7

Bump height matching[1]

The spatial impression the observer gets from an image (the *apparent layout*) does not necessarily match the geometry of the original scene as could be measured with a measuring rod (*Euclidean layout*). I will refer to such a mismatch between apparent and Euclidean layout as *distortion*. It is known that distortion is affected by the viewing position of the observer relative to the image (his *viewpoint*). For example, Smith and Gruber (1958) showed that apparent distances in a photograph of a corridor increase with viewing distance. In virtual window displays, where the displayed picture is adapted continuously to match the actual viewpoint of the observer, such distortions may be corrected. Whether distortions occur in virtual window displays will depend on the relation between the viewpoint of the observer and the camera settings (its position, rotation and viewing angle). I will refer to this relation between the observer's viewpoint and the camera settings as the *coupling method*. In virtual window displays, distortion may also be caused by inaccurate implementation of the coupling, for example due to delays in the update cycle or errors in the measurement of the viewpoint. The experiment described in this chapter investigates the effect of the coupling method and viewpoint measurement inaccuracies on distortion, using a height matching task. Distortions in virtual window displays may reduce task performance, and may therefore be important for eliminating such distortions in a baggage inspection system based on the DVWS.

If the visual angle subtended by objects affects the apparent depth of the scene, it can be expected that the optical angles in the photograph should match the optical angles in the real scene in order to give the same apparent depth (Pirenne, 1975). Only one viewpoint fulfils this constraint (is *geometrically equivalent*). This geometrically equivalent viewpoint matches the position of the lens of the camera relative to the recording plate, and scales with enlargement of the picture. Figure 7.1 illustrates this geometrically equivalent viewpoint for viewing a picture made with a normal camera. As the lens was 5 cm in front of the centre of the picture when the photograph was made, and the photograph was enlarged two times, the geometrically equivalent viewpoint lies 2*5=10 cm in front of the centre of the picture.

[1]Part of this work was presented in Pasman, Stappers and Smets (1997).

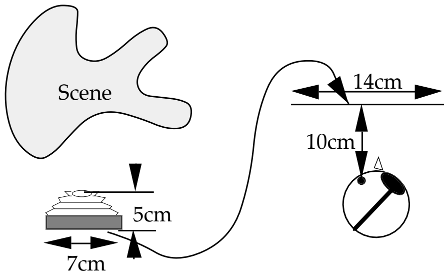


Figure 7.1. A geometrically equivalent viewpoint is a viewpoint of the observer where all visual angles subtended by the displayed objects match those of the real objects. See text.

Figure 7.2 illustrates that a geometrically inequivalent viewpoint can lead to distortion. The geometrically equivalent viewpoint lies at about 8 cm in front and 2 cm above the middle of the picture. You should slant the book 45° to get the horizon of the picture (actually above the photo) at eye level. If the picture is viewed more from the bottom (but at an equal perpendicular distance from the picture) the 2 barrels at the right, in particular, appear shorter. This effect is called apparent depth compression. If the picture is viewed more from the left, the tops of the barrels appear to move leftward relative to the floor, especially when the picture is also viewed from below. This effect is called apparent shear.

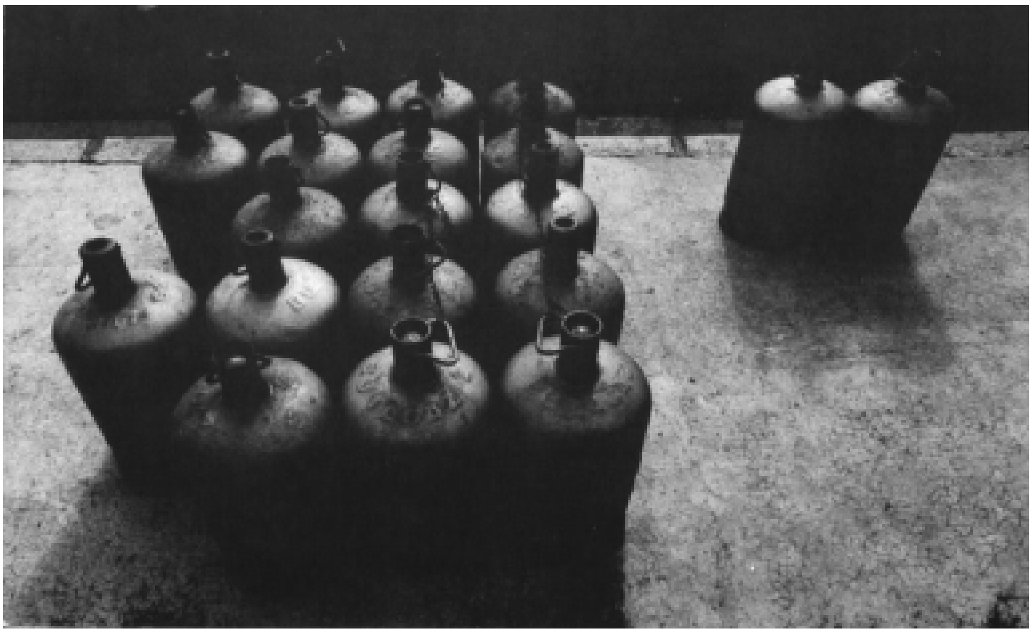


Figure 7.2. The barrels appear distorted at geometrically inequivalent viewpoints. See text. (From Petzold, 1973).

Similar effects are known for portraits, where the eyes of the depicted person seem to follow the observer, and in pictures of someone pointing his finger out of the picture, where his finger appears to keep pointing at the observer as he moves (Zorin, 1995).

For most images, such distortions are inconspicuous (Cutting, 1987). This may be related to the task: for example recognition of your family members or admiration of a beautiful landscape need not be hindered by such distortions. There is evidence that human observers actively compensate for the distortions caused by a geometrically inequivalent viewpoint, but only if the cues informing the observer that he is looking at a picture are strong enough. Such cues may consist of cues to the flatness of the image, such as a clearly visible frame around the picture, and a raster in front of the image (e.g., when the picture is displayed with pixels on a monitor).

In virtual window displays the displayed image is adapted to the viewpoint of the observer. This may reduce or even eliminate the distortions. To do this, the viewpoint has to be measured accurately, and a failure to do so (a *viewpoint measurement error*) may cause distortion. Furthermore, if the observer actively compensates for his non-perpendicular (*oblique*) viewpoint, adaptation of the display to his viewpoint might be counterproductive.

In the sections below, I will describe the various coupling methods and previous work that has been done on distortion in picture perception. Next, an experiment investigating distortions in virtual window displays will be described. Finally, the results and their implications for virtual window design will be discussed.

Coupling methods

Virtual window displays can be implemented by different coupling methods. Here, on-axis and off-axis coupling methods are considered (Figure 7.3). As the coupling method is similar for both eyes, it is sufficient here to discuss the case of only one eye. This coupling causes objects to shift relative to each other as the observer moves (*movement parallax*), and movement parallax is an important depth cue (Chapter 2).

With *on-axis coupling* (Figure 7.3a), as in the Delft Virtual Window System (Chapter 1), the lens of the camera always stays before the middle of the recording plate (thus making *on-axis images*). For the DVWS case considered here, the camera's motions are constrained so that it rotates at a fixed distance around the point of interest in the 3D scene, the *fixation point*. The angular position of the camera is controlled by the angular position of the observer. As the camera keeps aimed at the fixation point, the fixation point is always in the centre of the image.

With *off-axis coupling* (Figure 7.3b), such as the fishtank VR system (Ware, Arthur and Booth, 1993; Castle, 1995) both the camera and the lens position are coupled to the position of the observer. The camera translates with the observer, but does not rotate (thus making *off-axis images*). The lens translates relative to the recording plate as the observer translates relative to the display. Consequently, all visual angles subtended by objects in the image correspond to their angles in the real scene when viewed from the same viewpoint. With real cameras, off-axis images might also be acquired with a wide-angle image of which only a small part is used at a time, or with a camera making on-axis images and real-time image processing. With computer generated perspective images, both on-axis and off-axis coupling are easy to achieve.

