Perception, 2004, volume 33, pages 513 – 530

DOI:10.1068/p3454

Pointing out of the picture

Jan J Koenderink, Andrea J van Doorn, Astrid M L Kappers
Buys Ballot Laboratory, University of Utrecht, Princetonplein 5, PO Box 80 000, NL 3508 TA, Utrecht, The Netherlands; e-mail: j.j.koenderink@phys.uu.nl

James T Todd
Department of Psychology, Ohio State University, 1885 Neil Avenue, Columbus, OH 43210, USA
Received 9 September 2002, in revised form 24 November 2003

Abstract. The eyes of portrayed people are often noticed to 'follow you' when you move with respect to a flat painting or photograph. We investigated this well-known effect through extensive measurements of pictorial relief and apparent orientation of the picture surface for a number of viewing conditions, including frontal and oblique views. We conclude that cases of both oblique and frontal viewing are very similar in that perception simply follows what is indicated by the proximal stimulus, even though this may imply that the (perceived) physical and pictorial spaces segregate. The effect of foreshortening then causes an apparent narrowing of pictorial objects. We find no evidence for any 'correction' mechanisms that might be specifically active in oblique viewing conditions.

1 Introduction

It is a common observation that the eyes of portrayed people follow you around the room from their position within the frame of a flat painting or photograph. The effect has often been spontaneously remarked upon by naïve observers (people without a professional interest in visual perception) and is often described as striking or even eerie. Some paintings have been ascribed almost magical powers because of it. The effect has been known for centuries (eg Leonardo 1585, but it can be traced to much earlier—Arabic—sources). It is well described by the late art historian Sir Ernst Gombrich (1960) who uses the striking example of a well-known WWI recruitment poster (of 1915, by Alfred Leete, depicting Lord Kitchener looking sternly at the observer and pointing out of the picture 'straight at the observer', with the accompanying text "Your Country Needs YOU") to illustrate the effect. The poster shows the effect strikingly (even in print) and one easily verifies that Lord Kitchener addresses one from whatever angle one happens to look at the picture. It has been imitated widely (eg in the well known 1917 recruiting poster "I want YOU for the U.S. Army"). It is easy to imagine that the artist must have been especially talented in order to be able to achieve such a feat. Yet the fact is that no special gift or training is involved, amateurish photographic snapshots show exactly the same effect. The only 'trick' is to have the subject of the picture to look (and/or point) directly to the lens.

There has been much discussion on the question how it is possible that the *depicted person* is able to 'adjust' fixation or viewing direction as the observer looks at the picture from different angles. In a slightly different framing of the question, one asks how the *observer* is able to change the orientation of the depicted person in pictorial space such as to 'rotate along' with changes in viewing direction. The latter question makes more sense than the former, since the depicted person doesn't exist (except in the mind of the observer) and thus can hardly be expected to perform *any* action (no doubt the origin of the 'magical powers' ascribed to certain portraits derives from this confusion).

We suggest that *both* questions are spurious. The portrait consists of a planar surface covered with pigments in a certain simultaneous order. Changing the viewing direction has no effect on this order. If the portrait shows the front of the face, then no change of viewing direction is going to reveal the back of the head, simply because the back of the head was never depicted in the first place. The observer will invariably see an *en face* rendering because that is all there is to see. The observer sees of the depicted object what the observer would see of the actual object if that object were to rotate along with the observer (as in the setting of certain movie scenes where two actors move about facing each other, connected by a rope tied about their waists; see Arijon 1976). This might go some way in the direction of an explanation of the experience that the depicted person rotates along with the observer. The difference is that, in the case of the portrait, the retinal image suffers an additional deformation (depending upon the viewing position) due to perspective foreshortening. After all, the depiction is flat and does not physically rotate (not in the extrinsic frame of the observer, that is; it evidently counterrotates in the observer-centered frame). This, finally, suggests a valid research question: Do these additional deformations affect the percept, or do human visual observers actively (but perhaps subconsciously) 'correct' for them?

The final research question has remained largely unanswered. The general topic of visual spatial perceptions in pictorial perception is discussed at length in Gibson (1954, 1960, 1971), Gombrich (1960), Hochberg (1962), Pirenne (1970), Hagen (1986), and Kubovy (1986). Substantive empirical material is available from Cutting (1986, 1987) (on deformations seen in cinema theaters from the front row, side aisle) and Goldstein (1979, 1987, 1988). There exist apparently a number of categorically different approaches to what is being perceived as 'the problem' implied by the effect that pictorial objects keep pointing at the observer. One (perhaps dominant) view is that observers somehow actively (though perhaps unconsciously) 'correct' or 'compensate' for the perspective distortions of the retinal image due to oblique viewing. This typically involves a simultaneous awareness of the pictorial cues *and* the cues that reveal the flatness (and spatial attitude) of the picture surface [as discussed by almost all authors—see Topper (2000)]. Authors then seek to show that the picture surface is or isn't perceived, that observers actually compensate (typically in judgments of directions, angles, or spatial attitudes), and so forth. Theoretical ideas typically involve the geometry of monocular perspective [almost singularly following La Gournerie's (1859) analysis]. Good examples are Perkins (1973), Hagen (1976), Rosinski et al (1980), Wallach and Marshall (1986), Ellis et al (1987), or Cutting (1986). A quite different (and also quite common) view is that observers simply don't care, that is to say, disregard (but not in any *active* sense) the distortions of the retinal image, since the monocular cues as to the structure of the pictorial space are rich enough anyway. In that case, a subsidiary awareness of the picture surface is irrelevant. Authors seek to show that observers are generally unaware of the picture surface and that their perceptions don't suffer from that, eg that they are equally good at some (typically recognition) task whether they are aware of the picture surface or not [eg Gibson (1960), or see the title of Busey et al's (1990) paper].

Theoretical issues do not focus in perspective but on more general issues of monocular depth vision. Some of Cutting's (1986) work falls into this category. Sedgwick (1991) comes close but retains the La Gournerie-type analysis. Finally, there is the issue whether (or why) observers experience their ('double') perceptions as 'illusory' or, at least, surprising. This typically leads to discussions on 'double' perceptions, eg of something rotating and not rotating at the same time. Sometimes this is split into different things rotating, like Goldstein (1979) "rotation is perceived in terms of the pictorial space outside the picture, and spatial layout is perceived in terms of the pictorial space inside the picture" (page 78). Here 'rotation' generally remains devoid of a solid operational definition. Kubovy (1986) puts the problem in the generic form "my vantage point is changing" ... but "the scene isn't changing" (page 85). Although the former observation is without problems, the latter is not. Whether "the scene is changing" or not is not an objective fact until suitably operationalized in a scientific experiment. This is a problem with much of the available literature: whereas there exists a wealth of theoretical ideas (though mostly not of a very formal nature) there is an unfortunate scarcity of somewhat extensive bodies of coherent, quantitative data.

This paper is mainly about operationalization and quantification. We try to establish, first, in which cases and to what extent observers are simultaneously aware of both the picture surface and a pictorial space, and, second, whether and if so to which extent the scene is changing as the observer changes viewing angle. We attempt to operationalize these issues and to collect a coherent, quantitative body of data.

2 Formulation of the problem

With regard to the question whether human visual observers actively *correct* for the effects of perspective foreshortening, one can frame two extreme hypotheses: observers either do correct fully, or they don't bother at all. It is a priori likely that the actual state of affairs will lie somewhere in the middle, of course. In order to bring this problem in the realm of empirical enquiry one needs to specify the nature of 'correction' formally and one has to devise ways to operationalize the issue.

