Chapter 2

Perception of Pictures

One of the goals of image synthesis is to achieve the *realism* in pictures. *Realism* is hard to define, although most people have an intuitive notion of realism. We will call this intuitive notion *perceptual realism*. For computer graphics applications a precise definition is required. The most common solution is to define realism as equivalent to photorealism. Photographs are considered to be a standard of truth in images and the modeling of an idealized photographic process becomes the main task of the image synthesis. This solution has a number of advantages, but inevitably suffers from a fundamental flaw: the image synthesis algorithms are constructed and evaluated as models of a photographic process, while the output of these algorithms (pictures) is evaluated primarily by a perceptual process. If we use this approach, the perceptual quality of computer-generated images is inherently limited by that of photographs.

It is well known that an arbitrary photograph does not necessarily look good. Although photographers use some types of perceptual distortions occurring in photographs to achieve various artistic effects, in many cases distortions are undesirable and must be avoided. Avoiding these distortions by purely photographic means (choice of the viewing point or of the field of view) results in restrictions on the scenes that can be realistically reproduced in photographs. Computer generation of images is more flexible, and instead of faithfully imitating the photographic process, we can try to use perceptual principles directly: instead of modeling a camera we can redefine our working concept of realism formalizing some facts about picture perception. Moreover, photographic images can be processed digitally and we can eliminate distortion in already exiting images.

These goals require a more precise definition of the class of pictures that we are going to consider –after all, any pattern on a flat surface can be a picture. Section 2.1.2 describes the class of pictures in which we are interested.

There are three main sources of information about picture perception available to us: psychophysical research, art history and theory, and our everyday experience. Quite a few facts about perception appear to be so obvious that nobody ever states them explicitly or bothers to test them experimentally. We will see that some of these obvious facts are quite important (Section 2.1.3.)

There is a considerable difference between our approach to the problems of perception and the approach that is typically used in psychophysical research. Our main goal is not to prove or disprove some particular theory of perception but to try to collect facts about perception that can be used for making better pictures. We wish to avoid relying on any particular theory as much as possible. We cannot avoid making some theoretical assumptions altogether; we describe our assumptions in Section 2.1.

Most experimental studies consider linear perspective images, but their results apply to any projection producing similar images. Whenever possible, we will emphasize two-dimensional characteristics of images, rather than the characteristics of the central projection used for their construction. Then we can extend our conclusions to any mappings from three-dimensional space into the plane.

Linear perspective nevertheless is extremely important: it doesn't produce any distortion in most cases and our main practical application is correction of distortions in the cases in which it does. In Section 2.5 we discuss linear perspective in greater detail.

2.1 Assumptions and Definitions

2.1.1 Perceptual Nomenclature

One of the problems with using data from the psychological literature is the lack of consistent and clear terminology. We attempted to define and use consistently a limited number of terms for frequently-used concepts and quantities. Some of our definitions are nonstandard, such as parallel and perpendicular foreshortening. Our motivation in these cases was to create an unambiguous and concise way to describe a particular concept.

The word *quantity* here refers to numerical characteristics of three-dimensional objects and scenes represented in a picture, such as slant, size, distance or orientation. The numerical value of a *quantity* can be:

depicted A *depicted* quantity (size, slant, distance etc.) is the actual value for the object or scene represented in the picture.

apparent An *apparent* quantity is the value as perceived or deduced from the picture. Apparent quantity cannot be measureed directly — it can be estimated from the *judged quantity* or from some other response (such as ball-throwing in [Smith, 1961].)

judged A *judged quantity* is the value as reported by a subject; in a good experiment it is determined by the apparent measure, but can be biased by the mechanism of reporting or by the specified task. This bias, however, should be the same for identical apparent quantities.

The totality of apparent quantities can be thought of as the *perceptual space*. The *perceptual space* need not possess a consistent geometry in mathematical sense: the value of an apparent quantity can depend on the task, attention focus and many other factors. In most cases, the *perceptual space* of a picture is not connected to the real space around the observer. A possible reason for this is the scalability of the pictures: in most cases familiar objects in the pictures are smaller or larger than their real size, and they cannot be directly placed in the real space. There are several notable exceptions: the perceptual space of *trompe-l'oeil* pictures [Milman, 1986] is merged with the real space; the *orientation* of objects in the perceptual space can be judged with respect to the viewer.

Most names of quantities (e.g. size, slant, distance) are unambiguous. The definition of *orientation* requires some comments.

orientation We shall distinguish between two type of *orientation*: relative to the viewer and relative to the other objects in the picture. We can estimate the angle between the viewing direction and some direction specified by an elongated object in a picture. We shall call this angle orientation with respect to the viewer. It is also possible to estimate the angle between two directions in a picture. We shall call this angle *relative orientation.*

Information about the geometry of three-dimensional objects is partly contained in the two-dimensional geometry of images. Most terms that we use to specify the two-dimensional geometry of images are just general geometric terms (e.g. angles, lines, curves, areas etc.) A specific characteristic of images which is inherently two-dimensional, but cannot be described without referring to the three-dimensional object is *foreshortening*. For perspective projections we shall distinguish two types of foreshortening: *parallel* and *perpendicular*.

foreshortening If the size of the image of an object depends on its position in the scene, we shall call this effect *foreshortening*.

parallel foreshortening The decrease in the size of the image of an object when it is placed at increasing distances from the picture plane. It is also called *perspective convergence*.

perpendicular foreshortening Suppose it is possible to destinguish a feature of the object perpendicular to the projection plane and another feature which is parallel to the projection plane, such as faces in a cube or the lateral and frontal side of the head. The difference between the ratio of the sizes of these features and the ratio of the sizes of their images is the *perpendicular foreshortening*.

It is possible to extend these definitions to the case of nonperspective projections, if we can somehow define the distance to the projection plane (which need not coincide with the physical distance).

It is important to make a clear distinction between several related perceptual phenomena:

distortion Some images of familiar objects can be identified as *distorted*, if perceptual information in the picture is sufficient to identify the object with a high degree of confidence, yet some part of this information results in conclusions about the object that contradict experience. Returning to the example in Figure 1.1, one is reasonably confident that the depicted object is a sphere, because a sphere is more likely to be a part of architectural decoration than a tilted ellipsoid. At the same time, the shape of the image suggests that the object is not spherical, which results in a contradiction.

misperception Some images of objects can convey incorrect information about the objects, without causing distortion. If an object is part of a class of objects which vary in size, some images result in an apparent size which is close to the actual size or shape, while others will result in a different apparent size. We will call the later case *misperception.* For example, the right monitor in Figure 2.1 appears to have an aspect ratio significantly greater than the actual one (no more than 1:1.4) and than that of the left monitor (the monitors are known to be identical.)

illusion Some pictures under special conditions can be mistaken for the real scene that they represent. We shall call this condition *illusion.*

