures of real world scenes were used. Although of practical importance, the use of real world pictures provides to many cues to investigate the fundaments of the visual system. The purpose of this study is to isolate perspective information to get an insight in the fundamental mechanism of the visual system.

Cutting studied the different viewpoints when enjoying a movie in the cinema (Cutting, 1987). He proposed that the distortions are too small to notice and are therefore of unimportance. As in the study on IBR, this study is based on cue-rich stimuli and therefore does not provide insight into the fundamental mechanism of the visual system.

In a first experiment subjects were presented asymmetric stimuli, in which the projections of slanted rectangles were modified. This experiment was conducted to test whether subjects still perceive slant when the stimuli is asymmetric and whether subjects rely on a certain part of the stimulus to indicate slant.

In the second experiment subjects were presented with rectangles that are virtually slanted about the horizontal axis. The screen was slanted about the vertical axis. The symmetry plane of this image is the vertical plane through the centre of the image. Because the viewpoint of the subjects is kept stationary, the viewpoint lies outside the symmetry plane of the image when the screen is rotated. This experiment was conducted to test if subjects still perceive slant when the proximal stimulus is asymmetric because of screen slant, and whether this slant can be predicted.

[3] IBR is a technique that relies on a set of two-dimensional images of a scene to generate a 3-D model of that scene. Google street view is a popular example that uses this technique.

2. Experiment 1

2.1 Computations

Asymmetric stimuli were constructed to investigate whether subjects are able to perceive slants when the image is asymmetric. Based on geometry there are a few possible options what subjects may perceive.

Firstly, they could see the image simply as a 2-D image, without any depth at all. However studies suggest that humans are likely to perceive depth, even when depth cues are minimal (Farran et al., 2009).

Secondly they could choose the upper or lower converging line as 'starting point'. When subjects assume that only one pair of angles is 90°, this means that the corresponding virtual object has one converging side. If this is the case, the question arises which side humans prefer which will be discussed in the discussion section.

Thirdly, subjects could prefer all vertices of the virtual image to be 90°, which would imply that the virtual stimulus is 'twisted', so it is not a planar surface (figure 8). Note that by changing the convergence of a line, but keeping the distance between the two vertical sides the same, not only the slant of the virtual object will change, but also its shape. In this case the width of the virtual object increases when convergence increases (the upper rectangle in figure 8 is wider than the lower). Geometrically the aspect ratio

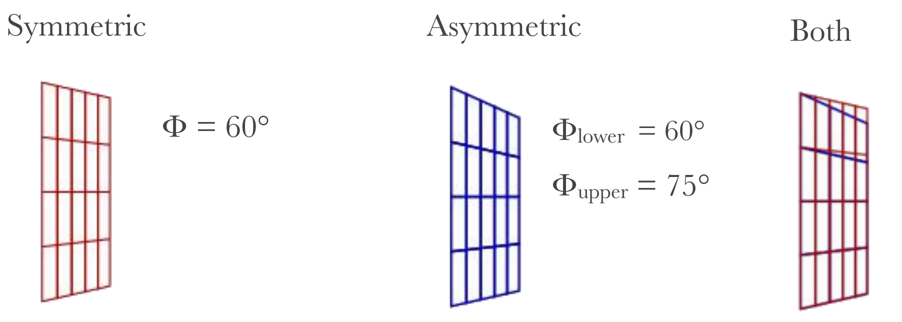


Figure 7. The modification of the original slanted rectangle. The upper edge of the stimulus is more convergent than the lower edge in the asymmetric stimulus. Note that the size at which the stimuli were presented was bigger (15 cm x 10 cm). Scaling an image affects the virtual slant, so the images are not representative for the stimuli shown to the subjects.

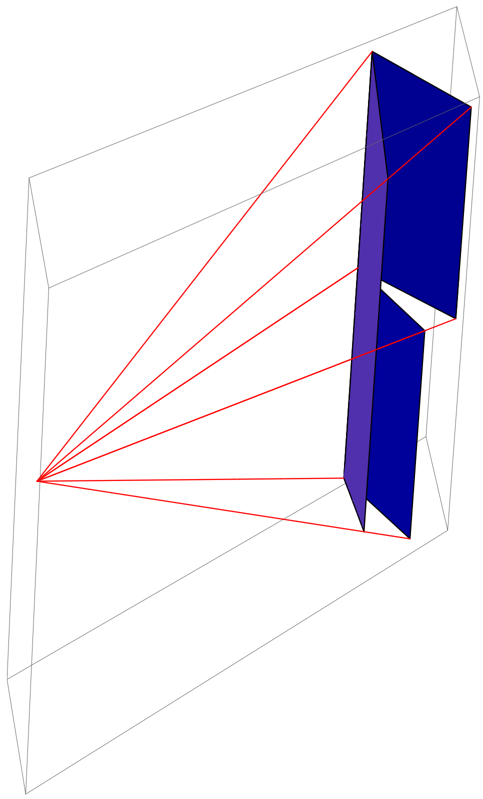


Figure 8. A twisted virtual object. This virtual object corresponds to an asymmetric stimulus, when the assumption of four 90° angles is made.

(the ratio between the lengths of the two vertical sides) only affects the shape of the virtual object, the slant remains constant. Studies have shown that for small changes in aspect ratio (as is the case here), perspective convergence being constant, perceived slant is constant (Braunstein & Payne, 1969; Saunders & Backus, 2006). In this experiment subjects are instructed to indicate perceived slant, the subjects are free to interpret the shape of the image as they wish.

The stimuli presented in this experiment contain perspective information only, which has been proven to be a strong cue to slant (Erkelens, 2013). In this experiment sets of physical stimuli were created (see appendix for method) such that the corresponding virtual object is a rectangle that is slanted about the vertical axis. The stimuli were modified such that the bottom or top converging lines become less converging or more converging than the opposite line and thus corresponding to a lower/higher slant when 90° angles are assumed (figure 7).

Three stimuli were constructed with each a different depicted slant (40°, 50°, 60°). For each of these three stimuli the bottom and the upper converging line were modified in nine steps, so that the new corresponding upper/lower depicted slant differed (-20°, -15°, -10°, -5°, 0°, +5°, +10°, +15°, +20°) from the original slant (see figure 7).

2.2 Method

2.2.1 *Experimental setup*

The stimuli were presented on a 21-inch LaCie 321 (1600 x 1200, 75 Hz). A chin rest was placed at a distance of 57 cm in front of the screen, corresponding to the centre of projection at which the stimuli were projected. At this distance the visual angle of the screen was approximately 43° x 32°. The experiment took place in a normally lit room. A vertical pad (10 cm x 15 cm) was placed between the monitor and the chin rest, at a distance of 21 cm from the chin rest. The pad is placed on a turntable which can be rotated to adjust the pad to the perceived slant.

2.2.2 *Stimuli*

The physical stimuli were rectangular grids of width 10 cm and length 15 cm (10° x 15°), which project on the screen as trapezoids when slanted. The rectangle consists of 6 vertical lines and 5 horizontal lines. The lines are white (93 cd/m^2) against a black background (± 2 cd/m^2). The depicted slant of the unmodified stimuli is abbreviated as Φ and the slant of the modified part of the stimulus as Φ_{upper} or Φ_{lower}, respectively when the upper part or the lower part is modified. Three stimuli with varying Φ (40°, 50°, 60°) were presented and both the lower and upper parts were modified varying from -20° to 20° in steps of 5°. The upper and lower part were always modified independently, so either the upper or lower part was modified, not both. Furthermore symmetric stimuli were presented as a reference. The depicted slant of the symmetric stimuli was varied between 20° and 80° in steps of 5°.

2.2.3 *Procedure*

Four subjects were instructed to indicate the perceived slant by adjusting the pad until the slants matched. The subjects did not receive any further instructions, so they were free to focus either on the upper or lower part of the stimulus. For the more asymmetric stimuli most of the subjects noticed a conflict: concentrating on the upper part indicated a different slant than the lower part. When this was the case, the subjects were asked to make two slant judgements, one for the upper and another for the lower part. All subjects were naive to the purpose of the study. All subjects had normal or corrected to normal vision. Before indicating the slants of the stimuli, the screen (without stimulus) was slanted about the vertical axis in steps of 10° to check whether the subjects were able to indicate slant accurately. After a bit of practice the subjects were able to indicate the slant of the screen with margins of 6°, with no bias toward under- or overestimation. All subjects conducted the experiment binocularly. Stimuli were presented in a random order.