Consider the actual situation (figures 1 and 2). The observer looks at a picture (for the sake of concreteness we think of a flat picture hung on a wall) and perceives the situation veridically (ie is aware of the wall, and the fact that a flat picture is hung on the wall). In the experimental setup we make sure that the wall is perceived as a wall (through an obvious brick texture) and the picture as a picture on the wall (through a carved 3-D frame that throws a cast shadow on the wall). The observer

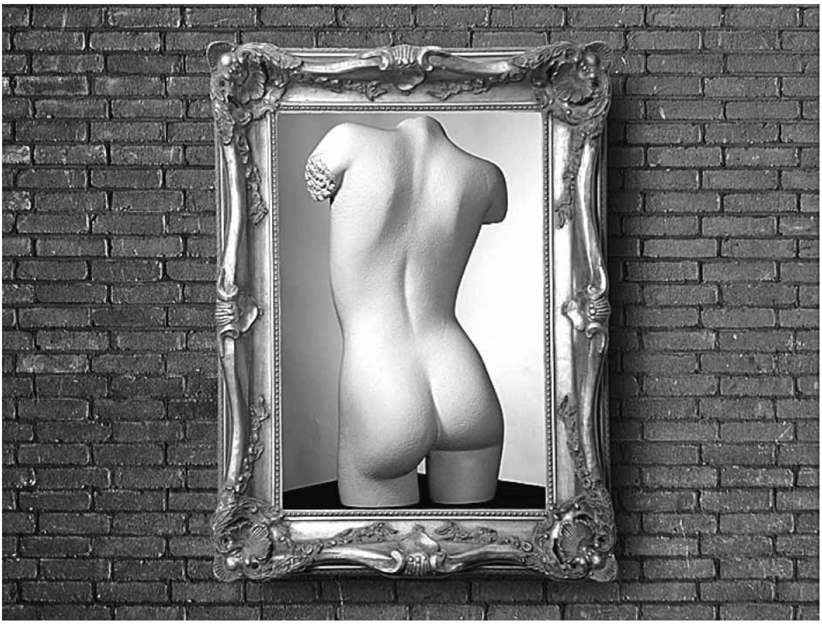

Figure 1. The fiducial stimulus. A picture of a torso (in monochrome) hangs on a brick wall in a richly sculpted gilded frame (this scene in color). The frame is 'attached' to the wall by means of a diffuse drop shadow.

Figure 2. An oblique view of the stimulus shown in figure 1. This is a perspective transformation of the fiducial picture; thus there are (minor) differences with an oblique photograph from the actual scene (not shown here for the sake of space). Note that the scene (wall and frame) looks slanted, though the torso still appears frontoparallel (though shrunk in the horizontal dimension). This is the effect often discussed in the literature: the depicted object (torso) 'turns with the observer'.

also looks *into* the picture and is aware of a pictorial space[(1)] containing an object (for concreteness we think of a portrait or a human torso) of which 'the front' (this should be unambiguous in the case of a portrait or a torso) is visible. (Notice that we don't assume that the pictorial object appears in a frontal view, ie we don't assume that the forehead or chest is in a frontal plane of pictorial space.) We make sure that the picture is obviously 'framed'; for instance, we use an elaborate (baroque style) gilded frame that doesn't escape visual attention easily. The ontological status of the frame is ambiguous, of course, since it might belong equally well to the scene (it is an object attached to the wall) as to pictorial space (it is immediately adjacent to pictorial objects). The latter is especially likely in cases the frame is painted along with the painting as in many modern painting since the late impressionistic period.

Whether the observer 'corrects' or not, the frame *as an object in physical space* should appear coplanar with the wall, that is to say, in oblique view when the wall is viewed obliquely (see figure 3). The frame as an object in pictorial space should appear coplanar with the wall if the observer indeed 'corrects', but could well have another spatial attitude (for instance frontoparallel in the observer-centered frame) when the observer doesn't 'correct'. The frontal plane of the pictorial object (the plane of the forehead or chest) will always be coplanar with the frame in pictorial space, since we will not assume the pictorial space to disintegrate. Thus the frontal plane of the pictorial object will appear in the plane of the wall if the observer does correct, but not necessarily so if the observer does not. We may assume that the frontal plane of the object will appear as a frontoparallel plane in pictorial space if the picture

[(1)] It is perhaps not entirely superfluous to define our terminology in somewhat more detail. *Picture* denotes a simultaneous order of pigments on a planar surface, *image* a mental entity. When looking *at* a picture, one sees the *picture surface* in the *visual world*. The latter is a 3-D image, which is our current perception of the physical world in front of us. The *visual field* contains an impressionistic mosaic of colors that is due to the picture surface. When looking *into* the picture, one is aware of a 3-D *pictorial space*. This is a 3-D image that is in many respects similar to the 'visual world', except for the fact that it does not correspond to a 'physical world', but is a mental entity due to 'pictorial cues'. In this paper, the situation is slightly more complicated because we consider such things as pictures in pictures, but this is the basic terminology.

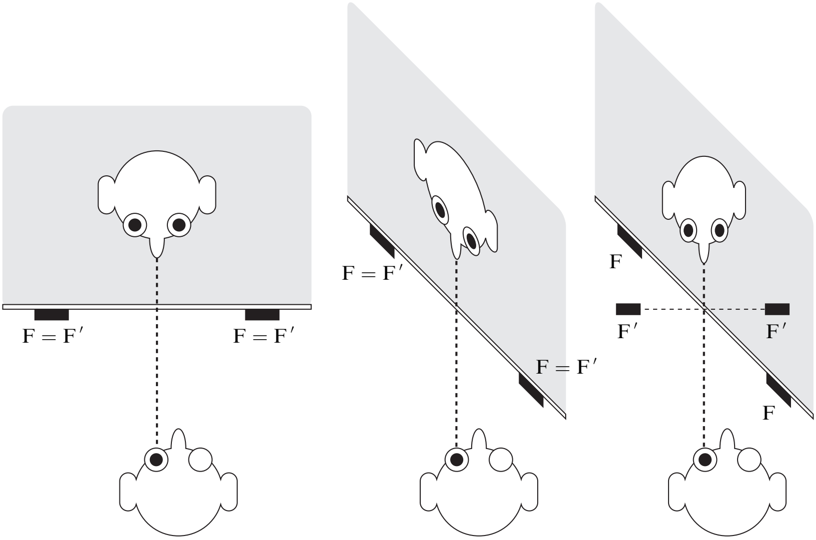

Figure 3. On the left is the fiducial situation. With the gray area we suggest 'pictorial space', distinct from physical space which contains the wall (thick horizontal line), frame F, and the picture (in the plane of the frame). The frame is also in pictorial space (denoted F′); the depicted object appears as a three-dimensional pictorial object (of course, only the front is really seen). In the middle and on the right are two extreme expectations for the case of oblique viewing. In the middle is shown the case of full 'correction'. The frames in physical and pictorial space coincide (F = F′), the pictorial object looks skewed but still faces the observer. The forehead is still parallel to the picture frame; in that sense it is not 'turned'. On the right is shown the case of no correction at all. Here the frames in physical space (F) and in pictorial space (F′) don't coincide. Again, the forehead is parallel to the frame F′; however, the frame (as F) is also seen to be slanted. The pictorial object again faces the observer.

is viewed normally, or if the frame and wall are invisible (eg when only the picture is presented in a reduced laboratory setting).

The effect of foreshortening is mainly to contract the horizontal scale of the retinal image, but not the vertical scale. (The much smaller nonlinear perspective effects occur mainly for very wide fields of view.) In the reduced situation the observer obtains the same stimulus in the normally and obliquely viewed cases, except for this horizontal scaling. All monocular cues being the same, we must assume that the pictorial objects will be the same too, except for a horizontal scaling. In all cases the observer is expected to 'see' the object in a frontoparallel attitude, since there is no cue (except for familiar size) to the contrary.