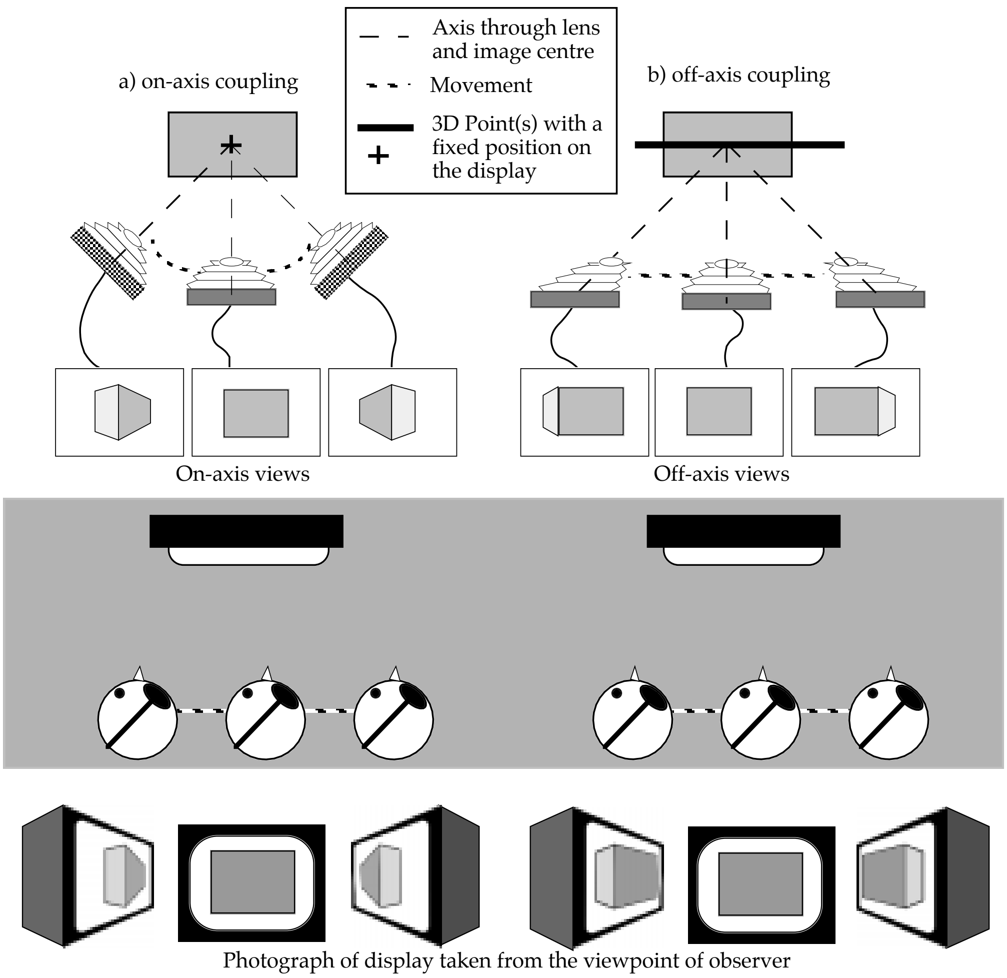


Figure 7.3. Two different coupling methods used for virtual window displays. With on-axis coupling the camera rotates around the fixation point as the observer rotates around the middle of the display. With off-axis coupling the camera translates with the movements of the observer. To select the appropriate part of the scene, the lens has to shift relative to the image sensor inside the camera.

Previous work

There are many visual cues that can cause distortion if presented inappropriately. Some general guidelines for avoiding distortions give a general idea about what cues are important to consider. Gibson (1971) suggested that for optimally viewing perspective pictures, "the observer should look with one eye at the correct viewing distance, the picture should be upright and perpendicular to the line of sight, and there should be a reduction screen in front of the picture hiding the rest of the world" (p. 30). Pirenne (1970), Gourneri (1859, see Cutting, 1986) and Jones and Hagen (1978) note that many artists have strategies to minimise distortions: "(1) stress the depiction of the texture of surfaces, (2) avoid extreme perspective convergence and (3) choose a station point [viewpoint, in our terminology] that presents a characteristic view of the scene." (Jones and Hagen, 1978, pp. 191-192). As discussed in the introduction, the coupling method, the geometric equivalence of the viewpoint, the accuracy of the measurement of the observer's viewpoint and cues to the flatness of the image may all have an impact on the apparent layout of the scene. If the observer mentally compensates for distortions expected at an oblique viewpoint, this may also affect the distortion. Furthermore, several other factors, such as the layout of the scene and delays in the system, may cause distortion. Although a lot of experiments have been done on distortion in picture perception due to geometric inequivalence (Smith and Gruber, 1958; Pirenne, 1970, 1975; Rosinski and Farber, 1980; Cutting, 1986; Halloran, 1989), little is known of what happens when the depicted image is coupled to the movements of the observer, as in virtual window displays. The following sections discuss previous work that has been done on these issues.

Geometric inequivalence

Much research has been done on the effect of geometric inequivalence on the apparent layout of static, on-axis images. Nearly all of these studies assume that humans make a 3D reconstruction of the viewed scene (see Chapter 2) using only the rules of geometry. This is tempting, as perspective cues from a single view and a few assumptions allow a precise depth reconstruction, but this is not how actual human perception operates, and none of the theories discussed in Chapter 2 suggests that the apparent layout depends on perspective cues only. Nevertheless, geometry is a useful explanation for distortion because it allows prediction of distortions in a mathematical way, and because it connects to most literature concerning distortions.

Figure 7.4 illustrates how geometric inequivalence may lead to distortion. If the viewing distance is smaller than geometrically equivalent (Figure 7.4b), a 3D reconstruction using geometry will be compressed in depth as compared to the original scene.

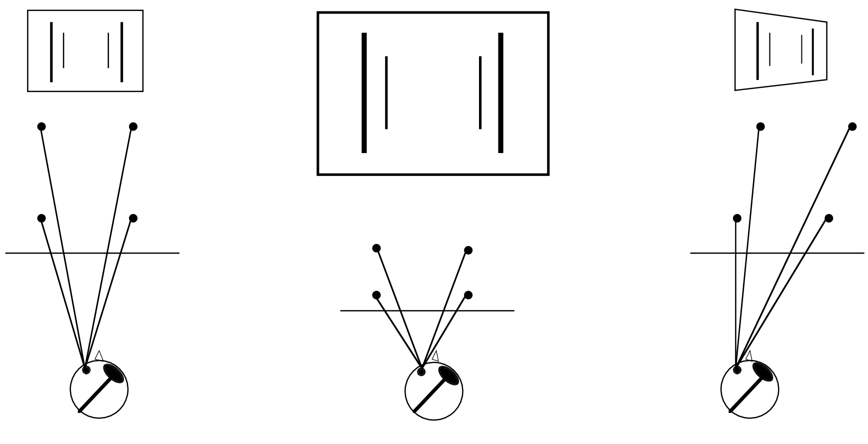


Figure 7.4a. Top: a view. Bottom: 3D reconstruction from perspective, at the geometrically equivalent viewpoint.

Figure 7.4b. If the viewpoint is too close, the view appears twice as large.
A geometric reconstruction becomes compressed in depth.

Figure 7.4c. Geometric reconstruction gives shear distortion if the viewpoint is displaced parallel to the display (from Cutting, 1986)

If the observer looks from a distance d' rather than the camera distance d, geometry suggests that the apparent depth will be scaled by a factor d'/d. If the viewpoint has been displaced parallel to the display, the 3D reconstruction from perspective cues will be sheared (Figure 7.4c) (Rosinski and Farber, 1980; Adams, 1972; Lumsden, 1980).

Numerous experiments show that apparent depth increases with increasing viewing distance, i.e. that human perception is sensitive to geometric cues. For example, Smith and Gruber (1958) showed that apparent distances in a photograph of a corridor increase with increasing viewing distance, though the effect is smaller than predicted by geometry. Adams (1972) displayed a floor and background of tiles. Participants matched the width-height ratio of tiles on the background to those on the wall. The effects matched geometric predictions, although the size of the effect is smaller than predicted. Purdy (1960), Smith and Gruber (see Rosinski and Farber, 1980) and Braunstein and Payne (1969) found similar results for the perception of a projected grid of fine opaque lines, that was magnified and slanted. An inconvenient point about this evidence for the effect of geometric inequivalence on apparent depth is that most of the stimuli used did not contain other depth cues (such as texture, shading, stereoscopic, shadow and accommodation cues), thus disabling the participants to use cues other than perspective. Therefore, these results may not be generalized to practical applications such as x-ray baggage inspection via a virtual window display.