2.3 Results

The mean results for four observers are presented in figure 9. The six graphs show perceived slant as a function of the slant of the lower half of the stimulus (figure 9 a,b,c), or as a function of slant of the upper half of the stimulus (figure 9 d,e,f). As a reference, the perceived slant of the symmetric stimuli is shown and the perceived slant of the symmetric stimulus which has the same slant as the side of the stimulus which is held constant (dashed line in figure 9). Some subjects perceived two slants in one stimulus, only the slant they perceived when focusing on the lower half of the stimulus is included, which will be discussed further on. A number of characteristics of the data stand out:

Firstly, the data is not consistent between subjects. Two subjects were very similar, they noticed the asymmetry for differences between the upper and lower part for deviations of ≥ 10°. In that case they were asked what they saw in the presented image. They noticed it could be either a twisted object or an asymmetric slanted planar object (options 2 and 3 in section 2.1). They were asked if they could indicate the slant independently for the upper and lower half. The other two subject did not really pay attention to the asymmetries and indicated slanted by focusing on the lower half.

The subjects were able to indicate the slant by focusing on the lower part. When the lower half of the stimulus was modified, the indicated slants match the indicated slants of the symmetric stimuli with the same slant as Φ_{lower} (see figure 9 a,b,c). When the upper part of the stimulus was modified (and thus the lower part had a constant slant), the perceived slant was nearly constant and again matched the perceived slants for the symmetric stimuli (figure 9 d,e,f). In contrast, in the occasion when the subjects noticed the asymmetry and were asked to judge the slant of the upper and lower part, they had great difficulty with indicating the slant of the upper half of the stimulus. Approximately three out of four times they weren't able to indicate the slant. Often they noted a bistability: the upper slant could be either

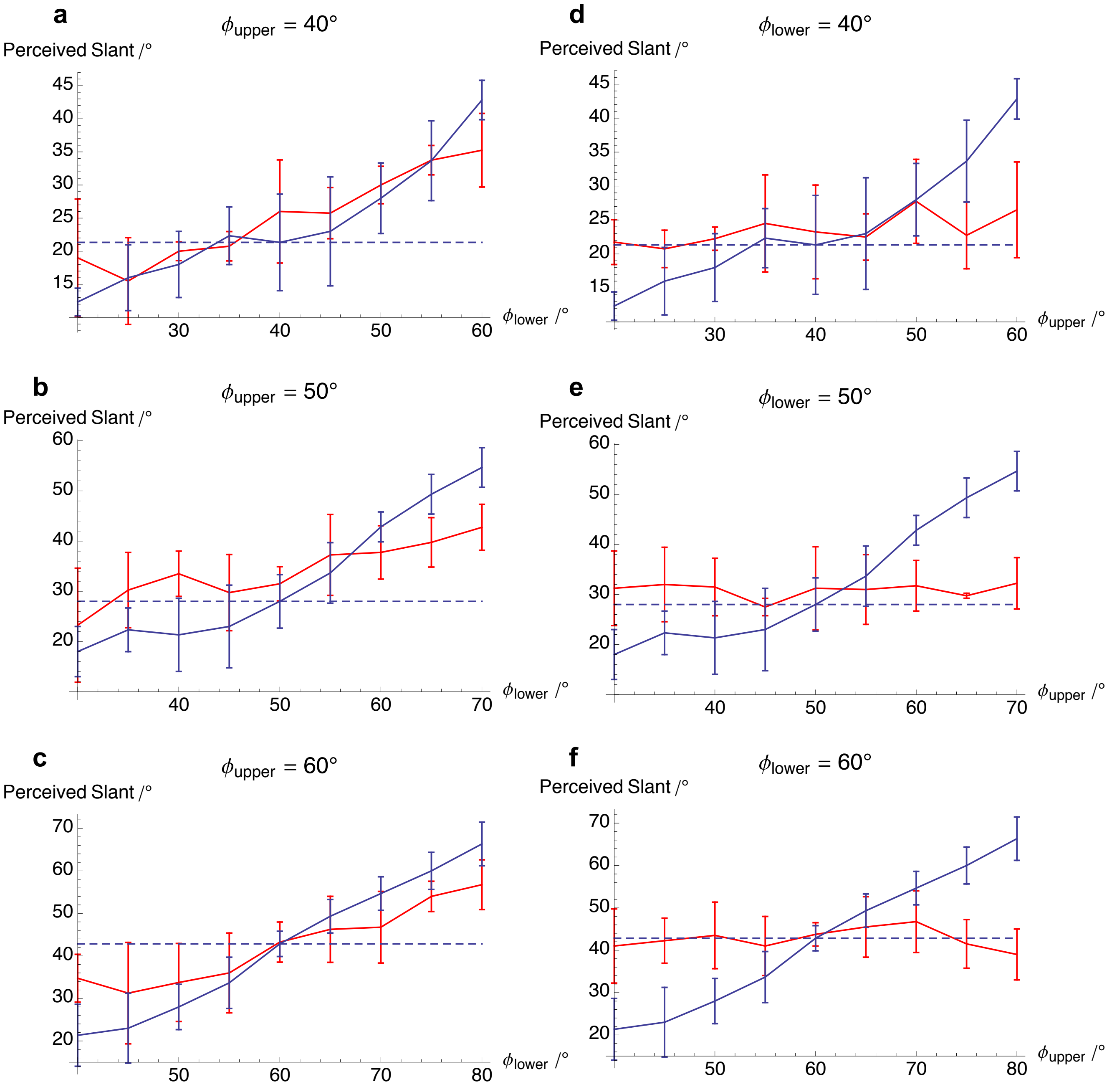
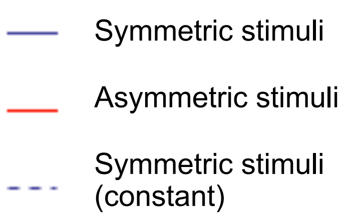


Figure 9. Perceived slant as function of the depicted slant. The graphs represent the means of four subjects. **(a,b,c)** Blue line: perceived slant as a function of depicted slant for symmetric stimuli. Red line: perceived slant as function of depicted slant of the lower half of the stimulus. The slant of the upper half was held constant. Dashed line: perceived slant of the symmetric stimulus with the same slant as the upper part of the stimulus. **(d,e,f)** Blue line: perceived slant as a function of depicted slant for the symmetric stimuli (upper half and lower half specify the same slant). Red line: perceived slant as function of depicted slant of the upper half of the stimulus. The slant of the lower half was held constant. Dashed line: perceived slant of symmetric stimulus with the same slant as the lower half of the stimulus.



protruding in the screen, or out of the screen. A phenomenon they never noticed when indicating the lower slant. In the case they did indicate the upper slant, results were very inconsistent. For example, an upper slant of 50° was indicated as 13°, 28° and 35° for one subject.

Subjects always experienced a sense of depth, even in the most asymmetric stimuli, which contradicts the first option stated in the method that subjects might not perceive depth at all.

The slant was systematically underestimated, both in the asymmetric and symmetric stimuli. The underestimation is in the same order as previous studies (Andersen et al., 1998; Erkelens, 2013). Slant underestimation seems a general result and remains unexplained.

3. Experiment 2

3.1 Computations

In this experiment rectangular grids (10 cm x 15 cm) were virtually slanted about the horizontal axis (figure 10). Subsequently the screen was slanted about the vertical axis. Because the stimulus has a vertical symmetry plane, slanting the screen about the vertical axis will place the viewer outside of the symmetry plane of the image. The proximal stimulus on the retina will therefore change from symmetric (frontoparallel screen) to asymmetric (the slanted screen) (figure 11). As explained in section 1.5, there exists no rectangular virtual object corresponding to the asymmetric proximal stimulus. Planar virtual objects corresponding to the proximal stimulus were computed with one pair of 90° angles and one pair of angles that increasingly deviates from 90° with increasing screen slant. The first class of solutions consists of surfaces with

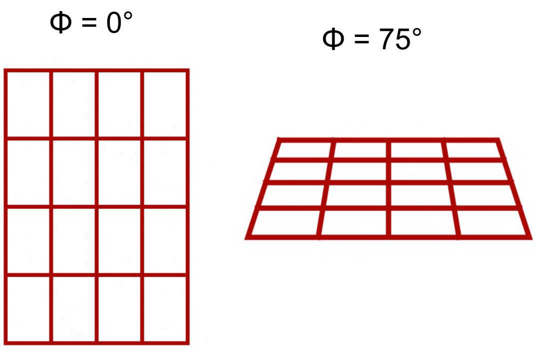


Figure 10. An example of the stimuli presented to the subjects. The rectangle (10 cm x 15 cm) is slanted about the horizontal axis, projecting as a trapezoid on the screen.