If the observer were to 'correct', then a transformation has to be applied to pictorial space that will move the frame in pictorial space such as to coincide with the frame in physical space. (We assume that the observer perceives physical space veridically; thus there is no need to distinguish between physical and visual space at this point.) This transformation cannot be a rotation in the Euclidian sense, since the back side of the object can never be revealed (it is not depicted). Fortunately, the type of transformation available to the observer (and indeed applied on frequent occasions as the empirical evidence shows) is currently well understood. It is a transformation that may skew planes (*any* plane can be transformed to the frontoparallel attitude), but cannot show their opposite sides. [This type of non-Euclidian 'rotation' is not periodic (Klein 1872; Strubecker 1941; Jaglom 1979; Sachs 1990). The 'rotations' are like shears, but categorically different (because in pictorial space) from those commonly mentioned in discussions on picture perception, originally described by La Gournerie (1859).]

Thus we may visualize the two extreme possibilities as shown in figure 3.

The next step is to devise methods to operationalize the various entities. This is a very difficult problem. In order to measure spatial attitudes (these are the major properties of interest) we have to introduce probes. Probes have to be implemented in physical space (there is no option since we have only indirect access to mental spaces), but they will appear either in visual space (which in many—though by no means all—cases roughly corresponds to physical space, although the two are categorically different) or in pictorial space of which we know very little. A generic measurement will either use a probe as a marker (eg a point), or as a gauge figure, that is a spatial structure to be compared with spatial structures in the space to be probed. The former type of probe is the least problematic. In monocular view, a point (presented as a tiny object in physical space) determines a visual direction in both visual and pictorial space. The major ambiguity here is the location of the point in depth. In a structured scene, a point will typically adhere to a surface. A gauge figure is much more problematic. It will usually adhere to the nearest surface all right, but its spatial structure will suffer the 'correction' that applies to the space it is perceived to be in. We cannot a priori be certain whether that will be visual space (the space that contains the perception of the wall on which the picture hangs) or the pictorial space of the picture.

In this experiment we use both (point) markers and a gauge figure that we have used frequently to sample surface attitude in pictorial space. We use this gauge figure to probe the wall, the interior of the frame, and the relief of the depicted object. The point markers are used to indicate near and far points along horizontal lines in the visual field. These measurements yield rather more information than has been available in studies reported in the literature.

3 Design of the experiment

Our basic stimulus is a picture presented on a computer screen in a dimly illuminated room. The picture fills most of the screen. A large fraction of the picture shows the photograph of a brick wall. On the wall hangs a framed picture (all in the picture, of course). The frame is an elaborately carved and gilded one; it clearly hangs in front of the wall (as is evident from a diffuse cast shadow of the frame on the wall). The plane of the interior of the frame is parallel to the wall. The frame contains the picture of a human torso. Although the scene (brick wall and gilded frame) is in color, the picture in the frame is monochrome (neutral gray). Thus the picture looks like a *picture*, not like a hole in the wall revealing another vista. The scene looks three-dimensional (frame in front of the wall, frame in high relief) but the picture looks like a flat, framed object. The observer is subsidiarily aware of a pictorial space though. The picture looks flat as a picture, but the torso is also seen as three-dimensional in a space ('pictorial space') that is distinct from the space of the scene.

On this basic stimulus we have prepared a number of variations. First of all, the computer monitor can be turned (about the vertical axis through the center of the front of the screen) into oblique positions. Next, we can turn the monitor into an oblique position, but pre-deform the image (brick wall, frame, picture and all) so as to produce the same retinal image as the observer would obtain in the fiducial situation (see figure 4). Then we created the situation where the monitor is in the normal (fronto-parallel) position, but the image has been pre-deformed [a technique that has been used before, eg by Busey et al (1990), and derives essentially from ancient techniques of anamorphosis] so as to produce the same retinal image as the observer would obtain if the normal picture was presented on the obliquely positioned monitor. Finally, we repeated all these situations with a picture in which the frame contained only a uniformly gray (blank) picture.

When doing experiments with the gauge figure (see figure 5) we ran into the problem whether to pre-deform the gauge figure so as to remain invariant on the retina or on

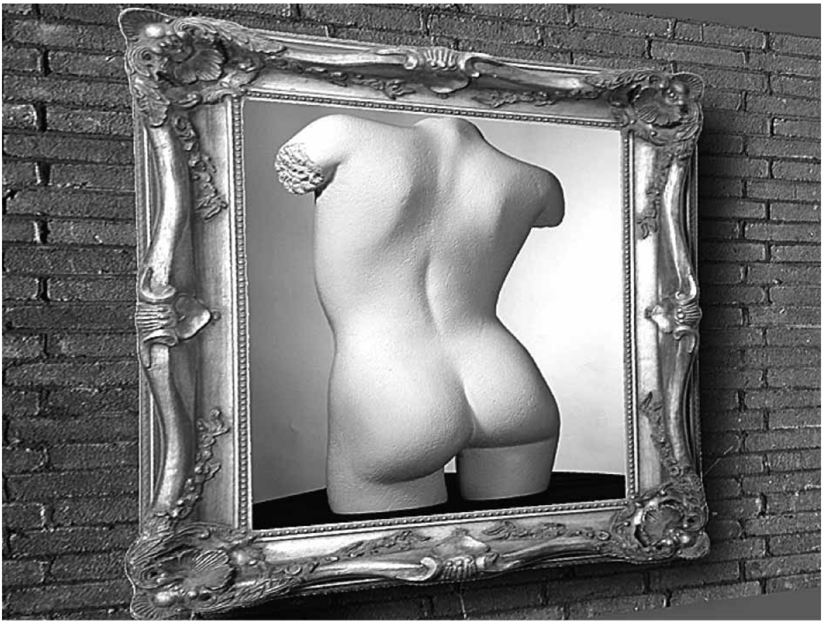

Figure 4. 'Pre-deformed' picture for oblique presentation. When viewed from the correct vantage point this picture yields the same retinal irradiance pattern as the fiducial picture (see figure 1) seen frontally.

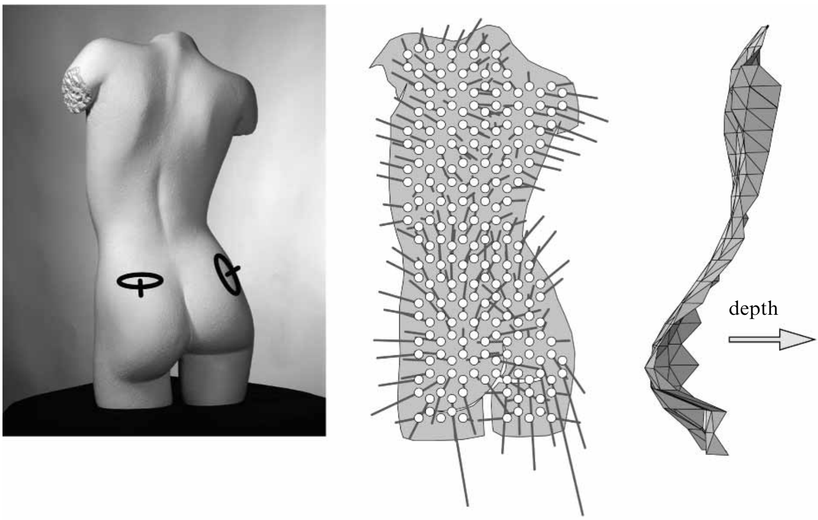

Figure 5. The 'gauge figure' method of determining pictorial surface attitude. On the left is the stimulus (a picture) with two gauge figures superimposed on it. The right gauge figure should look like a circle on the pictorial surface with a needle sticking out at right angles, unlike the left one. In the experiment observers adjust gauge figures to a visual 'fit' (one at a time). In the middle is shown the immediate result of a session. The vertices have been visited in random order; the line segments indicate direction and magnitude of the depth gradient. On the right is shown the best-fitting surface with respect to the empirically determined gradients. The surface is seen 'in profile' as compared with the stimulus; depth increases towards the right.

the screen. This is a point of major interest, since a deformed gauge figure that is *seen* as deformed will look 'wrong' and cannot properly function as a probe (in this case for surface attitude). The gauge figure we used was the projection (rendered in wire frame) of a circle with a line segment erected at the center of the circle, normal to its plane, with the length equal to the radius of the circle. Consider a view from a direction normal to the plane of the circle. Then the projection is a circle with a central dot (the singular image of the line segment). A deformation would transform this into an ellipse with a central dot, a planar figure that cannot be interpreted as the orthogonal projection of the fiducial gauge figure at all. Only if the observer 'corrected' the gauge figure (turning the ellipse into a circle again) would the gauge figure look 'right'. This is an issue that can only be resolved at the conclusion of the experiments. If a gauge figure is pre-deformed for the wrong space it looks 'wrong' (typically deformed) to the observer and the measurement should be regarded very critically.

