Correction of Geometric Perceptual Distortions in Pictures

Thesis by

Denis N. Zorin

In Partial Fulfillment of the Requirements

for the Degree of

Master of Science

California Institute of Technology

Pasadena, California

1995

Abstract

We suggest an approach for correcting several types of perceived geometric distortions in computer-generated and photographic images. The approach is based on a mathematical formalization of desirable properties of pictures.

We provide a review of perception of pictures and identify perceptually important geometric properties of images.

From a small set of simple assumptions we obtain perceptually preferable projections of three-dimensional space into the plane and show that these projections can be decomposed into a perspective or parallel projection followed by a planar transformation. The decomposition is easily implemented and provides a convenient framework for further analysis of the image mapping.

We prove that two perceptually important properties are incompatible and cannot be satisfied simultaneously. It is impossible to construct a projection such that the images of all lines are straight and the images of all spheres are exact circles. Perceptually preferable tradeoffs between these two types of distortions can depend on the content of the picture. We construct parametric families of projections with parameters representing the relative importance of the perceptual characteristics. By adjusting the settings of the parameters we can minimize the overall distortion of the picture.

It turns out that a simple family of transformations produces results that are sufficiently close to optimal. We implement the proposed transformations and apply them to computer-generated and photographic perspective projection images. Our transformations can considerably reduce distortion in wide-angle motion pictures and computer-generated animations.

Contents

Chapter 1

Introduction

1.1 Motivation

The pinhole photograph from [Pirenne, 1970] (Figure 1.1) demonstrates the main problem that is considered in this thesis. The sphere in the picture appears to be deformed and only the fact that ellipsoidal architectural decorations are very uncommon in ancient buildings makes us assume that the object in the picture is indeed a sphere. If we strip the picture of additional cues, we are likely to perceive the object in the picture as an ellipsoid-like shape rather than a sphere.

The shape of two-dimensional images in a photograph is determined by the projection of three-dimensional space into the plane produced by the camera. For a pinhole camera and for most standard lenses this projection is the *perspective projection*. The distortion in Figure 1.1 is one of several types of distortion that are quite common for wide-angle perspective images.

There are several questions that we can ask about these distortions: Why does perspective projection result in distorted images? Why do we perceive perspective projection images as realistic in most cases? How do we obtain a picture similar to Figure 1.1 which looks less distorted? (The possibility of decreasing destortion is demonstrated in Figure 1.2.)

These questions are important for image synthesis, because most existing rendering systems

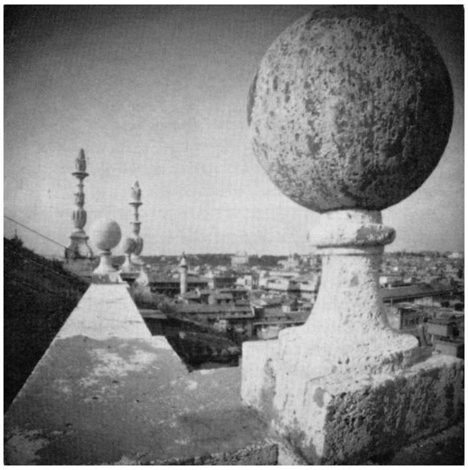


Figure 1.1: (Left) Wide-angle pinhole photograph taken on the roof of the Church of St. Ignatzio in Rome, classical example of perspective distortions from [Pirenne, 1970]. □

Figure 1.2: (Right) Corrected version of the picture with transformation described in Section 4.3.5 applied. □

model some aspects of photographic process; practically all of them use perspective projection as the *viewing transformation*[1]

Answering the questions that we have posed requires analysis of some features of human perception of pictures, for example, the perception of shape. Such analysis can be used to evaluate commonly used projections and create new ones.

The main goals of the work described in this thesis were:

– to collect and systematize available data on the perception of pictures of three-dimensional scenes;

– to develop a framework for evaluation and construction of viewing transformation based on perceptual considerations;

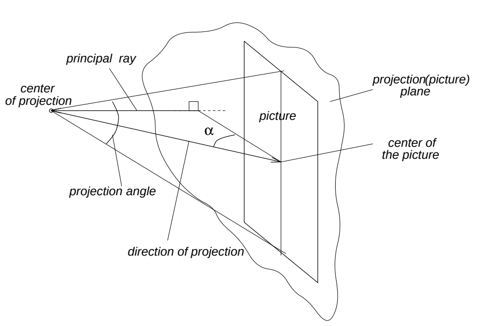
– to construct sample families of transformations in this framework.

1.2 Contributions

- We formalize a number of perceptual requirements as mathematical constraints imposed on the viewing transformation.

- We prove that a viewing transformation can produce perceptually acceptable images for arbi-

[1] *Viewing transformation* is the computer graphics term for the projection of three-dimensional space into the plane.



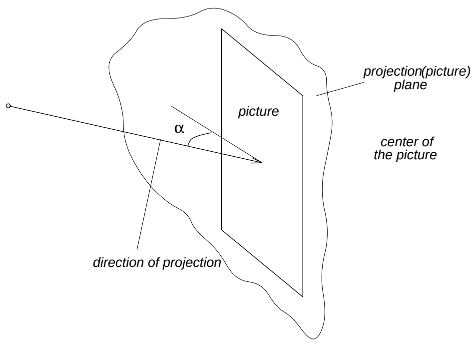


Figure 1.3: (Left) Central projection. □

Figure 1.4: (Right) Parallel projection. □

trary scene geometry only if it can be represented as a composition of a linear projection and planar transformation for each disjoint part of the scene.

- We introduce quantitative measures for two types of perceptual distortion that can be used for evaluation of viewing transformations.
- We demonstrate that two important types of distortion cannot be eliminated simultaneously.
- We construct a useful family of axially symmetric viewing transformations which allow the reduction perceptual distortion.
- We describe a non-symmetric extension of this family which allows further reduction of geometric distortion.

1.3 Perspective Projection and Viewing Position Nomenclature

We shall use a number of terms describing various characteristics of perspective projections and viewing positions. There is a lot of variation in usage of some of these terms; we try to use a consistent nomenclature whenever possible, even when describing results of authors who originally used a different one. Terminology more specific to *perception*, can be found in Section 2.1.2 and definitions related to more general types of projection are given in Section 3.2.1 and Section 3.3.1.

central projection Let C be a point in $\mathbf{R}^3$, let p be a plane such that $C \notin p$. For any point $x \in \mathbf{R}^3$, $x \neq C$, define l_x to be the line through C and x. Then a *central projection* with the center C to the plane p is the mapping $\mathbf{R}^3 \setminus C \to p$ which maps any point $x \in \mathbf{R}^3$ to $l_x \cap p$.

parallel projection Let p a plane, $\mathbf{v}$ a vector which is not parallel to the plane p. For any point $x \in \mathbf{R}^3$ define l_x to be the line through x parallel to $\mathbf{v}$. Then a *parallel projection* in the direction $\mathbf{v}$ to the plane p is the mapping $\mathbf{R}^3 \to p$ wqhich maps any point $x \in \mathbf{R}^3$ to $l_x \cap p$.

perspective projection Either *central* or *parallel projection* of a part of the three-dimensional space into a plane (the *projection* or *picture plane.*)

orthogonal projection A perspective projection is *orthogonal* if the direction of projection is perependicular to the plane of projection (this definition applies both to the parallel and central projections.)

principal ray In central projection, the *principal ray* is the ray through the center of projection which intersects the projection plane perpendicularly.

direction of projection For the parallel projection *the direction of projection* is a part of the standard definition. For the central projection, we define *the direction of projection* to be the direction from the center of projection to the center of the picture. This definition is precise for objects with central symmetry, such as a rectangle. For a picture of arbitrary shape, the center can be defined to be the geometric center (the center of mass.)

When we say "45° direction of projection" we mean that the angle α between the projection plane and the direction of projection (Figure 1.3, Figure 1.4) is 45°.

projection angle For a central projection with the center C, we choose a characteristic line segment $[AB]$ in the picture passing through the center of the picture and endpoints A,B on the boundary of the picture. Define *the projection angle* to be $\angle ACB$. For rectangular pictures, we can use the horizontal or vertical line segment going through the center of the picture. *The projection angle* is not defined uniquely for a given picture, even if it is rectangular. If the ratio of the minimal linear dimension of the picture to the maximal is not too small, *the projection angle* doesn't vary much with the choice of AB.

For the parallel projection, we define the projection angle to be zero.

projection distance For the central projection, the distance from the center of projection to the center of the picture is the *projection distance.*

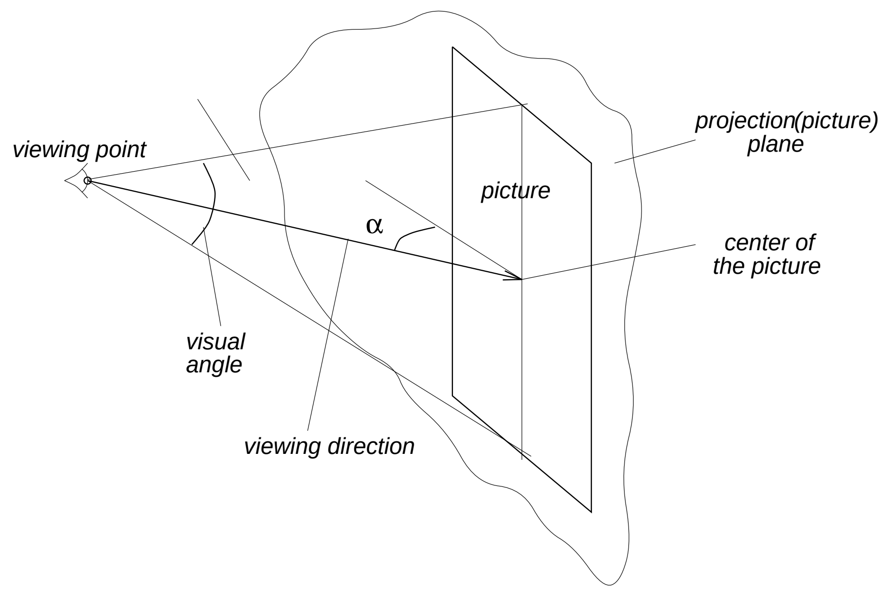


Figure 1.5: □

viewing point For monocular viewing, the *viewing point* is defined as the center of the lens of the eye. For binocular viewing, the *viewing point* is the middle of the line segment between the eyes.

viewing angle The viewing angle is equal to *the projection angle* of a central projection from the viewing point to the picture plane.

viewing direction *The viewing direction* is the direction from the viewing point to the center of the picture.

viewing distance The distance from the viewing point to the center of the picture.

visual field *visual field* is the solid angle with the apex at at the viewing point such the light from all directions within this angle reaches the eye(s).

field of view If an aperture is used for viewing, a large part of the *visual field* is occupied by the aperture. T he remaining part is called the *field of view*.

1.4 Summary

We start with an analysis of the perception of geometry in pictures (Chapter 2.) There is no general theory that explains all the aspects of perception of pictures. Instead of using any particular theory, we describe some minimal requirements of perceptual realism that arise from everyday experience and some experimental work in psychology. This part of the thesis, together with Appendix A, attempts to provide a concise but sufficient description of facts that are known about the perception of pictures. The most important topics that are discussed in Chapter 2 are the robustness of pictures, the perception of common geometric shapes such as rectangular corners and spheres, the perception of curvature and verticality and the perception of rigidity in motion pictures.

Perceptual requirements provide us with a set of constraints that we impose on the projection. Violation of some of these constraints results in unacceptable artifacts in the images, such as loops in the images of lines or twisted images of planes (Section 2.1.) Other constraints are more flexible and some deviation from the exact satisfaction of these constraints can be tolerated. An example of such a constraint is the curvature of the image of a straight line: it can be nonzero for relatively short line segments.

The first group of perceptual constraints (structural constraints) allows us to limit the set of viewing transformations that we consider (Section 3.2.) Using additional practical considerations, we prove that any projection satisfying these constraints is a composition of a perspective projection and a two-dimensional transformation.

Some constraints in the second group cannot be satisfied simultaneously: the curvature of lines and the distortion of shape cannot be eliminated completely at the same time (Section 3.3.) We introduce error functions for each constraint which characterize the quality of a projection with respect to this constraint at a point. We can use the average or the maximum error to obtain an overall characteristic of a projection.

Using global error functionals for the direct view constraint and the zero curvature constraint, we formulate a parametric optimization problem which allows one to find projections with minimal global errors for a given tradeoff between the two types of the error (Section 3.3.5.)

We consider a simple parametric family of projections (Section 4.3.5) and show that this family approximates exact solutions of the optimization problem in the axially-symmetric case well enough for practical purposes. The resulting family of transformations has a very simple form.

If we allow different tradeoffs between two types of constraints in different parts of the picture we can produce non-symmetric projections which can be better adjusted to the specific scene configuration. The resulting extended family of transformations is still quite easy to use, but is much more flexible.

Our families of projections have two principal components that can be chosen independently: the perspective projection and the two-dimensional transformation. We describe how the choice of appropriate projection is influenced by the geometry of the scene (Section 5.2.)

The fact that the first part of our decomposition of projections is the standard perspective projection, makes it possible to implement the two-dimensional transformation as a postprocessing step of rendering. It is also possible to apply this transformation to photographs if the angle of view is known or if it can be computed from the information in the picture. (Section 5.3; Appendix B.) In the wide-angle motion pictures varying distortion of shape results in an impression of nonrigidity which may be undesirable; our transformations allow to reduce this effect.

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.*

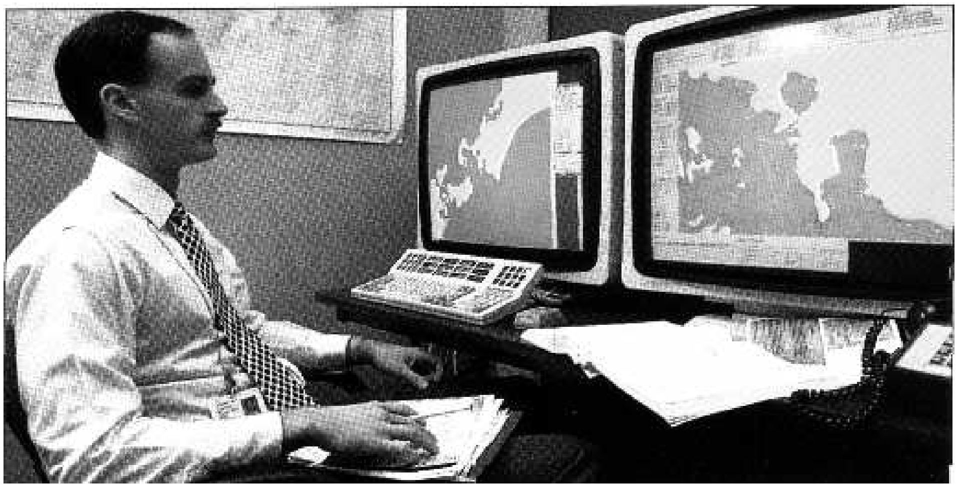


Figure 2.1: Photo from the article "Navigating Close to Shore" by Dave Dooling ("IEEE Spectrum", Dec. 1994), © 1994 IEEE, photo by Intergraph Corp. 92° viewing angle. The two monitors have the same size. But the left one appears to be wider (misperception). □

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.

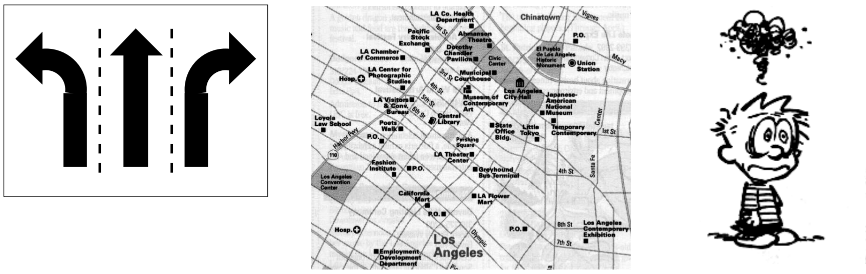


Figure 2.2: Examples of highly conventional pictures: a road sign,a map, a cartoon. □

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.

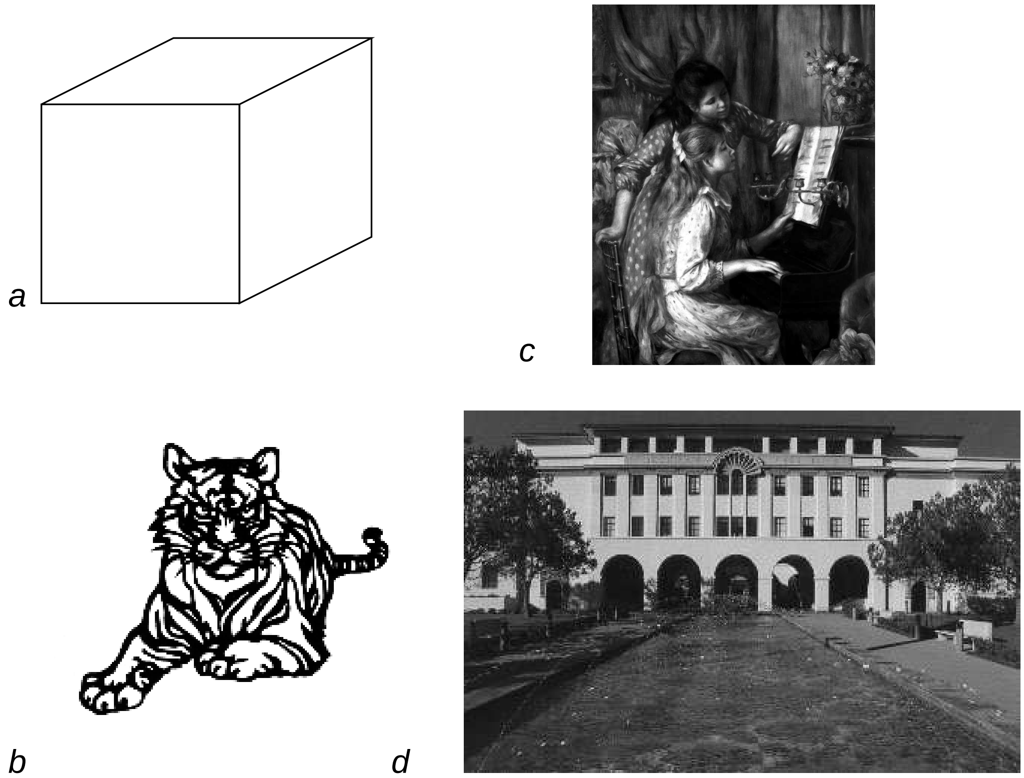


Figure 2.3: Examples of pictures that use symbols in a limited way, or don't use them: line drawings, a painting ("Piano lesson" by A. Renoir), a photograph (Beckman Institute at Caltech.) □

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.

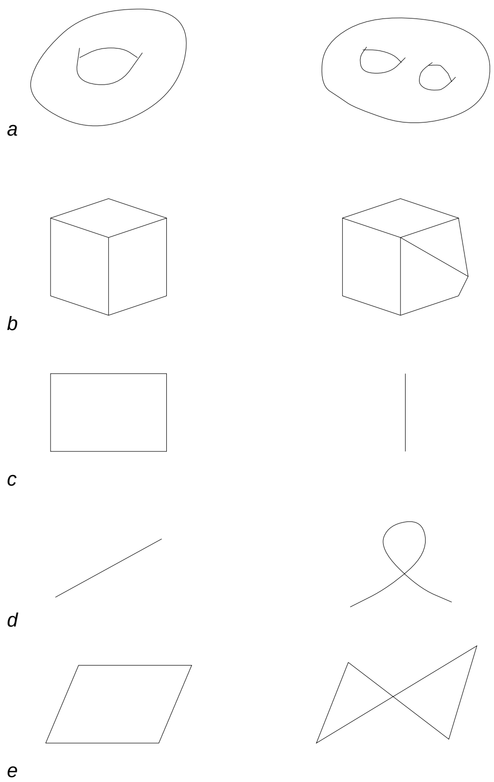


Figure 2.4: Examples of structural differences in images: a) the number of holes b) number of edges meeting at a vertex c) dimension d) number of self-intersections, local structure e) number of intersections of edges, local structure □

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.

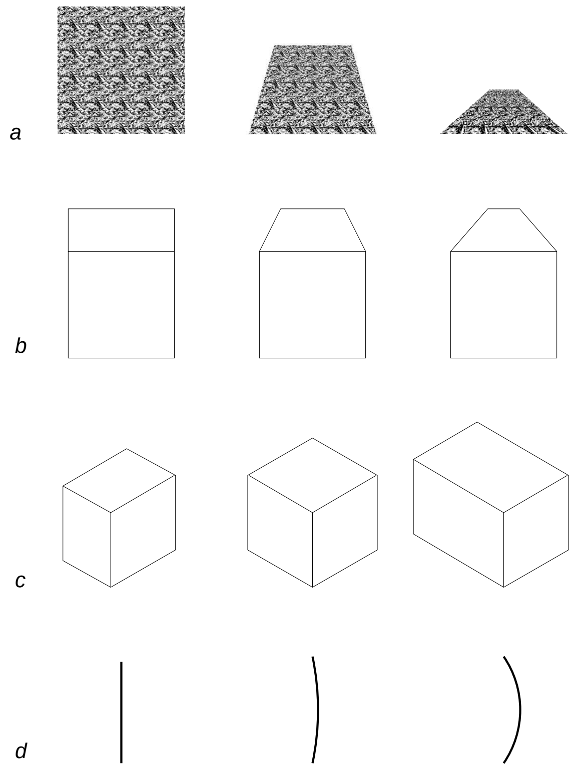


Figure 2.5: Examples of nonstructural differences in images: a) texture gradients b) parallel foreshortening c) perpendicular foreshortening d) curvature □

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.)

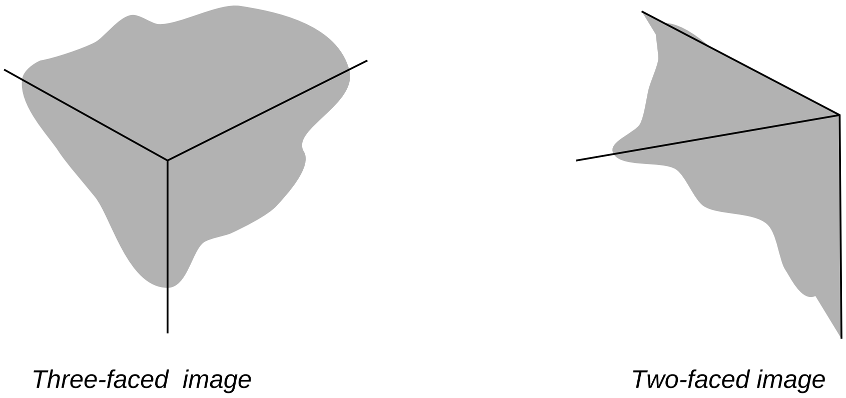


Figure 2.7: Two types of images of rectangular corners □

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.