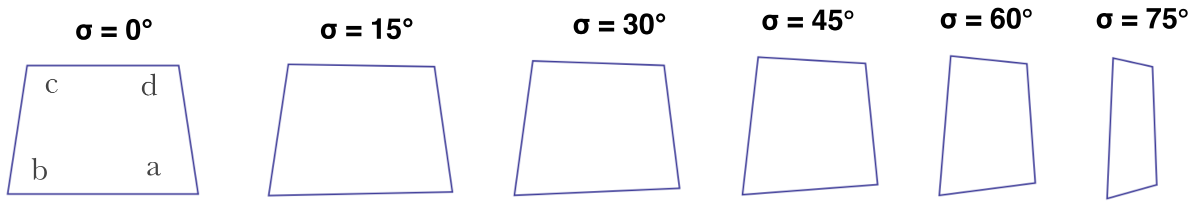


Figure 11. Proximal stimuli corresponding to different screen slants. The depicted slant in this figure is 60°.

parallel side edges, which can be divided in solutions with top or bottom 90° angles (a and b or c and d in figure 11). These solutions will be abbreviated as the t-model and b-model respectively. The second class consists of surfaces with parallel upper/lower edges, which can be divided in solutions with left and right 90° angles (a and d or b and c in figure 11). These solutions will be abbreviated as the l-model and r-model respectively.

The two different axes of rotation that define the orientation of the virtual slants are defined as v_{hor} and v_{vert}. V_{hor} is the horizontal line through the image and v_{vert} is the vertical line through the image (figure 12). Note that the solutions are slanted about the axes in the image and thus not about the stationary horizontal and vertical axis (figure 12). It is important to keep in mind that the order of rotation is not interchangeable, so it is of importance to state the order of rotation. The two classes of virtual objects computed here are slanted about these axes in reverse order, as will be explained in section 3.1.1 and 3.1.2.

3.1.1 *90° angles top/bottom*

One class of solutions has two bottom or top 90° angles (a and b or c and d in figure 11). The computed slants of these virtual objects are presented in (figure 13 a,b,c). The corresponding virtual object is slanted about the horizontal and subsequently about the vertical axis (figure 12a). The horizontal slant is constant as a function of screen slant for both the t- and b-model (figure 13a). The vertical slant increases as a function of screen slant (σ) for the t-model and decreases as a function of σ for the b-model. (figure 13 b,c). Slopes for vertical slant are steeper for smaller depicted slants (Φ). In other words, large depicted slants ‘follow’ you, while small depicted slants rotate away from the viewer.

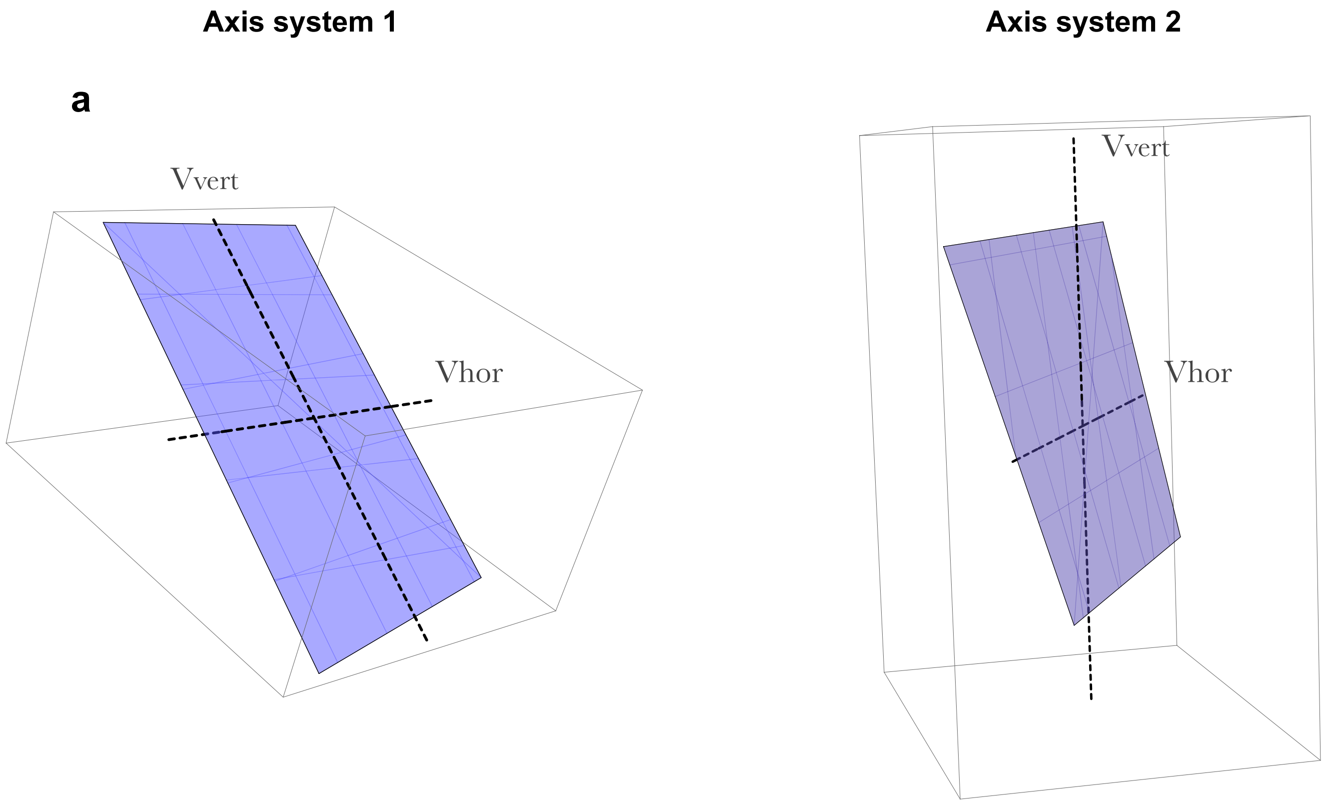


Figure 12. Rotation axes for the computed virtual objects. (**a**) Bottom/top model, which is slanted about the horizontal axis first and subsequently about the vertical axis. Note that because of the order of rotation the vertical axis lies in the plane of the object, the horizontal lies outside the plane. (**b**) Left/right model, which is slanted about the vertical axis first and subsequently about the horizontal axis. Note that because of the order of rotation the horizontal axis lies in the plane of the object, the vertical axis lies outside of the plane.

3.1.2 *90° angles left/right*

The second class of solutions has two 90° side angles (b and c or a and d in figure 11). The corresponding slants of these solutions are presented in figure 13 d,e,f. The solution is slanted about the vertical axis and subsequently about the horizontal axis (note that this order is opposite to the top/bottom model). For both the l- and r-model the computations predict a linear relation with a slope of 1 between σ and vertical slant. The slant about the horizontal axis decreases as a function of σ. In other words, these solutions completely rotate along with the screen, with a decreasing horizontal slant. Notice that the solutions (left and right) yield horizontal slants with an opposite direction (clockwise and counterclockwise). The sudden drop in figure 13e can be explained by the definition used for slant. Slant is defined as a positive or negative angle between 0 and 90°. As a result -90 and 90° correspond to the same slant. The solution presented in (figure 13e) flips from the positive to the negative region, and therefore flips from 90 to -90°.