Thus the full stimulus situation is parameterized by a number of parameters: whether the monitor is actually frontal or oblique, whether the picture on the monitor is the fiducial one or a deformed one, whether the irradiance pattern on the retina is fiducial (as in the case of normal viewing of the undeformed picture on the monitor in frontal

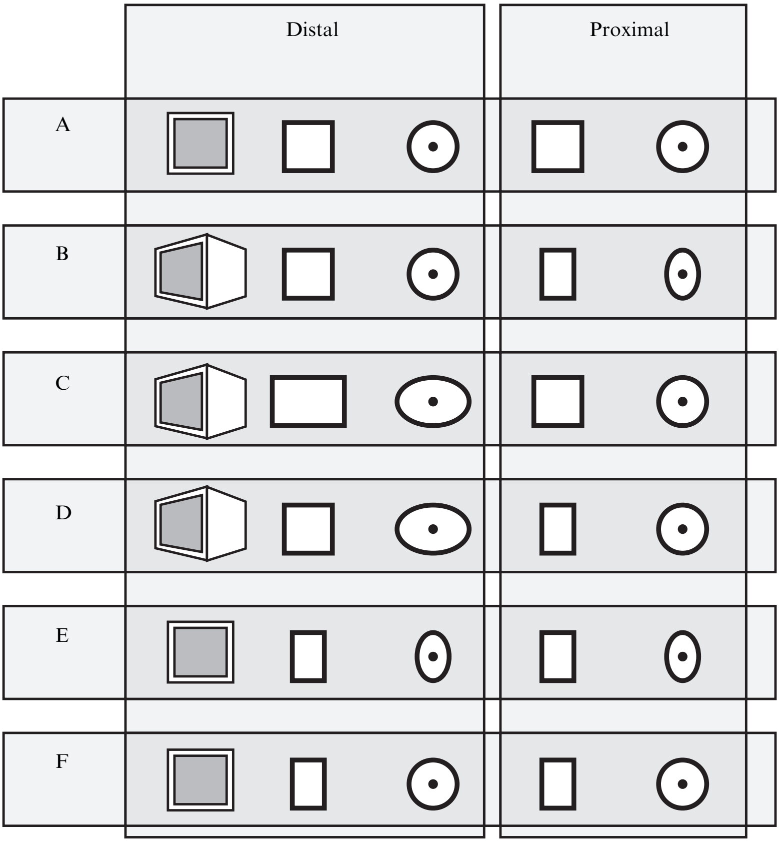

Figure 6. An overview of all stimuli (labelled A ... F as used throughout the text) used in the experiment. In the three left columns the distal stimulus is described: in the first column the attitude of the monitor (frontoparallel or oblique), in the second column the deformation of the stimulus on the screen (a square means undeformed, a wide rectangle denotes a pre-deformation that yields an undeformed proximal stimulus in the case of an oblique monitor, a narrow rectangle indicates a simulation of the proximal view of the fiducial stimulus for an oblique monitor), in the third column the shape of the gauge figure in frontoparallel attitude (circular means undeformed on the screen, the wide and narrow ellipses denote deformations similar to those of the picture in the second column). In the two right columns the proximal stimulus is described. Note that stimulus and gauge figure are fiducial for the distal stimulus in cases A and B, and for the proximal stimulus in cases A and C.

attitude) or not, and whether the picture of the gauge figure formally fits the (possibly deformed) picture or is undeformed on the retina. All in all, we measured six different conditions, denoted here A ... F. In figure 6 we present the various conditions for easy reference. In this figure we list both the physical conditions and the retinal irradiance distributions (the 'distal' and the 'proximal' stimuli in old-fashioned terms).

4 Methods

4.1 *Stimuli*

The brick wall and the gilded frame were photographed under controlled conditions, taking care that the direction of illumination was the same for both. The frame was cut out in Photoshop and pasted over the wall, a transparent, fuzzy shadow being pasted in between (conforming to the illumination used in the photographs). The picture was photographed in the studio and was pasted in the frame in Photoshop. For the stimuli without the picture, a uniform gray rectangle was substituted. The deformed stimuli were produced with a C-program that implemented the exact projective transformation that corresponded to a rotation of the monitor as seen from the observer's position (there is a well-defined perspective center since we used monocular viewing).

4.2 *Setup*

The stimuli on the screen were 400 mm $\times$ 300 mm (aspect ratio 1.33); the monitor was at 800 mm from the eye. The position of the head was fixed with a chin-and-forehead rest. Observers wore an eye patch over their non-preferred eye. During the sessions, the room was dimly illuminated, and the monitor could be seen. Observers controlled the attitude of the gauge figure or the position of a marker through a computer trackball device.

4.3 *Observers*

The observers were the authors. All had long experience with psychophysical experiments; none had any experience with the present experiments though. There were no prior expectations as to the outcome of the experiment since there exist virtually no prior data and no theory. During the experiments the data were not discussed or compared. All observers had normal or corrected-to-normal vision. Observation was monocular and observers used their preferred eye. Monocular vision was used to optimize pictorial cues.(2)

5 Procedure and results

5.1 *Initial observations*

The scene looked very realistic to all observers. The brick wall was perceived as perfectly frontoparallel in the fiducial position. With the monitor turned on the brick wall looked oblique to the observers. When the stimulus was pre-deformed to produce the same retinal irradiance pattern as in the fiducial situation, the picture (brick wall and all) appeared to detach from the monitor and hover in the air in a frontoparallel attitude. (This is a case of severe discrepancy between the layout of the visual space and the structure of the physical space: in the physical space, the surface of the monitor and the picture plane coincide; in the visual space, these planes subtend a subjective angle of tens of degrees.) When the oblique view was presented on the frontoparallel monitor, the impression was that of a picture of an oblique brick wall presented on a frontoparallel monitor.

(2) A reviewer asked us to "speculate on the differences binocular vision would make". Since we didn't do the experiment we can, indeed, only speculate. Binocular vision tends to counteract (although rarely destroy) monocular stereopsis in the case of pictorial perception (Koenderink et al 1994). We expect similar results, though much weakened and less clear.

All observers remarked spontaneously on the impression of the torso in the case of the deformed pictures viewed frontally. When the view that would produce the fiducial retinal irradiance pattern on the oblique monitor was presented on the monitor in fiducial attitude, the figure was described as "fat" or "gross". When the oblique view was simulated on the frontoparallel monitor, the figure was described as "rather slender" or "skinny". The same remark was made when the fiducial stimulus was presented on the oblique monitor.

These informal observations perhaps suggest that observers did not make corrections (at least of appreciable magnitude) for foreshortening, but reacted on the basis of the dimensions in the visual field.

5.2 *Surface attitude measurements*

We probed the surface attitude of
—the brick wall,
—the interior of the empty frame,
—many points on the (pictorial) surface of the torso,
with the objective of determining the apparent spatial attitude of the wall, whether the frame was perceived as coplanar with the wall, and, finally, whether the pictorial relief of the torso formed a coherent structure together with the frame.

In the first two cases the stimulus contained a uniform gray rectangle in the frame instead of the torso image.