A geometric reconstruction of an on-axis image viewed obliquely will be sheared. There is some evidence that shear distortion of this kind also occurs in human perception. Hochberg (1986) indicated in the cinema situation that distortion is usually not noticeable from oblique viewpoints, but that when relative sizes and velocities become important, as in dance, the distortions become obtrusive. Cutting (1987) displayed rotating wireframe cubes on a slanted display. With a slant of 45° the apparent rigidity of cubes is lower than with 67° or 90° (head on) display slant. For television viewing, McVey (1970) indicated that oblique viewing positions up to 45° oblique produce images with an 'acceptable' amount of distortion, but that informational displays (e.g. graphs) become problematic. For movies, telelenses may reduce distortion for viewers looking obliquely at a film (Cutting, 1989).

In our literature search, I did not encounter any experiment dealing with the effect of geometric inequivalence in virtual window displays. In all cases, only static viewpoints were considered.

Human insensitivity to geometric inequivalence

In contrast with the evidence above, many researchers report that viewing from a geometrically inequivalent viewpoint usually does not lead to distortions. For example paintings usually do not appear distorted when one looks at them from an oblique viewpoint, and viewing distances also have little effect (Pirenne, 1970, 1975; Rosinski and Farber, 1980). Again, objects on a television do not appear to deform when the camera moves (Cutting, 1989). Perkins (1973) showed that observers looking obliquely at 26° and at 41° from the image normal can usually distinguish depicted rectangular corners from nonrectangular corners.

At least three mechanisms have been proposed that might allow such undistorted information to be extracted from a geometrically inequivalent viewpoint. The first mechanism (Rosinski and Farber, 1980; Pirenne, 1970; Pirenne, 1975) assumes that the image has been made with an on-axis camera, and thus an estimation can be made of the geometric equivalent viewpoint. The geometrically equivalent viewpoint is in front of the middle of the image, and Rosinski and Farber (1980) suggest that the viewing distance may be estimated from the width and height of the image, or from image clarity. The distance between the actual and the estimated geometrically equivalent viewpoint can be used to correct for the oblique viewpoint[2]. The second mechanism suggests that the geometrically equivalent viewpoint can be recovered from perspective cues in the image, perhaps using assumptions about the Euclidean layout. As with the first mechanism, this allows a correction for the oblique viewpoint. The third mechanism (Cutting, 1987) proposes that only depth cues that are relatively unaffected by the viewpoint be used.

There is evidence that humans may use the first mechanism, which assumes that the geometrically equivalent viewpoint lies perpendicular to the display. Rosinski and Farber (1980) suggest that Perkins (1973) enabled this mechanism by not using a reduction screen, but that Smith and Gruber (1958) and Purdy (1960) (see 'Geometric inequivalence') did not enable it by using a reduction screen. Rosinski and Farber (1980) did two experiments. In the first, observers making slant judgements in photographs of a slanted, striped surface did not compensate for geometric inequivalence when looking with one eye and through an aperture, but they did when they looked with both eyes, without aperture and in a well-lit environment. In the second, they made computer rendered slanted lattices with a correct viewpoint at 112 cm. The observers made their slant judgements from various distances. The apparent depth matched that of a geometric reconstruction of the scene from a viewpoint 56 cm in front of the display. Rosinski and Farber suggest that human observers estimate the geometrically equivalent viewing distance to be twice the height of the display[3] (27 cm in their setup).

According to mechanism 2, the geometrically equivalent viewpoint is reconstructed from perspective cues in the image. Geometrically, such a reconstruction can be made (Halloran, 1989; Ramsey and Sleeper, 1988 (p. 798); Sedgwick 1980; 3D builder, 1996). However, there is not much evidence for the use of this mechanism in human perception. If human perception used such a mechanism, one would not expect apparent depth to depend on the presence of a frame, as described for the first mechanism. Then again, most of the stimuli used there contained only perspective cues, and therefore results may not hold for richer pictures.

The third mechanism does not need to compensate for a geometrically inequivalent viewpoint, as it uses only depth cues that are more or less viewpoint independent. For example Hagen (1974) suggests that "presumably the high-order information specifying an individual's face consists of such relations as the length or sharpness of the forehead curve relative to the length or sharpness of the nose curve, the width of the eyes relative to the length of the nose, and so on" (p. 475). Since about 1980, there has been growing experimental evidence supporting such a mechanism, and against the other two mechanisms. Cutting (1987) displayed rotating cubes projected onto a simulated plane which was slanted relative to the visible monitor display. Rigidity judgements on slanted planes without cues to slant did not differ from judgements with unslanted planes. He concludes that the orientation of the observer relative to the projection plane does not influence apparent rigidity, and that local distortions are so small as to be unregistered by

---

[2]For pictures in pictures, this mechanism would require a double correction.

[3]Rosinski and Farber (1980) argue that "this viewing distance is (. . .) optimal in terms of providing a maximal signal-to-noise ratio for conventional TV displays. It may be that with video displays, an observer learns an assumed correct viewing point on the basis of optical picture clarity" (p. 171). However, McVey (1970) suggests that the optimal viewing distance for TV displays is 6.25 times the width of the display, which is much further away than the twice the height of the display as suggested by Rosinski and Farber.

the visual system with a 67° slanted projection plane. Halloran (1989) tested judgements of the orientation of parts of a rowing-boat in pictures seen from geometrically inequivalent viewpoints, and concludes that pictorial relations between the image contents and the border influence the judgements. Furthermore, he replicated the experiment with Rosinski and Farber (1980), but with even more extreme oblique views, and showed that a geometrically inequivalent viewpoint does have an effect at extreme angles.

Finally, the ability to compensate for a geometrically inequivalent viewpoint may depend on age (Hagen and Elliott, 1976; Jones and Hagen, 1978). Haber (1980) points out that an adult looking at a picture sees both its flatness and the depicted spatial scene, but that 3 year-olds treat pictures as if they were looking at real scenes.

Concluding, for static images it seems that distortions usually do not occur for oblique views up to about 70° from head on. Humans may compensate for geometrically inequivalent viewpoints. The distortions predicted by the three mechanisms differ for different coupling methods. If humans use the first mechanism (thus assuming an on-axis image), huge distortions can be expected with off-axis coupling, and no distortion at all with on-axis coupling. If humans used the second mechanism (thus completely compensating for oblique viewing), no difference would be expected between on- and off-axis coupling. With the third mechanism, distortions will be as predicted with geometry.

Cues to the flatness of the image

Cues to the flatness of the image may trigger a compensation mechanism for oblique viewing (as discussed above). Also, they may flatten the 3D reconstruction made by the observer (see Chapter 2). Such flatness cues can consist of binocular parallax cues, the presence of a visible frame emphasizing the image borders, the grain of the surface (e.g. the pixel size of the display), reflections of the image surface, colour flatness, and convergence and accommodation cues. Following Gibson's advice (see above, under 'Previous work') will reduce these flatness cues.

Hochberg (1986) suggested that the presence of flatness cues reduces the apparent depth in pictures. For example, accommodation is such a cue, especially at small distances (Gooding, Miller, Moore and Kim, 1991; Schlosberg, 1941). For static images, Koenderink, van Doorn and Kappers (1994) found that monocular observers perceive a larger pictorial depth than binocular observers. When viewing through a synopter, an optical device that places both eyes optically at the same viewpoint, they paradoxically found that even larger pictorial depth is perceived. Reduction of the visible area with the synopter as compared to the other viewing conditions may have caused this effect. The results of the experiment described in Chapter 6 suggested a tendency towards monocular observers working faster and making fewer mistakes than binocular observers in tracing wires through a knot on a virtual window display. In the pilot research for the present experiment (see below, 'Pilot experiments') I noticed that, with off-axis coupling and binocular instead of monocular viewing while only a monocular image is displayed, a sphere appears as an egg, resembling the distortions described by Pirenne (1970). The presence of a frame around the image may be another flatness cue, and may trigger some compensation mechanism. The experiment of Rosinski and Farber (1980), described under 'Human insensitivity to geometric inequivalence', showed a difference between judgements with and without a frame indicating the slant of the display. Eby and Braunstein (1995) found compression of the apparent layout of a spatial scene caused by a frame in front of the image.

On the other hand, for virtual window displays Arthur, Booth and Ware (1993) compared performance in a tree tracing task on a virtual window display. They found no significant differences between monocular and binocular viewing, and 71% of the subjects preferred non-stereoscopic viewing (but using both eyes) to monocular viewing. For a depth-height matching task in a static image, Adams (1972) too found no differences between monocular and binocular viewing. This is surprising, as binocular viewing gives the observer a cue about the flatness of the image, and this cue conflicts with the other cues. Possible explanations for this are (a) that monocular viewing improved the apparent depth, but another factor reduced the apparent depth (monocular viewing may be less comfortable because subjects were asked to close or cover the left eye, or alternatively some subjects may have used their non-dominant eye), and the summed effect was zero, or (b) that it depends on the task whether binocular cues are being used by the observer.