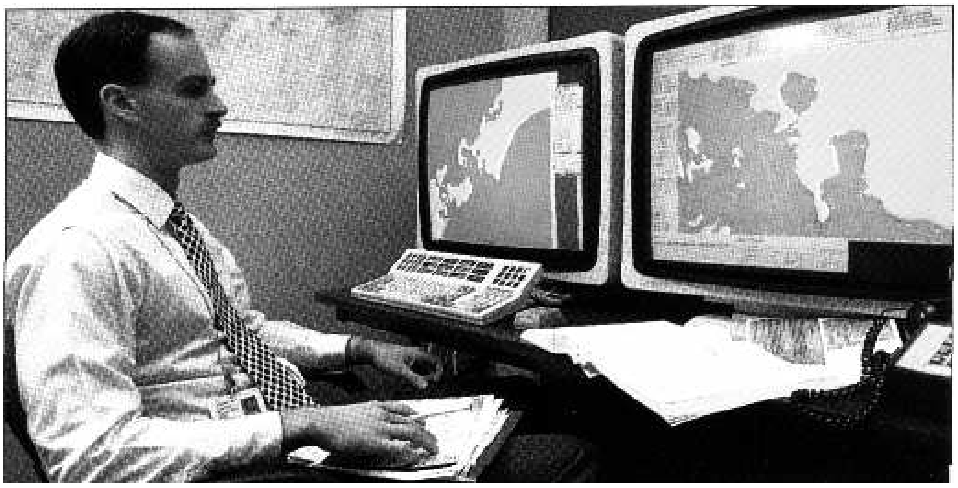


Figure 2.1: Photo from the article "Navigating Close to Shore" by Dave Dooling ("IEEE Spectrum", Dec. 1994), © 1994 IEEE, photo by Intergraph Corp. 92° viewing angle. The two monitors have the same size. But the left one appears to be wider (misperception). □

2.1.2 Defining Pictures

This section defines more precisely what we mean by a picture. It is also important to describe assumptions about viewing conditions –the perception of a picture depends on the position of the viewer with respect to the picture, presence or absence of apertures, frames, etc.

In this thesis our main object of study will be flat or nearly flat pictures on paper, a projection screen, or any other flat surface, observed binocularly, without any restrictions on the position of the head and without any special devices. Pictures of this type include book illustrations, photos, posters, screen projections of slides, motion pictures and pictures on computer displays. Excluded are stereograms of all types, pictures that are designed for observation through a fixed small aperture, pictures in head-mounted displays and anamorphic pictures.

Further, we are interested in pictures that are representations of three-dimensional objects. There are many different types of representation, from purely symbolic, like a verbal description or a *kanji* character, to a highly realistic photographic image which, when viewed from a correct position monocularly, can be confused for a real object. Linear perspective is used to some extent in most pictures of the latter type.

Some authors argued that essentially all forms of pictorial representation are based on convention [Arnheim, 1954], [Goodman, 1976]. For example, it has sometimes been claimed that linear perspective is a matter of convention and perspective images can be understood only within the context of a particular culture. It appears that such radical approach cannot be accepted. Cross-cultural [Deregowski, 1991], developmental [Hagen, 1976a], [Hagen and Jones, 1978] studies contain indications that no culture-dependent learning is required for adequate perception of perspective images. The existence of various phenomena in perception of pictures that we will discuss below, also demonstrate that there is a fundamental difference between the perception of most pictures and the perception of purely conventional representations, such as text.

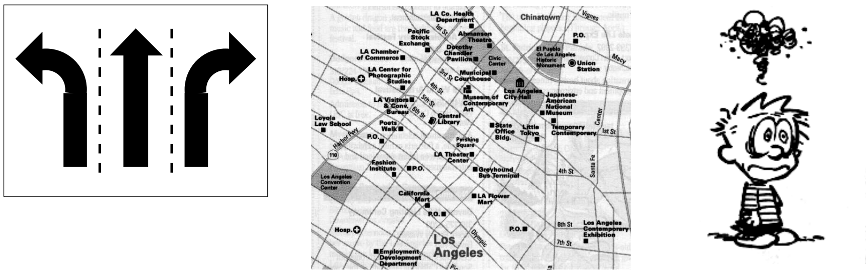


Figure 2.2: Examples of highly conventional pictures: a road sign,a map, a cartoon. □

The presence or absence of a conventional component in a picture is difficult to determine. In some cases (Figure 2.2, a) it is easy to classify a drawing as almost purely conventional. We will be most interested in drawings that don't use a symbolic method of representation or use it in a very limited way. This includes a wide range of pictures (Figure 2.3.)

2.1.3 Assumptions about Perception

We will try to make as few assumptions about perception as possible. We have, however, to assume that pictorial representation of information is mostly nonsymbolic. Our assumptions can be most clearly formulated in terms of retinal images and internal representations.

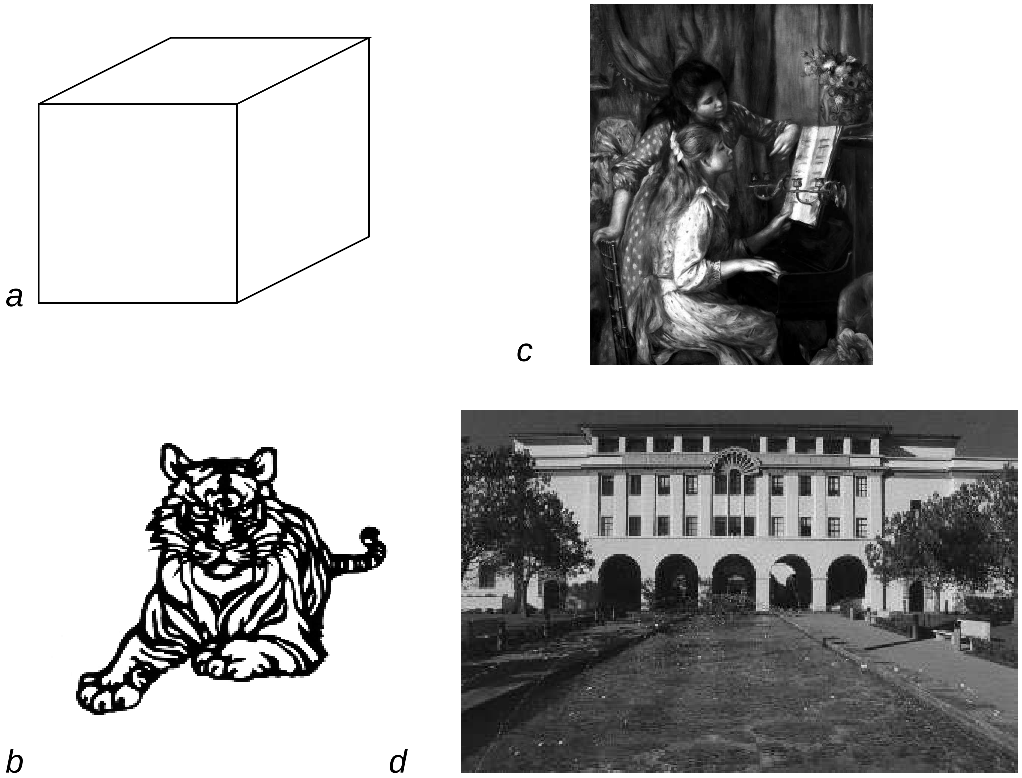


Figure 2.3: Examples of pictures that use symbols in a limited way, or don't use them: line drawings, a painting ("Piano lesson" by A. Renoir), a photograph (Beckman Institute at Caltech.) □

Our ability to recognize objects and textures that we have seen before is crucial both to perception of the three-dimensional world and pictures. We have to assume

existence of some internal representation of objects, and a process that compares information extracted from the retinal image with this representation.

All the visual information about our environment is contained in the images formed on the retina of the eye.

Therefore,

our internal representation of three-dimensional objects should be based on the information contained in these images.

We don't know what part of the information contained in retinal images is utilized in this internal representation.

As retinal images are two-dimensional central projections, there is one-to-one correspondence between a retinal image of an object and a projection of the object onto a plane –a linear perspective picture of an object. Therefore, we can consider geometry of perspective projections instead of geometry of retinal images.

We can distinguish at least two types of geometric information in retinal images: structural and nonstructural.