In the last case, measurements were taken on the barycenters of the faces of a triangulation of the interior of the bounding contour of the torso in the picture (one face at a time, in random order). Each face was visited once per session. We did three sessions for each stimulus, yielding a total of three settings per face. This yields an estimate of both slant and tilt per face and a notion of the spread in the data. In this paper, we show only averages over all three settings. Since we show the reliefs as three-dimensional surfaces, the spread in the data (which is small and plays virtually no role in the conclusions drawn from the data) is readily apparent from the lack of complete smoothness of the relief.

5.3 *Results of the surface attitude measurements*

The settings on the wall and in the empty picture plane did not differ. We show only the averages here (figure 7).

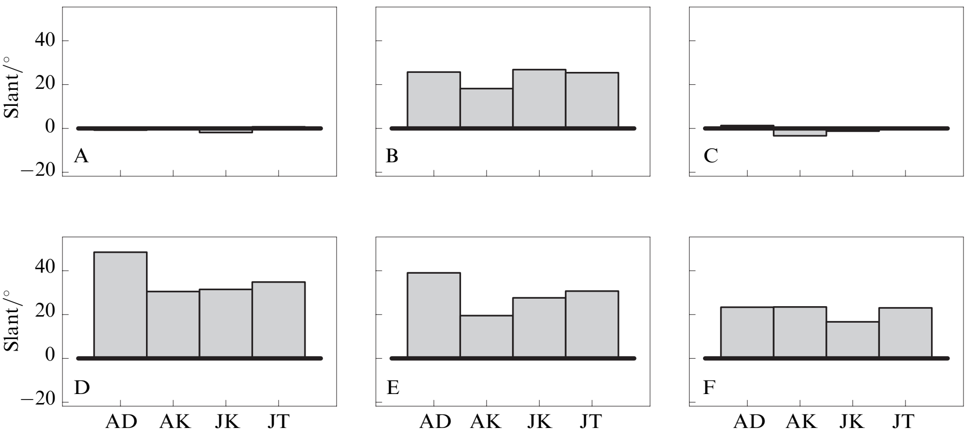

Figure 7. The slant of the picture plane for all stimulus conditions and all observers. The physical slant was 0° for cases A, E, and F; and 45° for cases B, C, and D. The proximal stimulus simulated slant 0° for cases A and C; and 45° for cases B, D, E, and F. Note that the results follow the proximal stimulus, though the magnitude of the slant is underestimated.

Note that the physical slant was 0° for the stimulus conditions A, E, and F; and 45° for the stimulus conditions B, C, and D. On the other hand, the proximal stimulus had a 'simulated slant' of 0° for the stimulus conditions A and C, and 45° for the stimulus conditions B, D, E, and F.

A cursory view of the results (figure 7) reveals that in the stimulus conditions A and C the picture plane was perceived as frontoparallel, whereas in all other cases it was perceived slanted. Cases A and C are, of course, the only cases where the proximal stimulus was consistent with frontoparallelity. We conclude from this that the observers invariably follow the proximal stimulus. The magnitude of the slant is underestimated though, except in rare cases (observer AD in stimulus condition D). [This underestimation has been frequently described as "regression to the picture plane"—eg see Goldstein (1979).] In the stimulus conditions A and C, the slants set by the observers are not significantly different from frontoparallel; in all other cases they are very significantly different from frontoparallel. Thus the result is as clear-cut as can be.

The pictorial reliefs for all observers were calculated as the best fits to their slant and tilt settings (the raw data). This can be regarded as an operational definition of the pictorial relief. It is convenient to use these reliefs instead of the raw data, because it allows us to use depth (at the vertices of the triangulation) instead of slant and tilt; otherwise it makes little difference. Note that the depth is determined only up to a constant offset (which is not a problem). Results are shown only for one observer (AK: figures 8, 9, and 10). We consider this amply sufficient, since the data of all four observers are very similar indeed. When we compare conditions through depth scatter plots (figure 8), we find that the reliefs are virtually the same in all six stimulus conditions. This impression is corroborated by a closer study of the reliefs (profiles shown in figure 9, top views in figure 10). Especially in the top view, an oblique position would be immediately evident (remember that we are looking for 45° differences). The conclusion is that the relief does not depend upon the stimulus conditions at all.

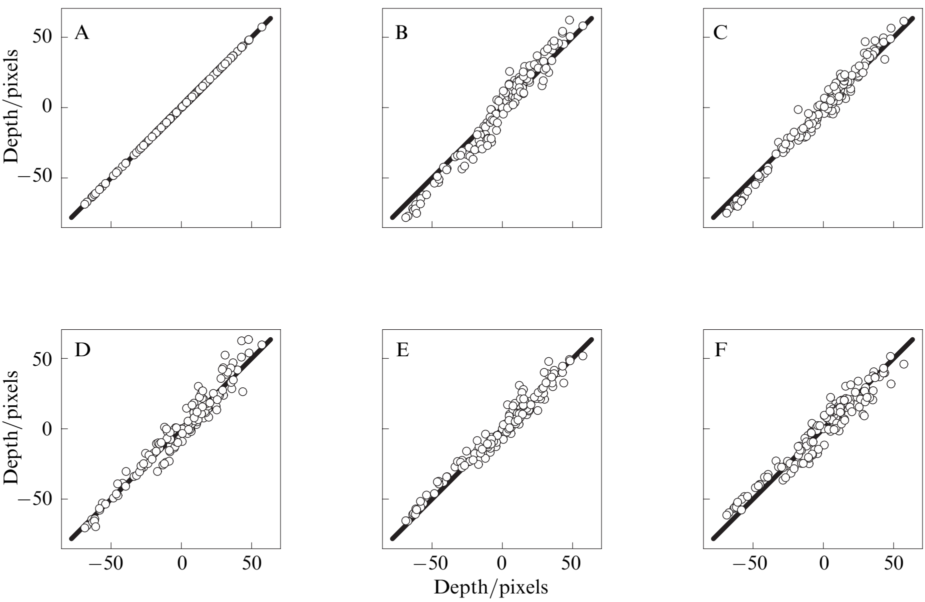

Figure 8. Scatter plots of the depths at the vertices for any stimulus condition against the depth for case A (frontoparallel monitor, fiducial stimulus, and gauge figure) for observer AK. The result is representative for all four observers. The settings are virtually identical. [The depth scale is in frontoparallel (picture plane) pixels. The total depth is about a quarter of the width of the picture.]

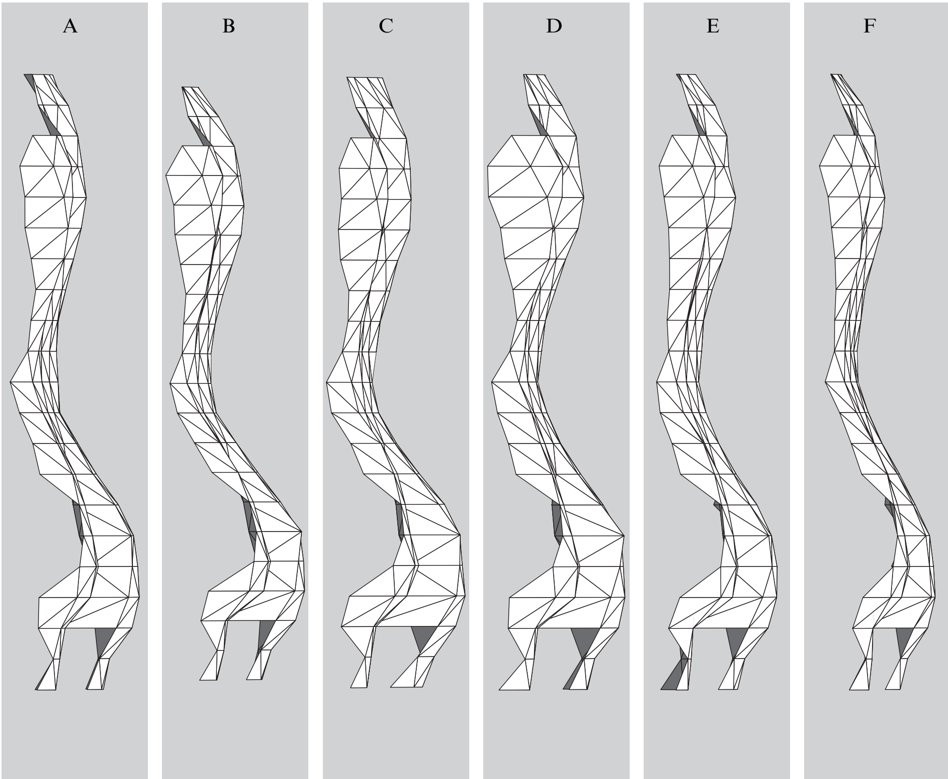

Figure 9. Profile views of the pictorial reliefs for observer AK in all stimulus conditions. The result is representative for all observers. The profiles are virtually identical.