Cues from parallax shifts

There are a number of theories which seek to recover the spatial layout from motion parallax cues geometrically. These theories differ in the specific parallax information that is used. Most theories recover the spatial layout from a few views taken from the infinite number of available views. Theoretically, two perspective views from different viewpoints provide sufficient cues for a geometrical reconstruction up to a scaling factor (Ullman, 1979; Longuet-Higgins, 1981), but more views may be useful in less restricted situations (Braunstein, Hoffman, Shapiro, Andersen and Bennett, 1987).

It is not clear to what extent human perception uses parallax cues. Human depth perception from parallax seems to use other information besides a few 'snapshots'. Braunstein et al. (1987) showed that, for same-or-different judgements of a few spatial points, human performance still increases when the observer is presented more viewpoints than geometrically required. A knot tracing task (Chapter 6) showed similar evidence. Todd, Akerstrom, Reichel and Hayes (1988) found, for shifting random dot clouds, complex effects of the number of available views and timing aspects of the image on apparent rigidity. Thus, the theories seem not to fit actual human performance. Furthermore, these theories are unclear about how human perception resolves conflicts. This issue will be discussed below, under 'Coupling method'.

Most experiments with parallax shifts have been done with shifting random dot clouds. These suggest that parallax shifts are an important cue for human perception. Tittle, Todd, Perotti and Norman (1995) investigated such clouds viewed from a small viewing distance and with the shifts not coupled to the viewpoint of the observer. They found that for matching the width and the depth of a cylinder and for setting two planes perpendicular, apparent depth appears expanded as compared with geometric depth. Rogers and Graham (1983, 1985) suggest that apparent depth from motion parallax is about the same as apparent depth from stereoscopic cues. Norman and Todd (1995) showed that apparent depth from motion parallax easily overrules depth from stereoscopic cues, and stereoscopic cues are usually considered to be strong depth cues (that is, they have a major impact on a 3D reconstruction). Therefore, the viewpoint dependency of apparent depth as found for static images may be completely different for virtual window systems.

However, other results suggest that motion parallax is a weak cue, probably about as strong as accommodation cues (Gogel and Tietz, 1979), and perspective cues may take precedence over motion parallax (Wickens, 1990). Braunstein, Hoffman, Shapiro, Andersen and Bennett (1987) showed that, for same-or-different judgements of a few spatial points, human performance still increases when the observer is presented more viewpoints than geometrically required. For motion parallax, there may be a tendency for apparent depth to match apparent width. For example, Durgin, Proffitt, Olson and Reinke

(1995) found for oscillating real cones a tendency of observers to judge the height equal to the width. This is in line with similar findings for depth judgements of shifting contours of concentric circles (Caudek and Proffitt, 1993). The apparent depth may be influenced by the shape of objects, for example Ono and Steinbach (1990) found for random dot patterns that a sine shape is perceived to be of higher amplitude than a sawtooth-shape. These results may well be affected by the relatively poor stimuli used in these experiments, and therefore it may be inappropriate to generalize them to complex images as encountered in x-ray baggage inspection. The strength of parallax cues may depend on the richness of the scene.

Concluding, it is not clear to what extent motion parallax cues will influence apparent depth, and how they interact with perspective cues. Furthermore, for human perception motion parallax cues may require support from other depth cues. Therefore it is difficult to predict apparent depth from motion for real scenes presented with the DVWS.

Other factors possibly causing distortion

Several other factors may cause distortion. These factors have to be considered in order to make appropriate stimuli for our experiment, as the experiment described in this chapter will compare performance with Euclidean measures.

The scene layout may affect distortion. Kjelldahl and Prime (1995) showed for relative depth estimation of computer rendered shaded dishes that people make larger errors in judging distance if the separation between the dishes is vertical rather than horizontal. The texture used in the scene also influences depth perception, especially in the case of local details (Cumming, Johnston and Parker, 1993). Texture density has an effect on apparent depth (Börjesson and Lind, 1996). The colour also influences depth, as Claessen (1996) showed that a real blue object appears to protrude further than a yellow one. Kappers, van Doorn and Koenderink (1994) showed that the shading of real objects can also cause distortion.

If the camera image is presented on a television monitor, both the thickness of the glass and the curvature of the monitor will deform the image. Glass in front of a picture, for example in front of the television monitor, will magnify the picture. In television monitors, the picture is not flat but curved. This will effectively reduce the picture. Deering (1992) indicates that both reduction and magnification can displace the geometrically equivalent viewpoint up to 2 cm, but that the combined effects will not cancel out. He derives a formula giving a correction factor, given some point of interest on the display.

The delay and the refresh rate of the system can cause distortion. If the displacement of objects gets larger than 0.15° per update (Padmos and Milders, 1992), the image movement will appear shaking. Furthermore, for a moving observer a delayed image does not fit his actual viewpoint. This causes striking deformations of objects when the observer changes the speed of his movement. In the pilot tests I noticed with off-axis coupling that shear became more striking with increasing observer speed and with decreasing refresh rate. This shear distortion was also noticed in a prototype setup of a virtual window system with three displays instead of one (Djajadiningrat, Smets and Overbeeke, 1997), and seems typical for virtual window systems. The delay effect is discussed in more detail in Chapter 9.

Finally, the task of the observer, and the relation between the observer and his environment, determine whether distortion is a problem or not. As mentioned earlier, paintings and objects on a television screen do not appear distorted when the viewer is dislocated. But in these cases there are few consequences for the spectator if he misperceives the depth in the image. But if the observer has to manipulate objects in such a distorted environment, he needs some time to adapt to the distortions. For example, one needs to get used to a new pair of spectacles. More extremely, Kohler (1962) reported that observers wearing distorting goggles will correct many distortions such as colour fringes, line curvature and reversed left and right. In contrast with the immediate correction that takes place when looking at pictures, the time required to get used to these distortions takes from days to some weeks.

Many other factors may exist that influence the depth in virtual window displays. For example, Drascic (1991) found slow learning effects for parallax displays. The factors discussed here will be taken into account in the setup of the following experiment, but one can only hope not to miss an essential cue.

Coupling method

As far as I know, no research has been done on the effect of on- and off-axis coupling on apparent layout in virtual window displays. Both for perspective cues and for parallax cues, a single geometric spatial reconstruction given multiple views from a virtual window system seems possible only for off-axis coupling. For on-axis coupling, several inconsistencies may occur, such as vertical movement of parts of objects while the observer moves only horizontally, geometrically suggesting viewpoint dependent (nonrigid) distortions. Therefore, distortions are expected with on-axis coupling.

With off-axis coupling, the viewpoint is always geometrically equivalent to the displayed view, provided that the viewpoint is measured accurately. In the presence of a viewpoint measurement error, distortions can occur. If human observers use perspective cues in a geometric way, the distortions will cause shear and depth compression as discussed under ‘Geometric inequivalence’. With on-axis coupling, the geometrically equivalent viewing position is always in front of the middle of the display, at the fixed camera distance. However, with on-axis coupling the observer needs to move away from this position in order to acquire another view. Therefore, geometry suggests that distortion will occur for most viewpoints, and this distortion is expected to increase with increasing distance from the geometrically equivalent viewpoint. However, if the observer compensates for his geometrically inequivalent viewpoint, the effects for on- and off-axis coupling may be different, as discussed under ‘Human insensitivity to geometric inequivalence’.

For on-axis coupling, motion parallax cues again do not fulfil geometric assumptions given a rigid spatial scene. For example, a viewpoint movement along a line through the centre of the display does not cause a different image to be displayed, as the camera stays at a fixed distance from the fixation point. This is a cue to the flatness of the image. Furthermore, with on-axis coupling it is possible that the displayed objects move downwards while the observer moves only sidewards, and this is in conflict with geometric models of motion parallax. It is not clear how human perception deals with these inconsistencies, and to what extent theoretical models match human performance. Therefore it is difficult to predict apparent depth in on-axis virtual window displays.