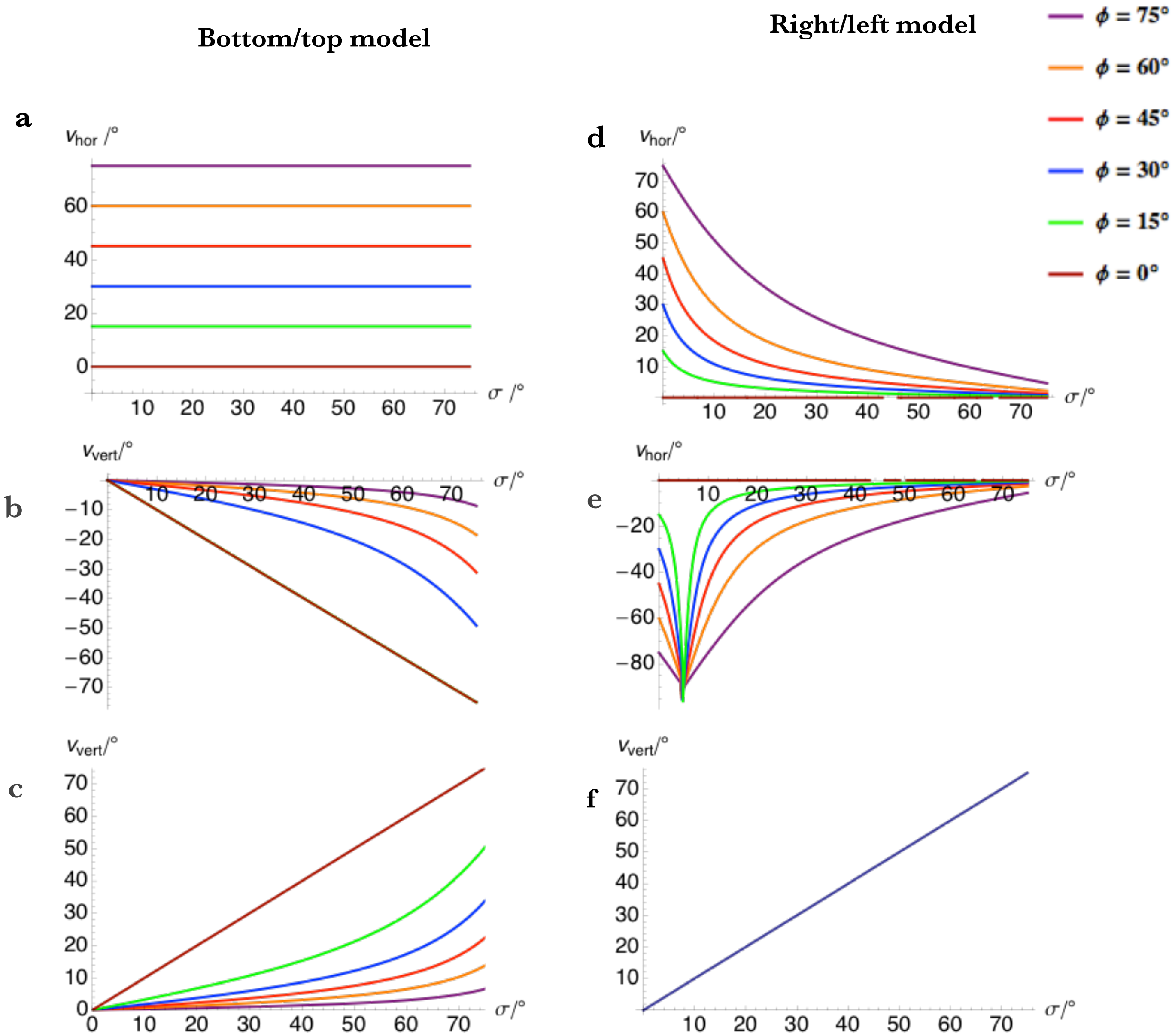


Figure 13. Horizontal and vertical virtual slant as a function of screen slant (σ). (**a**) Computed v_{hor} for top/bottom 90° angles. (**b,c**) Computed v_{ver} for the bottom model (b) and the top model (c). (**d,e**) Computed v_{hor} for right (d) and left (e) model. (**f**) Computed v_{ver} for the r- and l-model, both models predict the same vertical slant, which is equal to the screen slant.

3.2 Method

3.2.1 *Experimental setup*

The stimuli were presented on a 21-inch LaCie 321 (1600 x 1200, 75 Hz). The screen was placed on a turntable in which it could be slanted about the vertical axis in steps of 5° (-90° to +90°). A chin rest was placed at a distance 57 cm in front of the screen, which corresponds to the centre of projection at which the stimuli were computed. At this distance the visual angle of the screen was approximately 43° x 32°. A vertical pad (17 x 23 cm) was placed in between the chin rest and the screen at a distance of 25 cm. The vertical pad could be slanted about three axes: about the vertical axis, about the horizontal axis and about the depth axis.

3.2.2 *Stimuli*

The physical stimuli were rectangular grids of width 10 cm and length 15 cm, which project on the screen as trapezoids when slanted (figure 10). The rectangle consists of 5 horizontal lines and 5 vertical lines. The lines are white (± 93 cd/m^2) against a black background (±2 cd/m^2). The rectangles were slanted about the horizontal axis with a depicted slant (Φ) varying from 0° to 75° in steps of 15°. The stimuli were presented on a screen that was slanted about the vertical axis from 0° to 75° in steps of 15°.

3.2.3 *Procedure*

Six subjects were asked to indicate the perceived slant with the pad. All but the author were naive to the purpose of the study. All the subjects ran pre-trials without a stimulus on the screen to familiarise with the task. The trials were evaluated and settings were adjusted when results were inaccurate. The subjects were able to indicate slants within error margins of 5° after some practice. When indicating the horizontal slant for a frontoparallel screen, most subjects indicated the slant a few degrees higher than the expected 0°. By repeating this measure several times, a correction factor was obtained in the order of 4° . The subjects were instructed to focus on the stimulus on the screen, and not on the screen itself. The subjects viewed the screen binocularly. Combinations of depicted slant (Φ) and screen slant (σ) were presented in a random order.

3.3 Results

Figure 14 presents the mean results of experiment 2, in which the slant of the slanted rectangles on a slanted screen were indicated. The results of 5 subjects were rather similar (indicated by the standard deviations). The results of one subject deviated significantly from the other 5 subjects, these results are presented in appendix B. Note that by comparing axis system 1 and 2 (figure 11), v_{vert} and v_{hor} are slants about different axes for the l/r-model and the b/t-model. However, axis system 1 can easily be defined in terms of axis system 2. Vhor1 = Vhor2 (horizontal slant in system 1 and 2). Vvert2 (v_{vert} in axis system 2) = Vvert 1+ spin. Where spin is the rotation in the plane of the object. However this spin was indicated as zero for all observers. Therefore axis system 2 is chosen for convenience, so the slants as presented in figure 14 are defined by axis system 2.

3.3.1 *Horizontal slant*

Figure 14a shows that perceived v_{hor} is almost invariant as a function of σ for the different values of Φ. There is however a small decline for 15° ≤ Φ ≤ 60° for σ ≥ 45°. For large screen slants subjects did not perceive depth in some cases, and thus perceived the stimulus as a 2-D image. This results in large deviations in the results for σ = 75°. Although the pattern of the perceived data resembles that of the predicted v_{hor} of the b/t model (figure 14a), slant was systematically underestimated. This underestimation was also measured in

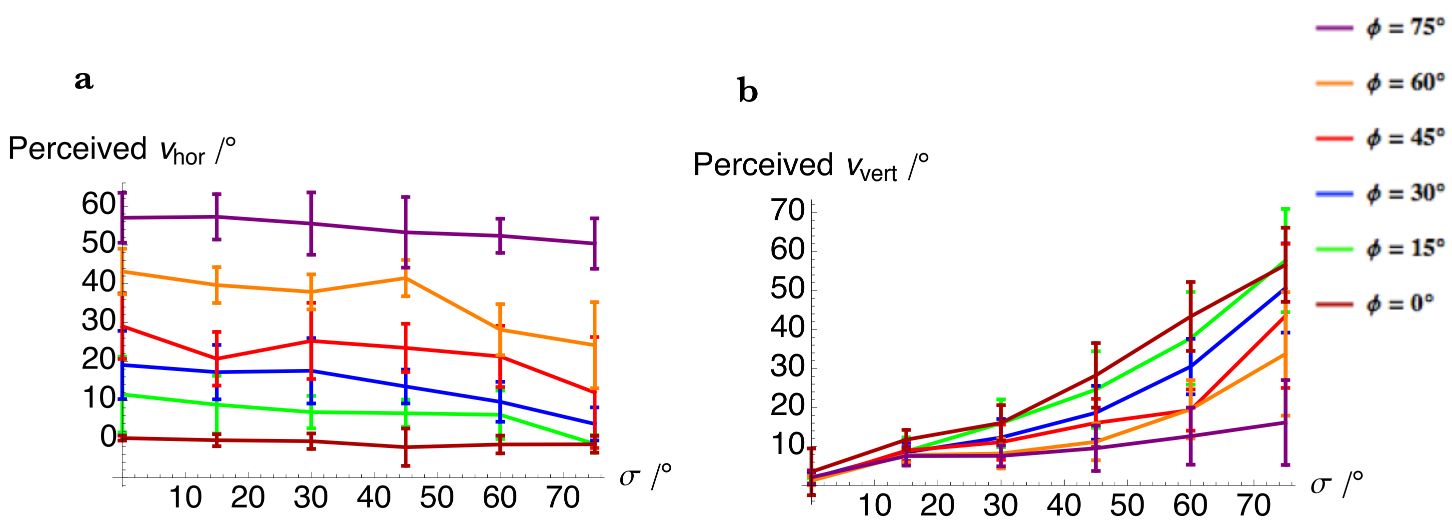


Figure 14. Mean data for 5 of the subjects, presenting perceived horizontal slant (**a**) and perceived vertical slant (**b**).The perceived slant is plotted as a function of screen slant (σ). The bars indicate ± 1 SD.

experiment 1, and seems to be a general feature in slant estimation. For higher slants ($v_{hor} \geq 60°$) the underestimation is ±25%. For lower slants ($v_{hor} \leq 45°$) the underestimation is ±33%.

3.3.2 *Vertical Slant*

The vertical slant judgments show a less obvious resemblance to either of the models. Contrary to the slants predicted by the l/r-model (figure 13f), the perceived slants differ per presented depicted slant (Φ). The t-model predicts that v_{vert} is dependent on Φ, but the perceived slants are several magnitudes larger than predicted by the t-model. This might be explained by compensation for the screen slant. This hypothesis will be statistically tested in the following section.

3.3.3 *RSS/TSS*

To quantitatively estimate the predictive quality of the different models, the residual sum of squares as a fraction of the total sum of squares (RSS/TSS) was computed for fits of the models to the mean indicated slants. Because slant was underestimated systematically, the predicted model values were given a variable weight. This way the weight factor can be found for which the model explains the data optimally. There is a significant variance between subjects for $\sigma = 75°$. However the predicted RSS/TSS barely differed when including $\sigma = 75°$, therefore this data is included for the sake of completeness.