Structural features have more qualitative nature and are the easiest to detect. Examples of structural information are the number of holes in an image of an object, dimension of the image or its parts, the number of edges meeting at a vertex (Figure 2.4.) Mathematically this information corresponds to the topological properties of the image. We assume that

structural information about projections of an object is present in the internal representation of the object. If an image of an object has structural features that don't match those of any perspective projection, it is perceived as contradictory, distorted or it is not recognized at all.

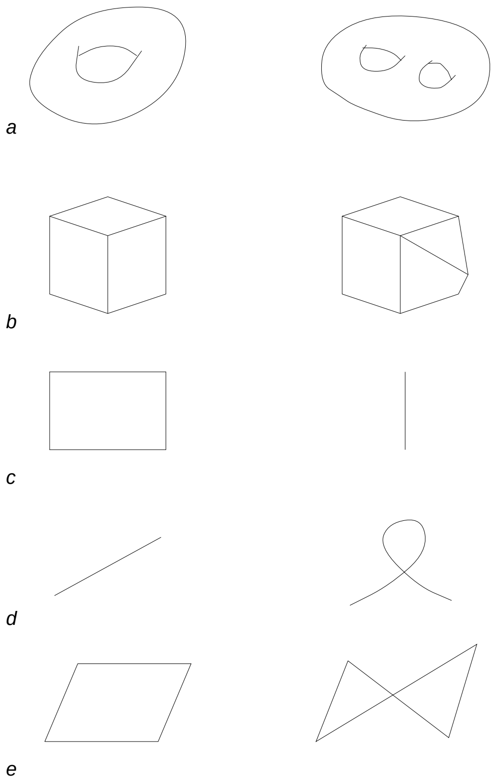


Figure 2.4: Examples of structural differences in images: a) the number of holes b) number of edges meeting at a vertex c) dimension d) number of self-intersections, local structure e) number of intersections of edges, local structure □

We should point out that our assumptions about the importance of structural features are mostly motivated by common sense and intuitive ideas about perception due to the lack of experimental data. It is also important to note that the precision of the visual system is finite [Finke and Kurtzman, 1981] and the above statement is an idealization - for example, if a hole in an image is too small, it is not detected by the visual system.

Nonstructural features typically can vary without causing any significant change in perception. Examples of nonstructural features are the degree of convergence, foreshortening, gradients of texture, angles and curvature of edges (Figure 2.5.) The important difference from structural features is that variation in nonstructural features can be registered by the visual system without producing considerable change in perception.

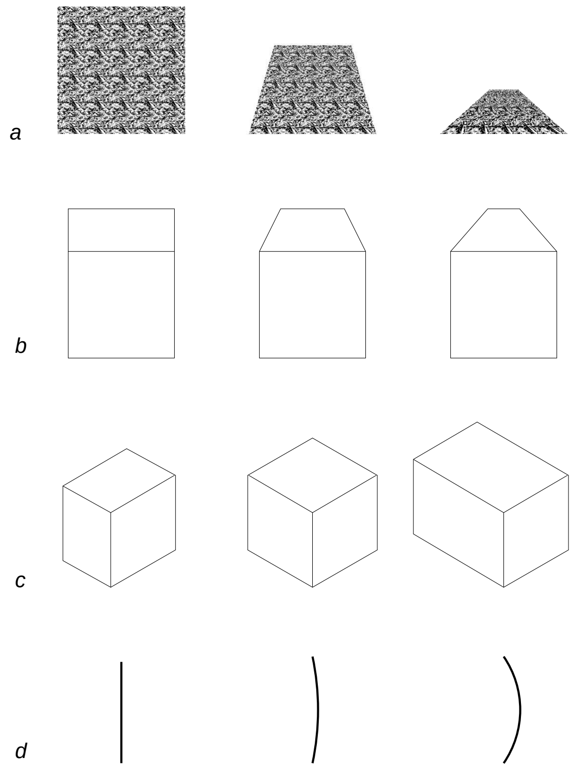


Figure 2.5: Examples of nonstructural differences in images: a) texture gradients b) parallel foreshortening c) perpendicular foreshortening d) curvature □

2.2 Robustness of Pictures

The term *robustness of pictures* was introduced by [Kubovy, 1986]. The perception of a picture typically doesn't depend on the viewing point - we can walk past a painting, tilt a book, move our head away, closer, or to the side of a computer display, and the objects in the picture are unlikely to change their shape or position. This is one of the most important properties of pictures; were it not true, a viewing apparatus or correct choice of viewing position would be required for correct perception. Robustness is not total — some perceptual variables are less robust then others.

A number of particular perceptual variables were confirmed to be independent of the viewing point.

Previous studies [Rosinski and Faber, 1980] [Rosinski et al., 1980] showed that for normal viewing conditions perception of slant doesn't depend on viewing position. This study was extended by [Halloran, 1989] who showed that for extremely oblique viewing directions with the angle between the viewing direction and the picture surface less then $20°$, robustness of the perception of slant breaks down.

There is some evidence [Hagen, 1976b][Smith, 1958b][Rosinski and Faber, 1980, Experiment 1], that relative size judgements are robust as well; only two viewing directions ($90°$ and $45°$) were used in the first study, and only the viewing distance was varied in the last two.

Robustness is very important for moving pictures: for example, inverse perspective constructions, such as described in [Deregowski, 1989] and [La Gournerie, 1859], applied to the images of rotating bodies might result in non-rigid objects if the projection point used for reconstruction is different from the actual projection point. However, this doesn't happen in most cases. As it was shown by [Cutting, 1987] for images of rotating near-rectangular solids, there is a general tendency to ignore small deformations, especially if they preserve angles between edges, and that rigidity is preserved at least up to a $45°$ viewing direction.

Relative orientations and layout are quite robust, ([Halloran, 1989, Experiment 5], [Goldstein, 1979, Experiment 4], [Goldstein, 1987, Experiment 1]) although relative distances between objects in the direction perpendicular to the picture plane depend somewhat on the viewing distance (see discussion below).

Not all perceptual variables are robust, and the range of viewing points within which robustness is preserved might depend on the depicted object.

We have already mentioned that relative distance between objects in the direction perpendicular to the picture plane is not very robust. Experiments show that the viewing position affects judgements of relative distance [Smith, 1958b][Smith, 1958a], although the effect is less than predicted by geometric reconstruction. It should be noted however, that reduced viewing conditions were used in these studies (monocular viewing through a peephole) and the absence of robustness can be attributed to the lack of information about the surface of the picture (see below).

The orientation of the objects with respect to the observer is not robust at all in some cases. It is the most apparent nonrobust perceptual variable: it was noted by many authors that objects pointing perpendicular to the plane of the picture such as gun barrels or fingers (Figure 2.6) appear to be following the observer as he moves past the picture. Portraits often appear to be following the observer with their eyes. It turns out [Goldstein, 1979][Goldstein, 1987] [Halloran, 1989] that this effect is more pronounced for orientations close to the perpendicular to the picture and decreases with the angle of orientation. This results in the following perceptual paradox: while spatial layout is perceived as more or less invariant, orientations of different objects change in different ways (for example, the rut in the road in Figure A.8 doesn't rotate much, while the direction of the road rotates considerably. Thus the difference between apparent orientations changes, while the relative orientation of the rut and the road doesn't change. This is an example of a more general phenomenon: the structure of perceived space need not be consistent; different perceptual mechanisms might produce contradictory information without causing any perceptual problems.

Robustness of perception may depend on the contents of the picture. For example, parallel projections of rotating rigid parallelepipeds are perceived as rigid in a wider range of viewing directions than central projections [Cutting, 1987, Experiments 1 and 2].