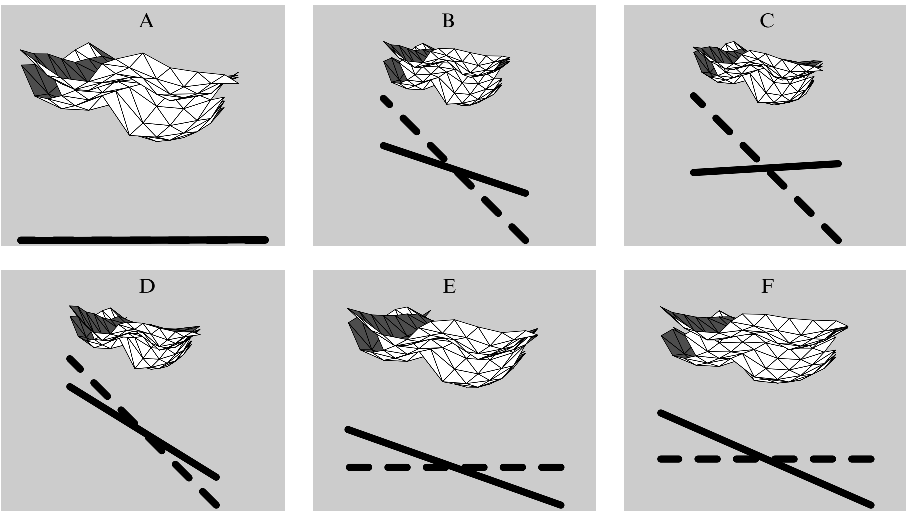

Figure 10. An overview of the results for observer AK. This figure shows top views of the pictorial relief with the orientations of the picture plane in physical space (broken line piece) and pictorial space (drawn line piece). Note that the orientations of the picture plane in physical and pictorial space are in general distinct (the fiducial case being an exception). The pictorial relief is not strongly related to either attitude (note that from this viewing direction an oblique attitude should show up most clearly). This is representative for all four observers.

We show an overview for observer AK in figure 10 (this is representative for all four observers). This figure shows the top views with the orientations of both the physical screen (drawn with respect to the viewing direction in physical space) and the perceived picture plane (drawn in pictorial space). Note that the picture plane is quite differently oriented in physical and pictorial space. The attitude of the top view is unrelated to either. Even when the observers perceive the picture plane as slanted, their pictorial relief for the object in the image hardly changes.

5.4 *Near and far point settings*

The torso was perceived as an undulating surface in pictorial space ('pictorial relief'). Observers were required to place a marker on a horizontal line at places where the relief (measured along the line) was locally nearest or farthest from the observer. The way this was implemented is that a marker was constrained to move on a horizontal line. The type of depth extremum can either be a far or a near point. The observer was left free to indicate as many near and far points as he/she deemed necessary. (In almost all cases near and far point indications alternated as would be expected.) The procedure was repeated for many heights in the picture, visited in randomized order. This task felt quite natural to all observers and took only little time (roughly limited by the handling of the mouse or track ball).

After the session, the results were sorted with respect to height in the picture. The indicated points neatly line up to reveal ridges and ruts of the relief. From the scatter of settings at adjacent heights one readily obtains an indication of the spread in these data. We find that all observers are able to indicate the ruts and ridges of the relief with remarkable precision (in the sense of repeatability).

The results can be interpreted in terms of changes of the apparent frontoparallel plane as a function of the parameters of the presentation (attitude of the monitor, deformation, if any, of the picture).

5.5 *Results of the near and far point settings*

In this case only the stimulus conditions A, B (or D), C, and E (or F) are relevant. Conditions B, D, and E, F are mutually equivalent because the gauge figure is not used; thus its deformation is not a relevant parameter. We show the ridges and ruts of the relief for observer AK (figure 11). This is sufficient because these results are fully representative for those of all four observers.

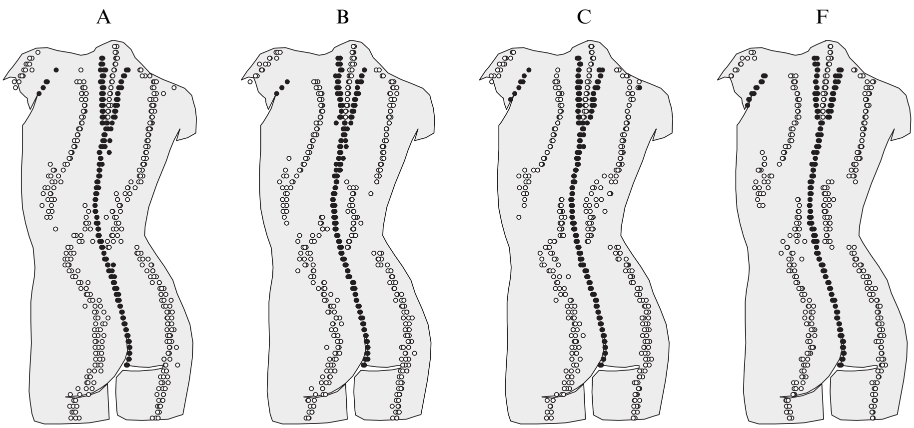

Figure 11. Ridges and ruts of the relief for observer AK. These results are representative for all four observers. Note that only stimulus conditions A, B (or D), C, and E (or F) are relevant, because the gauge figure is not used.

The results are very clear: there is no significant difference between the stimulus conditions. Thus the frontoparallel points on horizontal sections do not at all depend upon the stimulus condition, not on the orientation of the monitor, nor on the (pre-)deformation of the stimulus. Apparently the results are fully determined by the monocular cues which are identical, or equivalent, in all cases.

6 Conclusions

We don't find that the experiments serve to change (our) interpretation of the data and views available from the literature to any large extent. However, it has to be said that the picture that arises from a study of the literature is not entirely coherent. There appears to be some (weak) consensus though that no 'correction' is applied to pictorial space due to obliquely viewed pictures. With this we agree. We find ourselves in a more comfortable position than many earlier authors, because we can draw upon a much larger and consistent body of empirical data (empirically determined pictorial reliefs along with estimates of the spatial attitudes of picture planes) than has been available before.