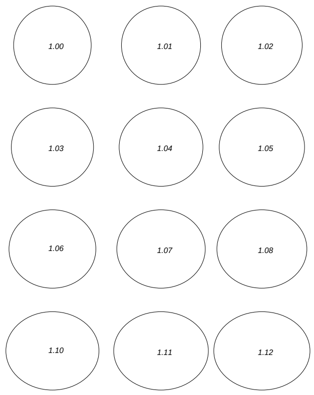


Figure 2.8: Ellipses with different aspect ratios (the aspect ratio is shown in the center of each ellipse.) □

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. □

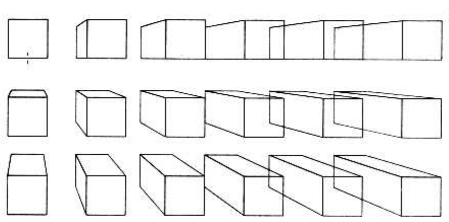
There are types of distortion that are more likely to produce deformed, not just misleading images. An important feature of acceptable frontal pictures of humans is their axial symmetry. When this symmetry is broken, the image becomes unacceptable.

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



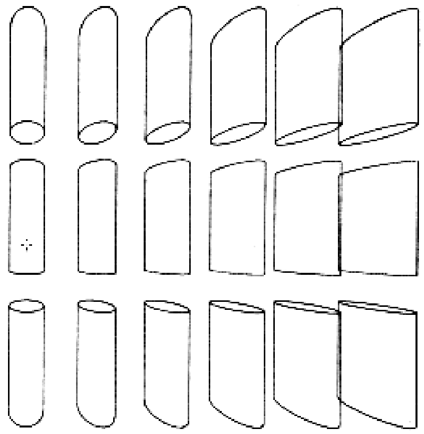


Figure 2.10: (Left) projections of cubes from [Pirenne, 1970] □

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.

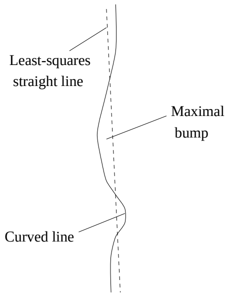


Figure 2.12: Watt's measure of perceptual curvature □

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.

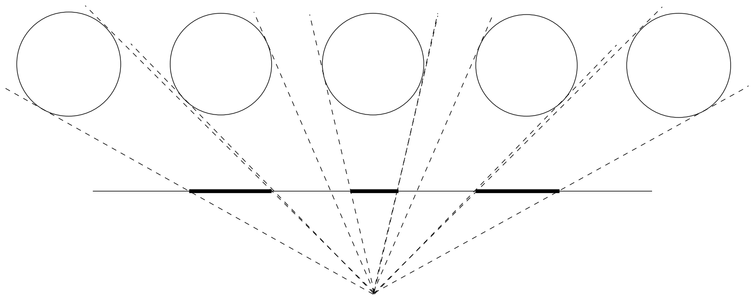


Figure 2.13: Leonardo's example of a perspective distortion □

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

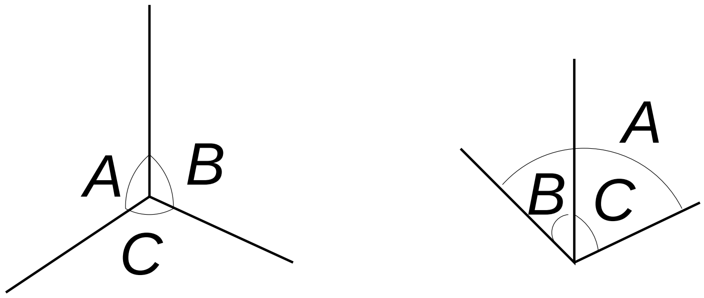


Figure 2.14: Three-sided and two-sided images of a rectangular corner □

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

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.

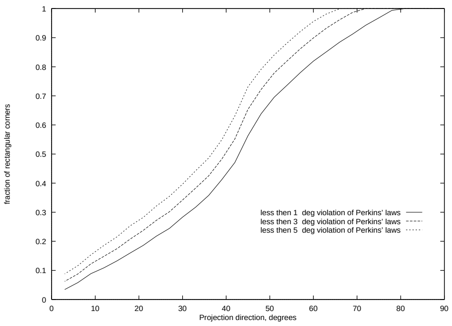
It can be easily shown that our definition is equivalent to the following simpler one:

> 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.



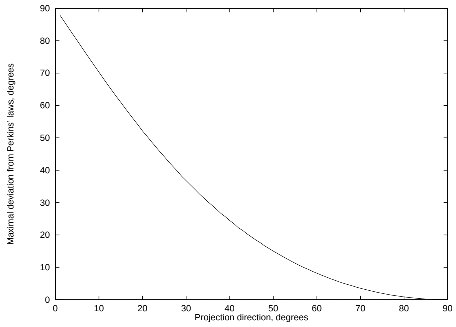


Figure 2.15: (Left) Fraction of "good" rectangular corners *vs* projection angle □

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].

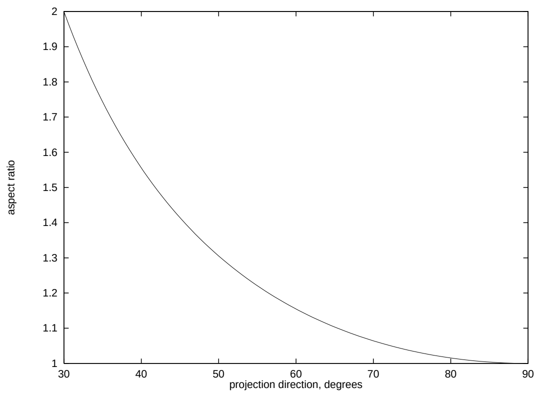


Figure 2.17: Aspect ratio of the image of a small sphere *vs* projection direction angle. □

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

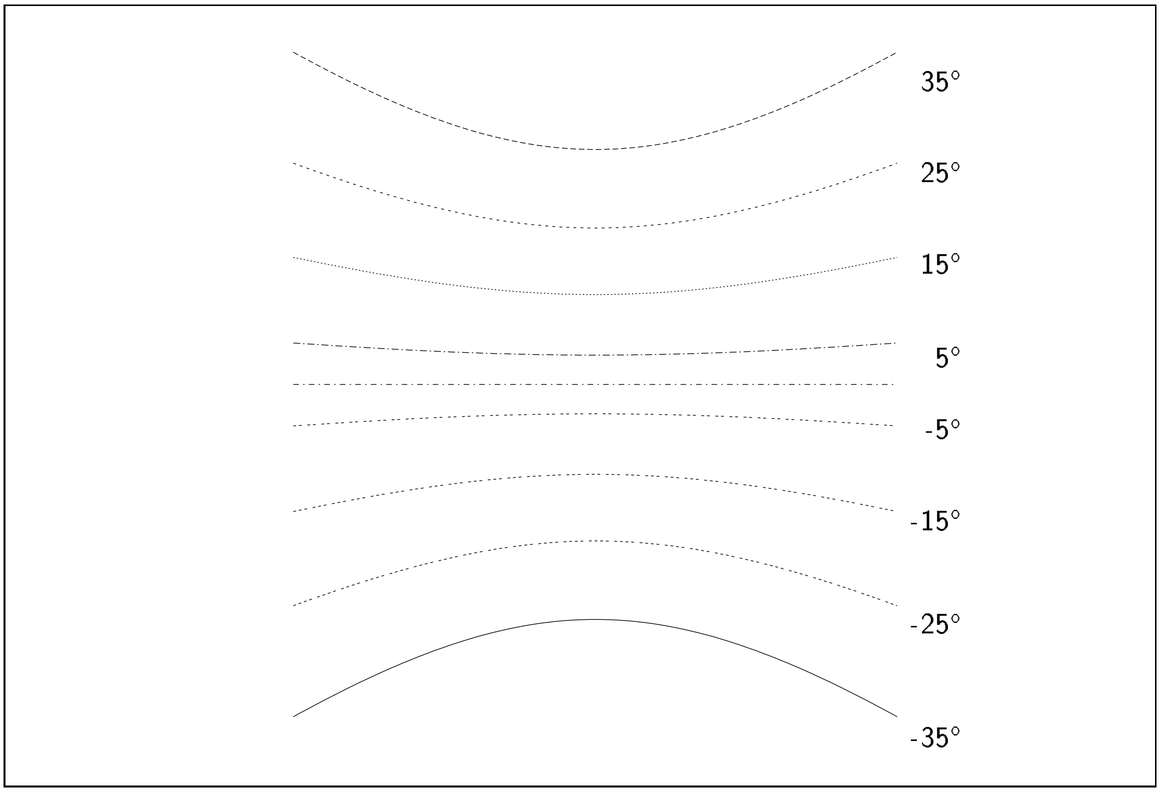


Figure 2.22: Trajectories of points in a wide-angle motion picture when the camera turns. The projection angle is 90° in vertical and horizontal directions. □

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]) □

Chapter 3

Formalization of Perceptual Requirements

In this chapter we show how it is possible to formalize perceptual requirements described in Chapter 2, use them together with additional technical restrictions to restrict the class of projections that we consider, and derive perceptual metrics that can help to evaluate projections from perceptual point of view.

3.1 Technical Requirements

To narrow down the area of the search for perceptually acceptable projections we are going to specify several additional design considerations. They don't have any perceptual basis and some of them are quite restrictive; however, they make the task of constructing projections manageable. Resulting projections can be applied to a wide class of images and the choice of correct tranformation can be made simple.

- We want a parametric family of viewing transformations so that an appropriate one can be chosen for each image.
- The number of parameters should be small, and they must have a clear intuitive meaning.
- Construction of the family of projections should not depend on the scene. Matching the scene to a particular projection in the family should be achieved by choosing an appropriate set of parameters.
- Pictures should be scalable: any part of a picture can be made a separate picture. Therefore, no area in a picture can be considered small enough to be ignored.

3.2 Structural Requirements

Structural distortions, as described in Section 2.1, are in general more apparent than nonstructural distortions. It is not difficult to find exceptions to this rule, but given the additional condition of scalability (Section 3.1), almost any structural distortion would be highly objectionable.

This idea allows us to use structural and nonstructural requirements in different ways: we postulate that some structural requirements should be satisfied exactly, considerably reducing the variety of projections that we consider. Then nonstructural distortions can be minimized within the class of projections satisfying structural requirements.

Even given the scalability assumption, some structural requirements appear to be more important than others: if a requirement is associated with unusual and complicated object geometry, it is less important than those associated with simpler and more common objects.

In this section we choose several formalizations of certain structural requirements and apply them to images of *all* possible objects in *all* possible positions: we want our family of projections to be independent of the objects that we want to depict, and we want the resulting images to satisfy a given set of structural requirements exactly.

More specifically, we will show that three sets of structural requirements result in the same structure for the projections:

the fibers of the projection should be subsets of the fibers of a perspective projection (see Section 3.2.1 for the definition of a fiber.)

It doesn't mean that any perceptually acceptable projection necessarily has this structure. It rather means that if a projection doesn't have this structure, there are objects that will have unacceptable images if they are in certain positions in space.

For each set of structural requirements, we will prove the statement above by assuming the contrary (i.e. presence of curved or non-coplanar fibers) and construct an object that will have an image violating some structural requirement.

As the objects that we consider in Section 3.2.4 are quite simple and common (object with planar faces and straight edges) it seems reasonable to assume that most mappings with curved fibers are likely to produce considerable structural distortions for at least some scenes, and to concentrate in the rest of our study on the projections with the fiber structure of a perspective projection which guarantees absence of structural distortions.

3.2.1 Definitions and Preliminary Lemmas

We will use $x, y, ..$ for the points in the domain of a projection (a volume in 3D space), and $\xi, \psi ...$ for the points in the range (a point in the picture plane).

∂A denotes the boundary of the set A.

$\overline{A}$ denotes the closure of the set A.

Definition 3.1 *By a* **line segment** *we mean any connected subset of a straight line.*

We will use the following notation for line segments:

– $[a, b]$ denotes the closed line segment between the points a and b;

– $]a, b[$ denotes the open line segment between the points a and b;

– (a, b) denotes the straight line containing the points a and b.

Definition 3.2 *We will call a mapping $P : \mathbf{R}^n \to R^m$ smooth, if all the components of this mapping are at least twice continuously differentiable.*

Definition 3.3 *The set of all points of the domain of a mapping that map to a fixed point ξ is called the* **fiber** *of the mapping at the point ξ.*

If a fiber is a line segment and is not a point, we will call the straight line containing it **the fiber line**. For each point x in the domain of a mapping there is fiber going through this point. We will denote it $F(x)$.

Definition 3.4 *A* **projection** *is a smooth mapping $P : V \subset \mathbf{R}^3 \rightarrow \mathbf{R}^2$, where V is an open path-connected domain in $\mathbf{R}^3$, with the rank of the Jacobian of P equal to 2 in all points of V.*

Definition 3.5 *We will call a continuous (smooth) curve $\gamma_\epsilon : I \rightarrow \mathbf{R}^3$, where I is a line segment, an* **ϵ-perturbation** *of a continuous (smooth) curve $\gamma : I \rightarrow \mathbf{R}^3$ if there is a homeomorphism (diffeomorphism) $g : \gamma_\epsilon(I) \rightarrow \gamma(I)$ such that the distance $|g(x) - x|$ is less than ϵ for any $x \in \gamma(I)$.*

Definition 3.6 *A* **homotopy** *of a continuous curve $\gamma : I \rightarrow A$ to another curve $\gamma' : I \rightarrow A$ in $A \subset \mathbf{R}^n$ is a continuous mapping $h : [0,1] \times I \rightarrow A$, such that $h(0,t) = \gamma(t)$ and $h(1,t) = \gamma'(t)$ for all $t \in I$.*

Definition 3.7 *Two smooth planar curves intersect* **transversally** *at a point x, if their tangents at this point are not parallel. A smooth curve and a smooth surface intersect* **transversally** *if the tangent vector of the curve is not in the tangent plane of the surface.*

A **crossing** is an intersection of two curves that cannot be eliminated by an arbitrarily small perturbation of the curves. More precisely,

Definition 3.8 *Two planar curves in an open set $A \subset \mathbf{R}^2$ have a* **crossing** *inside A if they have an intersection point in A and there is ϵ such that for any ϵ'-perturbation of any of the intersecting curves, where $\epsilon' < \epsilon$, the perturbed curves still intersect inside A.*

If two curves intersect transversally in A they have a crossing in A. Some non-transversal intersections result in crossing; some don't (Figure 3.1.)

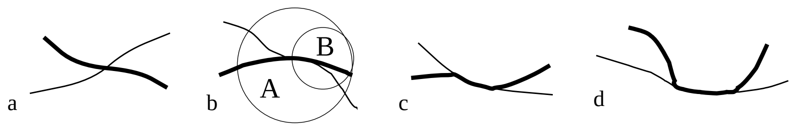


Figure 3.1: Examples of intersections. a. Transversal; curves γ and γ' have a crossing in any open set containing the intersection. b. Nontransversal; curves γ and γ' still have a crossing in any open set containing the intersection. c. nontransversal; γ and γ' have a crossing in A and don't have a crossing in B. d. nontransversal, γ and γ' don't have a crossing in any set. □

Two closed curves are homotopic in a set if one can be continuously deformed to the other without leaving the set.

The following definition is a mathematical expression for "having a different number of holes." We will use without proof the fact that a smooth closed curve without self-intersections separates the plane into two path-disconnected sets - one bounded ("inside") and the other unbounded ("outside"). The curve can be retracted to a point by a homotopy in the "inside" area and cannot be retracted to a point in the "outside" area.

Definition 3.9 *We will define a* **hole** *in the image $P(O)$ of an object O to be a path-connected set of points H that don't belong to $P(O)$, if in $P(O)$ there is a closed smooth curve without self-intersections such that H lies inside the curve. Two holes are different if there is a smooth closed curve without self-intersections such that one hole lies inside the curve and the other outside.*

Definition 3.10 *The* **convex hull** *of a set is the union of all line segments connecting two points from this set.* $\mathrm{CH}(A)$ *denotes the convex hull of the set A.*

Now we prove several lemmas that we will use later.

Lemma 3.1 *If the points of a continuous curve $\gamma : [0,1] \to \mathbf{R}^3$ are divided into two subsets S and S' so that $S \cap S' = \emptyset$, $S \cup S' = \gamma([0,1])$ then there is a common limit point of S and S'.*

Proof.

Suppose for each point of S and S' there is a neighborhood which doesn't contain any points of the other set. Let $x \in S, x' \in S'$. Then there is a continuous curve from x to x'. Consider the inverse images $I = \gamma^{-1}(S)$ and $I' = \gamma^{-1}(S')$. It follows from our assumption that for each point in I there is a neighborhood not containing points in I'. Let the union of all these neighborhoods be N and the union of neighborhoods constructed in the same way for S' be N'. Then N and N' are open sets which cover the whole interval. Their intersection is an open set that doesn't contain any points of I and I'; therefore it is empty ($I \cup I' = [0,1]$.) Thus we have a decomposition of the interval into two nonempty open sets which is impossible, as the interval is connected.

■

Lemma 3.2 *If $P : V \subset \mathbf{R}^3 \to \mathbf{R}^2$ is a continuously differentiable mapping with the rank of the Jacobian everywhere equal to 2, then*

- any fiber $P^{-1}(\xi)$ of the mapping is a 1-dimensional submanifold of $\mathbf{R}^3$;

- the unit tangent vector to this manifold at any point x is a continuously differentiable function of x in a neighborhood of x for some choice of orientations;

- the curvature of the fiber is a continuous function of x.

The first part of the statement immediately follows from Theorem 1.38 in [Warner, 1983]. Let's prove the second and the third part.

Consider a local one-to-one parametrization of $P^{-1}(\xi)$ in a neighborhood of a point $x_0 = (x_0^1, x_0^2, x_0^3)$, $x = x(t)$. The tangent vector at a point x is $\dot{x}(t) \neq 0$. $P(x(t)) = 0$. Differentiating this expression with respect to t, we get

$$\frac{dP(x)}{dx}\dot{x}(t) = 0$$

As the Jacobian $\frac{dP(x)}{dx}$ has rank 2, solution of this system exists and is unique up to a constant. Therefore the unit tangent vector τ is defined uniquely by the system and the formula

$$\tau = \frac{\dot{x}(t)}{|\dot{x}(t)|}$$

Components of τ are continous functions of the partial derivatives of the components of $\frac{dP(x)}{dx}$. As P is assumed to be continuously differentiable, the derivatives are continuous functions of x Therefore, $\dot{x}(t)$ is a continuous function of x.

As the curvature is simply the length of $\ddot{\tau}(t)$, it can be proven to be continuous by differentiating the expression for τ. It is important to note that $\dot{x}(t)$ doesn't depend explicitly on t.

■

Lemma 3.3 *If all fibers of a projection are line segments and any two fibers are coplanar the following is true: if any two fiber lines intersect but do not coincide, all fiber lines intersect at the same point. Otherwise, all fiber lines are parallel.*

Proof. Let f_1 and f_2 be two intersecting but not coinciding fiber lines. Let $x \in V$ be a point outside the plane p defined by f_1 and f_2, and f_3 be the fiber line of x. As f_3 should be coplanar to f_1 it can be either parallel to it or intersect it. But if it is parallel to f_1, it is parallel to the plane p, and therefore doesn't intersect f_2. But it isn't parallel to f_2 either; therefore, f_3 and f_2 are not coplanar. Thus, f_3 should intersect f_1 and f_2. But it can intersect the plane p only in one point, so it should be the common point of f_1 and f_2.

Consider a point x' in the plane p and the fiber line f_4 corresponding to it. The same reasoning applies to f_4 and the pair f_1, f_3.

Suppose f_1 and f_2 do not intersect. Suppose some fiber f_3 intersects f_1. If it is not in p, then it cannot be parallel to f_2 and it doesn't intersect it, which is impossible. If it is in p, then a fiber line f_4 which is not in the plane should intersect it in two points: $f_1 \cap f_3$ and $f_2 \cap f_3$, which is impossible. ■

3.2.2 Holes in Images

Given our definition of holes, description of example a on Figure 2.4 obtains exact meaning:

> **Condition H.** An image of an object should not have a number of holes different from the number of holes for any perspective projection of the object.

Lemma 3.4 *Given Condition H, all fibers of the projection should be line segments.*

Idea of the Proof. Suppose there is a fiber which is not a linear segment. We construct a tube around this fiber in such a way that no straight line can get through the tube. Therefore, no perspective projection of the tube can have holes in it. The fiber goes through the tube without intersecting the walls, therefore, there is at least one hole in the image of the tube under P, which contradicts the Condition H.

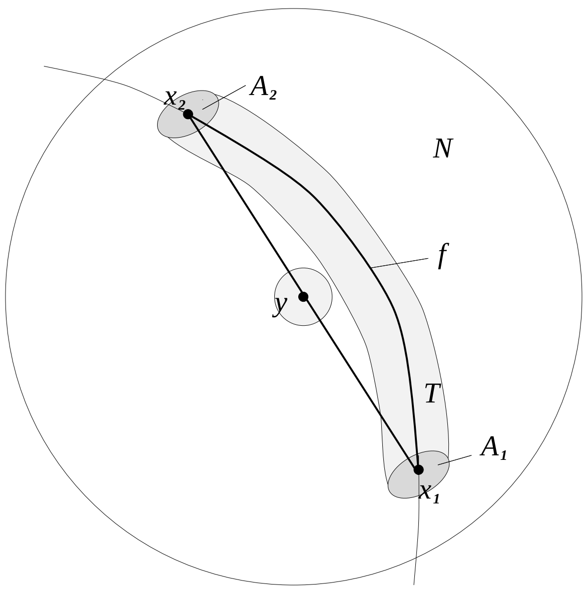


Figure 3.2: Construction of Lemma 3.4 □

Proof.

Suppose a fiber f is not a line segment. Choose a point x on the fiber where the curvature is not zero. There is a neighborhood of the point x where the curvature is not zero, because the curvature is continuous (Lemma 3.2.)

As the fiber f is a one-dimensional submanifold of $\mathbf{R}^3$ with induced topology, we can find an open neigborhhood N_f of x in f which is diffeomorphic to an open interval and $N_f = N \cap f$, where N is an open set in $\mathbf{R}^3$. We can choose N_f and N so that N is a ball.

Consider the endpoints x_1 and x_2 of N_f. If N_f is chosen in the interior of f, they should be outside N. As they are limit points of N_f, they should be on ∂N. These are the only intersections of f with ∂N.

Let f_c be the closure of N_f. Consider a cover of f_c with open balls centered at the points of f_c. Choose the balls so that for the points of N_f their closures are inside N. As f_c is bounded and closed, it is compact. Therefore, there is a finite subcover Cover_{f_c}, and there is ϵ such that any ball of the radius less than ϵ centered at a point of f_c is inside Cover_{f_c}. Let $T = \mathrm{Cover}_{f_c} \cap N$.