Robustness of pictures is related to the dual nature of pictures –we perceive both the picture surface and the three-dimensional scene represented in this surface. Surface texture, flatness, and visible frame are important factors in perception of pictures. Numerous experiments provide evidence that the absence of these factors results in a decrease in the robustness.

Perception of slant becomes completely unrobust when all the information about the picture surface is removed [Rosinski and Faber, 1980][Rosinski et al., 1980]. In unpublished experiments by W. Purdy [Lumsden, 1980] only the picture frame was absent and the information about the picture surface was not completely removed. Still, for the simple pictures used in the experiments (slanted striped surfaces) picture frame removal was sufficient for almost total suppression of robustness.

Figure 2.6: A World War I poster. (from [Taylor, 1963].) The finger appears to be following the viewer when he walks past the picture. Similar posters were created in the US ("Uncle Sam wants you!") and in Russia ("Did YOU volunteer?") during times when the government felt it necessary to point fingers at each citizen. □

2.3 Perception of Objects in Pictures.

The distortions that we mentioned in Chapter 1 are associated with images of particular objects. Two groups of objects which frequently appear to be distorted in photographs [Kubovy, 1986]. The first group included objects with rectangular three-dimensional corners (i.e. corners formed by three edges with all angles between them equal to 90°.) The second group included spherical and cylindrical objects and humans.

2.3.1 Rectangular Objects

Perception of pictures of rectangular corners is well described by simple rules [Perkins, 1972][Perkins, 1973] [Shepard, 1981].

These basic facts about perception of rectangular corners are usually called Perkins' laws. We will call three line segments meeting at a point a *three-star* [Perkins, 1968]. We will call an image of a rectangular corner *two-faced* if only two out of three boundary surfaces are visible, and *three-faced* if all three are visible (Figure 2.7.)

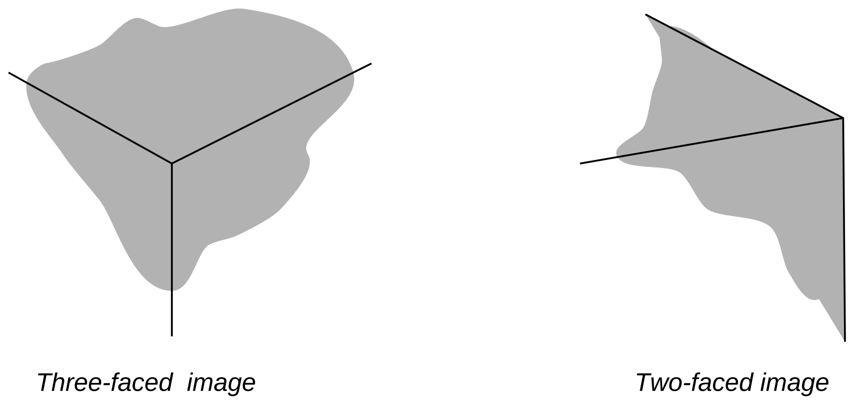


Figure 2.7: Two types of images of rectangular corners □

> **Perkins' first law**. A three-star is acceptable as a two-faced image of a rectangular corner if and only if it contains two angles less than or equal to 90°, whose sum is greater then 90°.

> **Perkins' second law**. A three-star is acceptable as a three-faced image of a rectangular corner if and only if all three angles are greater than 90°.

Small deviations from Perkins' laws can occur, but in general there is a good agreement between the experiment and theory [Perkins, 1972], [Shepard, 1981].

It is important to note that Perkins' laws have a very simple geometric interpretation: acceptable projections coincide with orthogonal perspective projections. There is no difference between central and parallel projections in this case, because only angles between lines are important. For the central orthogonal projection we have to assume that the principal ray goes through the center of the three-star.

From Perkins' data (Figure A.14) we can make a rough estimate of the deviations from Perkins' laws that don't result in considerable distortion: the three-stars deviating from Perkins' laws by less then 5° still can be perceived as rectangular corners. Of course, this is a rough estimate. We will discuss Perkins' laws in greater detail in Section 2.5.1.

2.3.2 Spheres, Cylinders, Humans and Animals

Perceptual requirements on the images of spheres are remarkably restrictive: only disks are generally accepted as good images of spheres [Pirenne, 1970][Kubovy, 1986]. No data on detection of "non-circularity" were available to us. It appears to be safe to assume that 1.1 aspect ratio is detectable (Figure 2.8) and accept it as a rough upper boundary. Spheres are not that common in real environments, although quite popular in computer graphics images. They are also convenient test objects, because the distortion of shape in the image of a sphere is easy to detect and describe.

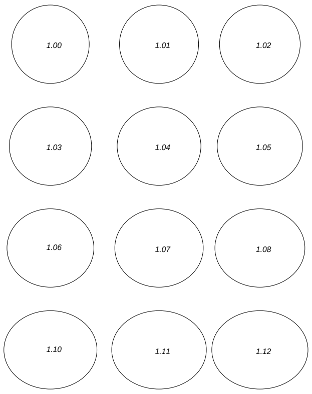


Figure 2.8: Ellipses with different aspect ratios (the aspect ratio is shown in the center of each ellipse.) □

There are two distinct problems associated with images of cylinders, such as columns: the one that is most often mentioned arises when a row of cylinders is depicted. It is not a problem associated with the image of any particular cylinder, but rather a problem of unacceptable relative sizes of images. We will discuss it in the Section 2.5.2.

The second problem can be considered a separate case of a more general class of problems that is associated with axially-symmetric objects.

The image of the foundation of the column or the upper part of a cup or a bowl is generally acceptable if it is an ellipse with major axis oriented perpendicular to the axis of symmetry (cf. Section 2.5.2, Figure 2.11).

Our perception of humans is likely to be very specialized. Humans come in many shapes, colors and sizes, so the result of perception of an inadequate image is more often a misperception (a false conclusion about the person in the picture) rather than direct perception of distortion as with spheres. If we stretch or shrink the image vertically or horizontally within a wide range, these changes are likely to produce acceptable although sometimes misleading pictures. For example, the men in Figure 2.9 are all identical but those close to the edges of the picture appear to be quite different from those in the middle. The part of the body which is most affected by deformation is the head, due to less variation in the shape of the head between individuals compared to other body parts.

Figure 2.9: Deformations of human figures in a wide-angle perspective picture: 140° horizontally. □

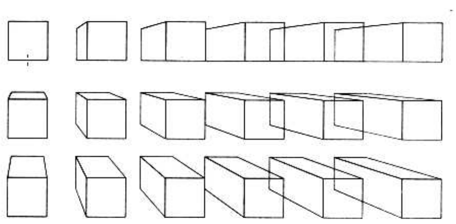
There are types of distortion that are more likely to produce deformed, not just misleading images. An important feature of acceptable frontal pictures of humans is their axial symmetry. When this symmetry is broken, the image becomes unacceptable.

Similar conclusions can be reached about images of animals, although larger deformations are tolerated.

2.3.3 Foreshortening

We consider two types of foreshortening - parallel and perpendicular. Our terms here are not quite standard: we believe it is useful to destinguish betweeen these two types of foreshortening.