Off-axis coupling may be more effective in giving a depth impression than on-axis coupling. Hayashibe (1991) showed that the apparent depth from shifting random-dot patterns is mainly caused by the relative speed between two shifting areas, and that the coupling to head movement only reduces the number of depth reversals. Similarly, Caudek and Proffitt (1993) argue that the depth perceived in a cone, protruding out of the display and oscillating around the vertical axis, is caused only by the relative shifts of object contours and not by foreshortening. This suggests that the rotational part of the on-axis coupling might not contribute to the apparent depth, but only to apparent rotational motion. As Ono and Steinbach (1990) found for random dot patterns that perceived object motion increases as perceived depth decreases, off-axis coupling may be more effective in suggesting depth.

Concluding, in virtual window displays with off-axis coupling, perspective cues and motion parallax cues agree geometrically, and the parallax shifts may be more effective than with on-axis coupling. With on-axis coupling, perspective cues and motion parallax cues are geometrically inconsistent with each other, and in themselves in most cases. Geometry suggests that even distortions that vary with the movement of the viewpoint (nonrigid distortions) may be expected for on-axis coupling. It is not clear how human perception deals with these inconsistencies, and therefore it is very difficult to prediction distortions in on-axis virtual window displays.

Pilot experiments

I did thirteen pilot tests, to gain an impression of the effects of the task and scene layout on the apparent layout and task performance. For example, in one pilot experiment observers adjusted one of three pillars to form an equilateral triangle (Figure 7.5a), and in another pilot experiment observers rotated a block-textured plane perpendicularly to another similarly textured plane (Figure 7.5b). For the last task, the effect of scene complexity was also investigated (Figure 7.5c). In another pilot test similar to that of Figure 7.5b, observers had to adjust the width of the textured plane to the width of the other plane.

Most effects could be explained from compression of the apparent depth with on-axis coupling. For example, the angles were adjusted smaller than 90˚ with on-axis coupling, and planes protruding out of the display were adjusted wider than planes parallel to the display. Paradoxically, shear distortion was noticeable only with *off*-axis coupling in spheres, triangles, and also of the smoother landscapes of the present experiment, especially when the objects protruded out of the display and were viewed with both eyes. With on-axis coupling no distortions were noted, neither changing with the viewpoint nor static.

Summarizing the long introduction, off-axis coupling seems necessary to avoid distortion, if human perception uses depth cues in a geometric way. It seems that humans notice only certain forms of distortion, and that these distortions do not occur with on-axis coupling but may occur with off-axis coupling. On the other hand, the behaviour as found from task performance can be explained with distortions as predicted by geometry. I am mainly interested in task performance, and less in noticeable distortions.

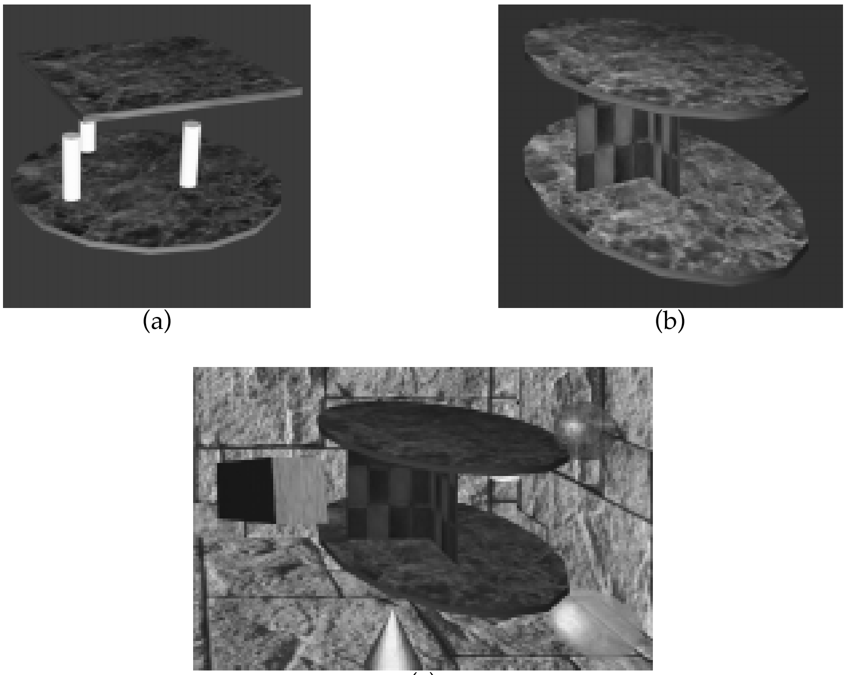

(a) (b) (c)

Figure 7.5. Views from three pilot experiments. See text. Distortion could be noticed with off-axis couplings (b, c) but not with on-axis couplings (a). On the other hand, the results could be explained from a compressed apparent depth with on-axis coupling.

Experiment

As discussed, this experiment is a first exploration of the effects of geometric inequivalence and cues to the flatness of the display on the apparent layout in virtual window displays. Although previous work (see above) is not conclusive about the mechanisms involved in depth perception, my working hypothesis was the use of a mechanism compensating for the effects of a geometric inequivalent viewpoint (mechanisms one and two in the section 'Human insensitivity to geometric inequivalence'). By diminishing other cues to the flatness of the display, I hoped to be able to manipulate the occurrence of compensation by having participants look with only one or with both eyes. With on-axis coupling, the geometric equivalent viewpoint always lies perpendicular to the display, but not with off-axis coupling. If the compensation mechanism supposes the viewpoint to lie perpendicular to the display, with on-axis coupling apparent depth would be unaffected by a geometric inequivalence. Therefore, we can distinguish from the experimental results what kind of compensation mechanism is being used.

Method

Variables, subjects, design

The independent variables (Table 7.1) were the coupling method *C* (on- or off-axis), the viewing condition *V* (monocular or binocular), a viewpoint measurement error parallel to the display plane *DXY* (0 or +10 cm in a random direction parallel to the display plane) and a viewpoint measurement error perpendicular to the screen plane *DZ* (0 or 10 cm out of the display plane). Due to a mistake, *DXY* was always to the right, relative to the observer (see Figure 7.13) instead of in all directions.

The dependent variables were the height ratio *H* (subject's height setting / measured height of foam model) and the total response time for each landscape *T*. All head positions were recorded during the experiment to check whether the average viewing distance matched the fixed camera distance from the fixation point for on-axis coupling (see 'Coupling method').

At random, six subjects were assigned to the monocular viewing condition and six to the binocular viewing condition. Each of the subjects judged 16 landscapes each of which contained 5 bumps (Figure 7.6a, b). The 2(*P*) x 2(*DXY*)x 2(*DZ*) x 2(repetitions)=16 conditions were randomized over these 16 landscapes, and the landscapes were presented in random order. For each subject, the 10 (2 landscapes x 5 bumps) height setting/measured height ratios he made in each condition were averaged, to get his height ratio *H*=virtual height/measured height. This gives a mixed analysis of variance (SPF-p.qru) design (Kirk, 1968).

The subjects were 12 naive volunteer students (6 male, 6 female) with normal or corrected-to normal vision. All subjects had stereoscopic vision, as they were able to recognize a figure hidden in a random-dot stereogram (Appendix A). They were paid NLG 7.50 (a loaf of bread costs about NLG 2).

Table 7.1. Independent variables and their levels.

Variable	Description	Levels
C	Coupling method	on-axis, off-axis
V	Viewing condition	monocular, binocular
DXY	viewpoint measurement error parallel to the display plane	0 cm, 10 cm in a random direction parallel to the display plane.
DZ	viewpoint measurement error perpendicular to the display plane	0 cm, 10 cm out of the display plane

Stimuli

The stimuli consisted of pairs of real (foam) and virtual (computer simulated) landscapes. Figure 7.6 shows a picture of both. The virtual stimuli were 228x228 mm and consisted of a gently sloping landscape (+5 .. -5 mm) with 5 bumps of height 10 .. 25 mm. They were rendered with 1280x1024 24-bit colour pixels by a Silicon Graphics Reality Engine. For on-axis coupling, the camera distance to the fixation point was 450 mm. This distance was chosen to match the expected average viewing distance of the participants. No spotlights were added in the virtual scene, so there was no shading information in the virtual stimuli. This was necessary to prevent the participants from adjusting the shadows instead of the bumps. The texture on the virtual bumps was made by scanning a real texture on a flat piece of foam. The real stimuli were milled in polyurethane foam from the computer data with an accuracy of 0.25 mm, and covered with a granite-effect structure spray (Plasti-kote Fleckstone 35: serpentine marble). They were lighted with a 20W halogen spot set to 26W to get a daylight-like colour. The spot was 1.8m higher and slightly from the back of the landscape, to get a light shading on the bumps at the side of the observer. The texture of the virtual stimuli was carefully matched to the real stimuli.