The results from the experiments are consistent and allow an obvious interpretation: the pictorial space in a painting hung on the wall and the visual space that contains the visual wall, picture frame, etc (in many respects very similar to the physical space containing wall, picture frame, and observer as physical objects) are virtually *independent*. Observers perceive the wall indeed as oblique with respect to their direction of view when they view it from an oblique angle, although they tend to underestimate the angle somewhat. They always see an object depicted in a frontoparallel pose (eg a portrait *en face* or the frontal view of a torso) as facing them squarely, whatever the angle of view (in our case frontal viewing and a 45° oblique angle). This is clearly borne out by the empirically determined pictorial reliefs and by the judgments of frontoparallel points along horizontal lines. Observers often remark spontaneously upon the unexpectedly slender appearance of pictorial objects seen in obliquely viewed pictures (see figure 2). This has been noticed by Goldstein (1987) and by Gombrich (1960), who in his well-known *Art and Illusion* argues repeatedly (chapters VIII and IX in particular) that observers are easily satisfied with depictions of persons, trees, etc, that are technically too thin or small. They notice such cases but see no particular need for any 'corrections'. This indicates that observers do not necessarily apply a foreshortening correction at all, but that their pictorial spaces are simply squashed in the foreshortening direction by the cosine of the oblique viewing angle. They do not remark on the aspect ratio of the frame or the bricks of the wall, so it seems likely, though we did not check this in the present experiment, that such a foreshortening 'correction' (often called 'shape invariance') might be applied to the wall and the picture plane (including the frame). Thus there appears to exist a perhaps remarkably *absolute segregation* between the pictorial space of the painting and the visual space that contains the painting as a visual object. This is remarkable in view of the fact that many authors connect the occurrence of 'compensation' with the absence of a strong awareness of the picture surface (eg Gombrich 1960; Pirenne 1970; Perkins 1973; Halloran 1993). This is in accord with Goldstein's (1987) distinction between "perceived projection" and "perceived spatial layout".

Does this mean that the pictorial object 'rotates along with the observer' as the observer assumes a series of oblique viewing positions by walking along the painting on the wall? Yes and no. (See figure 12.) *Yes* in terms of the physical space containing the scene, picture, and observer: the pictorial object always squarely faces the observer, it thus looks or points into the observer's visual direction. As the observer changes the visual direction with respect to the picture plane, the looking or pointing direction of the pictorial object in physical and visual space has to rotate with it.

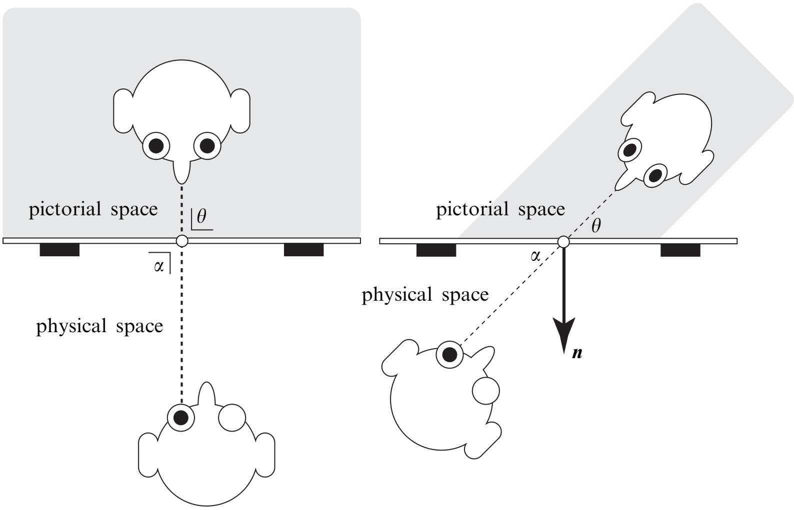

Figure 12. Analysis of the frontal and oblique cases. As the physical and pictorial spaces have been superimposed, the reader should take great care in interpreting these figures! The frontal case on the left is the simplest. The observer views the picture squarely, thus the angle α (in physical space) is a right angle ($\alpha = \pi/2$; in physical space angles are periodic with period 2π), the visual direction (in physical space) coincides with the surface normal $\boldsymbol{n}$ to the picture (again, in physical space). The angle θ denotes the angle in pictorial space subtended by the viewing direction and the frontoparallel plane (both in pictorial space). It is a 'right angle' ($\theta = +\infty$; in pictorial space angles are not periodic and range from $-\infty$ to $+\infty$). On the right we have the oblique case. Here $\alpha = \pi/4$; thus the viewing direction does not coincide with the normal to the picture plane (all in physical space). The angle θ in pictorial space still equals $+\infty$ (a right angle) though, thus the viewing direction is at a right angle to the frontoparallel plane (all in pictorial space). Hence there is no clash between the oblique view and the normal view of the pictorial object, for one angle is in physical, the other in pictorial space. Note that the object has 'shrunk' in the horizontal direction though.

No in terms of the pictorial space that contains the pictorial object. In this space, the pointing direction is invariant, and it is always orthogonal to the picture plane (as an object in pictorial space), exactly as the painter painted it. The 'angles' involved here are non-Euclidian, and so are the rotations: the reader should avoid confusion with the 'affine shears' commonly encountered in the literature (see Appendix). These affine shears derive from La Gournerie's (1859) analysis of the geometry of obliquely viewed perspective drawings and are categorically different from the transformations discussed here. Here the correct geometry is that of isotropic space (Strubecker 1941; Jaglom 1979; Sachs 1990). This geometry applies because of the essential ambiguity of fronto-parallelity in pictorial space, which is simply not constrained by the typical pictorial cues of a painting or photograph. This would be different for pictorial cues of a different type, eg the presence of a frontoparallel flat square in the picture ('frame in a frame'). Halloran (1993), using objects with various degrees of depth (flat to deep) and frames in frames, suggests that the depth is an important factor and that depicted frames also rotate. Perkins (1973), using line drawings of rectangular boxes as stimuli, finds that "compensation is the dominant trend in dealing with oblique views of depictions of rectangular solids". The pictorial content might well make a difference in experiments like these. This is a topic that invites further research.

Our experiment might be criticized on the grounds that the picture frame was not hung on a real brick wall, but on the picture of such a wall presented on a monitor. The monitor was clearly perceived as a physical object in physical space and the observers were always aware of its spatial attitude with respect to them.

The CRT screen is a slightly odd object though; it is different from a 'surface' (a rough canvas, say) but neither is it an 'aperture' (a window on a world), but rather something uncomfortably in between. Thus we have something like a picture of a picture in a picture of a scene presented in another scene (this difficulty in parsing the sentence indeed reflects the complexity of the situation very well). This is a valid objection. However, it is easy enough to try informal experiments in cases of physical pictures on physical walls. At least we find that the findings of the present experiment are corroborated by our experiences with such cases. It is certainly possible (though elaborate) to repeat the experiments in more realistic settings, for instance using a laser projector to superimpose gauge figures or markers on the wall or painting, and it is clearly of some interest to pursue this.

References

Arijon D, 1976 *Grammar of the Film Language* (Hollywood, CA: Silman-James)

Busey T A, Brady N P, Cutting J E, 1990 "Compensation is unnecessary for the perception of faces in slanted pictures" *Perception & Psychophysics* **48** 1 – 11

Cutting J E, 1986 "The shape and psychophysics of cinematic space" *Behavior Research Methods, Instruments & Computers* **18** 551 – 558

Cutting J E, 1987 "Rigidity in cinema seen from the front row, side aisle" *Journal of Experimental Psychology: Human Perception and Performance* **13** 323 – 334

Ellis S R, Smith S, McGreevy M W, 1987 "Distortions of perceived visual directions out of pictures" *Perception & Psychophysics* **42** 535 – 544

Gibson J J, 1954 "A theory of pictorial perception" *Audio-Visual Communication Review* **2** 3 – 23

Gibson J J, 1960 "Pictures, perspective, and perception" *Daedalus* **89** 216 – 227

Gibson J J, 1971 "The information available in pictures" *Leonardo* **4** 27 – 35

Goldstein E B, 1979 "Rotation of objects in pictures viewed at an angle" *Journal of Experimental Psychology: Human Perception and Performance* **5** 78 – 87

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** 256 – 266

Goldstein E B, 1988 "Geometry or not geometry—perceived orientation and spatial layout in pictures viewed at an angle" *Journal of Experimental Psychology: Human Perception and Performance* **14** 312 – 314

Gombrich E E, 1960 *Art and Illusion. A Study of the Psychology of Pictorial Representation* (London: Phaidon Press)

Gournerie J de La, 1859 *Traité de perspective linéaire contenant les tracés pour les tableaux plans et courbes, les bas-reliefs et les décorations théatrales, avec une théorie des effets de perspective* (Paris: Dalmont et Dunod; Mallet-Bachelier)