For rectangular solids the preferred amount of perpendicular foreshortening depends on the aspect ratio [Nicholls and Kennedy, 1993b]. it is almost constant for cubes: the preferred amount was close to 1:0.6–1:0.7 for all viewing conditions. It should be noted that the use of line drawings in this study could considerably affect the results: perception of the whole object could be important in this study, the absence of any distinctive features in the drawing, specifying absolute size (markings



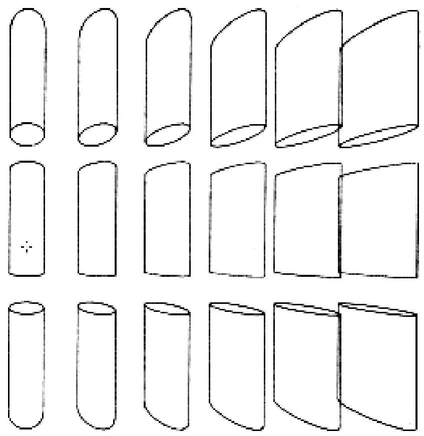


Figure 2.10: (Left) projections of cubes from [Pirenne, 1970] □

Figure 2.11: (Right) projections of cylinders from [Pirenne, 1970] □

on a match box or windows on a building) could affect the preferred amount of foreshortening in an unpredictable way.

The amount of perpendicular foreshortening deviating significantly from the preferred one results in distortion if it is known that the depicted objects are supposed to be cubes, and in misperception if no such information is available. (Figure 2.10.)

Perpendicular foreshortening is easy to define only for objects with distinct edges parallel and perpendicular to the picture plane; it can be also defined for familiar objects with distinctive features parallel or perpendicular to the picture plane. Some distortions of the human figures can be described in terms of the perpendicular foreshortening ratio.

Parallel foreshortening across objects can vary in a wide range without causing perceptual problems. In perspective images the ratio of parallel foreshortening is directly related to the viewing angle: the greater the viewing angle, the greater the ratio for an object of a given size. [Nicholls and Kennedy, 1993a] studied the effects of the change in the viewing angle on the perception of cubes; it was found that a moderate degree of parallel foreshortening (approximately 1:0.75) was consistently preferred independent of the viewing conditions.

2.4 Other Properties of Visual Perception

2.4.1 Visual Field

The total size of the visual field is quite large; it extends more than 160° horizontally and 150° vertically for each eye [Carterette and Friedman, 1975]. However, the density of the receptors is very nonuniform, and maximal resolution is achieved only in the fovea of the eye which has an angular size of only about 2°. Lateral vision is very limited in resolution; its main function is presumably motion detection.

As the resolution decreases, less and less detail and precision is available to the rest of the visual processing system. [Finke and Kurtzman, 1981] found that the size of the field where gratings 1° apart can be resolved is approximately 30°. For example, the angular size of the moon is 0.5°. Outside 30° field of view we cannot detect any details of the lunar surface and cannot even tell if the moon is round or not.

This also means that only low-resolution information in the retinal projections of the objects that are far enough off the viewing direction can be incorporated into the internal representation. The same study [Finke and Kurtzman, 1981] provides some evidence that indeed the resolution of the "internal imagery" corresponds to the resolution of the actual visual field.

2.4.2 Perception of Straightness and Curvature

Intuitively it is clear that the images of straight edges and lines should be straight. It is important to know the accuracy of the perception of straightness. Experimental studies that we know about considered the perception of straightness only for very short lines with angular size close to 1°. It was found that for small perturbations perceptual threshold is determined by the solid angle subtended by the maximal bump with respect to the least-squares straight line (Figure 2.12) [Watt et al., 1987]. The threshold value was determined to be close to 0.3 sq arc min.

In [Ogilvie and Daicar, 1967] a similar measure was used although a chord rather than a least-squares line was considered and there was only one bump. Resulting thresholds had the same order of magnitude (0.4–2.2 sq. arc. min.) These values are remarkably small: [Watt et al., 1987] points out that the receptor density is not sufficient for the detection of perturbations of that size and some type of higher-level processing should be involved to account for these thresholds (hyperacuity.) It appears that this threshold is likely to increase with the angular size of the line. However, it is bound to stay quite low and any considerable deviation from straightness will cause perceptual distortion.

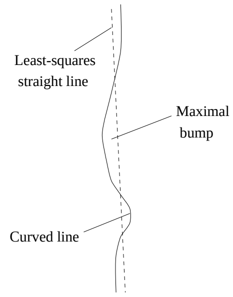


Figure 2.12: Watt's measure of perceptual curvature □

2.4.3 Perception of Verticality

The perception of verticality is influenced by two main factors: vestibular perception of the force of gravity and visual perception. Objects that we assume to be vertical (walls, trees, etc.) or horizontal (the surface of the ground) determine the visual vertical. A number of experiments ([Ong and Kessinger, 1971], [DiLorenzo and Rock, 1982]) show that when there is a conflict between vestibular and visual information, there is no clear preference for the objective vertical determined by the vestibular system. The "rod-and-frame" effect ([DiLorenzo and Rock, 1982]) indicates that a visible frame affects perception of actual vertical: when a vertical rod is viewed inside a tilted frame it appears to be tilted in the direction opposite to the direction of the frame.

In pictures, lines that are not parallel to the edges of a rectangular frame are perceived as non-vertical. If these lines represent vertical or horizontal edges of objects (buildings or furniture), the picture creates a feeling of instability which is sometimes undesirable. The "rod-and-frame" effect explains to some extent the origin of this feeling.

2.5 Linear Perspective

Linear perspective was introduced as a systematic tool during the Renaissance. To some extent linear perspective was known to the Romans (frescoes of Pompeii) and Chinese (see discussion of other perspective systems in Section 2.6) but during the Renaissance it was introduced as a rigorous system which was considered to be the foundation of painting and drawing till the second half of the nineteenth century. The first artist to use perspective during the Renaissance was Filippo Bruneleschi, but the first written description and analysis was done by Leon Alberti in [Alberti, 1976] and Leonardo da Vinci in [da Vinci, 1970]. He was also the first to observe the limitations of the method. As it is discussed in [Kubovy, 1986], Leonardo was not convinced that perspective pictures are robust, hence his overly stringent requirements to the application of perspective ("Leonardo's rule"):

> ...do not trouble yourself about representing anything, unless you take your view-point at a distance of at least twenty times the maximum width and height of the thing that you represent; and this will satisfy every beholder who places himself in front of the work at any angle whatsoever.

This means that angular size of any object in a picture should be less than 3° — clearly an overkill. Leonardo offers the following example of perspective distortion (Figure 2.13): in the image of a row of columns depicted from a close projection point, the width of the images of columns closer to the edges of a picture would increase, which is quite counterintuitive.

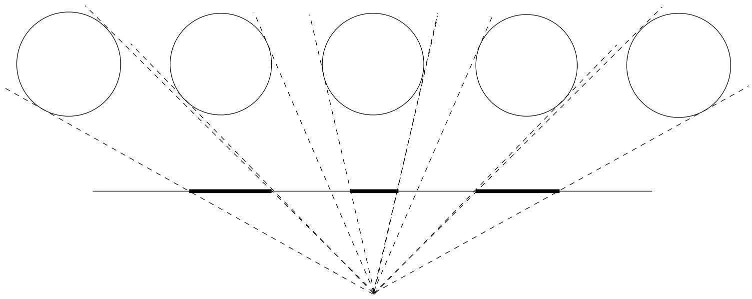


Figure 2.13: Leonardo's example of a perspective distortion □

Statements similar to Leonardo's rule are quite common. For example, [Glaeser, 1994] does not recommend using viewing angles above 40–50°. [Olmer, 1949] recommends viewing angles no more than 37° horizontally and 28° vertically.

Thus, the limitation of linear perspective were recognized simultaneously with its discovery. As it is pointed out by [Kubovy, 1986] it is not correct to identify artistic linear perspective with central projection of a halfspace onto a plane: only a limited range of viewing angles is used, and, as we will see, the images of many objects are often painted with deviations from linear perspective.