T has the following properties: $T \cap f = N_f$, $f \cap \partial T = \{x_1, x_2\}$. If a point y is inside T, there is a point y_f of f such that $|y - y_f| < \epsilon$. As N is a ball, the convex hull CH(T) will be contained in N and CH($\overline{T}$) will be contained in $\overline{N}$. Therefore, CH$(T) \cap f = N_f$.

Consider the line segment $[x_1, x_2]$. As the curvature of f in N is not zero, ther is a point y on this line segment which is not on f_c. there are neigborhoods M_f of f_c and M_y of y which don't intersect. Choose ϵ small enough for T to be inside M_f and for the ball of radius ϵ to be inside M_y.

Define the object O' to be CH($\overline{T}$). Consider $\partial T \cap \partial N$. By construction, the only part of ∂T that can intersect ∂N should belong to the boundary of the balls $B(x_1)$ and $B(x_2)$ centered at x_1 and x_2. Let $A_1 = B(x_1) \cap \partial N$, $A_2 = B(x_2) \cap \partial N$. For sufficiently small ϵ, A_1 and A_2 are disjoint, i.e. $\overline{T} \cap \partial N$ is a disjoint union of A_1 and A_2. As A_1 and A_2 are intersections of a ball and a sphere, ∂A_1 and ∂A_2 in ∂N are circles. Define the object O as $(O' \setminus T \setminus (A_1 \cap A_2)) \cup \partial A_1 \cup \partial A_2$. (We throw away T, and the interior of the intersection of $\overline{T}$ with ∂N.)

Consider a perspective projection Q with center outside O. Let l be a straight line going through the center of projection (a fiber line of the perspective projection) and let ξ be the corresponding point in the picture plane. We want to show that there are no holes in the image of O. As O' is convex, its perspective projection image is convex too, and it is easy to show that it doesn't have holes. As $Q(O) \subset Q(O')$, no point that doesn't belong to $Q(O')$ can be a point of a hole in $Q(O)$.

We have to consider only the points of $Q(O') \setminus Q(O)$. If $\xi \in Q(O') \setminus Q(O)$, the line l intersects O' but doesn't intersect O, i.e. intersects O' inside T.

No point outside T can be connected by a continuous curve to a point in T without crossing ∂T. All the points of ∂T, except those that are in ∂N, are in O. If the straight line l intersects O', but doesn't intersect T, it should go through the points of A_1 or A_2.

Suppose l intersects only one of these sets. Consider a plane through l which intersects ∂N only inside A_1. Then all the points of O are on one side of this plane (the only points of O' that are on the other side belong to T.) Therefore, for any curve in O there is a homotopy in $\mathbf{R}^3$ that contracts it to a point, without intersecting l. The image of O has the same property with respect to the image ξ of l, which is a point. Therefore, there is no closed curve in $Q(O)$ that has ξ inside, and ξ cannot belong to a hole in $Q(O)$.

Suppose l intersects both sets. Any line going through the balls of radius ϵ centered at x_1 and x_2 goes through a ball of radius ϵ centered at y. But this ball doesn't intersect T, therefore, there is a point on l which is outside T and inside ∂N. Then the line should intersect $\partial T \setminus \partial N$, which is part of O. Thus, ξ is in the image of O.

We conclude that $Q(O)$ has no holes in it. $P(O)$, however, has a hole. Fiber f doesn't intersect O by construction. It intersects the set A_1 at the point x_1. Consider the image of $A_1 \cup \partial A_1$. ∂A_1 is a circle, therefore, its image can be considered as a smooth closed curve. Again, if ϵ is sufficiently small, we can assume that the directions of the fibers going through the points of A_1 are within a small solid angle and are transversal to the sphere ∂N. The image of $A_1 \cup \partial A_1$ is one-to-one, and the curve $S^1 \to P(\partial(A_1))$ doesn't have self-intersections.

Clearly, $x_1 \in P(A_1)$ is in the interior of this curve, and therefore, there is a hole in the image of O.

Threfore, $P(O)$ and $Q(O)$ have different number of holes for any choice of Q and Condition H is violated.

■

Lemma 3.5 *Given Condition H, all fibers should be coplanar.*

Idea of the Proof. The idea is similar to the proof of the previous lemma. If there are two non-coplanar fibers, we can construct an object with tubes through the object along the fibers in such a way that one cannot see through both tubes at the same time. There will be at least two holes in the image of the object under P, and no more than one under P.

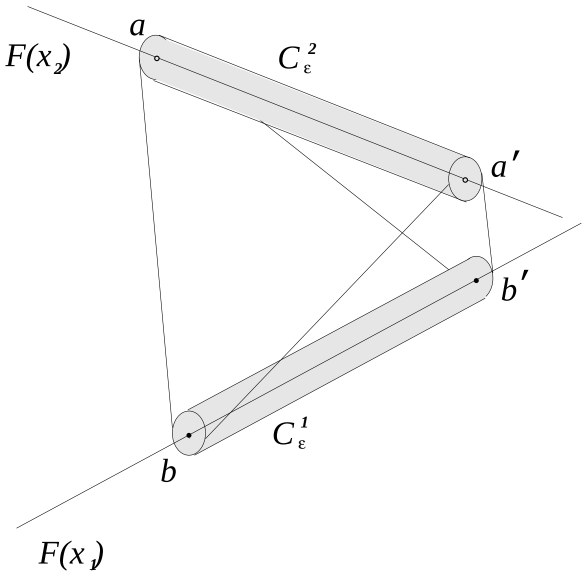


Figure 3.3: Construction from Lemma 3.5 □

Proof.

Suppose we have two non-coplanar fibers, f_1 and f_2.

As we have proven, all the fibers are line segments. Choose a curve γ connecting two points on the fibers f_1 and f_2. We can separate all the points on the path into two sets: S, the set of all points for which the fiber $F(x)$ is coplanar with f_1, and S', the rest of the path. Using Lemma 3.1 we can find a point x such that in any neighborhood of this point there are points from S and S'. Let $B(x)$ be a ball contained in V. Choosing x_1 to be a point from S and x_2 to be a point from S', we obtain fibers $F(x_1)$ and $F(x_2)$ that are not coplanar and both intersect $B(x)$. We can pick points $a, a' \in F(x_1)$ and $b, b' \in F(x_2)$ inside $B(x)$.

Consider two cylinders C_ϵ^1 and C_ϵ^2 of radius ϵ and with axes $[a, a']$ and $[b, b']$. Let l_1 and l_2 be line segments connecting points on the opposite faces of C_ϵ^1 and C_ϵ^2 respectively. We can choose ϵ in such a way that for any choice of l_1, l_2 they are not coplanar and C_ϵ^1 and C_ϵ^1 are inside $B(x)$. Appropriate choice of a, a', b and b' can also ensure that the faces of the cylinders are inside the boundary of the convex hull $CH(C_\epsilon^1 \cup C_\epsilon^2)$.

Consider an object O defined as $CH(C^1_\epsilon \cup C^2_\epsilon) \setminus (Int(C^1_\epsilon) \cup Int(C^2_\epsilon))$. Any perspective projection of this object has at most one hole: a straight line which intersects the convex hull, but doesn't intersect O, should enter at one of the cylinders' faces and exit at the opposite face of the same cylinder. If there are two disconnected regions in the image in which fibers of the perspective projection behave in this way, we have two coplanar straight lines (fibers of a perspective projection are coplanar) that intersect $CH(C^1_\epsilon \cup C^2_\epsilon)$ only inside C^1_ϵ and C^2_ϵ respectively. But by construction this is impossible. It can be shown that each cylinder can produce no more than one hole in a perspective projection.

On the other hand, as $[a, a']$ and $[b, b']$ belong to the fibers, and the boundaries of the faces of cylinders map to curves enclosing the points $P(a)$ and $P(b)$, the image $P(O)$ has at least two holes. ■

By Lemma 3.3, all fiber lines are either parallel or intersect at one point. Therefore, fibers of P are subsets of fibers of a perspective projection.

The main problem with the argument above is that the objects that we construct are rather unusual and the resulting structural distortion is not very significant: it is not obvious when we are supposed to be able to see through a curved pipe and when we are not. In the following subsections we will to consider simpler objects.

3.2.3 Loops, Folds and Twists

In this section we consider the following conditions, which formalize (Figure 2.4 d,e), but are more restricitve. We describe this set of conditions mostly for historical reasons –this is the set described in [Zorin and Barr, 1995]. The conditions used in the next section formalize essentially the same requirements in a better way.

The main advantage of this formalization is that we don't have to deal with smoothness properties and the proofs become simpler.

> **Condition L.** The mapping of any line segment l to its image $P(l)$ is one-to-one everywhere or it is a point.

This condition prevents "loops" in the images of lines. It is more restrictive: not only it doesn't allow "loops" but also "folds", that is situations when the image is not one-to-one, but is still simply connected.

Condition T. The mapping of a subset of a plane m to the image $P(m)$ is one-to-one everywhere or nowhere.

This condition prevents "twists" in the images of planes.

Lemma 3.6 *Given condition L, all fibers of the mapping P are line segments.*

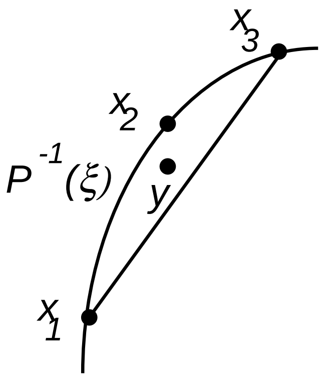


Figure 3.4: Construction from Lemma 3.2.3 □

Proof.

Consider a fiber $P^{-1}(\xi)$. Suppose it is not a line segment. Let x_1 and x_2 be two points of the fiber. Then $P^{-1}(\xi) \not\subset (x_1, x_2)$.

Let $S = P^{-1}(\xi) \cap (x_1, x_2)$, $S' = P^{-1}(\xi) \setminus (x_1, x_2)$. It is possible to show that there is a point $x \in P^{-1}(\xi)$ such that in any neighborhood of this point there are points of S and S'. Let $B(x)$ be a ball contained in V (the domain of the mapping). We can choose x_1, x_2 and $x_3 \in P^{-1}(\xi)$ to be three non-collinear points in $B(x)$.

Consider the image of the triangle $\triangle x_1\ x_2\ x_3$. As the whole triangle cannot map into ξ (this would imply that $\dim P^{-1}(\xi) > 1$, contradicting Condition L, there is a point y in the triangle such that $P(y) = \xi' \neq \xi$. Consider the segment $[x_1, y'] \ni y$, where $y' \in [x_2, x_3]$. If $P(y') = \xi$ then the image of $[x_1, y']$ is not one-to-one but contains at least two points. If $P(y') \neq \xi$, then the same is true for $[x_2, x_3]$. Thus the assumption that a fiber can contain non-collinear points contradicts Condition L.

■

Note that this lemma doesn't rely on the Condition T.

Lemma 3.7 *Any two fibers are coplanar.*

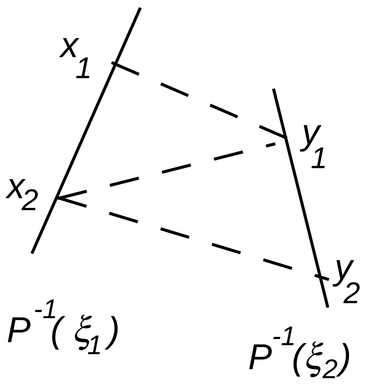


Figure 3.5: Construction from Lemma 3.7 □

Proof.

Suppose two fibers are not coplanar. Choose two points on these fibers. Consider a continuous path connecting these two points that is contained in V. Such a path exists because V is assumed to be path-connected. As in the previous lemma, we can separate all the points on the path into two sets: S, the set of all points x for which the fiber $P^{-1}(P(x))$ is coplanar with $P^{-1}(P(x_1))$, and S', the rest of the path. In the same manner we can find a point x such that in any neighborhood of this point there are points from S and S'. Again, let $B(x)$ be a ball contained in V. Let $P^{-1}(\xi_1)$ and $P^{-1}(\xi_2)$ be two non-coplanar fibers corresponding to points in $B(x)$. Consider two points x_1 and x_2 on $P^{-1}(\xi_1)$ and y_1 and y_2 on $P^{-1}(\xi_2)$. The image of $[x_1, x_2]$ is not one-to-one; it follows from Condition T, that the image of $\triangle x_1\ x_2\ y_1$ is not one-to-one anywhere. Let ξ' be a point in the image. As the intersection of the fiber $P^{-1}(\xi')$ with the plane of the triangle consists of more than one point, and the fiber is a line segment, the whole fiber belongs to the plane of the triangle. But the same thing will be true for $\triangle y_1\ y_2\ x_1$. Choose $\xi' \in P([x_1, y_1])$. As two triangles share the edge $[x_1 y_1]$, $\xi' \in P(\triangle y_1\ y_2\ x_1)$ and $\xi' \in P(\triangle x_1\ x_2\ y_1)$, and the fiber $P^{-1}(\xi')$ should be in the planes of both triangles, i.e. in (x_1, y_1). As it follows from continuity, should contain $[x_1, y_1]$. This is impossible, because x_1 and y_1 map to different points. Therefore, there can be no two fibers that are not coplanar.

■

As before, from the two lemmas that we have proven, it follows that the fibers of$P$ are subsets of fibers of a persepective projection.

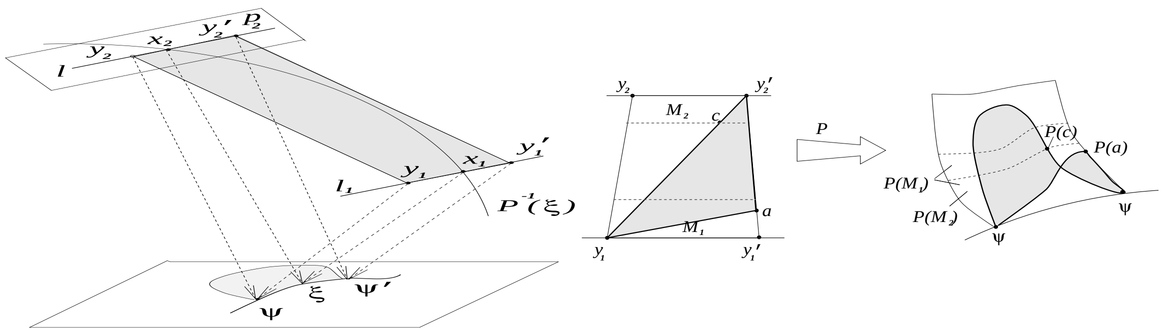


Figure 3.6: Constructions from Lemma 3.8: the image of the triangle $y_1\ a\ y_2'$ is twisted □

3.2.4 Loops and Twists

The following statement is the precise expression of the idea that the image of a line-segment should be a curve with no loops (although not necessarily one-to-one.)

Condition L′. The image of a line segment should be a simply connnected curve.

"Twists" in the image of a plane can be described more precisely in the following way: suppose an object has a flat face with several straight edges intersecting only at the endpoints. In this case no perspective projection produces images of edges crossing each other at some point in the middle unless they coincide. This leads us to the following formalization:

Condition T′. The image of a pair of coplanar line segments should have no crossing points that are not images of an intersection point of the segments.

Again, we prove the same two types of lemmas as in previous sections.

Lemma 3.8 *Given Conditions L' and T', every fiber is a line segment.*

Proof.

Suppose a fiber $P^{-1}(\xi)$ has non-zero curvature in a neighborhood of a point.

Choose two points x_1, x_2 on the fiber $P^{-1}(\xi)$ inside this neighborhhood. Consider two parallel line segments l_1 and l_2 going through the points x_1 and x_2 respectively, $l_1, l_2 \subset V$. Choose these intervals so that x_1 and x_2 are their interior points and the plane through l_1 and l_2 is transversal to all the fibers going through l_1 and l_2.

The images of the line segments l_1 and l_2 intersect at the point ξ. Suppose ξ is an intersection point such that there is no (possibly one-sided) neighborhood of ξ where the curves $P(l_1)$ and $P(l_2)$ coincide.

There should not be a crossing (Condition T$'$), so the intersection at ξ shouldn't be transversal. Consider a plane p_2 through the line segment l_2, which is crossed by $P^{-1}(\xi)$ transversally. The fibers of the mapping P going through the line segment l_1 map it into a smooth curve in the plane p_2 which is a one-dimensional submanifold (Condition L$'$.) Therefore, the tangent vector to the image of l_1 is unique at all points. As the plane p_2 is transversal to the fiber $P^{-1}(\xi)$, we can choose a one-dimensional neighborhood of x_1 in l_1 in such a way that for each fiber going through a point in this neighborhood there is only one intersection with the plane p_2 (this is possible due to continuity of the tangent vector, Lemma 3.2.) Restriction $P|_{p_2}$ of the mapping $P : \mathbf{R}^3 \to \mathbf{R}^2$ to the plane p_2 is differentiable and the point x_2 is nonsingular, because of the transversality assumption. Therefore, in a neighborhood of x_2 it can be inverted and the inverse mapping will have the same smoothness (inverse function theorem.) Consider $T = P|_{p_2}^{-1}$. It maps the neighborhhood N of ξ to a neighboorhood N_2 of x_2, the curve $P(l_1)$ to a smooth curve in the plane p_2. As ξ is an intersection point such that there is no (possibly one-sided) neighborhood of ξ where the curves $P(l_1)$ and $P(l_2)$ coincide, the same is true for x_2, $T(l_1)$ and l_2 in p_2. Then there is a point x_2 in N_2 on $T(P(l_1))$ where the tangent vector is not parallel to l_2. Consider l_2' which lies in the plane p_2 and goes through the point x_2'. The curve $T(P(l_1))$ intersects l_2' at that point and is transversal to it. Therefore, the same is true for the images $P(l_1)$ and $P(l_2)$, contradicting Condition T$'$. We conclude that any intersection point of $P(l_1)$ and $P(l_2)$ must have a neighborhood, possibly one-sided, where the curves coincide.

Pick two points ψ and ψ' in this neighborhood. Let $y_1 = P^{-1}(\psi) \cup l_1$, $y_1' = P^{-1}(\psi') \cup l_1$, $y_2 = P^{-1}(\psi) \cup l_2$, $y_2' = P^{-1}(\psi') \cup l_2$.

Consider the planar quadrangle $y_1\ y_1'\ y_2'\ y_2$. The images of the edges $[y_1, y_1']$ and $[y_2, y_2']$ coincide by construction.

For every point of l_1 there is a planar neighborhood in the plane p_q of the quadrangle $y_1\ y_1'\ y_2'\ y_2$ where the mapping $P|_q$ is one-to-one because at the points of l_1 fibers are transversal to the plane by construction. Due to the compactness of l_1, there is an ϵ such that there are no such points that are closer than $\epsilon/2$ to $[y_1, y_1'] \subset l_1$. Pick a point a on $[y_1', y_2']$ which is closer to l_1 than ϵ. Denote points on $[y_1, y_2]$ and $[y_1', y_2']$ which are at the distance $\epsilon/2$ from y_1 and y_1' as b_1 and b_1'. Denote the quadrangle y_1, y_1', b_1, b_1' as M_1. In the same way we can construct the quadrangle M_2 near l_2. We can choose M_1, M_2 and a in such a way that $P(M_2) \subset P(M_1)$ and $P([y_1, a]) \in P(M_2)$.

Consider the line segments $[y_1, y_2']$ and $[y_1, a]$.

The image of $[y_1, y_2']$ is a curve connecting $P(y_1)$ and $P(y_2') = P(y_2)$. $P([y_1, y_2'])$ crosses $P([b_2, b_2'])$ at a unique internal point $P(c)$ because the mapping of M_2 to the picture plane is one-to-one and $[y_1, y_2']$ intersects $[b_2, b_2']$ at one interior point c. $P(]c, y_2'[)$ lies in the interior of $P(M_2)$ and connects two points on the boundary. $P(]y_1, a[)$ lies in the interior of $P(M_2)$ and connects two points on the boundary.

We can easily show that $M_1 \setminus [y_1, a]$ is separated by $[y_1, a]$ into two sets S and S' with the following property: there is a path connecting two points of M_1 and lying inside M_1 that doesn't intersect $[y_1, a]$ if and only if the points belong to the same set. Define these sets to be the sets of points on the different sides of the line (y_1, a) in the plane of the quadrangle. Consider two points s and s' from different sets and a connecting path $\gamma(t)$. Choose a unit normal vector $\mathbf{n}$ to the line (y_1, a). Let $\mathbf{d}$ be the vector from y_1 to $\gamma(t)$. Consider the scalar product $(\mathbf{d}, \mathbf{n})$. If it is negative for s, it should be positive for s'. Therefore, there is a point on the curve where it is zero (continuity) and so the curve intersects $[y_1, a]$.

This property transfers to $P(M_1)$ and to $P(M_2)$, due to the inclusion $P(M_2) \subset P(M_1)$. The point $P(c)$ belongs to $P([b_2, b_2'])$. As $P([b_2, b_2'])$ is part of the boundary, and the only boundary points of $P([y_1, a])$ are $P(y_1)$ and $P(a)$, they don't intersect, as $P|_{M_2}$ is one-to-one. Therefore, $P([b_2, b_2'])$ lies entirely in one of the sets $P(S)$ or $P(S')$. It also lies in the same set as $P(b)$ because $P(b)$ and $P(b_2)$ can be connected by a path that doesn't cross $P([y_1, a])$. The point $P(y_2')$ lies in the other set, as y_1 and b_1 are in different sets and $P(y_1') = P(y_2')$. We conclude that $P([y_1, a])$ should intersect $P([c, y_2'])$ at an interior point. This intersection is a crossing: take interior points s' and s of $P(S')$ and $P(S)$ lying on the curve $P(y_1, a])$. There are neighborhoods of these points that are contained in $P(S')$ and $P(S)$ respectively. Thus, if the perturbation doesn't take s' and s of the curve out of these neighborhoods and doesn't take the part of the curve between s and s' out of $P(M_2)$, then the intersection is not eliminated. This is the case for sufficiently small perturbations, because the part of the curve between s and s' lies entirely inside $P(M_2)$. The same is true for $P([c, y_2'])$. Therefore, $P(y_1, a])$ and $P([c, y_2'])$ violate Condition T$'$, and we get a contradiction. Therefore, all fibers should have zero curvature, i.e.all fibers should be line segments.

■

Lemma 3.9 *Given Conditions L' and T', all fibers are coplanar.*

Proof.

Assume the that the opposite is true. In the same way as in Lemma 3.2.3 we construct two continuous parts of non-coplanar fibers f_1 and f_2 inside a ball which lies completely in V. Choose two points x_1 and x_2, on fibers f_1 and f_2 respectively.

Choose a point x on $[x_1, x_2]$ so that if we draw a plane p through the fiber going through x and $[x_1, x_2]$ the fibers f_1 and f_2 won't lie in the plane.