<u>Vhor</u>

For the b/t-model RSS/TSS shows a minimum of 0.05 at $w_{b/t} = 0.60$ (figure 15a). For the r-model RSS/TSS shows a minimum of 0.40 at $w_r = 1.05$ (figure 15a). Note that the analysis for the l-model was left out because it predicts a slant opposite to the perceived slant. A combination of the bottom/top and left model shows a minimum of 0.04 at $w_{b/t} = 0.55$ and $w_r = 0.20$ (figure 15b)

<u>Vvert</u>

For the t-model RSS/TSS shows a minimum of 0.08 at $w_t = 1.45$ (figure 15c). For the l/r -model RSS/TSS shows a minimum of 0.14 at $w_{r/l} = 0.45$ (figure 15c). Combination of the t-model with screen slant (σ) shows a minimum of 0.02 at $w_t = 0.95$ and $w_s = 0.20$ (figure 15d). Note that this combination is equivalent to the combination of the l/r-model and the t-model, because the

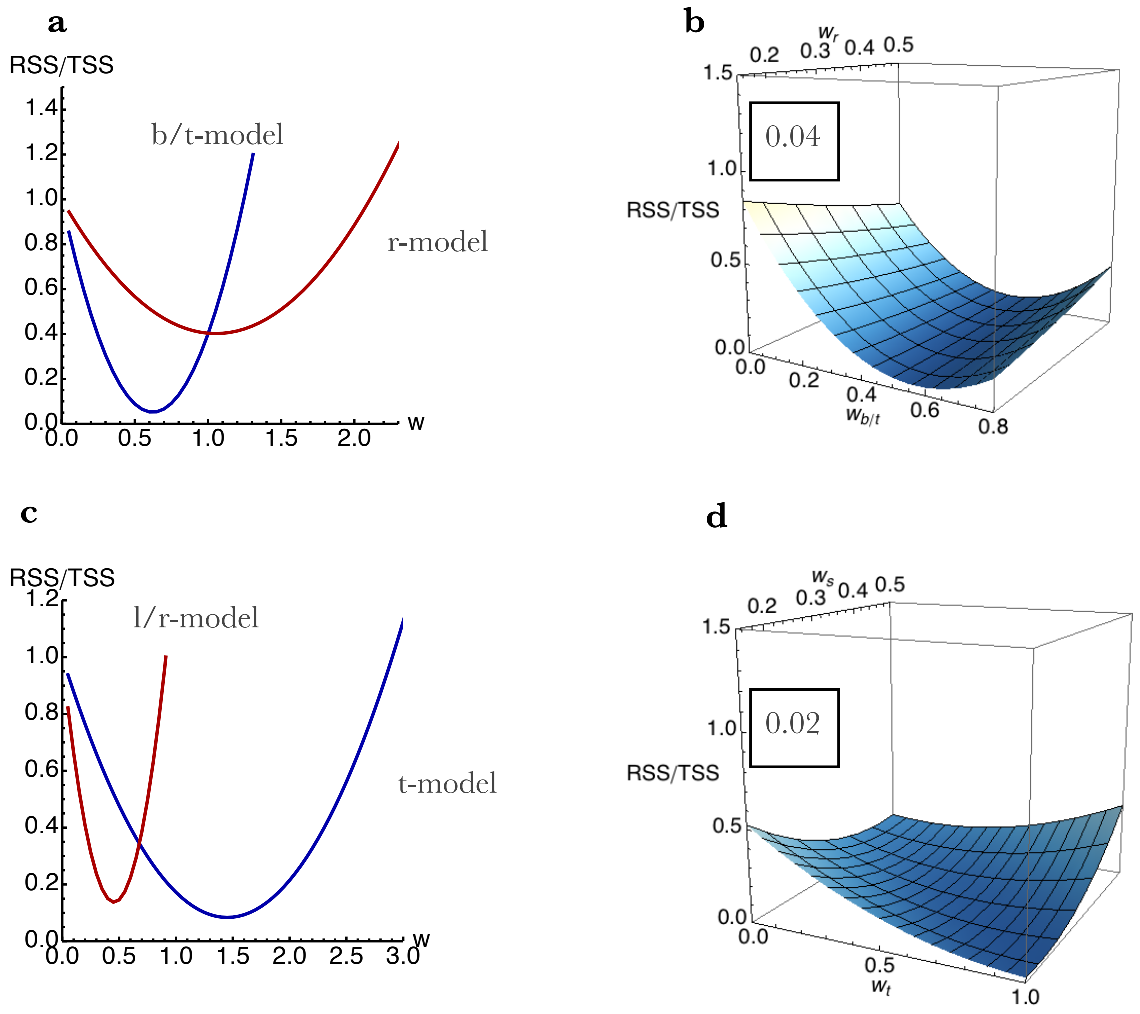


Figure 15. RSS/TSS analysis of the predictive quality of the different models. RSS/TSS is plotted as a function of the model weight factor (w). (**a**) RSS/TSS for fits of the data to b/t model and r-model. Note that the l-model is not included because it predicts a negative slant (opposite to perceived slant). (**b**) Combination of the b/t-model and r-model. RSS/TSS is plotted as a function of the weight factors (w_r and $w_{b/t}$). (**c**) RSS/TSS for vertical slant. The t-model and l/r - model are presented. RSS/TSS is plotted a function of cue weight (w). Note that the b-model is not presented because it predicts a negative horizontal slant, opposite to perceived slant. (**d**) Combination of r-model and l/r-model for perceived horizontal slant. RSS/TSS is plotted as a function of the weight factor of both models. The boxed numbers

predicted v_{vert} by the l/r-model is equal to the screen slant (σ).

These values imply that the different models explain different percentages of the perceived slants. The b/t-model explains 95% of the perceived horizontal slant. This percentage increases slightly when the b/t model is combined with the r/l model, explaining a percentage of 96%.

For the vertical slant the models explain a percentage of 92% (t-model) and 86% (l/r-model). The two models combined explain 98% of the data.

Repeating the RSS/TSS analysis for each subject individually tested the accuracy of the obtained RSS/TSS values. The obtained means with standard deviation are presented in table 1. Note that the computed residual errors are larger than the residual errors in figure 15. This difference can be explained by the reverse order of fitting and averaging. The standard deviations of the model weights are relatively large, which indicates the intersubject variability. The analysis shows that for v_{hor}, the b/t- model is far more powerful than the r/l -models. The combination of the two models results in a small increase in explanatory power (from 91% to 93%). The associated standard deviation is relatively small, so this combination is powerful in explaining the data with small intersubject differences. For perceived v_{vert} the both the b/t and l/r models are moderately explanatory (89 and 82%). However, the combination of the models explains a large part of the data (96%), with a relatively small standard deviation.

These percentages show that the b/t-model individually is a powerful predictor for perceived $v_{hor,}$ and the t-model individually is a powerful predictor for $v_{vert.}$ The l/r-model is not very predictive for perceived v_{hor} (60%), however it leads to an acceptable prediction for v_{vert} (88%).

Combination of the two models leads to a good prediction for vertical slant (96%).

Model	w	RSS/TSS minimum
Vhor		
Bottom/top	0.62 ± 0.06	0.09 ± 0.04
Right	1.06 ± 0.09	0.42 ± 0.03
Combination	0.54 ± 0.06 (B/T) 0.24 ± 0.08 (R)	0.07 ± 0.04
Vvert		
Top	1.46 ± 0.35	0.11 ± 0.05
Left/Right	0.45 ± 0.12	0.18 ± 0.08
Combination	0.85 ± 0.19 (T) 0.23 ± 0.13 (L/R)	0.04 ± 0.02

Table 1. Mean (± SD) model weights and corresponding minimum RSS/TSS values. The letters in the parentheses indicate to which model the weight corresponds.

4. Discussion

4.1 Conclusions Experiment 1

4.1.1 *General conclusion*

The first experiment showed a few significant results.

First, intersubject differences were large for an asymmetric stimulus presented on a frontoparallel screen. In contrary to similar symmetric stimuli where intersubject differences are small (Erkelens, 2013a; Erkelens, 2013b).

Second, all of the subjects perceived depth in the presented stimuli, suggesting that humans are apt to perceive depth in perspective images. This result is in accordance with previous studies (Farran et al., 2009).

Third, slant was underestimated systematically. The underestimation is in the same order as previous experiments (Erkelens, 2013a; Erkelens, 2013b; Andersen et al., 1998). See (Erkelens, 2013a) for possible explanations for the systematic underestimation of slant.

Fourth, subjects were generally not able to indicate slant by focusing on the upper part of the stimulus. Frequently subjects noted that the upper converging line could be either protruding in or out of the screen. Remarkably this bistability was only present for the upper half of the stimulus. Subjects were in all cases able to indicate slant by focusing on the lower half of the stimulus, in which case the entire stimulus is seen as either a twisted object or an asymmetric planar object. In other words, although an upper and lower converging line correspond to exactly the same slant, the visual system does not perceive the upper and lower part the same. In the following section a possible explanation for this phenomenon is proposed.