In photographs and computer generated images linear perspective is practically identical to central projection (exact linear perspective): wide-angle images are not uncommon, and, due to the physics of the process or nature of the algorithms, resulting images follow the rules of the central projection exactly. Intuitive compensation used by the artists is not available in this case. Two methods are helpful in the analysis of exact linear perspective: first, we can check how well linear perspective images satisfy various perceptually desirable requirements described in the previous sections; second, we can examine deliberate deviations from linear perspective that are common in painting.

Structural features. Linear perspective images have no undesirable structural features (see Section 2.1.3.)

Robustness. The studies of robustness were done using perspective images. Summarizing their results, we can say that under normal viewing conditions perspective images are robust when the viewing angle is sufficiently small or the viewing direction is not too oblique. In general robustness is a concern only for moving images.

2.5.1 Rectangular Objects in Linear Perspective

In perspective images Perkins' laws are always violated to some extent, with the only exception of orthogonal parallel projection. The extent of this violation depends for the parallel projection on the direction of projection and for the central projection on the direction from the center to the apex of the rectangular corner. We will call both directions directions of projection.

We have selected two parameters to characterize the violation of Perkins' laws: fraction of rectangular corners with projections violating Perkins laws for a given direction of projection and maximal deviation of the *offending angle* from 90°.

There are infinitely many possible positions of rectangular corners in space, and each of them can be represented by three angles: two angles θ and ψ determining the orientation of one of the edges, and the angle of rotation ϕ of the remaining pair of edges around the first one.

Define a triple (θ, ψ, ϕ) to be "good" if the projection of the corner in the given direction satisfies Perkins laws.

Then we can define the fraction of "good" projections to be

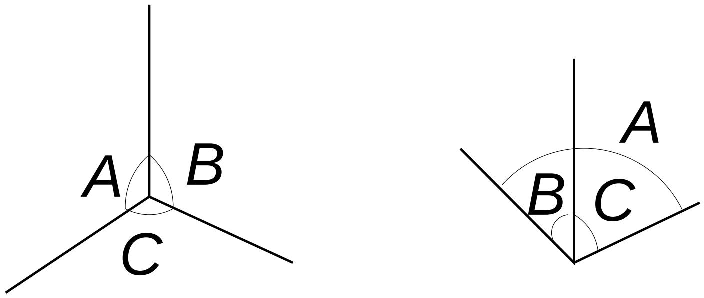


Figure 2.14: Three-sided and two-sided images of a rectangular corner □

$$\frac{\int_{\text{all "good" } (\theta,\psi,\phi)} d\theta d\psi d\phi}{\int_{\text{all } (\theta,\psi,\phi)} d\theta d\psi d\phi} = \frac{1}{4\pi^3}\int_{\text{all "good" } (\theta,\psi,\phi)} d\theta d\psi d\phi$$

We define the offending angle in the following way. Let A, B, C be the angles between the edges of the three-star (always less then 180°.) Let $A \geq B \geq C$. Then there are two possible violations of Perkins' laws: either all three angles are less than 90°, or only one of them is less than 90° (C). In the first case $A = B + C$, and A should be greater than 90°, while B and C can be less than 90°. In this case we define A (the largest angle) to be offending angle. In the second case, there are two possibilities: $A = B + C$ or $A = 360° - B - C$. In the first case B should be less than 90° (second Perkins law applies) and we define it to be the offending angle. In the second case all angles should be greater than 90° and we define C to be the offending angle.

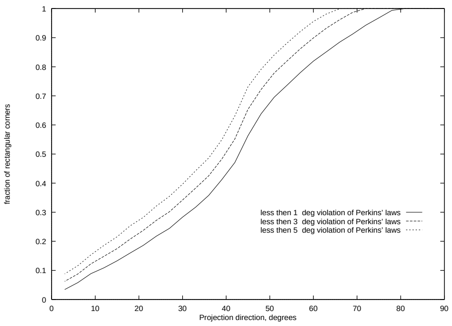
It can be easily shown that our definition is equivalent to the following simpler one:

> the **offending angle** is the angle between edges of the three-star which is the closest to 90°.

Figure 2.15 shows the fraction of "good" corners as a function of the angle between the projection direction and the picture plane (Monte-Carlo evaluation). It also shows the fractions of "almost good" corners - those with deviations less than 1, 3, 5 degrees.

Figure 2.16 shows the maximal offending angle as a function of the angle between the projection direction and the picture plane.

We can see that all factors are negligibly small for small angles between the projection direction and the projection plane. From [Perkins, 1972] we can see that deviations of approximately 5 degrees are still tolerated in almost 50% of the cases. Therefore, in most cases pictures with projection angles that are small enough not to create distortions of more than 5° can be considered acceptable.



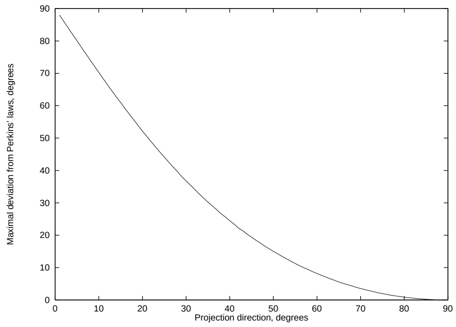


Figure 2.15: (Left) Fraction of "good" rectangular corners *vs* projection angle □

Figure 2.16: (Right) Maximal offending angle *vs* projection angle □

[Hagen and Jones, 1978], and [Nicholls and Kennedy, 1993a] examined preferred parallel foreshortening for parallelepipeds and other types of prisms. Again, the general trend of their results is in the direction of lesser projection angles, although [Nicholls and Kennedy, 1993a] argues that a moderate degree of parallel foreshortening is preferred to parallel projection, while [Hagen and Jones, 1978] claims that parallel projection has the highest rating. For reasons explained in our comments to [Hagen and Jones, 1978] in Appendix A we believe that the former paper is more reliable.

The perpendicular foreshortening in perspective images of rectangles might also create problems (Figure 2.10.) Comparing to the results of [Nicholls and Kennedy, 1993b], we can see that the ratio of foreshortening can be quite far from the perceptually acceptable range. It should be noted that cubes are not that common and much wider range of foreshortening ratios can be tolerated for arbitrary parallelepipeds, unless this is something familiar like a computer terminal or a TV.

2.5.2 Spheres, Cylinders, Humans

All nonorthogonal projections of spheres are ellipses. The aspect ratio of these ellipses is a useful measure of deviation of the image from perceptually acceptable. Figure 2.17 shows the aspect ratio of the image of a small sphere as a function of the projection direction. Assuming 1.1 upper bound for the acceptable aspect ratio, we can see that the projection direction should not be less than 75-77° (projection angle 45-50°.)

Images of spheres are not that common in painting. When they are present, they are depicted as circles regardless of their position. The most famous example of this type is "The School of Athens" by Rafael (Figure 2.18) [Pirenne, 1970][La Gournerie, 1859].