The tangent vector to the fiber is a differentiable function of the point in $\mathbf{R}^3$. Therefore, the angle between a fixed direction and the fiber $F(x(t))$, where $x(t)$ is a parametrization of $[x_1, x_2]$, is a differentiable function. We can choose the direction so that it is not a constant. Then there is a line segment inside $[x_1, x_2]$ where the derivative of the angle with respect to t has a constant sign. Then on this line segment all fibers have different angles with the fixed direction.

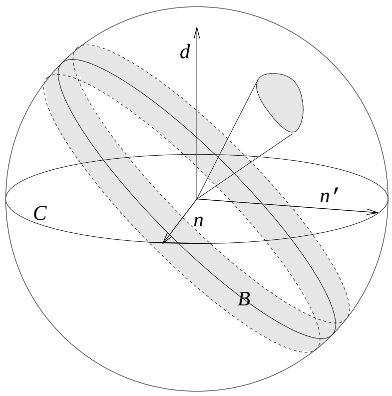


Figure 3.7: Unit sphere construction from Lemma 3.9 □

In addition, we can choose x_1 and x_2 close enough to ensure that no fibers along $[x_1, x_2]$ lie in the plane p' which is perpendicular to p and goes through $[x_1, x_2]$. The following construction on the unit sphere where points correspond to the directions of the fibers helps to explain why we can make such choice of x_1, x_2, p and p'.

The directions of fibers fill a closed area A on the unit sphere. This area doesn't include the point corresponding to the direction of $[x_1, x_2]$: x_1 and x_2 belong to different fibers. Let $\mathbf{d}$ be a unit vector corresponding to the direction of (x_1, x_2), with an arbitrary choice of orientation.

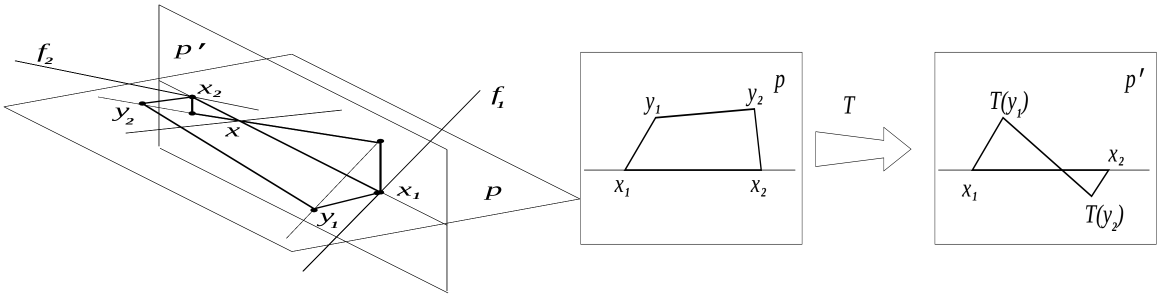


Figure 3.8: Construction from Lemma 3.9: the image of the quadrangle $x_1\ x_2\ y_2\ y_1$ is twisted. □

Possible directions of the normals $\mathbf{n}$ and $\mathbf{n}'$ of the planes p and p' form a circle C in the plane perpendicular to $\mathbf{d}$ with arbitrary choice of orientation.

Directions perpendicular to $\mathbf{d}$ form a "band" B –a union of the large circles of the sphere. This band doesn't include the circle of possible directions of p, p' because A doesn't include $\mathbf{d}$. As a result, the intersection of B and C can be made arbitrarily small. This intersection consists of two connected parts coresponding to sets of opposite directions. By making them sufficiently small, we can guarantee that the directions perpendicular to those in $C \cap B$ are not in B. Then we can choose $\mathbf{n}$ in $C \cap B$ and $\mathbf{n}'$ in C perpendicular to $\mathbf{n}$. This will satisfy the conditions stated above.

Choose a half-space with respect to the plane p and project the parts of the fiber contained in this half-space into p. We can always choose p in such a way that these projections are on different sides of (x_1, x_2): the point corresponding to n in the circle C should be between the points of intersection of large circles corresponding to the perpendiculars to the directions of f_1 and f_2 with C.

The plane p' is transversal to the fibers in some neighborhood N' of $[x_1, x_2]$. Therefore, the mapping of this neighborhood to the picture plane will be one-to-one and smooth. On the plane p we can choose a neighborhood N of $[x_1, x_2]$ which maps into the image of the neighborhood N'.

Choose the point $y_1 \in N$ on the projection of the fiber f_1 into the plane p and $y_2 \in N$ on the projection of the fiber f_2, so that y_1 and y_2 are on one side of (x_1, x_2). Choose the pairs of points in such a way that the intervals $[y_1, x_1]$ and $[y_2, x_2]$ don't intersect.

Let's show that there is a neighborhood of $[x_1, x_2]$ such that all fibers that lie in the plane p and intersect this neighborhood, also intersect $[x_1, x_2]$. Suppose for any ϵ there is a fiber f that lies in p, intersects (x_1, x_2) at a point outside $[x_1, x_2]$ and goes through a point which is closer than ϵ to $[x_1, x_2]$. By construction, there is a fiber f' intersecting but not containing $[x_1, x_2]$ that lies in the plane p. By choosing ϵ small we can make the lines containing f and f' intersect arbitrarily close to

$[x_1, x_2]$, i.e. inside V. This is impossible – fibers cannot intersect. Therefore, we can choose y_1 and y_2 in such a way that the only fibers in p intersecting the quadrangle $y_1\ x_1\ x_2\ y_2$ intersect $[x_1, x_2]$. The mapping $T = (P|_{N'})^{-1} \circ P$ maps this quadrangle to the plane p'. Consider the fiber going through y_1. If y_1 is sufficiently close to x_1, the angle between the fiber going through y_1 and the plane p is close to the angle between f_1 and p.

The plane p subdivides the space into two half-spaces H and H'. Let H be the half-space that contains part of the fiber f_1 whose projection into the plane p lies in the same half-plane with respect to (x_1, x_2) as y_1 and y_2. By our choice of p the projection of $f_2 \cup H$ lies in the other half-plane.

Therefore, the projection into p of the part of $F(y_1)$ which lies in H will be close to the projection of f_1. By choosing y_1 close enough to $[x_1, x_2]$, we can guarantee that it lies entirely in the same half-plane of p as the projection of f_1, and doesn't intersect $[x_1, x_2]$. As $p' \perp p$, $F(y_2) \cap p' = T(y_2)$ projects to a point on $[x_1, x_2]$. We conclude that $T(y_1)$ is not in H and is, therefore, in H'. In the same way we can show that $T(y_2)$ is in H. Thus, $T(y_1)$ and $T(y_2)$ are in different half-planes of p' with respect to (x_1, x_2). $T([y_1, y_2])$ intersects $[x_1, x_2]$ at the point x, as the fiber $F(x)$ lies in p and doesn't belong to (x_1, x_2). If there is a path inside the image which goes from a point of the image on one side of the line (x_1, x_2) to a point on the other side, there should be a point $z \in [x_1, x_2]$ on the path, because no point in $x_1\ x_2\ y_2\ y_1$ is mapped to a point of (x_1, x_2) outside $[x_1, x_2]$. Therefore, the set $(x_1\ x_2\ y_2\ y_1)$ is subdivided into two disconnected parts by $[x_1, x_2]$.

Let D be an open neighborhood of $T(x_1\ x_2\ y_2\ y_1)$ with $(x_1, x_2) \setminus [x_1, x_2]$ excluded. It has the same property with respect to $[x_1, x_2]$, and $T(y_1, y_2)$ is contained in the interior of D. As $F(x_1) \not\subset p$ and $F(x_2) \not\subset p$, there are neighborhoods of the endpoints of $[x_1, x_2]$ in (x_1, x_2) that don't contain any points of of fibers that lie entirely in p. Therefore, we can pick two points x_1' and x_2' on $[x_1, x_2]$ that don't coincide with x_1 or x_2 and such that for all $x \in [x_1, x_2] \setminus [x_1', x_2']$ $F(x) \not\subset p$.

Define $D' = D \setminus ([x_1, x_2] \setminus (x_1', x_2')$. D' has the same property as D. Consider perturbations of $[x_1, x_2]$ and $T(y_1, y_2)$. Sufficiently small perturbation doesn't move $[x_1', x_2']$ out of a band around (x_1, x_2). It also doesn't move x_1 and x_2 into D', and the endpoints of $T(y_1, y_2)$ out of their neighborhoods that have empty intersection with the band $T(y_1, y_2)$. Then perturbed $[x_1, x_2]$ retains the dividing property for D' and the endpoints of $T(y_1, y_2)$ are still in different sets. Therefore, the intersection is preserved under small perturbations and is a crossing. This contradicts Condition T$'$. We conclude that the fibers should be coplanar.

This is sufficient (Lemma 3.3) to prove that the fibers of the projection P coincide with the fibers of a perspective projection.

3.2.5 Decomposition

We have proved in previous sections that various sets of structural requirements result in the following property for projections:

Theorem 3.1 *Given Condition H from Section 3.2.2, or Conditions L and T from Section 3.2.3, or Conditions L' and T' from Section 3.2.4, a projection P satisfying these conditions should have fibers that are subsets of the fibers of a perspective projection Π.*

From this theorem it immediately follows that we can decompose the projection P in two ways, described in the following

Corollary 3.2 *Assume that the conditions of Theorem 3.1 hold. Let Π be the perspective projection coresponding to P.*

1. *Π is a central projection with center C. Assume, in addition, that the region V lies entirely in one half-space with respect to a plane going through the center of Π. Then P can be decomposed in two ways:*

   – as a composition of a central projection Π_{plane} into a plane followed by a smooth one-to-one transformation T_{plane} of the plane, or

   – as a central projection Π_{sphere} into a sphere with center C followed by one-to-one smooth mapping T_{sphere} of the sphere into the picture plane.

2. *Π is a parallel projection. Then P can be decomposed as a composition of an orthogonal central projection Π_{plane} into a plane followed by a smooth one-to-one transformation T_{plane} of the plane.*

It is important to point out that the theorems were proven for a path-connected domain V. Therefore, if all the objects in a picture that are subject to the structural conditions described above can be separated into sets with each set lying in a separate path-connected domain (for example, foreground, middleground and background) a different perspective projection can be chosen in the decomposition for each set. It is interesting to see actual realization of similar principles in some Western and Japanese paintings (Section 2.6, Section 4.1).

3.3 Nonstructural Requirements

As we have mentioned in the beginning of this chapter, we will use nonstructural requirements to choose projections in the class defined by the structural requirements. Due to the more quantitative nature of these requirements, the appropriate way to formalize them is different. Rather than defining which images satisfy the requirement and which images don't, we define *error functions* that give us a measure of a nonstructural distortion. Formally speaking, these are functions on the set of all possible images. Functions of this type are difficult to deal with. Therefore, it will be more convenient to define error functions on the point of the picture plane, in such a way that the error function at each point gives an estimate of the maximal distortion of a small image at this point.

In this section we consider only projections on a path-connected open domain V satisfying Theorem 3.1. Therefore, for each projection we can define the center of projection (unless all fibers are parallel) and decompositions as described in Section 3.2.5. We will assume that the fiber structure of the projection is fixed, i.e. projections Π_{sphere} and Π_{plane} in the decompositons are fixed.

3.3.1 Definitions

Let's establish some notation for the projections that satisfy the conditions of the theorem.

$$\begin{tikzcd}
V \subset \mathbf{R}^3 \arrow[r, "\Pi_{plane}"] \arrow[d, "\Pi_{sphere}"'] \arrow[dr, "P"] & W_{plane} \subset \mathbf{R}^2 = p \arrow[d, "T_{plane}"] \\
W_{sphere} \subset \mathbf{S}^2 \arrow[r, "T_{sphere}"] & \mathbf{R}^2 = \pi
\end{tikzcd}$$

We will consider projections $P: V \subset \mathbf{R}^3 \to \mathbf{R}^2$, from an open connected domain V into the *picture plane* π, which are compositions of the projection Π_{plane} from the center O into the *intermediate plane* p and a mapping $T_{plane} : W_{plane} \subset \mathbf{R}^2 \to \mathbf{R}^2$, $W_{plane} = \Pi_{plane}(V)$, We can choose the plane p so that that the distance from O to the plane is L.

P can be also represented as a composition of central projection from O into the sphere of radius L with center at O (*intermediate sphere*) $\Pi_{sphere} : V \subset \mathbf{R}^3 \to \mathbf{S}^2$ and some mapping

$\mathsf{T} : W_{sphere} \subset \mathbf{S}^2 \to \mathbf{R}^2$, $W_{sphere} = \Pi_{sphere}(V)$, We will assume that the image of V in the sphere belongs to a hemisphere.

If the fiber structure of the projection P is that of a parallel projection, we will use only the first decomposition.

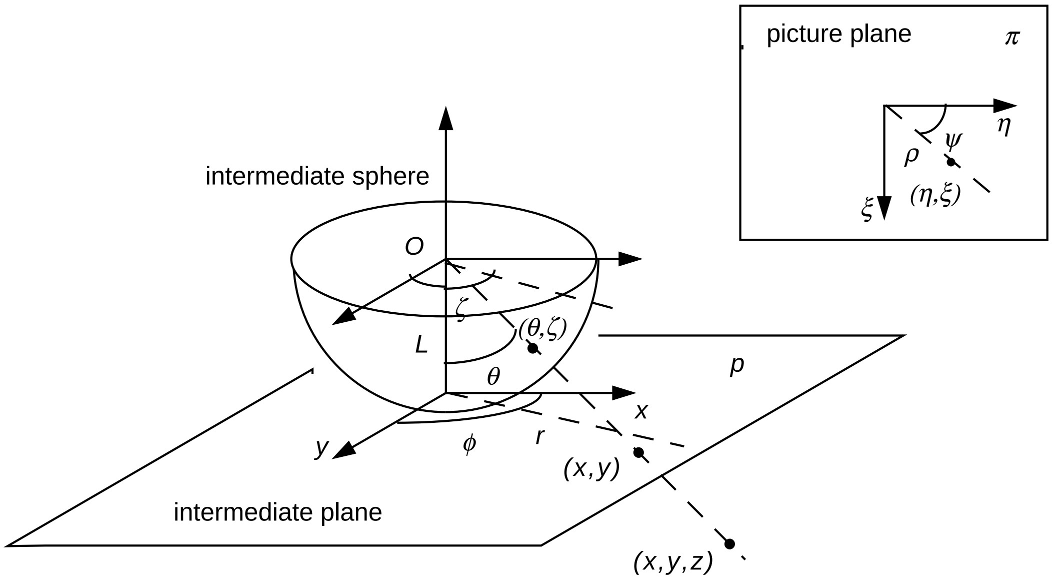


Figure 3.9: Coordinate systems. □

Let's introduce rectangular coordinates (x,y) and polar coordinates (r, ϕ) on the plane p; rectangular coordinates (η, ξ) and polar coordinates (ρ, ψ) in the plane π. On the sphere we will choose angular coordinate system (θ, ζ), and local coordinates in the neighborhood of a fixed point (θ_0, ζ_0): $u = L(\theta - \theta_0), v = L(\zeta - \zeta_0)\sin\theta_0$ (Figure 3.9).

The correspondence between T_{plane} and T is given by the stereographic projection $\mathbf{S}^2 \to \mathbf{R}^2$: $\phi = \zeta$, $r = \tan\theta$.

3.3.2 Zero Curvature Condition

The requirement from Section 2.4.2 is perhaps the easiest to formalize:

> **Zero-curvature condition.** Images of line segments should have zero curvature at each point.

It would be desirable to use a measure of deviation from this condition similar to Watt's measure (Section 2.4.2). Watt's measure, however, is defined for a particular viewing position, while our measure has to be independent of the viewing conditions. A natural candidate for such measure is the curvature of the image of a straight line at a point.

As the curvature has the dimension of inverse length, we need a scaling factor to make the measure independent of the size of the picture. As a scaling factor we can choose some characteristic length. The most appropriate choice appears to be either a typical size of a line segment that can be found in a picture, or the size of the picture. The later choice gives a conservative measure, assuming the worst case: there are images of straight lines crossing the whole field of view.

Let's qualitatively compare this measure with Watt's measure. Watt's measure depends on the viewing conditions, and therefore, we have to make assumptions about these conditions. Due to the spatial limits of acute vision described in Section 2.4.1, we can assume that the picture occupies no more than $40°$ of the field of view. Most pictures are viewed from a distance when their angular size is not less then $5°$. Given these estimates, we have the viewing distance l equal roughly to 4 to 20 times the size of the picture. Let $l = k_1 R_0$, where R_0 is the size of the picture. The area between a chord of length d and a segment of a curve with radius of curvature r is to the first order of approximation $d^3/6r$. Assuming d to be of the order of magnitude of the size of the picture, $d = k_2 R_0$, where $k_2 < 1$ and close to one by the order of magnitude. The Watt's functional can be written as $C(k_2^3/6k_1^2)R_0/r$, where C is less than unity, because we used a chord instead of a least-squares lines. Therefore, it is roughly proportional to the curvature times the size of the picture, with the coefficient of the order 10^{-2}.

If we consider the decomposition $P = T_{plane} \circ \Pi_{plane}$, we can observe that Π_{plane} satisfies the zero-curvature condition. Therefore, we have to consider only the mapping T_{plane}. We will denote components of $T_{plane}(x, y)$, which is a point in the picture plane, by $(\eta(x, y), \xi(x, y))$

The curvature depends not only on the point but also on the direction of the line whose image we are considering. As an error function for the zero-curvature condition at a point x we will use an estimate of the maximum curvature of the image of a line going through x.

Let's estimate the maximum curvature of the image of a line defined by $\mathbf{x}+\mathbf{v}t$ in the intermediate projection plane, in the point $T_{plane}(\mathbf{x})$.

By definition, $\kappa = \frac{d\alpha}{ds}$, where $d\alpha$ is the change in the direction of the tangent vector of the curve along the segment ds.

Let $\mathbf{w}$ be the tangent of the curve $\gamma = \gamma(t)$: $\mathbf{w} = \frac{d\gamma}{dt}$. Then

$$|\kappa| = \frac{\left|\mathbf{w} \times \frac{d\mathbf{w}}{dt}\right|}{|\mathbf{w}|^3} \leq \frac{|\mathbf{w}|\left|\frac{d\mathbf{w}}{dt}\right|}{|\mathbf{w}|^3} = \frac{\left|\frac{d\mathbf{w}}{dt}\right|}{|\mathbf{w}|^2}$$

In our case the curve is defined by the equation $\gamma(t) = T_{plane}(\mathbf{x}+\mathbf{v}t)$. To estimate the curvature from above, we will estimate the maximum of $\frac{\left|\frac{d\mathbf{w}}{dt}\right|}{|\mathbf{v}|^2}$ and the minimum of $\frac{|\mathbf{w}|^2}{|\mathbf{v}|^2}$ for a fixed point $\mathbf{x}$. Using $\frac{d}{dt}F(\mathbf{x}+\mathbf{v}t) = (\nabla F, \mathbf{v})$ for a scalar function F, we can write:

$$\begin{aligned}
&\left|\frac{d\mathbf{w}(t)}{dt}\right|^2 = \left|\frac{dw_\eta}{dt}\right|^2 + \left|\frac{dv_\xi}{dt}\right|^2 = (\nabla w_\eta, \mathbf{v})^2 + (\nabla w_\xi, \mathbf{v})^2 \\
\leq\;& (|\nabla w_\eta^2 + |\nabla w_\xi|^2)|\mathbf{v}|^2 = (|\nabla \frac{d\eta}{dt}|^2 + |\nabla \frac{d\xi}{dt}|^2)|\mathbf{v}|^2 \\
=\;& (|\nabla(\nabla\eta, \mathbf{v})|^2 + |\nabla(\nabla\xi, \mathbf{v})|^2)|\mathbf{v}|^2 = (|(\mathbf{v}, \nabla)\nabla\eta|^2 + |(\mathbf{v}, \nabla)\nabla\xi|^2)|\mathbf{v}|^2 \\
\leq\;& (|(\mathbf{e}_x, \nabla)\nabla\eta|^2 + |(\mathbf{e}_y, \nabla)\nabla\eta|^2 + (\mathbf{e}_x, \nabla)\nabla\xi|^2 + (\mathbf{e}_y, \nabla)\nabla\xi|^2)|\mathbf{v}|^4
\end{aligned}$$

Expanding $|(\mathbf{e}_x, \nabla)\nabla\eta| = |\eta_{xx}|^2 + |\eta_{xy}|^2$ and other terms in similar way, we get

$$\left|\frac{d\mathbf{w}(t)}{dt}\right| \leq \sqrt{|\eta_{xx}|^2 + |\eta_{yy}|^2 + 2\,|\eta_{xy}|^2 + |\xi_{xx}|^2 + |\xi_{yy}|^2 + 2\,|\xi_{xy}|^2}\,|\mathbf{v}|^2$$

Now let's estimate $|\mathbf{w}|^2$ from below.

$$\begin{aligned}
&|\mathbf{w}|^2 = (\nabla\eta, v)^2 + (\nabla\xi, v)^2 \\
\leq\;& ((\eta_x^2 + \xi_x^2)v_x^2 + 2(\eta_x\eta_y + \xi_x\xi_y)v_x v_y + (\eta_y^2 + \xi_y^2)v_y^2) \\
=\;& Av_x^2 + 2Bv_y v_x + Cv_y^2
\end{aligned}$$

We can find α such that

$$\cos\alpha = \frac{w_u}{\sqrt{w_u^2 + w_v^2}}, \quad \sin\alpha = \frac{w_v}{\sqrt{w_u^2 + w_v^2}},$$

Thus

$$|\mathbf{w}|^2 = A\cos^2\alpha + 2B\cos\alpha\sin\alpha + C\sin^2\alpha$$

.

Minimizing this expression w.r.t. α, we get

$$|\mathbf{w}|^2 \geq \frac{1}{2}((A+C) - \sqrt{(A-C)^2 + 4B^2})|\mathbf{v}|^2$$

Finally, we get the estimate for the curvature:

$$|\kappa| \leq \frac{\sqrt{|\eta_{xx}|^2 + |\eta_{yy}|^2 + 2\,|\eta_{xy}|^2 + |\xi_{xx}|^2 + |\xi_{yy}|^2 + 2\,|\xi_{xy}|^2}}{\frac{1}{2}((A+C) - \sqrt{(A-C)^2 + 4B^2})}$$

To simplify calculations, we will use the square of the curvature:

$$K(T_{plane}, x, y) = \frac{|\eta_{xx}|^2 + |\eta_{yy}|^2 + 2\,|\eta_{xy}|^2 + |\xi_{xx}|^2 + |\xi_{yy}|^2 + 2\,|\xi_{xy}|^2}{\frac{1}{4}((A+C) - \sqrt{(A-C)^2 + 4B^2})^2} \tag{3.1}$$

If we set $K(T_{plane}, x, y) = 0$, we see that all of the second derivatives of η and ξ should be equal to zero, and so T_{plane} should be a linear transformation. This coincides with the "main theorem of affine geometry", which says that the only transformations of the plane which map lines into lines are linear transformations.