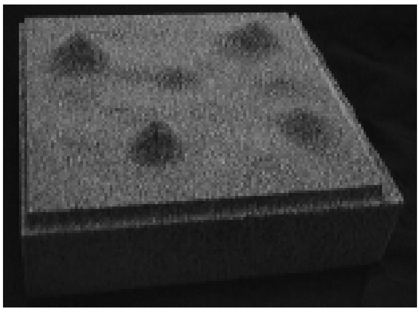


Figure 7.6a (**see colour figure on right cover flap**). Real landscapes were milled in foam and had textured paint sprayed on the surface.

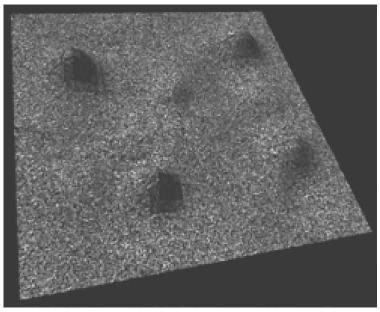


Figure 7.6b (**see colour figure on right cover flap**). Rendered virtual landscape. Shadows were added for this picture, to make the bumps visible.

Apparatus

Figure 7.7 shows an overview of the experimental setup. In the middle is the monitor display for the virtual landscape. On its left is the real landscape, on its right the button box and the mouse for adjusting the bumps. Behind the monitor is the eye position tracker. A trade-off was made between the image complexity and the update rate.

Figure 7.7. An overview of the experimental setup. The monitor with the virtual landscape is in front of the subject. The real landscape is on the left, the head tracker behind and the mouse and button box to the right of the monitor. A cardboard cover attached to the display served as a reduction screen.

I chose to use approximately 1800 triangles and an update rate of 29 Hz. The triangle density on the bumps was 2 times as high as the landscape between the bumps. While the subject was adjusting a bump, the update rate dropped to 14 Hz. The 21 inch screen (Taxan Ergovision 2100LR; 0.3 mm dot pitch; anti-glare and anti-reflection coating; refresh rate 100 Hz) was placed parallel to the floor. This setup was designed to elicit extremely oblique viewpoints from the subjects (see Chapter 2). The short side of the monitor was next to the subject, to prevent the bumps going off the display when the observer was looking at extreme viewpoints. In front of the screen, a reduction screen of 25 x 32 cm hid the borders of the screen to reduce conflicting visual cues between the views on the screen and the environment. The visual cues from the environment were further reduced by a dark background and dimmed lighting. The screen was warmed up at least 30 minutes in advance of each trial to prevent colour changes during the experiment. The 843x843 pixels of the virtual landscape matched the size of the real landscape. A box with two buttons, a green one labelled 'volgende bult' (next bump) and an orange one labelled 'landschap OK' (landscape finished), allowed the subject to select the next bump or to finish the landscape. An infrared tracker (Dynasight from Origin Instruments) tracked a small reflector on a spectacle frame that the subjects wore. Of monocular observers, the position of the eye they used was tracked, of binocular observers the average of the two eyes was tracked. As the reflector was not exactly in the eye position, the position of the eye contained an error of at most 3 cm, but for normal viewpoints the error was estimated to be about 0.5 cm.

Procedure

It was explained to the participants that they would be presented a real and a virtual landscape, and that they had to adjust the height of the 5 peaks on the virtual landscape to match the heights of the corresponding peaks on the real landscape. They were told to increase the height of the selected bump by moving the mouse away, and to decrease the height by pulling the mouse towards them. With the button labelled 'volgende bult' (next bump) they could select the next bump. If they pressed this button again at bump 5, bump 1 was selected. This way, they could jump through the bumps to select one for adjusting the height. If they were confident about their settings, they had to press the button labelled 'landschap OK' (landscape OK). Then the virtual landscape disappeared, the real landscape was replaced by a new landscape and the new landscape appeared on the screen.

About the head tracker they were told that it followed a small reflector on a spectacle frame. Subjects in the binocular viewing condition received a spectacle frame with a small reflector. Subjects in the monocular condition were told that they have to look with one eye, and their dominant eye was determined, as described in Chapter 6. These subjects received a spectacle frame with both a reflector and an eye patch covering their non-dominant eye. It was explained to the subjects that an LED on the tracker would turn from green to red if they moved too fast or out of range, and that the image on the screen would shake in that case because it could not be adjusted in the correct way.

Participants were told not to be obsessed with a perfect match, but to match the height in a few seconds. They were asked to try to set the correct height of each bump the first time they selected it and to take about 10 seconds for each bump, as it would take too much time to re-adjust all the bumps later. There was no strict time limit, but they were warned during the training (and in two cases during the experiment) if they took more than 20 seconds for each bump. Subjects were asked not to touch the screen or the real bumps.

The light was then turned off, except for the spotlight for the real landscape. Scattered light allowed the subjects to see the button box and the mouse. For training, they were asked to adjust a landscape twice. Some subjects noticed that they could look around the virtual landscape during the first presentation of the landscape, but not during the second (the virtual landscape was projected on-axis during the first presentation, and off-axis during the second presentation). If participants asked about this, they were told that one of the aims of the experiment was to test this difference in ability to look around the scene. They were not given feedback about the correctness of their settings, as this might have frustrated the subjects with on-axis coupling. After the training subjects were allowed to ask questions, and they were asked to work faster if necessary.

Before the experiment, participants were informed that they had to adjust 16 landscapes (none of the subjects noticed that there were only 8 landscapes, each presented twice but rotated 90°). In total, each trial lasted about 40 minutes.

Results

Initially it was planned to analyse the difference between the matched virtual height and the measured height of the bump. However, this would give negative values if the subject adjusted lower than the measured height. Furthermore, the variance of this value was found to depend linearly on the height of the real bump. Therefore, the height ratio *H* (virtual bump height as matched by the subject / measured bump height) was analysed. The ratio *H* solves both problems, and moreover it is a dimensionless ratio with an optical interpretation, allowing direct comparison of the depth scaling due to geometric inequivalence d'/d (see 'Geometric inequivalence'). For example, a setting giving a ratio *H*=2 indicates that the virtual bump was adjusted twice as high as the real bump to appear to be of equal height, or to put it in another way the apparent depth in the virtual scene was half as great as the apparent depth in the real scene.

The setup was successful in eliciting extreme viewpoints. Especially for inspecting the real landscape, a lot of subjects looked at a at nearly-zero height over the landscape to see the bumps in side view. They usually tried the same with the virtual landscape, although they had to take a slightly higher viewpoint to avoid clipping of the bumps by the screen border.

A repeated-measures analysis of variance (Kirk, 1968) on the height ratios (Table 7.2) shows significant main effects of all variables. The mean height ratio *H* for on-axis coupling is 1.79, for off-axis coupling 1.31. A post-hoc *t*-test (Hays, 1981) shows that the average *H* is larger than 1 with on-axis coupling ($p<0.01$). Another *t*-test showed that for off-axis coupling the average ratio *H* does not differ significantly from 1.

Table 7.2. Significant effects for the analysis of variance performed on the height ratio (height setting/measured height) *H*.

Interaction	*F*	*p*
V	$F(1,10) = 5.35$	<0.05
C	$F(1,10) = 119.21$	<0.001
DXY	$F(1,10) = 19.88$	<0.01
DZ	$F(1,10) = 58.81$	<0.001
C x *DZ*	$F(1,10) = 5.47$	<0.05
V x *DXY* x *DZ*	$F(1,10)=9.88$	<0.01

With monocular viewing, the mean height ratio *H* is 1.71, with binocular viewing 1.39. The hypothesis that binocular viewing reduces performance only for off-axis coupling is not confirmed: the *V* x *C* interaction is not significant, and binocular viewing improves the performance with both coupling systems.

A viewpoint measurement error perpendicular to the screen *DZ* gives an increase from 1.405 to 1.694 of the mean height ratio as compared to the condition without measurement error. An error parallel to the display plane *DXY decreases* the mean ratio from 1.601 to 1.497. Figure 7.8 shows the effects of the independent variables on the mean height ratio *H*. A post-hoc analysis of variance was done to check the absence of interactions with off-axis coupling as suggested by Figure 7.8. The analysis showed that this is indeed the case, as for off-axis coupling only the viewpoint measurement errors have a significant effect on the height ratio. For a viewpoint measurement error parallel to the display *DXY* $F(1,10)=9.08$, $p<0.05$ and for a viewpoint measurement error perpendicular to the display *DZ* $F(1,10)=63.1$, $p<0.001$.