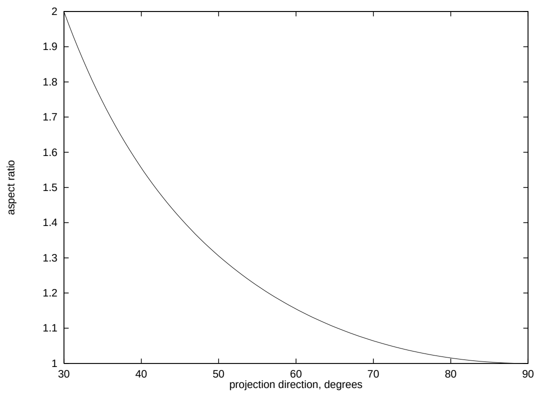


Figure 2.17: Aspect ratio of the image of a small sphere *vs* projection direction angle. □

In the beginning of this chapter we pointed out the problem with rows of cylinders that is specific to the linear perspective. Another, more general problem is the tilt of the axes of the image of a cross-section (Figure 2.11): major axis of the horizontal cross-section is always oriented towards the center of the image. The most apparent distortions occur in the images close to the lines through the center of the picture tilted at 45°.

Images of humans in exact linear perspective may look grossly distorted because of extreme changes of the aspect ratios and violations of symmetry (Figure 5.5a.) No such asymmetry is observed in the art: human bodies are typically drawn as if the projection point was located directly in front of them.

Another problem, which is well-known to photographers, is the distortion of the features of a face resulting from high degrees of perpendicular foreshortening.

2.5.3 Straightness, Verticality, Texture Density, Movement

Linear perspective images of straight lines are straight. Linear perspective is the only projection that has this property [Klein, 1939].

A number of problems are associated with the perception of verticality in linear perspective images. Unless the picture plane is perpendicular to the ground, vertical lines converge or diverge. In some cases (Figure 2.20) this divergence creates desirable artistic effect, in other cases (Figure 2.19) it is undesirable. Even small tilts of the picture plane create considerable distortion in wide-angle images partly due to the violation of the Perkins laws and partly to the divergence of the vertical line and the conflict with the frame of the picture.

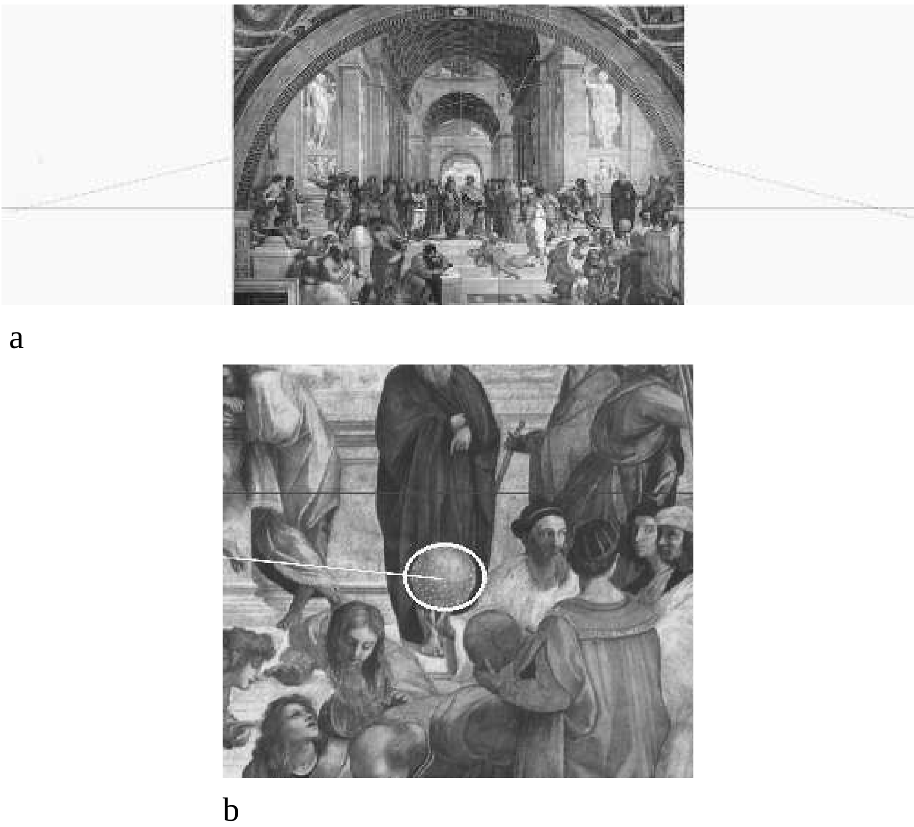


Figure 2.18: 'The School of Athens" by Rafael. a. General view; b. Detail; compare the image of the sphere as painted by Rafael with reconstruction of the central projection of the sphere (position of the center of projection were determined from the images of the columns and walls.) □

Another problem is related to the change in the texture density - texture elements of identical size have more extended projections when they are close to the edges of the image. This effect can be observed in Figure 2.21: the tree appears to be stretched to the left partly because of the texture density gradient.

For moving linear perspective images (computer animations, cinema) some of the problems associated with oblique and wide-angle projections are amplified.

For example, as it is described in [Cutting, 1987], rotating rectangular solids are more often perceived as non-rigid for oblique projections and wide-angle perspective. Cutting attributes this fact to the greater percentage of frames that don't satisfy Perkins' laws for a given viewing point.

Because of the changes in perpendicular foreshortening, many objects appear to stretch or shrink when they move across the field of view.

The height of remote objects, such as mountains or clouds, above the horizon changes when

Figure 2.19: Two views of New York City from [Feininger, 1953]; it is interesting to note that the author offers the second image as an example of "rather nonorthodox picture far more interesting than the conventional view" □

Figure 2.20: A picture of a building taken with a tilted camera from [Feininger, 1953]; we get the impression of looking up as if we were standing in the street in front of the building □

they move from the periphery of the picture to the center. These changes are quite apparent in wide-angle pictures and in many cases are perceptually undesirable (Figure 2.22,) although correct geometrically.

2.6 Other Traditional Perspective Systems

From the Renaissance until the second half of nineteenth century Western painting and drawing was based on linear perspective. As we have mentioned before, some deviations from its laws were common, and analysis of perception provides some explanations for them. Artistic perspective is mostly a tool for depiction of relative positions and sizes of objects in space; its role in depiction of separate objects is much less important.

Figure 2.21: Beckman Institute courtyard at Caltech, photo by the author □

In the Middle Ages European art was quite advanced, but didn't make use of any projection system in a consistent manner. The same can be said about Byzantine and ancient Russian art. Compositions created by the artists were often quite complicated, and some facts about their organization can be useful to understand. Other cultures produced evolved forms of pictorial art independently.

Perhaps Chinese and Japanese art is the most advanced system different from Western art. Their approach to the depiction of objects might differ in some details but similarities are considerable: the use of color and shading is quite different but geometrically there is little or no difference. Most importantly, the approach to the depiction of space adopted in Oriental art is quite consistent and somewhat different from the concept of linear perspective.

As it was pointed out by many authors (for example, [Hagen, 1986]) Japanese and Chinese artists typically used oblique parallel projection in their drawings. As in Western art, projection was used mostly for depiction of relative positions of objects rather than for separate objects.