4.1.2 *Height in the visual field*

Most objects in our environment are connected to the ground plane, far less frequent objects are floating or hanging from a ceiling. Research has shown that a background surface provides important information about the position of a 3-D object (Meng & Sedgwick, 2001). When a ground plane is presented in the image, or is assumed by the subject, height in the visual field is a potential cue to depth. Because humans have an eye position at a certain height above the ground, the angular declination an object has in our visual field provides information about its depth. When objects are in contact with the ground, higher objects are usually more distant. When a ceiling is present instead, lower objects appear more distant. This could possibly lead to a cue conflict with the depth information provided by linear perspective. In the stimuli presented, the lower converging line has it's right vertex located higher than the left vertex, in the upper converging line this is opposite: the right vertex is located lower than the left vertex. Based on perspective information both edges have the same orientation in depth. However, based on height in the visual field the lower edge should correspond to an edge receding in depth, while the upper edge should correspond to an edge pointing toward you. Previous research has shown that the visual system combines cues using a statistically optimal weighted average. Further studies have shown that when conflicts between cues are large, robustness behaviour is observed: slant is specified solely by one cue (Girshick & Banks, 2009). This possibly explains the bistability observed.

The stimuli as presented to the subjects did not provide information about the relative position of the rectangle, it could be hanging, resting on a ground plane, or even floating in the air. However, observers can base their observations on an implicit horizon, either implied by the eye height of the observer, or the (implied) content of the image (S. Rogers, 1996).

The pad used to indicate the slant was positioned on a ground plane (the table), which might suggest that the rectangle in the image is also positioned on a ground plane. To investigate the significance of this experimental feature the experiment could be repeated with an experimental set in which the position of the 'indication pad' could be varied. In order to get a indication of the significance of the experimental setup, subjects were showed three stimuli at the end of the experiment. The stimuli were: a symmetric rectangle with $\Phi = 60°$, an asymmetric stimulus with $\Phi_{lower} = 60°$ and $\Phi_{upper} = 50°$ and an asymmetric stimulus with $\Phi_{lower} = 50°$ and $\Phi_{upper} = 60°$. So the two asymmetric stimuli were reversed versions of each other. When the subjects were asked qualitatively what they perceived (without having to quantify their perceiving with the pad), they again noticed a difference between the upper and lower part of the stimulus. The lower converging line is clearly protruding into the screen, while the upper converging line could be either protruding in or out of the screen. Apparently this phenomenon is independent of the setup used.

To test whether this phenomenon is caused by an implicit ground plane an experiment could be conducted in which alternately a ground plane or a ceiling is drawn in the image. In the case of the ceiling the lower converging edge would protruding out of the screen based on the height cue, the opposite should be the case when a ground plane is presented.

4.2 Link between experiment 1 and experiment 2

The original idea of experiment 1 was to investigate whether observers prefer a specific part of the asymmetric stimuli. The results could subsequently be related to experiment 2, where a asymmetric stimulus is presented by rotating the screen. However, the stimuli used in the two experiments are slanted about different axes (the vertical axis in experiment 1 and the horizontal axis in experiment 2). The initial idea was to use rectangles slanted about the vertical axis for both experiments. The screen would then be slanted about the horizontal axis to create asymmetric retinal stimuli in experiment 2. Unfortunately this setup was nog possible with the material used, therefore horizontally slanted rectangles were used in experiment 2. Therefore the link between the two experiments is missing, the results of both experiments yield separate conclusions.

4.3 Conclusions experiment 2

4.3.1 *General conclusion*

Experiment 2 shows that the perceived slants can be strongly related to the virtual slant predicted by the t-model, either by itself or in combination with the other models. For the horizontal slant the optimal weight factor implies that slant was underestimated by about 40%. This underestimation is similar to the underestimation found in similar research (Erkelens 2013a, 2013b).

For perceived vertical slant, the weight factor for the t-model implies a overestimation of about 50%. When the t-model is combined with the screen slant, the weight factors imply that the overestimation is equal to 20% of the screen slant, when almost fully utilising the t-model (w_t = 0.95).

For the perceived horizontal slant, the screen slant seems to have only a small effect on the perceived slant: the perceived horizontal slant is largely invariant, except for a small decline for $15° \leq \Phi \leq 60°$ at $\sigma \geq 45°$.

Contrary to the perceived horizontal slant, the vertical slant can also be well explained by the l/r-model (same as screen slant). This raises the question whether the vertical slant depends on the depicted slant, because the t-model predicts a dependency on Φ, while the vertical slants predicted by the l/r-model do not depend on Φ. As can be seen in both the predictions (figure 13c) as the perceived slants (figure 14b), the vertical slants for different depicted slants lie close together, which makes it hard to determine whether there is a significant difference. This also explains why both models can explain a large part of the data. In other words, the models are insufficiently different to determine which model most accurately explains the data. A repeat experiment could be conducted where the predictions for different depicted slants more significantly differ (as is the case for rectangles with a smaller aspect ratio for example).

4.3.2 *Significance of screen slant*

In this experiment it cannot be decided whether the screen slant is significant in vertical slant judgments, because the l/r-model predicts the same vertical slant as the screen slant. However, because the l/r-model seems insignificant for the perception of the horizontal slant, it is questionable whether this model is significant for the vertical slant. This would imply that observers only partially base their slant indication on the virtual object predicted by the l/r model (only for vertical slant, not for horizontal slant).

An interesting future experiment would be presenting images with equal proximal stimuli on a frontoparallel (as in figure 11) and slanted screen. This way perspective information remains constant with increasing screen slant, and therefore the experiment would isolate cues related to screen slant. The results found in (Erkelens, 2013b) suggested that screen cues were insignificant and perspective information dominated. It is questionable whether this is also the case for asymmetric stimuli. Results of experiment 1 suggest that subjects are able to perceive asymmetries in the image, which subsequently leads to multiple possible interpretations of the proximal image. In experiment 2 subjects were generally able to indicate slants without noticing asymmetries and/or multiple interpretations. Apparently the asymmetries in experiment 2 did not impair slant judgements.

4.3.3 *Shape of the virtual object*

The experiments in this research focused on slant perception. Another important aspect of the image is its shape. The computed virtual slants belong to a virtual object with a shape that varies as a function of the screen slant. The depicted rectangles have an aspect ratio (height to width) of 1.5.

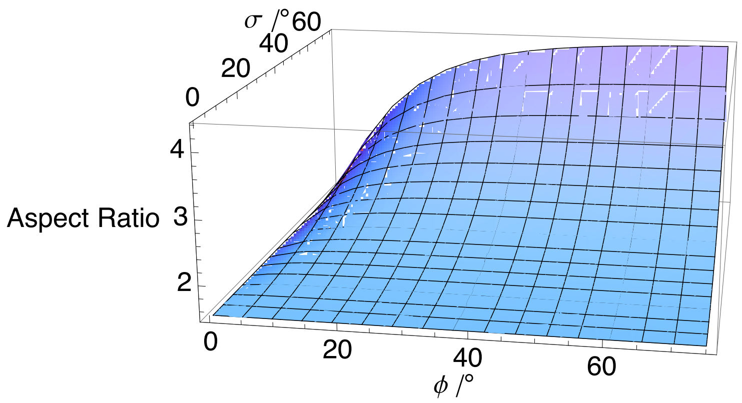


Figure 16. Aspect ratio of the virtual object as computed for the t-model as a function of Φ and σ.

As the screen slant increases, the aspect ratio of the corresponding virtual object also increases. Figure 16 shows the relation between the aspect ratio, the depicted slant and the screen slant for the t-model (the other models have similar plots). Note that the aspect ratio is defined as height/width, while the virtual object does not have one height, because the two side edges have a different size. The mean height of the two edges is taken as the height for this plot. As can be seen in the plot the aspect ratio increases from 1.5 to $\pm$ 4 for $\Phi \geq 20°$. So the aspect ratio increases by a factor of almost 3. The aspect ratio increases exponentially from $\pm$ σ =45°. Further research could investigate the perceived shape for different viewing angles/screen slants.

4.3.4 *Picture perception*

This study shows that the slant perception of rectangular grids can be strongly related to computed virtual slants. Erkelens showed similar results for rectangular grids in which the depicted slant and screen slant were slanted about the same axis. Both this study and the study by Erkelens used basic rectangles. Linear perspective is (albeit a strong one) one of the several cues that are generally present in 'real' pictures of natural scenes. Whereas geometric information changes when an image is viewed obliquely, other cues are viewpoint-invariant (e.g. shadowing and occlusion). Furthermore the rectangle grids used are shapes without a representation of a real world object. Therefore it is still a bridge too far to predict slants of natural scenes from stimuli as used in this experiment. A future study could gradually add cues to stimuli as used in this experiment. It could for example use rectangular grids as in experiment 1 and let it represent a door. This could provide interesting results concerning the role of knowledge of the depicted objects when they are viewed obliquely.