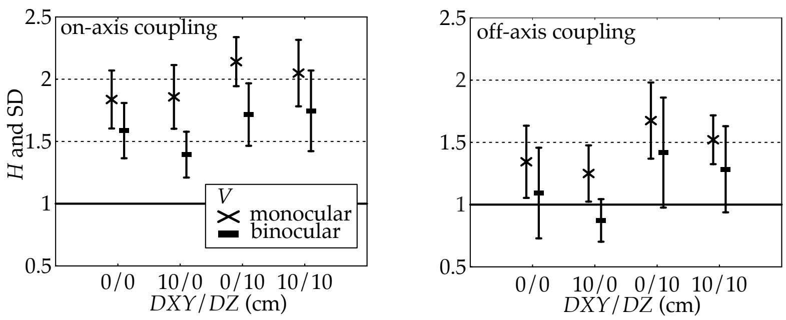
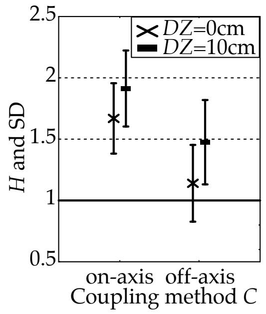


Figure 7.8. The effects of the coupling method *P*, the viewing condition *V*, a viewpoint measurement error parallel to the display plane *DXY* and perpendicular to the display plane *DZ* on the mean height ratio *H*.

The analysis of variance of the mean error on the height ratios *H* (Table 7.1) also showed an interaction between *DZ* and the coupling method *C*. Figure 7.9 shows this interaction: with on-axis coupling, with *DZ* = 0 cm *H*=1.67, and with *DZ* = 10 cm *H*=1.91. The increase of *H* with off-axis coupling is larger: from 1.14 to 1.48. Thus, off-axis coupling is more sensitive to viewpoint measurement error perpendicular to the display. For a viewpoint measurement error parallel to the display *DXY*, this sensitivity difference does not exist, as there is no significant interaction between the coupling method *P* and *DXY*. Furthermore, because *H* gets closer to 1 in the presence of a viewpoint measurement error parallel to the display plane *DXY* as compared to absence of such an error, it seems that the apparent layout gets closer to the Euclidean layout with *DXY* than without *DXY*.

The three-way interaction *V* x *DXY* x *DZ* is shown in Figure 7.10. It shows that the difference between monocular and binocular viewing is much larger at a geometrically inequivalent viewpoint than at the geometrically equivalent viewpoint.

A Pearson product-moment correlation (Norusis, 1993) did not show a correlation between the distance of the bump to the centre of the screen and the height ratio *H*.



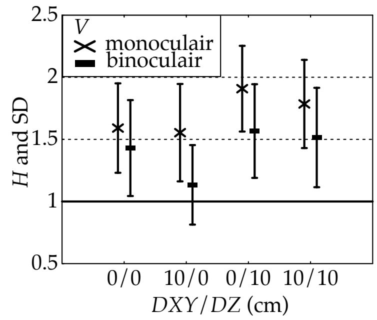
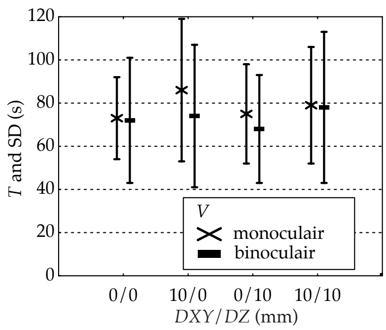


Figure 7.9. The effects on the mean height ratio *H* of the coupling method *C* and a viewpoint measurement error perpendicular to display plane *DZ* .

Figure 7.10. The effects on the mean height ratio *H* of the viewing condition *V*, a viewpoint measurement error in the display plane *DXY* and perpendicular to the display plane *DZ*.

A repeated-measures analysis of variance of the response time *T* shows a main effect for *DXY* (p<0.05) and for the three-way interaction *V* x *DXY* x *DZ* (p<0.05). Figure 7.11 shows the effect of the variables on the mean response time *T*. The average response time is 75.5 s (15.5 s for each bump). In contrast with the error on the height ratio *H*, the response time for the *DXY* = 10 cm condition is larger than in the *DXY* = 0 cm condition. This indicates that some difficulties with shear may be present. As with the effect of the same interaction on the height ratio *H*, the difference between monocular and binocular observers is much larger when a viewpoint measurement error *DXY* or *DZ* is present than when it is not.

The difference between the observers was checked graphically. Figure 7.12 shows that the average height ratios *H* are different for different observers, and that one observer has nearly the same response with on- and off-axis coupling.



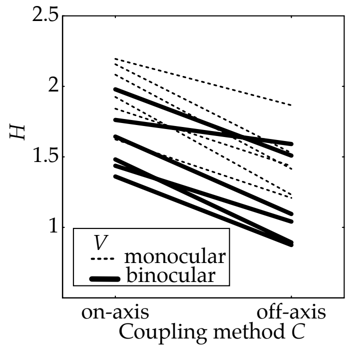


Figure 7.11. The effects of the viewing condition *V*, a viewpoint measurement error in the display plane *DXY* and perpendicular to the display plane *DZ* on the mean response time *T*.

Figure 7.12. The effects for each observer of the viewing condition *V* and the coupling method *C* on his mean height ratio *H*.

An analysis was done to test for an effect on the height ratio *H* of the direction of the viewpoint measurement error parallel to the display plane *DXY*. Figure 7.13 shows the errors on the height ratio *H* for each tested direction of the error. For an observer sitting normally in front of the display (Figure 7.7), the positive x direction is backwards and the positive y direction is to the right. Erroneously, the direction of each error was chosen randomly between 0 and π instead of between 0 and 2π.

The effects of the x- and y- components of the direction on the height ratios were tested separately. A Pearson product-moment correlation (Norusis, 1993) was done for both coupling methods. All correlations were found to be significant. For on-axis coupling the ratio decreases with increasing x (r=-0.2, p<0.01) and with increasing y (r=-0.33, p<0.001). For off-axis coupling the ratio decreases with increasing x (r=-0.34, p<0.001) and increases with increasing y (r=0.2, p<0.01).

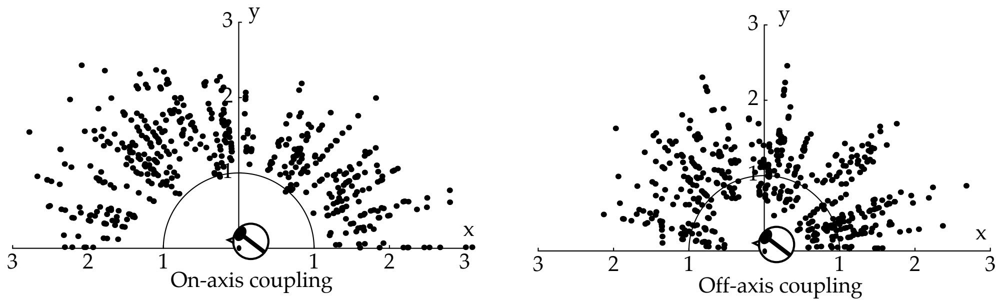


Figure 7.13. The effects on the mean height ratio H of the direction of a viewpoint measurement error parallel to the plane of the display DXY and the coupling method C. The angular position of a dot indicates the direction of the viewpoint measurement error, and H is represented by its distance from the origin. For on-axis coupling, H is outside 1 and elliptical. For off-axis coupling, H is around 1 and circular.

To check the fixed distance of the camera to the fixation point for on-axis coupling, the head positions of the subjects were averaged to find the average viewpoint. This was found to be 394 mm to the bottom, 44 mm to the right and 274 mm above the display, so the average viewing distance to the fixation point was 482 mm. Therefore, the fixed camera viewing distance for on-axis coupling (450 mm) had been well chosen.

Summarizing, with off-axis coupling the apparent height of the virtual bumps is equal to the apparent height of the real bumps. With on-axis coupling the virtual bumps have to be 1.79 times as high as the real bumps in order to give an equal apparent height, indicating that the apparent height is lower than the virtual height with on-axis coupling. As compared with on-axis coupling, with off-axis coupling the matching is more sensitive to viewpoint measurement errors perpendicular to the plane of the display. Similarly, monocular viewing reduces apparent depth as compared with binocular viewing in the presence of a viewpoint measurement error. For a viewpoint measurement error parallel to the display, the direction of the error also influences the apparent depth.