Hagen M A, 1976 "Influence of picture surface and station point on the ability to compensate for oblique view in pictorial perception" *Developmental Psychology* **12** 57 – 63

Hagen M A, 1986 *Varieties of Realism. Geometries of Representational Art* (Cambridge: Cambridge University Press)

Halloran T O, 1993 "The frame turns also—factors in differential rotation in pictures" *Perception & Psychophysics* **54** 496 – 508

Hochberg J, 1962 "The psychophysics of pictorial perception" *Audio-Visual Communication Review* **10** 22 – 54

Jaglom I M, 1979 "A simple non-Euclidian geometry and its physical basis: an elementary account of Galilean geometry and the Galilean principle of relativity" translated by A Shenitzer, Ed. Ass. B Gordon (New York: Springer)

Klein F, 1872 *Vergleichende Betrachtungen über neue geometrische Forschungen* (The "Erlangen Program") Programm zu Eintritt in die philosophische Fakultät und den Senat der Universität zu Erlangen (Erlangen: Deichert)

Koenderink J J, Doorn A J van, Kappers A M L, 1994 "On so-called paradoxical monocular stereoscopy" *Perception* **23** 583 – 594

Kubovy M, 1986 *The Psychology of Perspective and Renaissance Art* (Cambridge: Cambridge University Press)

Leonardo da Vinci, 1585/1989 "Trattato della pittura", in *Leonardo on Painting* Ed. M Kemp; selected and translated by M Kemp and M Walker (1989, New Haven, CT: Yale University Press)

Perkins D N, 1973 "Compensating for distortion in viewing pictures obliquely" *Perception & Psychophysics* **14** 13 – 18

Pirenne M H, 1970 *Optics, Painting and Photography* (Cambridge: Cambridge University Press)
Rosinski R R, Mulholland T, Degelman D, Farber J, 1980 "Picture perception: An analysis of visual compensation" *Perception & Psychophysics* **28** 521 – 526
Sachs H, 1990 *Isotrope Geometrie des Raumes* (Braunschweig/Wiesbaden: Vieweg)
Sedgwick H A, 1991 "The effects of viewpoint on the virtual space of pictures", in *Pictorial Communication in Virtual and Real Environments* Eds S R Ellis, A M K Kaiser, A C Grunwald (London: Taylor & Francis) pp 460 – 479
Strubecker K, 1941 "Differentialgeometrie des isotropen Raumes I" *Sitzungsberichte der Akademie der Wissenschaften Wien* **150** 1 – 43
Topper D, 2000 "On anamorphosis: Setting some things straight" *Leonardo* **33** 115 – 124
Wallach H, Marshall F J, 1986 "Shape constancy in pictorial representation" *Perception & Psychophysics* **39** 233 – 235

APPENDIX: The angle metric of pictorial space

In order to simplify the discussion, we consider only pictorial spaces due to orthographic projection here. (It is not hard to generalize to the general perspective case.)

Pictorial space is different from Euclidian space. For instance, in Euclidian space any oriented plane (a plane with 'front' and 'back' sides assigned) can be *reversed* (front and back interchanged) through suitable rotations. In pictorial space this is not possible: when you see the front, the back side is not in the picture and there is no way you could ever see it! The 'movements' or 'congruences' of pictorial space may only shift depth values along visual rays in combination with Euclidian movements in the frontal planes. Thus the only periodic rotations are in the frontoparallel planes, with a visual ray as axis. 'Rotations' (we use quotes because of the non-Euclidian character) involving the depth dimension cannot be periodic because you cannot bring the back sides of pictorial objects into view.

In pictorial space the depth dimension is 'virtual' (a completely mental thing), whereas the frontal plane dimensions derive from the dimensions of the picture plane and in that sense are 'real'. For the distance metric of pictorial space one simply takes the Euclidian metric of the picture plane. Note that this implies that points at zero distance may be distinct (eg points coinciding in the picture plane but located at different depths, for instance in the case of apparent transparency). Such points are 'parallel'. For parallel points one may use the depth difference as the distance. 'The' distance of any two points is then defined as either the distance in the picture plane, or (in the case of parallel points) the depth difference. This is a good definition because it implies that the proper movements (the transformations that conserve distances, ie the 'congruences') are movements in the picture plane combined with rigid depth shifts along visual rays. If we assume pictorial space to be a special case of a projective space (ie the movements conserve lines and planes with their intersections), the geometry is fully determined.

Rotations involving depth can change the slant and tilt of planes, but cannot 'turn them around'. Such rotations of pictorial space appear to the Euclidian eye as 'shears' along a family of parallel planes containing visual directions. The angle metric can be defined in an analogous fashion to the angle metric in Euclidian space, namely as arc length over the unit sphere (see figure A1). This works out slightly differently because the unit sphere in pictorial space (the locus of all points at unit distance from a given point) is made up of the visual rays that meet a unit circle in the picture plane. For the rotations in depth, the angle metric is thus simply distance along a visual ray. It runs from minus to plus infinity. A rotation of a given tilt changes the slants of all planes by the same amount, exactly as in the Euclidian case. These 'rotations' appear as the 'bas relief ambiguity' in the theory of shape-from-shading and as interobserver/intertask differences in the psychophysics of pictorial space.

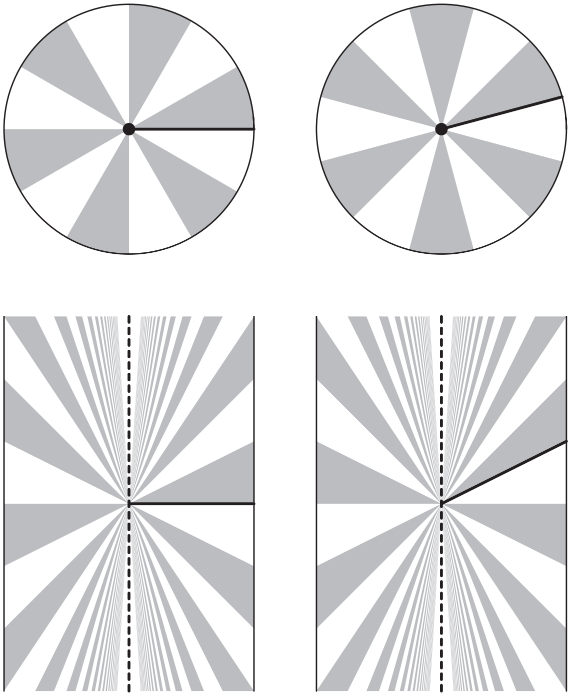

Figure A1. The top row shows a Euclidian and the bottom row a pictorial space protractor. The protractors in the right column have been rotated with respect to those in the left column. Note that the 'rotation' in pictorial space looks like a shear to the Euclidian eye. The protractors are unit disks. For the Euclidian case the center is a point, for the case of pictorial space it is the dotted line, which has the direction of a visual ray. The points on the circumferences have unit distance to the center in either case. The angle metric is arc length along the circumference in either case. The divisions illustrate equal angular increments. Note that the Euclidian angle metric is periodic, that of pictorial space is not.

The 'north pole' of the unit sphere (the point at equal distance from the 'equator', a unit circle in the frontoparallel—or picture—plane) is the vanishing point in the visual direction. The polar axis is clearly a plumbline in the frontoparallel plane. Note that it subtends an infinite angle with that plane. But the polar axis subtends the same (infinite) angle with any proper (not containing a visual direction) plane, for there exists a rotation that will turn the plane into frontoparallel position and that leaves the polar axis invariant. Since the polar axis is a visual ray, we conclude that all visual rays are plumblines on all proper planes in the geometry of pictorial space!

Different from Euclidian space, the 'normals' on all planes, irrespective of their slants and tilts, are parallel to each other and to the viewing direction. This is the geometrical property alluded to in the discussion.

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 33 2004

www.perceptionweb.com