3.3.3 Direct View Condition

Perceptual distortions of the shape of rectangular corners, spheres, bodies, cylinders, etc. (Section 2.3) don't occur when we use narrow-angle perspective projection, i.e. when the principal ray goes through the object and the angular subtense of the object when viewed from the projection point is small. Therefore, the following condition can be considered a generalization of a number of nonstructural perceptual requirements:

> **Direct-view condition.** It is desirable for the projection to be close in a neighborhood of each point to the perspective projection into a plane perpendicular to the fiber line going through this point.

First, we consider the case when the fiber structure of the mapping coincides with the fiber structure of a central projection.

We can observe that the projection Π_{sphere} into the sphere is locally close to the projection into the tangent plane of the sphere which is perpendicular to the fiber lines of P.

Therefore, if we use the decomposition $P = \mathsf{T} \circ \Pi_{sphere}$ we have to construct the mapping T_{sphere} which is locally as close to a similarity mapping as possible. Formally, it means that the differential of the mapping T_{sphere}, which maps the tangent plane of the sphere at each point x to the plane $T_{f(x)}\mathbf{R}^2 = \mathbf{R}^2$ coinciding with the picture plane π at the point $f(x)$, should be close to a similarity mapping. The differential $DT_{f(x)}$ can be represented by the Jacobian matrix J of the mapping T_{sphere} at the point x.

In coordinate form this can be written as

$$J_{f,x} = \alpha^2 I \tag{3.2}$$

for some $\alpha \neq 0$.

For local coordinates (u, v) on the sphere the matrix $J_{f,x}$ can be written as

$$J_{f,x} = \begin{pmatrix} \eta_u & \eta_v \\ \xi_u & \xi_v \end{pmatrix}$$

Then for (3.2) to be satisfied it is necessary and sufficient that

$$|J_{f,x}| \neq 0, \quad \eta_u = \xi_v, \quad \eta_v = -\xi_u$$

Note that these equations formally coincide with Cauchy-Riemann conditions. If we consider the sphere and the plane to be complex manifolds, then the equations above are necessary and sufficient conditions for T_{sphere} to be a conformal mapping.

For global spherical coordinates we have

$$\eta_u = \eta_\theta, \qquad \eta_v = \eta_\zeta \frac{1}{\sin\theta} \tag{3.3}$$

$$\xi_u = \xi_\theta, \qquad \xi_v = \xi_\zeta \frac{1}{\sin\theta} \tag{3.4}$$

We will say that mappings T_{sphere} and P satisfy direct view condition, if they satisfy the conditions 3.3.

We also need of measure of deviation from the direct view condition. A nondegenerate linear transformation J is a similarity transformation if and only if $|Jw|/|w|$ doesn't depend on w. If this ratio depends on w, then we define the direct view error function to be

$$D(T_{sphere}, \theta, \zeta) = \left| \max\left(\frac{|Jw|^2}{|w|^2}\right) / \min\left(\frac{|Jw|^2}{|w|^2}\right) - 1 \right| \tag{3.5}$$

$D(T_{sphere}, \theta, \zeta)$ can be used as the measure of "non-directness" of the transformation at the point.

This measure has a simple geometric meaning. Suppose we have a small sphere at some point in space. The image of the sphere on the intermediate sphere is a small disk. Then this disk is mapped by T_{sphere} into the picture plane. Ideally, this image should be a circle, and as it was noted in Section 2.3.2, the aspect ratio of the image is a reasonable measure of distortion in this case. The functional that we have defined is the deviation from 1 of the aspect ratio of the image of a small sphere. This measure can be also related to the maximal offending angle in the Perkins' laws (Section 2.3.1).

Let's find the explicit expression for $\min \frac{|Jw|^2}{|w|^2}$ and $\max \frac{|Jw|^2}{|w|^2}$. We will use the following notation:

$$\begin{aligned} E &= (\eta_u)^2 + (\xi_u)^2, \\ F &= \eta_u \eta_v + \xi_u \xi_v, \\ G &= (\eta_v)^2 + (\xi_v)^2. \end{aligned}$$

Then $|Jw|^2 = Ew_u^2 + 2Fw_v + Gw_v^2$.

This expression is identical to the expression in Section 3.3.2 so we immediately get

$$\begin{aligned} \min \frac{|Jw|^2}{|w|^2} &= \frac{1}{2}((E+G) - \sqrt{(E-G)^2 + 4F^2}) \\ \max \frac{|Jw|^2}{|w|^2} &= \frac{1}{2}((E+G) + \sqrt{(E-G)^2 + 4F^2}) \end{aligned}$$

Finally, for the direct projection error function D we get

$$D(T_{sphere}, \theta, \zeta) = \frac{(E+G) - \sqrt{(E-G)^2 + 4F^2}}{(E+G) + \sqrt{(E-G)^2 + 4F^2}} - 1 \qquad (3.6)$$

Using the correspondence between the intermediate sphere and the plane we can write D as a function of T_{plane}, x and y.

The only transformations that satisfy the direct view condition exactly are the conformal mappings from the sphere to the plane.

As it can be easily shown, if T_{plane} is linear, corresponding T_{sphere} is not conformal.

Now we consider the case of projections P with parallel fibers. In this case, the projection to the intermediate plane has the direct view property. All conformal transformations of the plane will preserve the direct view property. Then the conditions 3.3 become standard Cauchy-Riemann equations.

The same measure of deviation from the direct view condition (3.5) can be used, but in this case we take the local coordinates (u, v) to be the same as (η, ξ).

In this case any composition of orthogonal and similarity transformation of the plane will be conformal and linear and will satisfy the direct view condition. But any projection $T_{plane} \circ \Pi_{plane}$ where T_{plane} is a composition of orthogonal and similarity transformation is equivalent to an orthogonal parallel projection up to a scaling factor. Therefore, we arrive at an important conclusion:

> The direct view condition can be satisfied simultaneously with the zero-curvature condition within the class of projections described by Theorem 3.1 only if the projection coincides with an orthogonal parallel projection to the plane up to a similarity transformation of the plane.

3.3.4 Error Functionals

In view of the statement in the previous subsection, we can see that only orthogonal parallel projection satisfies both the direct-view and zero-curvature conditions simultaneously. It is quite obvious that parallel projection is not a satisfactory choice in many cases. In Section 5.2 we examine some reasons for that. Therefore, we are also interested in projections that have the fiber structure of a central projection.

Since for central projections the direct-view and zero-curvature conditions cannot be satisfied exactly, we should find a way to define the optimal tradeoff between these two types of distortions. In order to do this, it is convenient to use global error functionals which indicate the quality of the projection with respect to a given condition.

In order to estimate the total error for a projection, we have to choose a measure and integrate the error function over the domain of the T_{plane} (we will assume that the direct projection error function is converted to the intermediate projection plane coordinates: $D = D(T_{plane}, x, y)$.) The most conservative way to define the curvature and direct projection functionals is to use L_∞ measure, which results, for continuous D and K, in

$$\mathcal{K}[T_{plane}] = sup_{x,y} K(T_{plane}, x, y) \quad \mathcal{D}[T_{plane}] = sup_{x,y} D(T_{plane}, x, y). \tag{3.7}$$

Other possibilities are, for example, L^p measures:

$$\mathcal{K}[T_{plane}] = \left(\int\int_{x,y} K(T_{plane}, x, y)^p dxdy\right)^{\frac{1}{p}}, \mathcal{D}[T_{plane}] = \left(\int\int_{x,y} D(T_{plane}, x, y)^p dxdy\right)^{\frac{1}{p}}. \tag{3.8}$$

3.3.5 Optimization Problem

Initially we don't make any assumptions about the form of the functional that we choose, except that it should be some functional norm in the space of smooth functions. Assuming T_{plane} to be smooth, $\mathcal{K}[T_{plane}] = 0$ if and only if $K(T_{plane}, x, y) = 0$ almost everywhere. The same is true for $\mathcal{D}$. Then the only functions satisfying $K[T_{plane}] = 0$ are linear functions and the only functions satisfying $D[T_{plane}] = 0$ are those derived from conformal mappings of the sphere onto the plane. Therefore, we cannot make both functionals equal to zero for any choice of the norm. So we want for a given value of one functional to find a function T_{plane} such that the other functional has the smallest possible value.

$$\text{minimize} \quad \mathcal{K}[T_{plane}] = \nu, \text{ subject to } \quad \mathcal{D}[T_{plane}] = \mu \tag{3.9}$$

or,

$$\text{minimize} \quad \mathcal{D}[T_{plane}] = \mu, \text{ subject to } \quad \mathcal{K}[T_{plane}] = \nu \tag{3.10}$$

We want to find a family of functions $T_{plane}(\mu)$ or $T_{plane}(\nu)$ solving the constrained optimization problems stated above. In each case we get a function $\nu(\mu)$ or $\mu(\nu)$.

We can reduce the constrained problem stated above to an unconstrained one in the following way:

Consider a linear combination of the functionals $\mathcal{F}[T_{plane}] = \lambda\mathcal{K}[T_{plane}] + (1-\lambda)\mathcal{D}[T_{plane}])$. If T_{plane_0} minimizes $\mathcal{F}[T_{plane}]$ for a fixed λ, then $\mathcal{K}[T_{plane_0}]$ is minimal for the fixed value of $\mathcal{D}[T_{plane}] = \mathcal{D}[T_{plane_0}]$. As we change λ, the value of $\mathcal{D}[T_{plane_0}]$ will increase up to the maximal value, which is attained when $\lambda = 1$. Therefore, instead of minimizing $\mathcal{K}[T_{plane}]$ for fixed values of $\mathcal{D}[T_{plane}]$, we can minimize the functional $\mathcal{F}[T_{plane}]$ for λ varying from 0 to 1. The same is true for minimizing $\mathcal{D}[\mathsf{T}]$ for fixed values of $\mathcal{K}[T_{plane}]$. The problem can be restated as follows:

$$\text{minimize} \quad \mathsf{F} = \lambda\mathcal{D} + (1-\lambda)\mathcal{K} \quad \text{for } \lambda \in [0,1]$$

We also have to specify the boundary conditions in order to make the problem well-defined. This can be done by fixing the frame of the picture, that is, the values of T_{plane} on the boundary of W.

To ensure that the mapping is one-to-one and the inverse is smooth, the Jacobian of the mapping should not be zero anywhere.

It would also be desirable not to increase one distortion *anywhere* in the image beyond the level that is possible when the other distortion is 0. It is a somewhat restrictive requirement and potentially can be relaxed if it is possible to obtain considerable reductions in maximal distortion at the expense of a small increase of of small values of distortion. It is unlikely for the distortion of curvature — if the curvature is 0 in some part of the image, small increases in the curvature are likely to be observable, because of the threshold nature of the perception of straightness [Watt et al., 1987]. It is potentially possible to gain some improvement by relaxing the same requirement for the direct view condition: small distortions of shape are undetectable for most objects.

Mathematically the restrictions above can be formulated as

$$K(x,y) < K_{conf}(x,y), \quad D(x,y) < D_{lin}(x,y),$$

where $K_{conf}(x,y)$ is the value of the error function at the point (x,y) for a conformal mapping $(D(x,y) \equiv 0)$, and $D_{lin}(x,y)$ is the value of the error function for the identity mapping (the only one which has $K(x,y) \equiv 0$.) $K_{conf}(x,y)$ depends on the choice of the conformal mapping – this choice can be arbitrary, because all conformal mappings result in $D(x,y) \equiv 0$, and any is sufficient if condition (3.3.5) holds.

In Section 4.3, we will explore a particular case of this optimization problem for T_{plane} with central symmetry.

It is also interesting to consider weighted error functions, that is, $w_K(x,y)K(T_{plane},x,y)$ and $w_D(x,y)D(T_{plane},x,y)$, where $w_K(x,y)$ and $w_D(x,y)$ are nonnegative weights.

The weights of the error functions can be interpreted as the relative importance of a particular type of perceptual error close to the given point in the image. It can be user-specified or calculated using information about the object space.

In this case the total error can be defined as the norm $||w_K(x,y)K(T_{plane},x,y) + w_D(x,y)D(T_{plane},x,y)||$, and the optimal transformation.

$K - D$ **dependencies.** As we have shown, the original optimization problem can be reduced to minimization of one functional F. This functional doesn't have a good perceptual interpretation. A good way to characterize the optimal family of transformation is by the $(\mathcal{K},\mathcal{D})$ pairs corresponding to each member of the family.

If we solve the optimization problem (3.3.5), we get a one-parametric family parametrized by λ, which defines a parametric curve $(\mathcal{K}(\lambda),\mathcal{D}(\lambda))$. This curve allows to see to what extent distortions can be eliminated simultaneously.

It is necessary to emphasize that the optimization problem formulated in this section is not equivalent to the task of minimizing distortions. It is a specific formalization of this task with relatively low dependence on the scene contents. Projections can be improved perceptually by introducing more dependence on scene contents. The tradoff is between the difficulty of the construction of an appropriate projection and the quality of the resulting image. We tried to choose a path in between, in which the choice of projections is limited, but some improvement can still be achieved.

Chapter 4

Construction of Projections

4.1 Previous Work

A number of reasons lead to the constructions of picture projections that were different from the linear perspective. There were three main motivations: perceptual [Raushenbakh, 1986], [Reggini, 1975]), [Stark, 1928], artistic ([Ernst, 1976], [Inakage, 1991], [Flocon and Taton, 1963], and technical [Goncharenko et al., 1979] considerations.

4.2 Perceptually-based systems

Reggini's perspective. Reggini's system of perspective [Reggini, 1975] is based on Thouless' formula ([Thouless, 1931a], [Thouless, 1931b], [Thouless, 1932]), rewritten as

$$\frac{p}{r} = \left(\frac{t}{d}\right)^{1-i}, \tag{4.1}$$

where t is the distance from the projection point to the object, d is the distance from the projection point to the plane and i is a constant between 0 and 1. The projection of a point is computed in the following way: let r be the distance from the point to the principal ray (the ray through the projection point perpendicular to the projection plane.) Let t be the distance from the point to the projection plane. A point X in three-dimensional space can be specified by three numbers (r, t, ϕ) in the cylindrical coordinate system with the origin coinciding with the projection point and the axis coinciding with the principal ray.

In this system the rays (sets of points mapping to a single point in the projection plane) are curved. In the previous chapter we have discussed the problems that are associated with this curvature.

The main goal of Reggini's system is to provide for changing parallel foreshortening; as a result, close objects have smaller degree of parallel foreshortening. The effective viewpoint, determined by the intersection point of the tangent lines to the rays at a given distance from the picture plane, moves away from the projection plane as the distance to the object increases.

Raushenbakh's system. Raushenbakh's system [Raushenbakh, 1986] is based on an idea similar to Reggini's. Instead of using Thouless' formula, he derives similar curvilinear systems under more general assumptions about the relationship of the apparent size on the projected size. More specifically, Raushenbakh assumes that the apparent size is related to the projected size by the formula

$$s^* = sF(D)$$

where s^* is the apparent size, s is the size of projection onto a fixed plane, D is the distance from the plane to the object. He suggest a technique for measuring $F(D)$.

The result of his derivations is a family of projections; different systems correspond to different tradeoffs between vertical and horizontal deviation of the depicted size from the apparent size.

Both systems that were described above are based on the assumption that the apparent size should be depicted. We should note, however, that the size constancy works also for pictures, although less than for the real scenes. Still there is a difference in the perception of size in pictures and in the real world, and some intermediate values between apparent and projected size might create the most appropriate impression. These systems are also useful for understanding some deviations from linear perspective in paintings and naive art – some of the representations of objects that are consistently present in the pictures, could result from unconscious attempts to depict the apparent size.

We believe that main problem of these systems is that the distortion associated with relative size can be found mostly in wide-angle images, and a number of other distortions , such as the distortion of shape, in these image tend to be more prominent. The systems described above might result in reduction of these distortions but nothing in their construction guarantees that. The important problem of the curvature of images of straight lines is not addressed either.

We don't describe two more systems - Flocon's and Stark's - because we didn't have the original references available and we could not find satisfactory accounts of their work.

4.2.1 Artistic systems

In a sense, any projection system falls into this category, because any system can be used for artistic purposes. We describe several systems that were created specifically for artistic purposes.

Escher's perspective. M. C. Escher is famous for geometric construction in his paintings. He used unuasual perspective constructions in many works, and one of his experiments (Figure 4.1) is particularly relevant to our discussion. From Escher's notes that he had made when he was working on this woodcut, we know that he was trying to create a smooth transition between different views of a building - with the viewing direction changing from horizontal to almost vertical. Resulting construction is similar to the panoramic projection described below and allows to avoid distortion of rectangular corners in all parts of this complicated construction.

Systems described by M. Inakage. A number of projections were described in [Inakage, 1991].

Inakage defines *curvilinear perspective* as a linear projection of the three-dimensional space into a curved surface. This definition implies that the surface is projected into the plane to get the final image. Therefore, this definition is equivalent to the decomposition given by Theorem 3.1.

3D-warping extends the curvilinear perspective, allowing to change the metrics of space, producing curved rays. Any continuous system can be considered an instance of curvilinear perspective combined with 3D-warping.

Specific case of *inverse perspective* is described by the formulas

$$x' = x(\frac{z}{d} + 1); \quad y' = y(\frac{z}{d} + 1), \tag{4.2}$$

Figure 4.1: M. C. Escher. High and Low,lithograph, 1947 □

where (x, y, z) is a point in space (x', y') is it's projection and d is the distance from the "projection point" to the projection plane.

Another example of perspective constructions are the photographic collages by David Hockney [Hockney, 1988].

As we have mentioned in Chapter 2, deviations from linear perspective in the art are probably more common than rigorous perspective constructions. Rejection of linear perspective was one of the theoretical foundations of cubism. The systems that we have described above are just several examples of attempts to use alternative methods of projection.

4.2.2 Technical Projections

Extreme distortions of linear perspective for angles close to 180°, and impossibility of using it for projection angles greater than 180°, prompted several practical solutions, primarily in lens design. For extremely wide-angles (greater than 110-120 degrees,) linear projection is never used because of the distortions and difficulties in creating appropriate rectilinear lenses. Fish-eye lenses are based on projections that are different from linear perspective. All of these projections have straight rays (unless the light propagates in a medium with variable refraction coefficient.)

Several projections used in lens design are described in [Goncharenko et al., 1979]. The fiber structure of all these projections is identical to the central projection –all of the fibers are straight lines intersecting at the center of the lens, which can be called the projection point. Moreover, these projections have axial symmetry. Let p be the axis of symmetry, which corresponds to the ray going through the center of the lens and its focal points. Let β be the angle between the direction from the projection point to a point in space X. We will use a polar coordinate system (r, ϕ) on the projection plane with the origin at the intersection point of the axis and the plane. Then the image of the point X is determined by a function $r = F(\beta)$.

Three projections described in [Goncharenko et al., 1979] are shown in the table below (f is the focal length of the lens.)

name	formula	examples of lenses
orthographic	$r = f \sin \beta$	Fisheye-Nikkor OP (180°)
equidistant	$r = f\beta$	Kowa Fisheye (180°), Olympus Zuiko-Fisheye (180°), Nikkor-Fisheye (220°)
equalized	$r = 2f \sin \frac{\beta}{2}$	Asahi Super Takumar-Fisheye, Minolta MC Fisheye

Another example of a projection that is widely used in photography is the panoramic projection. If the film in a panoramic camera is placed on a cylindrical surface, the resulting projection resembles equidistant projection in one direction and regular linear perspective in the other. Panoramic projection was advocated by [Feininger, 1953] as the only "correct" perspective, on the grounds that parallel lines should always be seen as convergent to a point at infinity.

The projections that we have described in this section were mostly *ad hoc* solutions to the problems of wide-angle images. No consistent effort was made to minimize geometric distortions.

4.3 Axially Symmetric Projection

In general, the optimization problems formulated in Section 3.3.5 are very difficult. In this section we will consider the case of axially symmetric projections. The transformation T_{plane} for these projections can be written in polar coordinates as

$$\rho = \rho(r), \quad , \psi = \phi \tag{4.3}$$

In this case the problem can be reduced to a one-dimensional problem. Unfortunately, even in this form direct solution of the optimization problem is quite difficult. We shall make some estimates of the optimal values of the functional and show that some simple families of projections are close enough to optimal. We shall consider some applications of one of these families in Chapter 5.

4.3.1 Reformulation of the Problem

Using Equations 4.3 for the transformation T_{plane}, we can simplify the expressions for error functions. After substitutions, we obtain

$$K(\rho, r) = R^2 \frac{\frac{3}{r^2}\left(\frac{\rho}{r} - \rho'\right)^2 + \rho''^2}{\left(\frac{\rho^2}{r^2} + \rho'^2 - \left|\frac{\rho^2}{r^2} - \rho'^2\right|\right)^2} = R^2 \frac{\frac{3}{r^2}\left(\frac{\rho}{r} - \rho'\right)^2 + \rho''^2}{\min\left(\frac{\rho^2}{r^2}, \rho'^2\right)^2}$$

where R is the scaling constant.

The expression above is for the transformation $T_{plane} : \mathbf{R}^2 \rightarrow \mathbf{R}^2$. The direct projection functional was derived for $T_{sphere} : \mathbf{S}^2 \rightarrow \mathbf{R}^2$. The formulas relating T_{plane} and T_{sphere} are $\phi = \zeta,\ \ r = \tan\theta$.

The direct view functional has the form

$$\frac{\max(\frac{d\rho}{d\theta}^2, \frac{\rho^2}{\sin\theta})}{\min(\frac{d\rho}{d\theta}^2, \frac{\rho^2}{\sin\theta})} - 1$$

Using $r = \tan\theta$, we get

$$D(r, \rho) = \frac{\max(\rho'^2(1+r^2)^2, \rho^2(1+\frac{1}{r^2}))}{\min(\rho'^2(1+r^2)^2, \rho^2(1+\frac{1}{r^2}))} - 1$$

We can notice that in both cases the dependence on the angular coordinate completely disappeared, so now the problem is one-dimensional.

We should also note that if ρ doesn't depend on θ, the only consistent way to define the boundary conditions is to specify the value of ρ on a circle, i.e. for $r = R$, where R is the size of the picture (we will assume that we use it as the scaling factor in the curvature error function.)

As the transformation T_{plane} is required to be one-to-one and continuous everywhere, $\rho(0) = 0$ and $\rho'(r) \geq 0$ everywhere. In order for the error functions to be finite at 0, $\rho'(0) \neq 0$. In general, it is desirable to have $\rho'(r) > 0$ for all r, because otherwise we get an "infinite compression factor" at the points where $\rho' = 0$.

Moreover, if we want the transformation to be smooth at zero, $\rho''(0) = 0$.