Discussion

The finding that geometric inequivalence affects apparent depth with on-axis coupling is in conflict both with a human compensation mechanism that works by assuming the correct viewpoint perpendicular to the display and with a compensation mechanism that extracts the correct viewpoint from pictorial cues. Instead, the results suggest that no compensation mechanism is used at all.

Many results can be explained by geometry. The height ratios *H* can be interpreted as how much the apparent depth from the scene on the display is compressed as compared to the apparent depth of the real scene. For the coupling method, an important result is that on-axis coupling causes an apparent layout corresponding to a depth-compressed Euclidean layout, while with off-axis coupling the apparent layout of a displayed scene corresponds to the Euclidean layout, provided that the viewpoint is measured accurately. The average depth compression for on-axis coupling was 1.79. This compression can be explained geometrically, as follows. The average (perpendicular) distance of the observers to the plane of the display was 274 mm. As the camera for on-axis coupling was at a constant distance of 450 mm, the scaling of the depth in a geometric reconstruction is 274/450 = 0.61 (see 'Geometric inequivalence'). To compensate for this scaled depth, the bump would have to be 1/.61=1.64 times the normal height. This is slightly lower than the 1.79 from the experiment, so perhaps the observers move slightly closer to the display while comparing the precise heights of the bumps, and move away again during the adjustment. If a compensation mechanism assumed the correct viewing distance to be twice the height of the display (see 'Human insensitivity to geometric inequivalence'), 50 cm would be taken as the correct viewing distance. 50 cm is close to the actual camera distance of 45 cm, and therefore no distortion would be expected if human perception used such a mechanism. But this is not in agreement with our results. For off-axis projection, a viewpoint measurement error of 10 cm increases the average viewing distance from 27.4 cm to 37.4 cm. Geometry predicts a height ratio of 37.4/27.4 = 1.36, which is close to the actual results. For off-axis projection, geometry predicts independent effects of a viewpoint measurement parallel to the display and such an error perpendicular to the display, which also was shown to be the case for the experimental results. Concluding, the results, both for on- and for off-axis coupling, are in agreement with the compression predicted by uncompensated geometry.

A last argument against a compensation mechanism assuming a correct viewing position perpendicular to the display comes from the results for monocular and binocular viewing. It was hypothesized that binocular viewing provides a flatness cue that triggers a compensation which, in its turn, worsens performance with off-axis coupling but not with on-axis coupling. Furthermore, binocular viewing was expected to introduce a conflict between parallax shifts and stereoscopic cues. Experimentally, though, binocular viewing was found to improve performance as compared to monocular viewing, especially if the observer is dislocated. Why do binocular observers perform better? Consider the explanations mentioned in 'Cues to the flatness of the image'. In contrast with the experiment of Arthur, Booth and Ware (1993), observers in the monocular condition in the present experiment always looked with their dominant eye. Reduced comfort or conflicting information from the two eyes in the monocular condition (see 'Cues to the flatness of the image') is unlikely, given the results of the experiment described in Chapter 6. But there are other differences between the present experiment and the experiments that indicate an advantage for monocular observers. In both the experiment of Koenderink, van Doorn and Kappers (1994) and that of Chapter 6 the environment was clearly visible, although a reduction screen in front of the monitor was used in the experiment of Chapter 6. Koenderink et al. (1994) tested the depth perceived in photographs by stationary observers at the geometrically equivalent viewpoint. For the knot tracing task of Chapter 6, the observers could move, and the observer's angular position was scaled by a factor 4. Furthermore, extreme viewpoints were provoked by the present experiment, but not in the experiments of Koenderink et al. (1994) and that of Chapter 6. All these differences can affect the apparent depth, as was discussed in the 'Previous work' section.

Finally, one result suggests a compensation mechanism that assumes a correct viewing position perpendicular to the display: on-axis coupling was found to be less sensitive to a viewpoint measurement error perpendicular to the display than off-axis coupling. However, the sensitivity effect may also be caused by the indirect effect of a viewpoint measurement error perpendicular to the display for on-axis coupling: the camera is always at 450 mm from the fixation point and only its viewing direction is affected by a viewpoint measurement error.

There are two unexpected results. The first is that a viewpoint measurement error parallel to the display decreases the error on the height settings. Geometry indicates that this causes a shear distortion, which should not affect the height of the bumps. One explanation is that subjects may match the axis of the bump instead of the perpendicular height of the bumps (Figure 7.14), and that therefore a bump appears to be higher if a shear distortion away from the observer or sideways is present. At an average viewing distance of 27.4 cm, a viewpoint measurement error of 10 cm would enlarge the diagonal of the bump 1.065 times. This fits the actual enlargement of 1.601/1.497=1.069 amazingly well.

The last unexpected result is that for on-axis coupling the position of the bump relative to the fixation point does not influence the error on the height settings. Apparently, the movements of bumps far from the fixation point do not disturb the subjects.

Figure 7.14. Instead of the perpendicular height of the bumps (left), subjects may have used the axis of the bumps (right) in making their judgements.

Most subjects seemed not even to perceive movement of bumps, they just noticed that they were able to look further around the scene. Also, no effect of the coupling mechanism on response time was found. These results suggest that response time depends on distortions of which the observer is conscious, while observer performance is affected by other distortions of which the observer is less conscious.

Conclusions

Most results could be explained from distortions as predicted by geometry. The literature suggested that cues to the flatness of the display might trigger a perceptual mechanism compensating for the resulting distortions. But the major part of the present results could be explained by the absence of such a compensation. This result is in agreement with more recent literature. Absence of compensation geometrically implies that off-axis coupling gives an apparent depth matching Euclidean depth, and that on-axis coupling gives a compressed apparent depth if the observer is closer to the display than geometrically equivalent given the fixed camera distance to the fixation point. It can also explain why on-axis coupling is less sensitive than off-axis coupling to a viewpoint measurement error perpendicular to the display. The fact that with on-axis coupling observer performance is independent of the distance of the bumps from the fixation point (the centre of the display) is unexpected.

There are several factors to consider when choosing the appropriate coupling method. When the apparent depth of a displayed scene must match some Euclidean measures, for example for enhanced reality setups where a displayed world is projected over the real world, off-axis coupling is required[4]. On the other hand, on-axis coupling also has advantages over off-axis coupling. The technical implementation of an on-axis coupling is cheaper than that of an off-axis coupling, especially when a real camera is used. As the experiment described in Chapters 5 and 6 indicated, head movements can be scaled with on-axis coupling, allowing the observer to look further around the object. In contrast with off-axis coupling, with on-axis coupling other observers looking along with the moving observer do not see large shear distortion but an 'ordinary closed-circuit TV picture'. Furthermore, it may be possible to correct in a simple way for the reduced perceived depth, because the depth compression seems quite constant over a large range of viewpoints. Finally, on-axis coupling was found to be less sensitive to errors in the distance of the eye to the display. For many tasks, these advantages may be more important than geometric equivalence. For a given application and task, a trade-off has to be made on these points. It is possible to combine on- and off-axis coupling, by rotating the scene depending on the angular position of the observer and using an off-axis coupling, but the perceptual aspects of such a combination remain to be examined.

Concluding, we found that subjects' behaviour was very well predicted by the geometric model: people did not correct for a geometrically inequivalent view. Human perception notices some distortions, while other, unnoticed, distortions disturb task performance. The results support the conclusions of Halloran (1989), and are in perfect accord with our ideas about the task dependency of the required views: "when partial cues conflict, the choice among them will depend on the requirements of the perceptual task at hand. To account for performance, it seems necessary and sufficient to know two things: the partial geometries being projected, and the observer's perceptual decision strategy" (p. 478). If I had time for more research, I would concentrate on this 'observer's perceptual decision strategy' and its relation to his task.

---

[4]Another argument for choosing off-axis coupling might be that on-axis coupling gives an undesired vertical disparity when it is used to present separate pictures stereoscopically to the observer's two eyes (Castle, 1995). However, vertical disparity is no real problem as it also occurs with natural viewing. Vertical disparity may be even used to recover depth (Bishop, 1996), but it may be not used as such by humans (Cumming, Johnston and Parker, 1991) or only with large displays (Rogers and Bradshaw, 1993).