The model used in this research can also be used for rectangles slanted about other axes than the vertical/horizontal axis. This research shows that perceived slants can be related to linear perspective when the retinal stimulus is asymmetric. Further research could investigate the slants of rectangles about other axes. Research has shown that complex objects can be divided in a superposition of simple shapes (the so-called geons). These consists of shapes like rectangles, triangles and parallelograms. Superposition of these shapes is shown to be sufficient to represent complex objects (Biederman & Gerhardstein, 1993). In the future the gap between the slant perception of 'real' pictures and simple shapes might be closed by the superposition of simple shapes.

Appendix A

Constructing and reconstructing a perspective image

Constructing a perspective image

A 3-D object projects onto a 2-D surface by the array of lines that fall on the eyes of the viewer. The screen on which the image is projected can be mathematically defined as a plane. The array of lines between the eye and the object cuts through this plane, and thereby defines the points of projection (figure 17). The mathematics involved are performed in Wolfram Mathematica (version 9). Vectors are defined to represent the array of lines rays between the viewer and the object. A right-handed axis system is used, with the positive x-axis corresponding the right, and the positive y-axis corresponding to increasing height. First, a centre of projection is defined, corresponding to the point where the cyclopean eye of the viewer will be located. This point is defined in the origin of the axes-system, so: $\{x, y, z\} = \{0,0,0\}$. The 3-D object is a planar rectangle, and can be defined by four coordinates connected by straight lines. The coordinates of the rectangle correspond to the vectors between the viewpoint to the vertices of the rectangle. By multiplication the vectors with coefficients (c1, c2, c3, c4 in figure 17) , the lines can be elongated or shortened. The frontoparallel screen is defined as a vertical surface at a distance of 570 mm from the viewer. By equating the vectors to the surface, the coordinates are found at which the vectors ‘cut’ though the screen, and thus is the projection of the 3-D object on the 2-D screen.

Reconstructing a 3-D object from a perspective image

Reconstructing the 3-D object from the 2-D image is the reverse process of constructing a perspective image, except that the slanted rectangle is only one of the infinite number of possible virtual objects corresponding to the image. Constraints have to be implemented to arrive at a single solution.

Again the viewpoint is defined as the point in the origin {0,0,0} in a right-handed axes system. The screen is positioned at a distance of 570 mm of the viewer, so the centre of the screen has the coordinate (0,0, -570). The vectors v1, v2, v3, v4 represent the vectors from the viewpoint through the corners of the image on the screen. The coefficient elongates or shortens the vector. By subtracting the vectors (v2 – v1 etc.) the sides of the image are defined (a, b, c ,d in figure 17). The next step is to apply constraints to the sides to define the virtual object. In this case we ought to find the rectangular planar solution. The

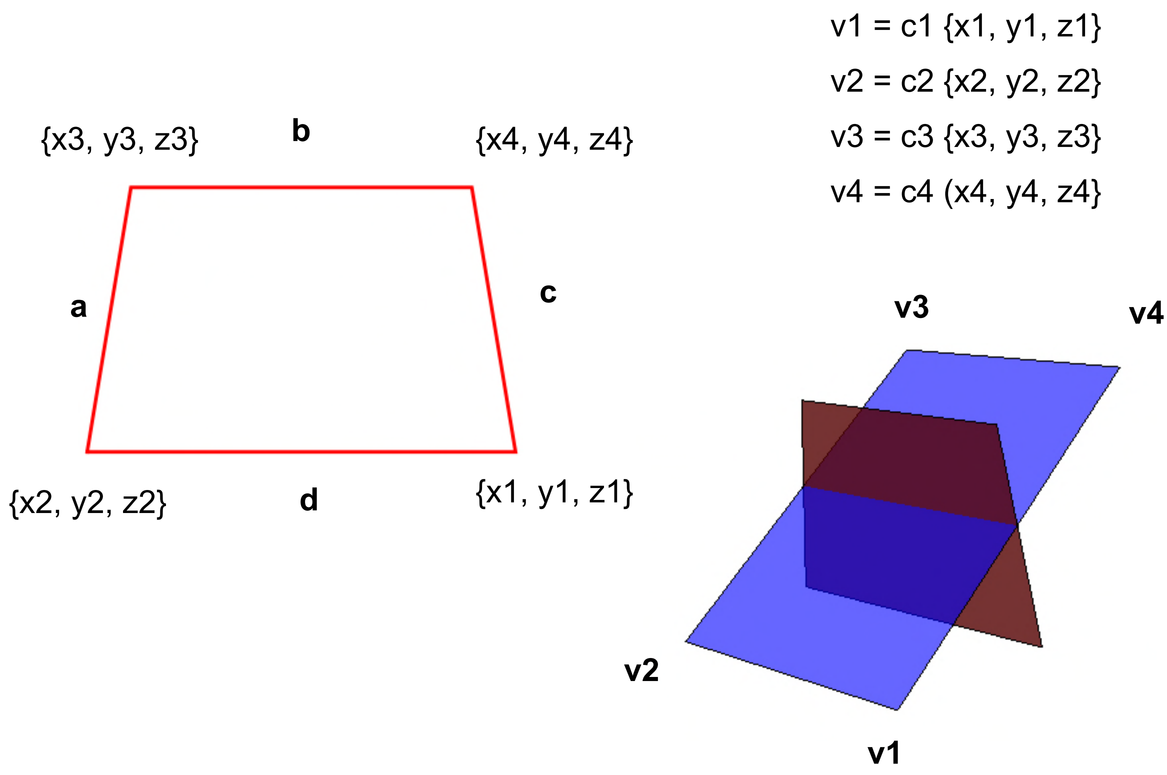


Figure 17. Representation of the stimulus as depicted on the screen.

properties are: orthogonal sides, the sides lie on the same surface and straight lines connecting the vertices. The latter property is already implicitly built-in by defining the sides as vector subtractions (which creates straight lines). Mathematically orthogonal sides can be found by solving the equations in which the inner product between the adjacent sides is zero. Furthermore the equations can be solved for which the adjacent sides have equal cross products (when normalised), and therefore lie on the same plane. For the case that the viewer coincides with the centre of projection these equations reduce four variables (c1, c2, c3, c4) to one variable (either c1, c2, c3 or c4). The remaining 'free' variable is not constrained, it can take any number. This represents the fact that perspective information alone is not sufficient to retrieve information about the size of the object. So the solutions fix the relationship between the vectors (i.e. the orientation/slant of the object) but not the absolute size. A rotation matrix can be applied to the vectors to represent the screen slant. The matrix can be defined to rotate coordinates in any direction about the origin of the coordinate system. In this case the rotation matrix which corresponds to a rotation about the vertical axis was used. The same process as described before can be used to extract a 3-D image from the rotated coordinates. However, the constraints of four 90° angles and a planar surface does not give a solution for the slanted stimulus. It is however possible to reduce the variables to one, with the constraints of two right angles and a planar surface. This provides the solutions as presented in experiment 2.

Appendix B

Deviating results for experiment 2

Figure 18 presents the perceived slants for the subject that was not included in the mean data of figure 14. Comparing the data with the mean data for the other 5 subjects shows a few significant differences. First, the decline in perceived v_{hor} is larger than the mean data. The perceived v_{vert} is remarkably similar for varying depicted slants, in contrast to the mean data where the slopes were significantly different for varying depicted slants. A RSS/TSS analysis was performed to statistically test whether this data can be explained by one of the models.

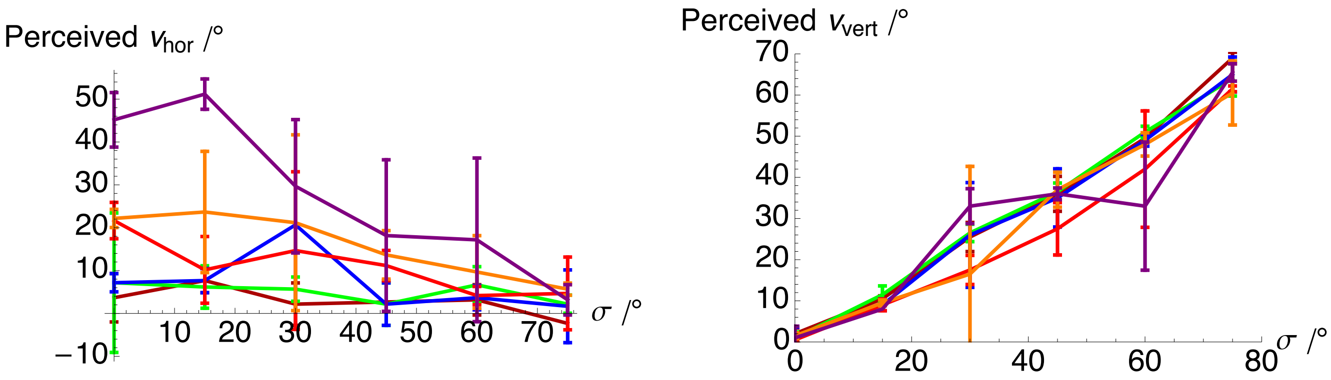


Figure 18. Perceived v_{hor} (a) and perceived v_{vert} (b) for one subject.