The optimization problem for the *sup*-norm now takes the form

$$\text{minimize} \quad \mathsf{F}[\rho] = \max_{r\in[0,1]} \lambda K(\rho, \rho', \rho'', r), +(1-\lambda)D(\rho, \rho', r) \tag{4.4}$$

$$\text{subject to} \quad \rho(0) = 0,\ \rho(R) = R,\ \rho'(r) > 0, \rho''(0) = 0 \tag{4.5}$$

4.3.2 Limiting Cases

Consider the cases $\lambda = 1$ and $\lambda = 0$. In the first case the optimal solution is $\rho(r) = r$. In the second case we obtain a differential equation

$$\frac{\rho'}{\rho} r\sqrt{1+r^2} = 1$$

which can be easily solved. The solution satisfying $\rho(0) = 0,\ \rho(R) = R$ is

$$\rho(r) = \frac{R^2(\sqrt{1+r^2}-1)}{r(\sqrt{1+R^2}-1)} \tag{4.6}$$

The expression for the composition of the stereographic projection and ρ, which can be obtained by substituting $r = \tan\theta$ into the expression for ρ, is more intuitive:

$$\rho(\tan\theta) = C\tan\theta/2,$$

where C is a normalization coefficient, and results in the following expression for ρ:

$$\rho(r) = \frac{R\tan\frac{1}{2}\arctan r}{\tan\frac{1}{2}\arctan R}$$

This mapping has a simple geometric interpretation (Figure 4.2): first, the plane is projected to the intermediate unit sphere, and then reprojected back from the pole.

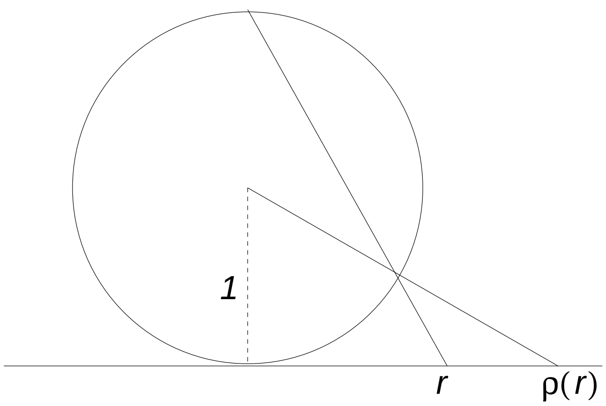


Figure 4.2: Construction of $DP(r)$ □

This mapping has an important role in the rest of the thesis. We will call it *direct view transformation* and denote it $DV(r, R)$.

In Figure 4.3 we compare this transformation with transformations implemented in the fisheye lenses.

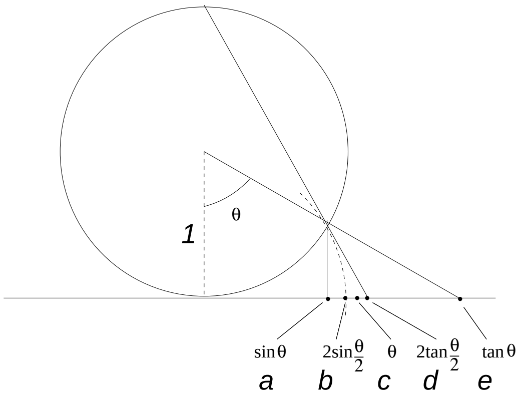


Figure 4.3: Projections: orthographic (a), equalized (b), equidistant (c), $DP(r)$ (d), linear (e). □

4.3.3 Lower Bound of the Optimal Values of F

In this section we will find a lower bound of the optimal values of the functional 4.4.

First, we observe that

$$\begin{aligned}
&\min_{\rho} \mathsf{F}[\rho] = \min_{\rho} \max_{[0\ R]} F(\rho, \rho', \rho'', r) \geq \\
&\min_{\rho} F(\rho(R), \rho'(R), \rho''(R), R) = \min_{\rho} F(R, \rho'(R), \rho''(R), R) \geq \\
&\min_{x_1, x_2} F(R, x_1, x_2, R)
\end{aligned}$$

Therefore, we can estimate the minimal value of the functional from below by minimizing the function

$$F(1, x_1, x_2, R) = \lambda \frac{\frac{3}{R^2}(1-x_1)^2 + x_2^2}{\min(1, x_1^2)^2} + (1-\lambda)\left(\max\left(\left(1+R^2\right)x_1^2, \frac{1}{\left(1+R^2\right)x_1^2}\right) - 1\right)$$

If $\lambda \neq 0$, x_2 should be zero at the point where the minimum is achieved. We can consider the function F of only one variable x_1. Assume $\lambda \neq 0$ (we shall consider the case $\lambda = 0$ case later.) Then we can divide the functional by λ and use $\alpha = (1-\lambda)/\lambda$.

Consider three ranges of x_1: $0 < x_1 \leq 1/\sqrt{R^2+1}$, $1/\sqrt{R^2+1} \leq x_1 \leq 1$, and $x_1 > 1$.

1. $0 \leq x_1 \leq 1/\sqrt{R^2+1}$. In this case the function can be rewritten as

$$F(1, x_1, x_2, R) = \frac{\frac{3}{R^2}\left(1-x_1\right)^2}{x_1^4} + \alpha\left(\frac{1}{\left(1+R^2\right)x_1^2} - 1\right) \tag{4.7}$$

The derivative of the function is

$$F'(1, x_1, x_2, R) = \frac{-\frac{6}{R^2}\left(1-x_1\right)\left(2-x_1\right)}{x_1^5} - \frac{2\alpha}{\left(1+R^2\right)x_1^3} \tag{4.8}$$

In this range of x_1 the derivative is clearly negative. Therefore, if the minimum is attained in this range, it must be attained at the right endpoint $1/\sqrt{1+R^2}$.

2. $x_1 > 1$. In this case the function takes the form

$$F(1, x_1, x_2, R) = \frac{3}{R^2}\left(1-x_1\right)^2 + \alpha\left(\left(1+R^2\right)x_1^2 - 1\right) \tag{4.9}$$

The derivative of the function is

$$F'(1, x_1, x_2, R) = \frac{6}{R^2}(x_1 - 1) + 2\alpha\left(1 + R^2\right)x_1 \tag{4.10}$$

The derivative is positive in this range, therefore, the function is increasing and the minimum can be attained only at the left endpoint $x_1 = 1$.

3. $1/\sqrt{R^2+1} \le x_1 \le 1$. In this case the function and the derivative are

$$F(1, x_1, x_2, R) = \frac{\frac{3}{R^2}(1 - x_1)^2}{x_1^4} + \alpha\left(\left(1 + R^2\right)x_1^2 - 1\right) \tag{4.11}$$

$$F'(1, x_1, x_2, R) = \frac{-\frac{6}{R^2}(1 - x_1)(2 - x_1)}{x_1^5} + 2\alpha\left(1 + R^2\right)x_1 \tag{4.12}$$

At the right endpoint of the interval the derivative is always nonnegative; at the left endpoint it might be positive or negative.

The second derivative of the function in this range is

$$F''(1, x_1, x_2, R) = \frac{\frac{30}{R^2}\,(1-x_1)\,(2-x_1)}{x_1^6} + \frac{\frac{6}{R^2}\,(2-x_1)}{x_1^5} + \frac{\frac{6}{R^2}\,(1-x_1)}{x_1^5} + 2\alpha\left(1+R^2\right) \tag{4.13}$$

It can be easily seen that each term is positive inside the interval. Therefore, the first derivative increases inside the interval and has no more than one root. If it is positive at the left endpoint, there are no roots, otherwise, there is one root.

Using the continuity of the function, we conclude that the minimum is attained at a point in the interval $1/\sqrt{1+R^2} \leq x_1 \leq 1$ for any value of α.

If $\lambda = 0$, it is easy to see that in this case the minmal value of F is 1 and is attained at $x_1 = 1/\sqrt{1+R^2}$.

If the value of D at R is known, we can find $\rho'(r)$.

The $K - D$ curve for the bound is shown in Figure 4.4 for $R = 1$ (field of view 90°.)

4.3.4 Second Bound

The bound that we have derived in the previous section is not very tight, which it can be easily seen for $\lambda = 0$. In this case the minimum of the functional is 0, and the optimal solution is unique (Section 4.3.2.) The values of K in this case are much higher then the bound found in the previous section. In this section we shall make some additional assumptions about ρ and find a better bound for the optimal pairs (K, D) in a narrower class of functions. We believe that this bound is still far from the optimal, it is, however, the best we could obtain. Although the optimization problem that we consider can be easily reduced to a problem of optimal control in the form of Mayer with inequality constraints [Pontriagin, 1962], the necessary conditions that we could find don't yield a good estimate of the optimal values of the functional F, or of the optimal solution, in particular because the second derivative of the Hamiltonian with respect to the control vanishes identically. The elementry estimate that we derive in this section supports the idea that the interpolations derived in the next section are close to optimal but, of course, doesn't prove it.

We make the following assumptions about ρ:

1. D,K are nondecreasing functions of r.

2. $D(\rho, \rho', r)$ is less than for $\rho(r) = r$ for all r, i.e. $D(\rho, \rho', r) < \sqrt{1 + r^2}$

3. $(\rho'/\rho)r\sqrt{1 + r^2} > 1$

4. The function ρ' is concave.

The first assumption is motivated by the desire to avoid larger distortions closer to the center of the picture, i.e. to preserve the same type of distribution of distortions as for the limiting cases: $\rho(r) = r$ and $\rho(r) = DP(r)$.

The second assumption follows from inequalities (3.3.5).

The third assumption means that there is no "overcompression" in the radial direction. Consider an image of a small sphere. For perspective projection it is an ellipse with its major axis oriented along the direction to the center of the picture. For $DP(r)$ the image is a circle. We would expect intermediate projections to produce images of intermediate shapes, i.e. ellipses with major axis still oriented along the direction to the center but with smaller aspect ratio.

The last assumption is motivated by the fact that in a regular perspective image the amount of the shape distortion increases towards the edges, and the compression $(1/\rho')$ should increase.

$K(\rho, \rho', \rho'', r) = 3K_1^2 + \frac{\rho''^2}{\rho'^4}$, where

$$K_1 = \frac{\frac{1}{r}\left(\frac{\rho}{r} - \rho'\right)}{\rho'^2}$$

Due to Assumption 1, K_1 is positive. Note that $\rho''/\rho'^2 = -(1/\rho')'$. As $\rho'' \leq 0$ (Assumption 4), $|\rho''|/\rho'^2 = (1/\rho')'$.

Let $u = \frac{1}{\rho'}$. Then

$$K_1 = \frac{\rho u^2}{r^2} - \frac{u}{r}; \quad K = 3K_1^2 + (u')^2$$

We have assumed that K increases. Suppose $u' \geq 0$ decreases at some point; than K_1 should increase. Therefore, at that point $\frac{dK_1}{dr} > 0$. Differentiating, we get the following expression for the derivative:

$$\frac{dK_1}{dr} = \frac{1}{r}\left(\frac{2}{r}\left(u - \frac{\rho}{r}u^2\right) + u'\left(2\frac{\rho}{r}u - 1\right)\right)$$

It is non-negative if

$$u' \geq \frac{2u}{r}\frac{\frac{\rho}{r}u - 1}{2\frac{\rho}{r}u - 1}$$

Suppose that u' increases on $]r_0, 1]$. Then

$$u' = \max_{[r_0,1]} u' \geq \frac{1}{1-r_0}\int_{r_0}^{1} u' du = \frac{(u(1) - u(r_0))}{(1-r_0)} \tag{4.14}$$

If we choose $[r_0, 1]$ to be the maximal interval on which u' doesn't decrease, it will decrease on an interval $[r_0 - \epsilon, r_0]$. As $u'(r)$ increases on $[r_0, 1]$, the following estimate can be made:

$$u'(1) \geq u'(r_0) \geq \max\left(\frac{u(1) - u(r_0)}{(1-r_0)}, \frac{2u(r_0)}{r_0}\frac{\frac{\rho(r_0)}{r_0}u(r_0) - 1}{2\frac{\rho(r_0)}{r_0}u(r_0) - 1}\right) \tag{4.15}$$

It remains to estimate $u(r) = 1/\rho'(r)$. First, we estimate ρ. Denote $d = \sqrt{D+1}$. Then

$$1 \leq \frac{\rho'}{\rho} r\sqrt{1+r^2} \leq d \tag{4.16}$$

Dividing by $r\sqrt{1+r^2}$ and integrating from r to 1, we get

$$\int_r^1 \frac{dr}{\sqrt{1+r^2}} \leq -\ln\rho \leq d\int_r^1 \frac{dr}{\sqrt{1+r^2}} \tag{4.17}$$

The integral in (4.17) is equal to the function $DP(r)$ defined in Section 4.3.2. Finally, we get

$$DP(r)^d \leq \rho \leq DP(r) \tag{4.18}$$

From (4.16) and (4.18) we obtain

$$\frac{r\sqrt{1+r^2}}{DPd} \leq u = \frac{1}{\rho'} \leq \frac{r\sqrt{1+r^2}}{DP^d} \tag{4.19}$$

These bounds can be directly used in (4.15), because the first bound is an increasing function of

$u(r_0)$, and the second bound is a decreasing function of $u(r_0)$. Due to the Assumption 2, the value of $u(1)$ is known: as $D(\rho', \rho, r)$ is nondecreasing, $D(\rho'(1), 1, 1) = d$, therefore, $\rho'(1) = d/\sqrt{2}$.

We can further improve the lower bound for $\mathcal{D}$ close to 0 in the following way: the estimate (4.14) still holds for maximum of u' on $[r_0, 1]$. Suppose the estimate is maximal for some r_0. Then the value of $\mathcal{K}$ will be greater than the lower estimate for K_1 on the interval $[r_0, 1]$ plus the maximum of the estimate of u'^2.

The plot of the resulting bound is shown in Figure 4.4

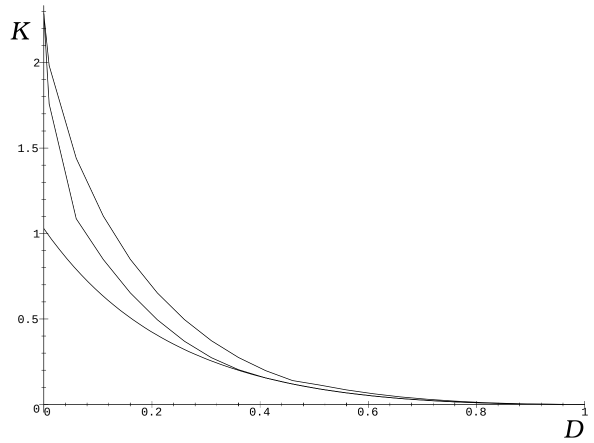


Figure 4.4: Bounds from Section 4.3.4 (upper 2 curves) and Section 4.3.3 (lower curve.) □

We can further narrow down the range of the functions that we are considering, and assume that u' increases. In this case the estimate can be further improved, as the value of $u'(1)$ will be greater than the maximum of (4.14) (Figure 4.4.)

4.3.5 Interpolations

We couldn't find a direct solution of the optimization problem (4.4), but given the estimates of the previous sections, we can evaluate interpolations between the solutions for the cases $\lambda = 1$ and $\lambda = 0$, described in Section 4.3.2.

We suggest two possible interpolations:

– *linear interpolation* described by the formula

$$\rho(r) = \lambda r/R + (1-\lambda)\frac{R(\sqrt{r^2+1}-1)}{r(\sqrt{R^2+1}-1)} \tag{4.20}$$

– *geometric interpolation* described by

$$\rho(r) = R\frac{\tan\frac{\arctan r}{2-\lambda}}{\tan\frac{\arctan R}{2-\lambda}} \tag{4.21}$$

The second case corresponds to the following geometric construction: $DP(r)$ is the projection of the plane to the unit sphere from the center of the sphere, followed by reprojection from the pole. If we move the point from which we reproject from the sphere to the plane from the pole to the center we get a smooth interpolation from r to $DP(r)$ described by (4.21).

If we compare the K–D curves for these interpolations with the estimates derived in the previous sections (both families of functions satisfy the assumptions for all estimates), we can see that not much improvement can be gained by finding better approximations to the solution of the optimization problem (Figure 4.5.)

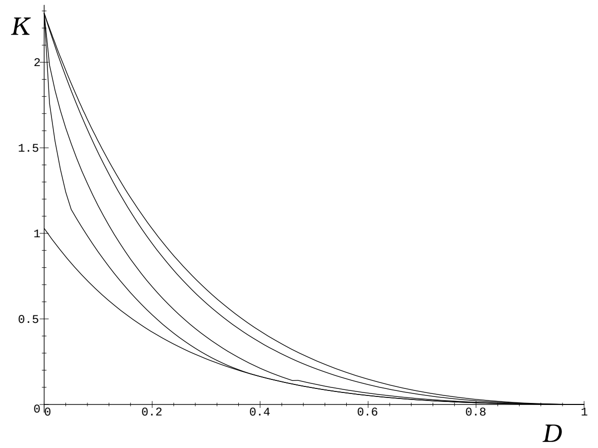


Figure 4.5: The K-D curves from top to bottom are linear interpolation, geometric interpolation, estimate from Section 4.3.3 estimate from Section 4.3.4 (lower curve.) □

It is important to remember that perception of curvature and of some shapes (such as spheres and rectangular corners) has a threshold nature. Therefore, once a perceptual requirement is violated by a sufficiently large margin, "perceptual saturation" is likely to occur and even large differences do not make much difference. Our first estimate was made without any assumptions about ρ, and it still indicates that for low values of $\mathcal{D}$, $\mathcal{K}$ should be quite large.

Nevertheless, the difference between the interpolations that we propose, and the estimates is still considerable, and it is at least of theoretical interest to obtain better estimates of the optimal (K, D) pairs and of the optimal solution.

4.3.6 Extentions

Both one-parameter families that we have described can be easily extended to larger families parametrized by the function $\lambda(\phi)$. By allowing λ depending on the angle, we obtain greater flexibility in the choice of transformation appropriate for a particular picture.

The equations for the projections take the form

$$\rho(r) = \lambda(\phi)r/R + (1 - \lambda(\phi))\frac{R(\sqrt{r^2+1} - 1)}{r(\sqrt{R^2+1} - 1)} \tag{4.22}$$

$$\rho(r) = R\frac{\tan\frac{\arctan r}{2-\lambda(\phi)}}{\tan\frac{\arctan R}{2-\lambda(\phi)}} \tag{4.23}$$

In this case, $\lambda(\phi)$ is similar to the user-specified weights described in Section 3.3.5. If we assume that λ changes slowly, the expressions for direct view and curvature error functions can be applied and for each direction θ the problem can be solved independently. Therefore, optimal values for each direction are close to those obtained for the axially-symmetric case.

For fast changes in λ, however, this does not apply, and two-dimensional optimization is neccessary to compare the family to the optimal solution as defined in Section 3.3.5.

Chapter 5

Implementation and Applications

5.1 Implementation.

We have implemented the transformations described in Section 4.3.5 and Section 4.3.6. The implementation of our projections is quite straightforward. The Π_{plane} projection of Theorem 3.1 practically coincides with the standard perspective/parallel projection. There is, however, an important implementation detail that is absent in some rendering systems. As we mentioned before, our center of projection need not coincide with the position of the camera or the eye. It is chosen according to perceptual requirements. For instance, it can happen that the most appropriate center of projection for an office scene is outside the room (Figure 5.1c.) In these cases it is necessary to have a mechanism for making parts of the model invisible. (These parts of the model should participate in lighting calculations but should be ignored by the viewing transformation.) This can be done using clipping planes.

The 2D part of the viewing transformation, T_{plane}, can be implemented as a separate postprocessing stage. In this way, the transformation can be applied to any perspective image, computer-generated or photographic. The only additional information required is the position of the center of projection relative to the image. The basic structure of the implementation is very simple:

```
for all output pixels (i, j) do
  r := sqrt(i^2 + j^2)
  setpixel( i, j,
     interpolated_color( ρ^{-1}(r)i/r, ρ^{-1}(r)j/r))
end
```

For the geometric interpolation (4.21) the inverse function ρ^{-1} can be calculated directly using the formula

$$\rho^{-1}(r) = \tan{(2-\lambda)}\arctan\left(\tan\left(\frac{\arctan R}{2-\lambda}\right)\frac{r}{R}\right) \tag{5.1}$$

Equation 4.21 can be used in the cases of constant and variable λ.

For linear interpolation (4.20), the inverse function ρ^{-1} can be computed numerically with any of the standard root-finding methods, such as those found in *Numerical Recipes* [Press et al., 1988] As we have mentioned in Section 4.3.5, for moderately wide angles the difference between any transformation from the family that we are considering and the identity transformation is quite small. That means that the initial approximation of identity for any numerical root-finding procedure is good, and even the dichotomy method works quite well. There is no particular advantage in using the linear interpolation (at least in our implementation), but it was the first one that we have implemented, and some of our images were generated using this transformation.

The interpolated_color(x, y) function computes the color for any point (x, y) with real coordinates in the original image by interpolating the colors of the integer pixels. This interpolation can be done using convolution with any reasonable filter. Using a wider filter slows down the computation, but virtually guarantees the absence of aliasing.

Implementations of (4.23) and (4.22) are practically identical to the implementation described above. There is a significant difference in the case of (4.22): the values of ρ^{-1} cannot be precomputed, because they depend on ϕ. Using (4.23) is preferable, because the inverse function can be computed explicitly.

We also need a mechanism that allows user to specify desired values of λ for various directions in the picture. One way to implement this is to allow user to specify the values of λ at a number of points, compute the direction from the center of the picture to each point, and use a closed spline curve to interpolate the values of λ for all directions (see, for example, [Hill, 1990].)

The algorithm is linear in n and should be about as fast as any other image transformation (change of the size, filtering or other geometric transformations.) Our non-optimized program (in which no attempt was made to eliminate extra function calls in inner loops, etc.) took about 30 sec on an HP9000/700 to transform a 600 by 400 image with a filter of width 4.

5.2 Deep and Shallow Scenes: When Do We Need a Wide-Angle Projection?

As we have discussed in detail in Chapter 2, narrow-angle perspective projections are likely to produce perceptually acceptable images, while wide-angle projections are likely to produce distortions and are practically unusable for extremely wide angles (Figure 5.5a.) As we have shown in Chapter 3 and Chapter 4, it is in general impossible to find a projection which would produce undistorted images for any scene, given a fixed angle of projection.

Before we try to use compromises suggested in Chapter 4, we have to answer the following question: when is it possible for a given scene to use a narrow-angle projection instead of a wide-angle one?

A general direction from the scene to the projection point is determined by our choice of the side of the objects that we want to see in the picture (we assume that our preferences are consistent with the fiber structure of the perspective projection, i.e. we don't want to see around the corners or oppposite sides of one object at the same time.) We are only free to choose the distance between the projection point and the scene.