Parallel projection is just a special case of perspective projection with the projection point

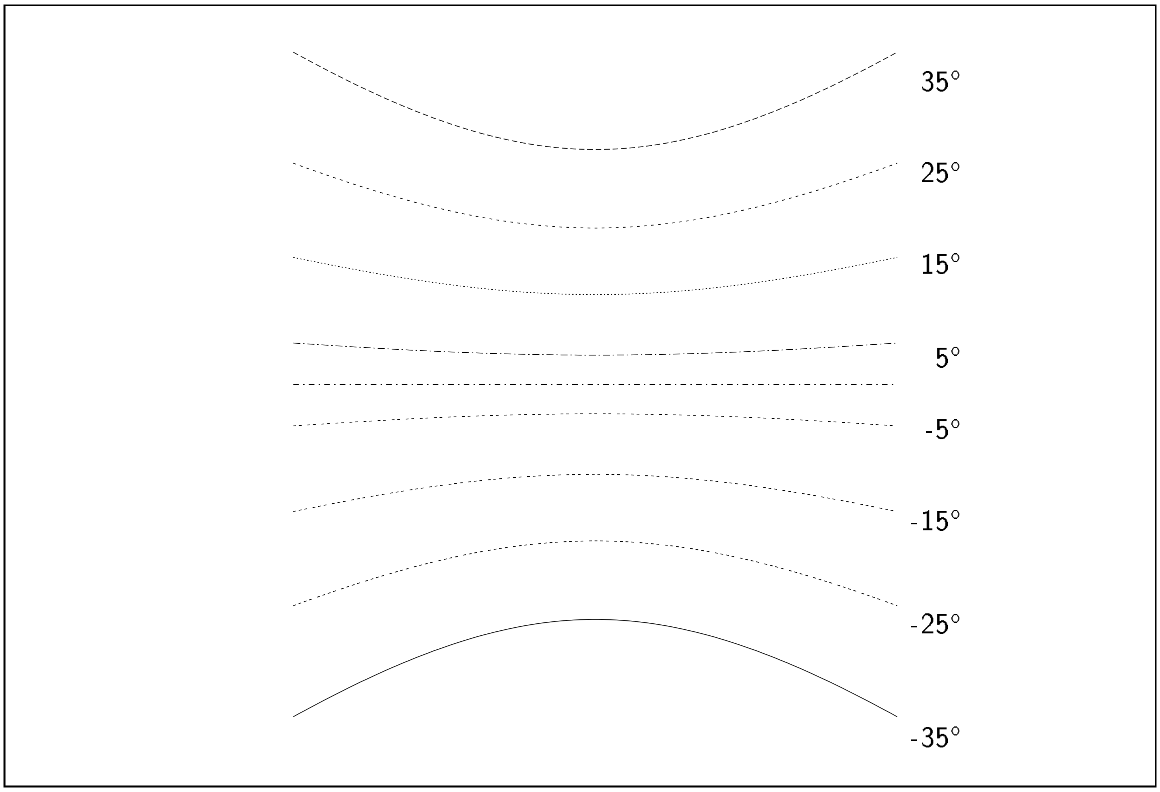


Figure 2.22: Trajectories of points in a wide-angle motion picture when the camera turns. The projection angle is 90° in vertical and horizontal directions. □

removed to infinity. The main problem of this projection is that there is no way to combine images of very large objects (mountains) and small objects (people) in the same picture. This problem was often solved in the following way: parallel projection was used for each part of the picture, but some amount of perspective deminution with distance was introduced for objects located further away. As a smooth transition between these separate areas was impossible, they were separated by symbolic clouds.

In a sense, inside the continuous areas Oriental practice is more consistent than Western. As we have pointed out,the shape of linear perspective images depends upon its position within the image and this fact is typically ignored by the painters. For parallel projection this is no longer true; therefore, a rigorous construction similar to an aerial photograph is possible. For oblique projections, however, it means not the absence of the distortion of shape but a constant distortion. If the angle between the direction of projection and the projection plane is greater than 70-75°, all types of

Figure 2.23: Chang Tse-tuan (c. 1100–1130) *Going Up River at Ch'ing-ming Festival Time*, Detail of a handscroll. Palace Museum, Peking. (from [Sullivan, 1989].) □

distortion are quite small and can be ignored. Of course, this is mostly theoretical observation and no precise geometric constructions were ever utilized.

Oriental artists didn't use linear perspective not because they were competely unfamiliar with it. A painting by Chang-Tse-tuan dating as early as XII century (Figure 2.23), clearly exhibits perspective deminution. [Sullivan, 1989] quotes a Sung dynasty critic Shen Kua who criticized a tenth-century landscapist Li Ch'eng for his use of linear perspective. "Why look at a building, said Shen Kua, from only one point of view? Li Ch'eng's 'angles and corners of buildings' and his 'eaves seen from below' are all very well, but only continually shifting perspective enables us to grasp the whole."

Another interesting aspect of Oriental art is occasional use of divergent perspective for rectangular objects. The same trend can be even more consistently observed in the Byzantine and ancient

Figure 2.24: *Eucharist* (XV century), State Russian Museum (from [Raushenbakh, 1980]) □

Russian art (Figure 2.24.) It cannot be attributed to the lack of skill; other aspects of images demonstrate very high skill and artistic talent. In Byzantine art in particular it became to large extent a matter of convention. But no religious or symbolic basis is firmly established for this particular convention, and instances of such images in Oriental and Persian art demonstrate that it might have some perceptual basis.

[Deregowski and Parker, 1992], [Deregowski et al., 1994] argue that laterally displaced objects (i.e. those located in the periphery of the visual field) are perceived in this way. A number of other explanations were attempted [Raushenbakh, 1986].

This phenomenon is mostly of historical interest, because it is likely that most contemporary people wouldn't find these images to be good representations of rectangular objects. It demonstrates that potentially the conventional element in representation is very significant: it could be that some of these images were perceived as quite acceptable by contemporaries.

The organization of space in Byzantine and ancient Russian paintings is of greater interest: in some sense, the general principle of the representation of space is similar to that of the Japanese art. The scene is subdivided into several parts without connecting elements. Each part of the scene is depicted separately and then they are overlayed. The difference from the Oriental art is in considerable occlusion and much less consistent size relationships.

Summarizing our observations on different artistic cultures that have attempted to depict space rather than separate objects, we can observe that they have at least one point in common: all straight lines are depicted as straight lines; This a major restriction on the images of rectangular objects. Indeed, in all systems parallel or central projection is used. Divergent perspective can be considered as central projection applied "backwards". A common trend can be observed in the depiction of space: the whole composition is divided into several parts, and each of this parts is drawn independently. The projection used for each part is close to parallel. There are quite a few exceptions to the above, but if any system is ever used consistently, it follows this pattern.

It should be noted that other cultures developed even more different systems of pictorial representation. For example, [Hagen, 1986] describes the system that is used in the art of Northwest Coast Indians. These images appear even less realistic than Byzantine art to a modern Western observer and are not particularly relevant to our discussion.

2.7 Summary

The overview of the perception of pictures that we have provided in this chapter will serve as a basis for the construction of projections, allowing us to achieve in some cases perceptually better results than regular linear perspective.

In this summary we repeat the main points of our discussion of perception which are relevant to the task of constructing projections of the three-dimensional world into the picture plane .

- Images of objects should be topologically similar to a linear perspective image of the object.

- To be perceptually acceptable, images of objects should satisfy certain criteria, such as Perkins' laws for rectangular corners, have aspect ratio close to 1 for the images of spheres, bounds on the perpendicular foreshortening ratio for parallelepipeds, cylinders and humans.

- Linear perspective works best for smaller projection angles. Estimates of the critical angle vary, but for projection angles less than $35°$ we are guaranteed that there will be practically no distortion. Within these bounds most of the perceptual criteria mentioned above are satisfied. This fact is confirmed by the artistic practice in different cultures.
- For wider projection angles distortions may occur. They become quite objectionable when the projection angle exceeds $80°$. Some familiar geometric shapes (rectangular corners, spheres, cylinders, human figures) are especially sensitive to such distortions. Distortions are amplified in moving pictures.

Figure 2.25: Funaki Screens, *Rakuchu rakugai za byobu*, 1614-15, right panel. Tokyo National Museum. (from [Mason, 1993]) □