<u>Vhor</u>

For the b/t-model RSS/TSS shows a minimum of 0.28 at $w_{b/t}$ = 0.30 . For the r-model RSS/TSS shows a minimum of 0.26 at w_r = 0.65. Note that the analysis for the l-model was left out because it predicts a slant opposite to the perceived slant. A combination of the bottom/top and left model shows a minimum of 0.15 at $w_{b/t}$ = 0.15 and w_r = 0.45.

<u>Vvert</u>

For the t-model RSS/TSS shows a minimum of 0.30 at w_t = 2.05. For the l/r -model RSS/TSS shows a minimum of 0.02 at $w_{r/l}$ = 0.80. Combination of the t-model with the l/r-model shows a minimum of 0.01 at w_t = 0.20 and $w_{l/r}$ = 0.75. Note that this combination is equivalent to the combination of the screen slant and the t-model, because the predicted v_{vert} by the l/r-model is equal to the screen slant (σ).

The analysis shows that neither of the models is a powerful predictor for perceived $v_{hor.}$ The combination of the two models explains 85% of the data. The t-model explains 70% of the perceived v_{vert}. However, the l/r-model (equivalent to screen slant) explains 98% of the data and increases slightly by adding the t-model, explaining 99%. It seems that this subject fully compensated for screen slant. Unfortunately there is no time left within this research for a repeat experiment. It would be interesting to repeat the experiment with this subject with a invisible screen (e.g in a dark room), to check whether the results differ when there is no information available of the screen slant.

References

1. Andersen, G. J., Braunstein, M. L., & Saidpour, A. (1998). The perception of depth and slant from texture in three-dimensional scenes. *Perception, 27*(9), 1087-1106.

2. Barlow, H. B., & Reeves, B. C. (1979). The versatility and absolute efficiency of detecting mirror symmetry in random dot displays. *Vision Research, 19*(7), 783-793.

3. Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance, 19*(6), 1162-1182.

4. Braunstein, M. L., & Payne, J. W. (1969). Perspective and form ratio as determinants of relative slant judgments. *Journal of Experimental Psychology, 81*(3), 584-590.

5. Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: The interaction, relative potency, and contextual use of different information about depth. In W. Epstein, & S. Rogers (Eds.), *Perception of space and motion* (pp. 69-117). San Diego, CA: Academic Press.

6. Cutting, J. E. (1987). Rigidity in cinema seen from the front row, side aisle. *Journal of Experimental Psychology: Human Perception and Performance, 13*(3), 323-334.

7. Epstein, W. (1966). Perceived depth as a function of relative height under three background conditions. *Journal of Experimental Psychology, 72*(3), 335-338.

8. Erkelens, C. J. (2013a). Computation and measurement of slant specified by linear perspective. *Journal of Vision, 13*(13)

9. Erkelens, C. J. (2013b). Virtual slant explains perceived slant, distortion, and motion in pictorial scenes. *Perception, 42*(3), 253-270.

10. Farran, E. K., Whitaker, A., & Patel, N. (2009). The effect of pictorial depth information on projected size judgments. *Attention, Perception, and Psychophysics, 71*(1), 207-214.

11. Girshick, A. R., & Banks, M. S. (2009). Probabilistic combination of slant information: Weighted averaging and robustness as optimal percepts. *Journal of Vision, 9*(9)

12. Goldstein, E. B. (1979). Rotation of objects in pictures viewed at an angle: Evidence for different properties of two types of pictorial space. *Journal of Experimental Psychology: Human Perception and Performance, 5*(1), 78-87.

13. Goldstein, E. B. (1987). Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *Journal of Experimental Psychology: Human Perception and Performance, 13*(2), 256-266.

14. Gombrich, E. H. (1972). The visual image. *Scientific American, 227*(3), 82-96.

15. Juricevic, I., & Kennedy, J. M. (2006). Looking at perspective pictures from too far, too close, and just right. *Journal of Experimental Psychology: General, 135*(3), 448-461.

16. King, M., Meyer, G. E., Tangney, J., & Biederman, I. (1976). Shape constancy and a perceptual bias towards symmetry. *Perception and Psychophysics, 19*(2), 129-136.

17. Koenderink, J. J., Van Doorn, A. J., Kappers, A. M. L., & Todd, J. T. (2004). Pointing out of the picture. *Perception, 33*(5), 513-530.

18. Ling, Y., Nefs, H. T., Brinkman, W. -., Qu, C., & Heynderickx, I. (2013). The effect of perspective on presence and space perception. *PLoS ONE, 8*(11)

19. Meng, J. C., & Sedgwick, H. A. (2001). Distance perception mediated through nested contact relations among surfaces. *Perception and Psychophysics, 63*(1), 1-15.

20. Norman, J. F., Todd, J. T., Norman, H. F., Clayton, A. M., & McBride, T. R. (2006). Visual discrimination of local surface structure: Slant, tilt, and curvedness. *Vision Research, 46*(6-7), 1057-1069.

21. Ozkan, K., & Braunstein, M. L. (2010). Background surface and horizon effects in the perception of relative size and distance. *Visual Cognition, 18*(2), 229-254.

22. Palmer, S. E. (1999). Perceiving surfaces oriented in depth. *Vision science, photons to phenomenology* (pp. 199-253). Massachusetts: MIT Press.

23. Papathomas, T. V. (2002). Experiments on the role of painted cues in hughes's reverspectives. *Perception, 31*(5), 521-530.

24. Papathomas, T. V. (2008). Art pieces that 'move' in our minds - an explanation of illusory motion based on depth reversal. *Spatial Vision, 21*(1-2), 79-95.

25. Perkins, D. N. (1973). Compensating for distortion in viewing pictures obliquely. *Perception and Psychophysics, 14*(1), 13-18.

26. Rogers, B., & Gyani, A. (2010). Binocular disparities, motion parallax, and geometric perspective in patrick hughes's 'reverspectives': Theoretical analysis and empirical findings. *Perception, 39*(3), 330-348.

27. Rogers, S. (1996). The horizon-ratio relation as information for relative size in pictures. *Perception and Psychophysics, 58*(1), 142-152.

28. Rosinski, R. R., Mulholland, T., Degelman, D., & Farber, J. (1980). Picture perception: An analysis of visual compensation. *Perception & Psychophysics, 28*(6), 521-526.

29. Saunders, J. A., & Backus, B. T. (2006). The accuracy and reliability of perceived depth from linear perspective as a function of image size. *Journal of Vision, 6*(9)

30. Saunders, J. A., & Backus, B. T. (2007). Both parallelism and orthogonality are used to perceive 3D slant of rectangles from 2D images. *Journal of Vision, 7*(6)

31. Saunders, J. A., & Knill, D. C. (2001). Perception of 3D surface orientation from skew symmetry. *Vision Research, 41*(24), 3163-3183.

32. Stevens, K. A. (1983). Slant-tilt: The visual encoding of surface orientation. *Biological Cybernetics, 46*(3), 183-195.

33. Todd, J. T., Thaler, L., & Dijkstra, T. M. H. (2005). The effects of field of view on the perception of 3D slant from texture. *Vision Research, 45*(12), 1501-1517.

34. Todorovic, D. (2005). Geometric and perceptual effects of the location of the observer vantage point for linear-perspective images. *Perception, 34*(5), 521-544.

35. Todorovic, D. (2008). Is pictorial perception robust? the effect of the observer vantage point on the perceived depth structure of linear-perspective images. *Perception, 37*(1), 106-125.

36. Todorovic, D. (2009). The effect of the observer vantage point on perceived distortions in linear perspective images. *Attention, Perception, and Psychophysics, 71*(1), 183-193.

37. Vangorp, P., Richardt, C., Cooper, E. A., Chaurasia, G., Banks, M. S., & Drettakis, G. (2013). Perception of perspective distortions in image-based rendering. *ACM Transactions on Graphics, 32*(4)

38. Vishwanath, D., Girshick, A. R., & Banks, M. S. (2005). Why pictures look right when viewed from the wrong place. *Nature Neuroscience, 8*(10), 1401-1410.

39. Wijntjes, M. W. A. (2014). A new view through alberti's window. *Journal of Experimental Psychology: Human Perception and Performance, 40*(2), 488-501.

40. Yang, T., & Kubovy, M. (1999). Weakening the robustness of perspective: Evidence for a modified theory of compensation in picture perception. *Perception and Psychophysics, 61*(3), 456-467.