Suppose we want an object of size s is located at the distance l from the center of projection to have approximately the same size as an object of size S located at the distance $l + L$, $S > s$. As the ratio of the sizes of their images for a perspective projection is $Sl/(l + L)s$, the distance l, determining the projection angle, can be estimated from $Sl/(l + L)s \approx 1$. The projection angle α can be found from the expression $\tan \alpha \approx S/(l + L)$. The resulting estimate fo r l is $S - s/L$. Of course, S need not be a size of an object; it can be a distance between objects and in general any characteristic length in the direction perpendicular to the principal ray. We can consider S to be the "width" of the scene in the background, and s to be the "width" of the scene in the foreground. L is just the distance between the closest and most remote objects in the scene ("depth" of the scene).

Therefore, the projection angle has to be large if the difference between the foreground width of the scene and the background width of the scene is large compared to the depth of the scene.

We shall call a scene *shallow*, if the ratio $S - s/L$ is small; if this ratio is large (> 1), then we shall call the scene *deep.*

Deep scenes require wide-angle projections, while narrow-angle projections can be used for shallow scenes.

Almost any indoor scene is inherently shallow, because the objects in such scenes typically have comparable sizes. Therefore, in most cases for closed spaces of relatively small size we can get away with using the perspective projection. It is important to note that the choice of the projection point for a scene doesn't have to reflect an actual physical position of the observer in the scene. For example, any real-life photo showing a whole wall of an office would require a very wide projection angle. However, in computer graphics, nothing prevents us from choosing a projection point in an unreachable position in or outside the room. Our choice can be based purely on the perceptual quality of the picture. Minimization of shape distortion requires narrow angles of projection. At the same time, extremely small angles of projection result in a lack of parallel foreshortening and excessive perpendicular foreshortening for some objects. Therefore, a moderate projection angle is most preferable. In Figure 5.1 we can see pictures of the same part of the office from three projection points, corresponding to the projection angles 90°, 36°and 3°. Most people prefer the projection in the middle. In fact, the number 36° was determined by showing to people a set of pictures with the projection angle varying from 3 to 90°, and asking them to choose the nicest picture. The variance of the results was small: nobody chose any image outside the limits 32-40°.

The effect of the change of the projection angle on the image of a deep scene is radically different. For a deep scene the general layout and relative sizes of the objects are unchanged only for a small range of angles. Consider Figure 5.2. The 90° image has the desired composition, but distortions are clearly visible: the men close to the edges of the picture appear to be "bulkier" than the men in the center of the picture, although their shape was identical. The shape of the head appears to be distorted in some cases, and the men close to the edge of the picture appear to be asymmetrical.

The two 60° images are radically different: in the first case, when all the objects are still present in the picture, the men in the picture almost disappear. In the second case, the pyramids in the background are absent, and relative sizes of the images of the men are changed.

a

b

c

Figure 5.1: Example of a shallow scene: a. 90° projection angle; b. 3° projection angle; c. 36° projection angle. Position of the camera is shown on the right for each picture. Note that for 36° image the camera is outside the room. □

For deep scenes, the projection angle is practically determined by the scene and composition of the picture, and cannot be varied. In this case transformations of the type described in Section 5.1 can be applied to decrease the distortions in the perspective image. The results of applying the transformation (4.20) to the $90°$ picture is shown in Figure 5.2b. Note the reduction in the distortion of shape of the men. As a tradeoff, the barrel distortion is introduced, but due to the absence of straight lines, its only manifestation is a slant of some figures.

Striking distortions occur in wide-angle motion pictures, especially when the camera rotates around a fixed vertical axis (pan.) They are primarily associated with two factors:

– when an object moves to from the edge to the center of the picture, the shape of its image changes from distorted to non-distorted. This change is noticeable and produces an impression of non-rigidity.

– the vertical coordinate of objects that are located at a large angle from the horizon changes considerably (cf. Figure 2.22). This effect appears unnatural.

These effects are reduced if transformations like (4.20) or (4.21) are applied to the frames.

Summarizing the discussion above, we suggest the following recipe for the choice of an appropriate projection for an image:

1. Determine if the scene is shallow or deep.

2. If the scene is shallow, choose a perceptually optimal projection angle which is likely to be in the range 15-45°.

3. If the scene is deep, determine the minimal projection angle that is acceptable for the scene. Then choose a transformation from (4.21) or (4.23) which produces the best results.

5.3 Applications of Correcting Transformations

In the previous section we described the conditions which require a wide-angle projection under the assumption that we are restricted in our choice of projection only by the composition of the scene. In the case of photography there are several additional restrictions: there should be no obstacles between the camera and the objects that must appear in the image, and the size of the camera cannot be made arbitrarily large. For example, it is impossible to make a picture of a room from

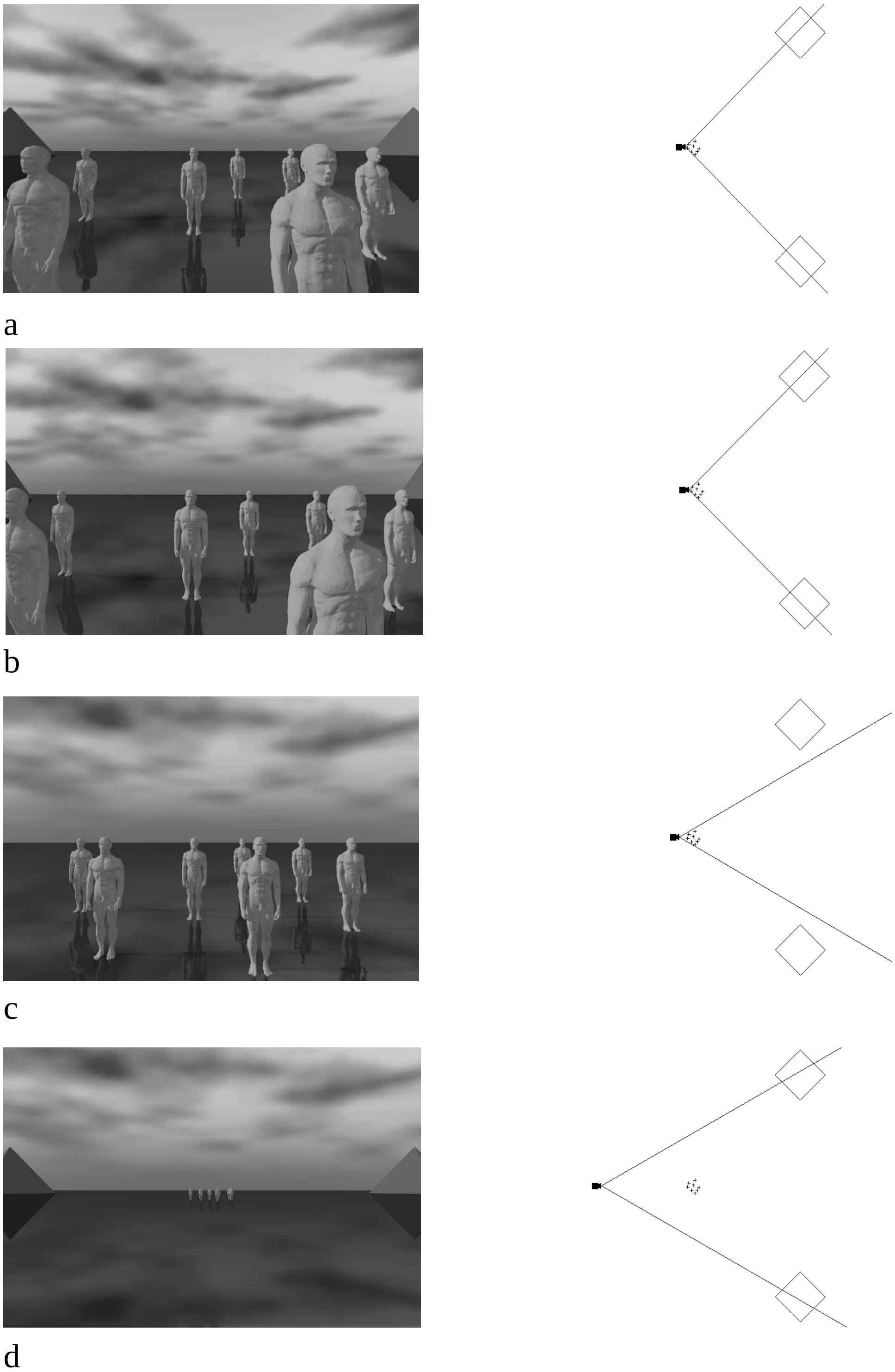


Figure 5.2: Example of a deep scene: a. 90° projection angle; b. 90° projection angle, correcting transformation with $\lambda = 0$ applied. c. 60° projection angle, keeping the figures of men approximately the right size; d. 60° projection angle, keeping the pyramids in the field of view; □

a projection point behind a wall. As a result, the projection angle can be chosen only in a limited range, and some pictures must use a wide-angle projection.

In many cases the best camera positions are physically inaccessible or there is no time to make a good choice of the position.

All these reasons explain why wide-angle images are inevitable in many cases even for shallow scenes. Transformations 4.21, 4.23, 4.21, 4.23 can be used to correct for distortions of shape in such images.

The position of the center of projection is usually known for computer-generated images, but is more difficult to obtain for photographs. For photographs it can be calculated if we know the size of the film and the focal distance of the lens used in the camera. Alternatively, it can be computed directly from the image if there is a rectangular object of known aspect ratio present in the picture (Appendix B.)

If our task is correcting a distorted photographic image, it is important to understand what type of distortion is present.

For example, distortions related to verticality (Section 2.4.3) can be corrected by a linear transformation, as it was known to photographers for a long time ([Feininger, 1953], Figure 5.3.)

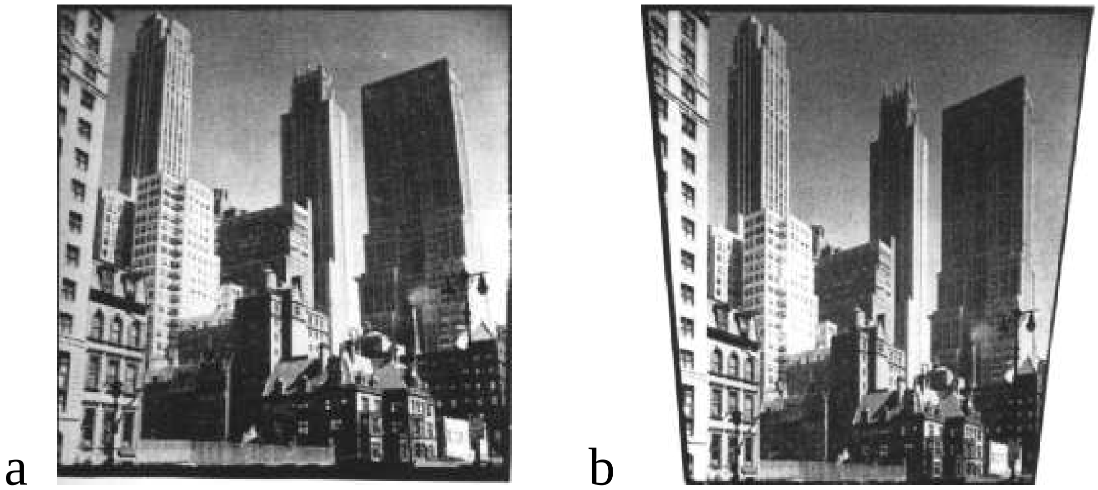


Figure 5.3: Photos from [Feininger, 1953]: a. The original photo with convergent verticals; b. Same photo printed with slanted easel, the slant is chosen so that the verticals become parallel. □

Distortion of shape can be corrected using our transformations. In some cases these transformations also decrease distortions of relative size due to the excessive perpendicular or parallel foreshortening.

Let's consider in greater detail an example of application of transformations 4.21 and 4.23.

The head of the author in Figure 5.4a is considerably distorted, and his right shoulder appears to be longer than the left. He also appears to have much more muscle than in real life. The shape of small objects on the table is distorted. The rectangular corner indicated in the left part of Figure 5.4a also looks distorted (cf. Section 2.3.1.)

After application of the transformation (4.21) for $\lambda = 0$, the author looks much more like himself and the rectangular corner looks better. The straight edges in the left part of the picture are curved.

If we use the transformation (4.23) with λ gradually increasing from 0 on the left to 1 on the right, we can eliminate the curvature of edges, while leaving the corrections in the left part o the picture intact. Unfortunately, the rectangular corner doesn't look right again - we cannot have both zero-curvature and correct shape at the same place.

Again, we give a short recipe for correction of wide-angle photos:

1. If vertical objects in the picture appear to be falling, apply a linear transformation which makes convergent verticals parallel.

2. If distortions of shape are present, apply (4.21) or (4.23). Transformation (4.21) is sufficient when all long straight lines in the picture pass close to the center of the picture; otherwise, choose $\lambda(\phi)$ close to 1 in 4.21 for the directions where there are straight lines and no objects of simple or familiar shape, and close to 0 in the directions where there are such objects and no long straight edges.

5.4 Other Applications

There are two more applications of our transformations that might be useful.

Zooming. If it is necessary to extract a part of a wide-angle image, the objects in the sub-image are likely to look quite distorted. In the context of the whole image this distortion appears to be more acceptable, but when a part of the picture is clipped, the distortions become more objectionable. If the clipped part is small, no long straight edges will appear in the picture, and transformation (4.21) with λ close to 0 can be applied before clipping. The results are demonstrated in Figure 5.4d and e.

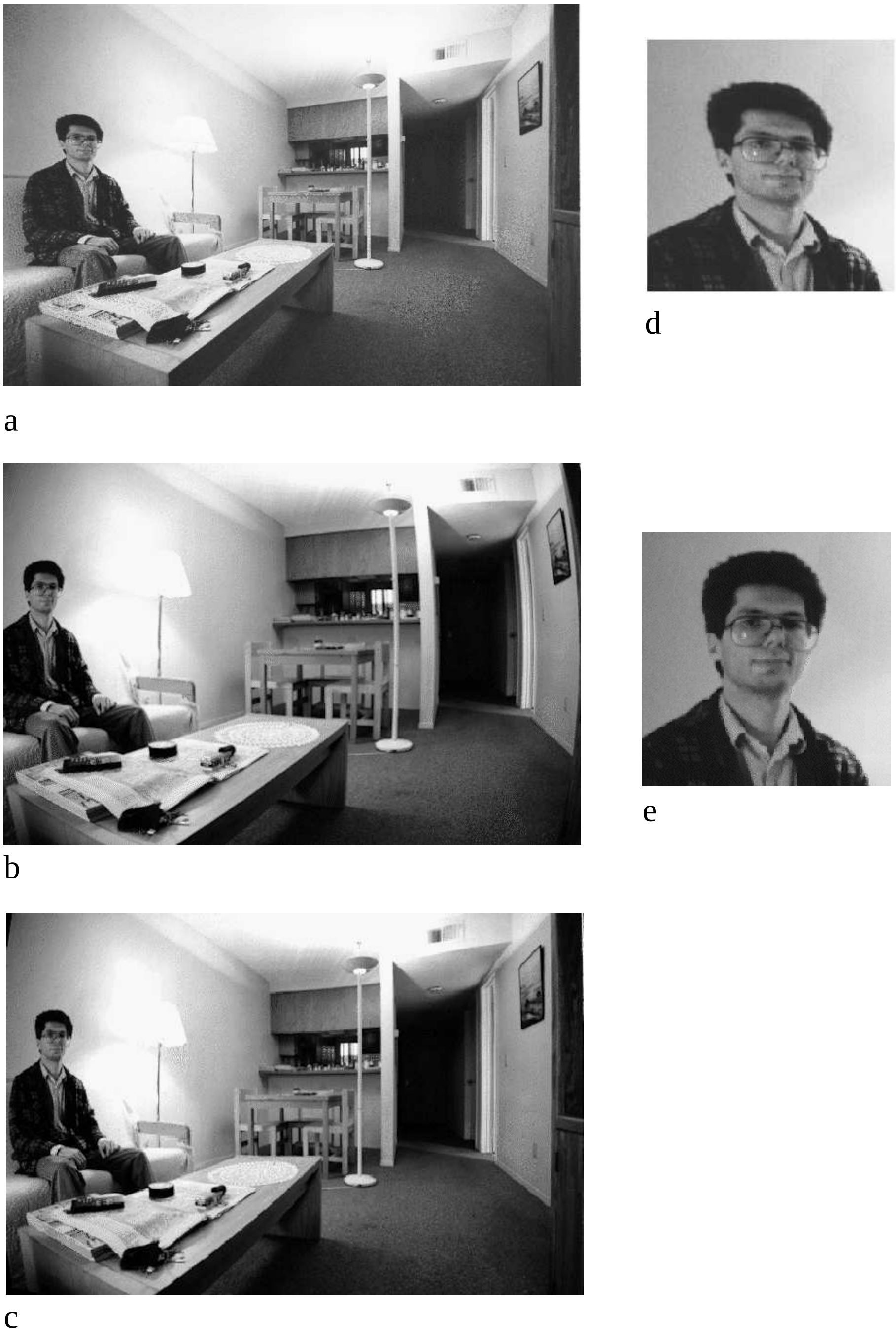


Figure 5.4: An example of application of (4.21), (4.23) to a wide-angle photo: a. Original picture, 180°; note the distortion of the shape of the head and of the marked rectangular corner in the right part of the picture. b. Transformation 4.21 with $\lambda = 0$ applied; the straight edges in the right part of the picture are curved. c. Transformation 4.23 applied; λ changes from 0 to 1. d. Zoom of the head in the original photo. e. Zoom of the head in the transformed photo (b). □

Extremely wide angle images. As we have seen, (Figure 5.5a), linear perspective is unsuitable for extremely wide angles. Rather than using traditional fisheye views, which "overcompensate" for the distortions of the perspective projection, we can use one of our transformations. Visible distortions are practically inevitable, but for our transformations they are likely to be less than for the perspective projection and fisheye views.

5.5 Summary

Here is a summary of possible applications of the transformations that we propose:

- Creation of wide-angle computer-generated pictures with minimal distortions.
- Reduction of distortions in photographic images.
- Creation of wide-angle motion pictures with reduced distortion.
- A better alternative to fisheye views.
- Zooming of parts of a wide-angle picture.

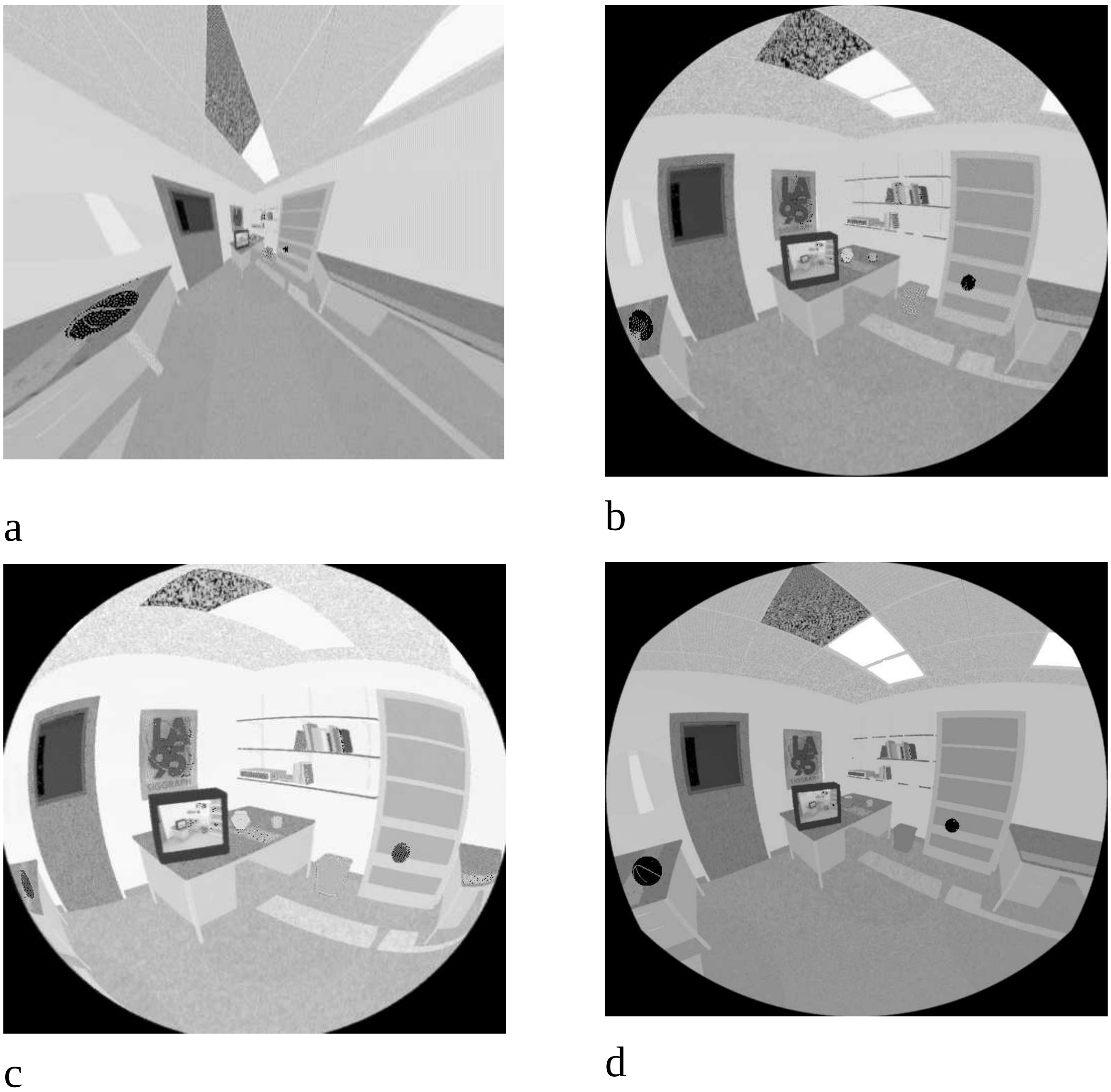


Figure 5.5: Extremely wide angle images (160°): a. Linear perspective. b. Angular fisheye (equidistant projection). c. Hemispherical fisheye (orthographic projection). d. Our direct projection. Images (b) and (d) are quite similar, but in (c) we can observe excessive compression of the objects in the margins –the tables appear to be narrower and the ball doesn't look spherical. □

Chapter 6

Conclusions and Future Work

6.1 Conclusions

Let's reiterate the main points of the thesis below.

- We have provided a review of perception of geometry in pictures and of the usage of perspective systems in the art (Chapter 2).

- We have suggested a classification of perceptually important geometric properties of images (structural and nonstructural) which is convenient for the study of general projections of the space into the plane from a perceptual point of view (Section 2.1).

- We have shown that a number of simple structural requirements (such as the absence of loops in the images of straight lines, the absence of twists in the images of planes,) when applied to all possible objects that can occur in a picture, lead to the conclusion that the fibers of the projection should be subsets of the fibers of a perspective projection. (Section 3.2).

- We have introduced two types of error functions which can be used as quantitative measures of distortions of curvature and shape. (Section 3.3).

- We have shown that the error functions cannot be simultaneously identically zero for any projection except orthogonal parallel. For each error function, however, there are projections for which the error function vanishes identically. (Section 4.3.2).

- We have described error functionals and formulated optimization problems that can be used to construct projections with minimized distortions. (Section 3.3.4.)

- We have found estimates of the optimal values of error functionals for the axially symmetric case and demonstrated that under certain assumptions simple interpolations have values of error functionals close to the optimal. (Section 4.3.3).

- We have described situations when distortions are likely occur and suggested methods to decrease them. (Chapter 5).

6.2 Future Work

There are a number of directions in which this work can be completed and extended.

The most immediate goals are

– to find better estimates of the optimal values of functionals and approximations to the solution of the optimization problem in Section 4.3;

– to consider other cases of simple symmetries; for example, projections similar to cylindrical.

A promising direction of work which we hope to pursue is suggested by the fact that conformal transformations preserve the direct view error function. That means that we can try to decrease curvature or achieve improvement in some other perceptual factor without causing additional distortion of shape, if we use conformal transformations of the picture plane.

The decomposition theorem from Chapter 3 applies to the case of connected regions of space. It is interesting to consider the ways of splitting the space into several disconnected parts and using a different projection for each of these parts.

We tried to avoid dependence on the contents of the scene. It is also possible to consider transformations that depend strongly on our knowledge of the geometry of the objects in the scene. In this case error functions have to be evaluated only for the actual images of objects rather than for all possible images and better results can be achieved at the expense of much greater complexity and lesser generality of the algorithms.

Appendix A

Review of the Experimental Studies of Perception of Pictures

This is a brief exposition of experiments described by various authors. It is not in any way complete. Our main purpose here is to give the reader a general idea about stimuli, viewing conditions, tasks and results of the experiments.

We tried to describe all the experiments using consistent terminology from Section 2.1.2. Thus, on numerous occasions we had to change the original terms to their equivalents.

In most cases we say nothing about theories of perception that the experiments were supposed to test. On the contrary, we tried to separate the experimental data and the theory and to get rid of the bias introduced by the theory into the interpretation of the results.

We subdivided the papers into several sections according to what we believe was the main experimental subject of the paper.

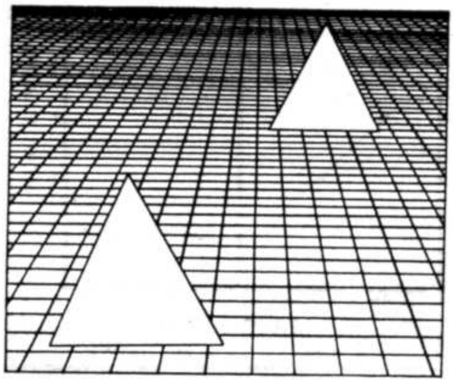


Figure A.1: Stimulus from [Hagen, 1976b] □

A.1 Robustness and Perception of the Picture Surface and Frame

[Hagen, 1976b] Hagen, M. A. "Influence of picture surface and station point on the ability to compensate for oblique view in pictorial perception.," *Developmental Psychology*, 12(1):57–63. 1976.

Experiment

Subjects. Children 4 years old, 7 years old and adults (40 in each group).

Stimuli. Photographs of triangles and squares against a textured background (horizontal plane) (Figure A.1). Three relative sizes of the images in the pictures: bigger was 100%, 85% and 75% of the smaller. wo types of stimuli were used: color prints and slides.

Viewing conditions. Peephole, monocular viewing, two viewing directions - 90 and 45°, viewing distance equal to the projection distance.

Task. The subjects were instructed to point to the bigger object.

Results. Correct station point increased the percent of correct choices for adults and 7-year-old children for slides and *decreased* for prints. For 4 year-old children the percent of correct answers was very small, indicating insufficient understanding of perspective pictures.

[Lumsden, 1980] Lumsden, E. A. "Problems of magnification and minification: An explanation of the distortions of distance, slant, shape, and velocity," In [Hagen, 1980], pages 91–135. 1980.

[Rosinski and Faber, 1980] Rosinski, R. R. and Faber, J. "Compensation for viewing point in the perception of pictured space," In [Hagen, 1980], pages 137–176. 1980.

E. A. Lumsden, R. R. Rosinski and J. Faber describe some unpublished experiments by W. C. Purdy.

Subjects. Unknown.

Stimuli. Texture gradients projected onto a screen. Visible portion of the projection was circular and filled with texture (oblique central projection of a piece of transparent film with a grid of lines approximately 1/8 inch apart.) Two positions of the light source and the film (source-to-film distance remained constant) were chosen so that they resulted in magnification 1.0 and 1.5 of the tilted film. Four projection directions of the film were used: 30°, 40.9°, 52.4°, 62.8° and 71.1°. The projection directions were chosen so that each image coincides with the image one step below if projected from the viewing point.

Viewing conditions. Fixed viewing position on the other side of the screen, binocular viewing (uncertain), viewing angle limited to 25° by an aperture.

Task. Subjects were asked to adjust a rectangular plate to a position parallel with the slant of the textured surface in the picture.

Results. For 1.5 magnification, judged slant corresponds to the depicted slant one step below, that is, the slant is perceived as if the projection was made from the viewing point. For 1.0 magnification (viewing point coincides with the projection point) judged slant was in correspondence with depicted slant with the same constant error.

Note: The equivalence of 1.5 magnification to 1.0 magnification one step below is not surprising, as the distance from the viewing point to the screen remained constant and the resulting stimulus was practically identical.

[Rosinski and Faber, 1980] Rosinski, R. R. and Faber, J. "Compensation for viewing point in the perception of pictured space," In [Hagen, 1980], pages 137–176. 1980.

Experiment 1.

Subjects. Unknown.

Stimuli. Photos and line drawings of gray square parallelepipeds, 6 cm in height and varying in width. Three faces of each parallelepiped were visible. Projection point was 25 cm away from the picture plane. Orthogonal central projection was used.

Viewing conditions. The viewing point was located 25 cm, 50 cm, 75 cm away from the picture on the principal ray. This corresponds to te viewing angles of approximately 55°, 30° and 20° and magnifications 1.0, 0.5, 0.33.

Task. Subjects were asked to estimate the size (more specifically, width) of the objects in the pictures. Estimate were "modulus-free" in abstract units.

Note: it is entirely unclear from the text what were the specific instructions to the subjects and what abstract units were to be used.

Results. There was no significant change in apparent (relative?) size of the objects when the viewing point was changed.

Experiment 2.

Subjects. Unknown.

Stimuli. Computer-generated images of slanted lattices displayed on a CRT. Projection point for the pictures was 28, 56, 84, 112, 225, 337, 450 cm away from the screen. Central orthogonal projection was used. Depicted orientation varied from 0 to 170° in 10° increments.

Viewing conditions. Viewing point was on the principal ray to the picture, 112 cm away; this corresponds to magnifications in the range $0.25 \ldots 4.0$. Viewing was binocular.

Task. Observers were asked to make direct judgement of the apparent slant in degrees.

Results. Judged depended on magnification/minification; this dependence was more significant for slants closer to 0 or 180° and less significant for the medium range. For magnified pictures dependence was more apparent than for minified.

Note: The task of judging the slant in degrees is quite complicated; it appears that the data were distorted by the absence of experience of matching numerical measure with a particular slant. The main point, however, is that the apparent slant depends to some extent on magnification. Purdy's experiment appears to be more reliable.

Experiment 3

Subjects. Unknown.

Stimuli. Computer-generated images of slanted lattices; in this experiment projection point was always at 112 cm from the screen. The range was 0 - 170° in 10° increments. (same as in Experiment 2.)

Viewing conditions. Viewing point was located on the principal ray 28, 56, 112, 225, 337, 450 cm away from the picture, which corresponded to magnification in the range 0.25 ... 4.0. Viewing was binocular.

Task. Same as in Experiment 2.

Results. Practically no dependence of the judged slant on the magnification was observed.

[Rosinski et al., 1980] Rosinski, R. R., Mulholland, T., Degelman, D., and Farber, J. "Picture perception: An analysis of visual compensation," *Perception & Psychophysics*, 28(6):521–526. 1980.

Same experiments are described in [Rosinski and Faber, 1980] in less detail.

Experiment 1.

Subjects. 10 college students.

Stimuli. Photographs of a slanted striped rectangle, slant varies in the range 30 ... 150° (orthogonal central projection.) Surface was rotated around a vertical axis (exact location of the axis was not specified. [Kubovy, 1986] in his account of the experiment specifies the vertical axis going through the center of the rectangle.) Photographs were cut into trapezoidal shapes coinciding with the outlines of the images of the rectangle.

Viewing conditions. Viewing was monocular, through an aperture, and the photograph was placed in a viewing box where no frame was visible. Viewing direction was either 90° or 135°; the distance from the picture plane was equal to the projection distance.

Task. Subjects were asked to set the inclination of a vertically pivoted palm board to match the apparent slant.

Results.

- Judged slant differed considerably for 90° and 135° viewing direction.
- Define *Projected slant* as equal to the slant of the rectangle that would produce the same image if projected from the viewing point. When the viewing point coincides with the projection point,

projected slant is equal to the depicted slant. It was found that the dependence of the judged slant on the projected slant for 135° viewing direction was very similar to the dependence of the judged slant on the depicted slant for 90° viewing direction. The authors conclude that under restricted viewing conditions apparent slant is determined by the projected slant.

Experiment 2

Subjects. Same.

Stimuli. Same.

Viewing conditions. Viewing was binocular, the photos were in a frame without a viewing box, no aperture was present. Viewing direction was 135° or 45°, viewing distance was equal to the projection distance.

Task. Same.

Results. There were practically no difference between judged slants. As projected slant in this experiment differed considerably, the authors conclude that the apparent slant is determined by the depicted slant for unrestricted viewing conditions.

[Hagen et al., 1978b] Hagen, M. A., Jones, R. K., and Reed, E. S. "On a neglected variable in theories of pictorial perception: Truncation of the visual field," *Perception & Psychophysics*, 23(4):326–330. 1978.

Experiment

Subjects. 60.

Stimuli. a) Isoceles triangles whose height and base were 2,3,4,5,6 in. The triangles were placed vertically on the table. Distance was measured from a round marker placed on the table 18 in away from the subject. The distance between the markers and the triangles was 6, 10.2, 17.3, 29.5 and 50 in. b) Slides of the scene (6.5×4.5 cm).

Viewing conditions. Viewing was monocular. Three viewing conditions were used for the real scene: fixed position with untruncated field of view, peephole (2 mm), rectangular truncation (6.5×4.5 cm slot 5 cm away). Slides were viewed from the projection point; conditions for the slides coincided with the rectangular truncation.

Task. Subjects were asked to reproduce with a tape the size of the triangles and the distance from the marker to the triangles.

Results.

• The distance was consistently underestimated for the truncated field of view; there was a considerable compression of the scene (i.e. the relative distances decreased.) The effect was similar for the three truncated conditions (slides, rectangular truncation, peephole.)

• Intercept (extrapolated apparent "zero" distance) was the greatest for slides, and significantly non-zero for all conditions except the untruncated viewing.

• Size estimates for the truncated field of view were lower than the estimates for the untruncateed field of view for the real scene; for slides, they were higher at smaller distances, but lower at longer distances.

A.2 Rigidity in Motion Pictures

Subjects. 6.

Stimuli. Orthogonal and oblique (direction of projection 67°) parallel and central projections (projection angle 5.8°) of nearly rectangular rotating rigid and nonrigid solids on a computer display. The images were wireframe with hidden lines removed. Two types of deformations were used for non-rigid solids: affine (angles between edges are unchanged) and non-affine (Figure A.2). Deformations were periodical, with the period equal to the period of rotation. Extent of deformation was 32%, 16%, 8% of the smallest dimension of the solid (?). For the oblique projection the aspect ratio of the objects was changed to make them appear less narrow.

Viewing conditions. Unrestricted viewing conditions; the viewing angle not specified, but constant. Duration of each trial was 10 sec (3 periods of rotation).

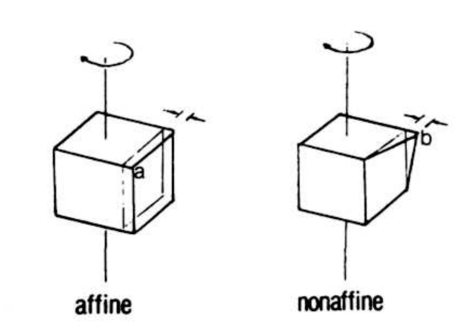
Task. Subjects were asked to rate the rigidity on the scale of 1 to 9, 9 indicating high confidence that the object was rigid.

Results.

• Parallel projection makes it more difficult to detect non-rigidity.

• The change of 23° in viewing direction doesn't make any difference.

• Affine deformations are more difficult to perceive.



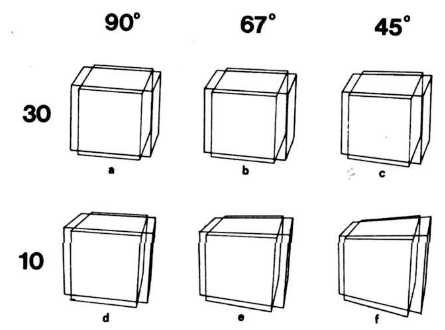


Figure A.2: (Left) Nonrigid stimuli from [Cutting, 1987] □

Figure A.3: (Right) Stimuli from Experiment 2, [Cutting, 1987] □

- Small deformations (8%) are not recognized very well.

Experiment 2

Subjects. Same.

Stimuli. Stimuli of the same type. Central projections were used. Projection angles were 11.5° and 3.8°, projection directions 90°, 67°, 45°, and a variable direction in the range 55°...80°, with the period equal to one half of the period of rotation (Figure A.3).

Viewing conditions. Two viewing angles were used: 11.5° and 3.8°.

Task. Same.

Results.

- For parallel projection we have less non-rigidity perceived for all slants of the screen.
- For perspective projection there is considerable non-rigidity for slant 45°.
- Simulated and physical distance (projection point position and viewing point position) did not affect judgements
- Among non-rigid stimuli non-affine transformations were much easier to perceive.
- Relation between wider angle perspective and viewing direction: 10 radii projection (just 12°!) looked bad from 45° viewing direction.

Experiment 3

Subjects.

Stimuli. Rigid/nonrigid stimuli, 90°, 67°, 45° projection direction was compensated by the viewing direction of the same magnitude, 37.5, 15 ,6 radii projection distances (corresponding viewing angles are 20°, 8°, 3°.)

Viewing conditions. Viewing distances were 37.5, 15, 6 radii.

Task. Same.

Results. Rigidity was preserved, no compensation for slant occurred.

A.3 Size and Distance in Pictures

[Smith, 1958b] Smith, O. W. "Judgements of size and distance in photographs," *The American Journal of Psychology*, 71:529–538. 1958.

Experiment

Subjects. 52 (26 per viewing condition).

Stimuli. 18 Photographs of a plowed field extending from the foreground to distant hills. 15 stakes in each photographs were located on a circular arc 14 yd away from the projection point, the height of the stakes varied in the range 27–83 in in 4 in increments. An 8 ft gap in the middle of the arc permitted subjects to see the 16th stake at a distance. 5 stakes were identical in all pictures; 16th stake was 63, 67, 71 or 75 in high and it was 28, 112, 224, or 448 yd. All 16 possible combinations were used. Two more photographs were obtained by air-brushing out all the detail except for the image of one stake 63 in high at 54 yd from the projection point. On one of them shadows were painted. All photos were 11.875 by 18.625 in.

Viewing conditions. Monocular viewing from a fixed position through an aperture. Two viewing boxes were used: one resulted in 1.33 magnification and the other in 0.4 magnification (the viewing distance was equal to 0.75 and 2.5 of the projection distance). The boxes were black inside, back-lit slides were used. Data for similar pictures from two different studies were used: restricted viewing from the projection point, binocular viewing and real scene viewing.

Task. Subjects were asked to judge the size of the distant stake by matching it with a near stake, estimate the distance to the distant stake and stakes in the foreground in yards and the height of the distant stake in feet. Impoverished photographs were judged either before or before and after the regular photographs.

Results.

• Matches of height and judgements of height in feet didn't vary much with the change of viewing conditions. Height judgements were overestimates.

• The judgements of distance increased for 0.4 magnification and decreased for 1.33 magnification. Standard deviation increased with the distance and was quite high, so only qualitative conclusions are valid.

• Qualitative trend for the impoverished photographs is the same (distance judgements depend on the viewing point) but distances were considerably underestimated if the impoverished photographs were viewed first. Size judgements in impoverished photographs were not significantly different.

[Smith, 1961] Smith, Patricia C., S. O. W. "Ball throwing responses to photographically portrayed targets," *Journal of Experimental Psychology*, 62(3):223–233. 1961.

Experiment 1-2.

Subjects. 50 for Experiment 1, 16 for Experiment 2.

Stimuli. In a room 36 by 30 ft, there were 25 marked target locations at five distances 3, 4.25, 5.5, 6.75, 8 m from the viewer (Figure A.4). The viewer was used with apertures which resulted in approximately 7°×5°, 24°×17°, 41°×30° fields of view. Below the viewer a curtain could be opened and through this opening the subjects could toss a ball at a target. Next to the viewer another curtain could be opened and through this opening the subjects could observe the targets with unrestricted head motion.

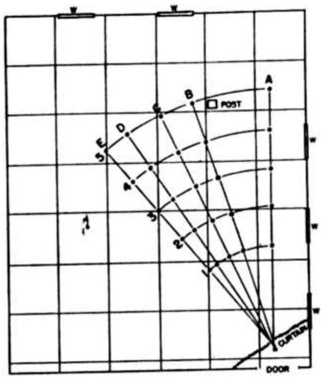


Figure A.4: Diagram of the room of the experiments, [Smith, 1961] □

Viewing conditions. 5 viewing conditions were used: monocular viewing through the viewer with one of the possible apertures, binocular and monocular unrestricted viewing.

Task. Subjects were asked to throw the ball at one of the targets for different viewing conditions (restricted conditions first in Experiment 1, unrestricted conditions first in Experiment 2). Immediately after the toss the viewer or the curtain was closed.

Results. Restricted field of view resulted in larger distance to the point of impact of the ball. There was no significant dependence on the size of the aperture and between monocular and binocular unrestricted viewing. The order of the unrestricted and restricted viewing conditions didn't affect the qualitative relationship between the results for restricted and unrestricted viewing conditions.

Experiment 3.

Subjects. 30.

Stimuli. Same, except that the subjects observed photographs of the scene in the viewing apparatus instead of the scene itself. For unrestricted viewing conditions, the real scene was used. Photographs were made from the correct viewing point; no attempt was made to match the colors or intensities; photographs were black-and-white, with yellow target markers painted on them.

Task. Same. Subjects were not told that photos are used; after the experiment they were asked if they saw anything unusual in the scene visible through the viewer.

Results.

• The data were adjusted in the following way: the difference between means in unrestricted viewing conditions (identical for Experiments 1–2 and 3) was added to the means for restricted conditions to correct for the difference between groups of subjects. Adjusted results were very close for photographs and real scene (restricted viewing).

• There was practically no difference between binocular and monocular unrestricted viewing.

• No subject reported that he had seen photos in the viewer or noted any unusual quality of what he had seen.

Experiment 4. We omit this experiment because we don't find its results relevant to our purpose.

[Smith, 1958a] Smith, O. W. "Comparison of apparent depth in a photograph viewed from two distances," *Perceptual and Motor Skills*, 8:79–81. 1958.

Experiment

Subjects. 52.

Stimuli. 36×45 in back-lit photograph of a corridor 360 ft. long

Viewing conditions. Monocular viewing through a peephole from two viewing distances: 0.53 and 2.53 of the projection distance from the viewing points on the principal ray to the picture (Viewing angle was 86° and 23°, magnification 1.875 and 0.395). A viewing box was used to restrict subject's field of view to the image area.

Task. Subjects were asked to estimate the distance to the front edge of the corridor and the length of the corridor in paces.

Results. The means for the greater viewing distance were more than two times greater than the means for the smaller viewing distance.

A.4 Orientation and Layout

[Halloran, 1989] Halloran, T. O. "Picture perception is array-specific: Viewing angle versus apparent orientation," *Perception & Psychophysics*, 45(5):467–482. 1989.

Experiment 1.

Subjects. 12.

Stimuli. Pictures similar to those used in [Rosinski et al., 1980]. Black-and-white striped rectangles, depicted at 15, 30, 45, 60, 90, 120, 135, 150 165° slant around a vertical axis; the axis was located at the right edge of the rectangle.

Viewing conditions. Unrestricted viewing from the viewing distance equal to the projection distance and the viewing direction of 30, 45, 90, 135, 150°.

Task. Subjects were asked to reproduce the slant with a pointer pivoted around a vertical axis.

Results. For large slants and large displacements (viewing direction of 30 or 150°), exceeding those of [Rosinski et al., 1980], there was a considerable dependence of the estimated slant on the viewing position.

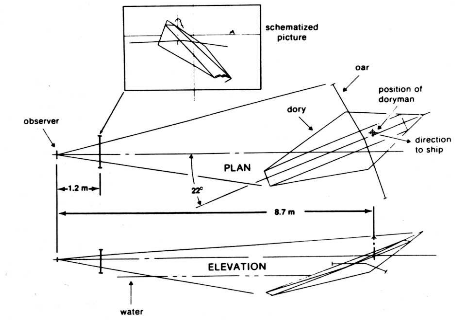


Figure A.5: (Left) Winslow Homer, "The Fog Warning", from [Halloran, 1989] □

Figure A.6: (Right) Scene reconstruction, from [Halloran, 1989] □

Experiment 2.

Subjects. Same.

Stimuli. Pictures from Experiment 1, which depicted 15° and 165° slant transformed in the manner that corresponds to the viewing positions of Experiment 1; two frames were present: the transformation of the frame of the original 15° and 165° pictures and a rectangular frame of the final picture.

Viewing conditions. Viewing point was fixed, viewing direction was 90°.

Task. Same.

Results. No difference from Experiment 1, except for a small nearly constant digression towards picture plane. Compare to the apparent distortions of the pictures of slanted pictures as demonstrated in [Pirenne, 1970].

Experiment 3.

Subjects. 23.

Stimuli. Multicolored reproduction of Winslow Homer's painting "The Fog Warning", 42 by 66 cm. (Figure A.5).

Viewing conditions. Unrestricted viewing from 7 positions; the following are pairs (viewing distance, viewing direction): (2.3 m, 14°), (2.4 m, 30°), (1.7 m, 45°), (1.2 m, 90°), (1.7 m, 135°), (2.4 m, 155°), (2.3 m, 166°).

Task. The subjects were instructed a) to orient the pointer in the direction that the dory in the picture was headed, with respect to the picture plane; b) to orient the pointer in the direction of the mother ship on the horizon from the man's position.

Results.

• There is a consistent shift in both judged orientations with the change in the viewing point. As the viewing direction angle increases, the judged angle of orientation of the boat and of the direction to the ship increases.

• The rotation of the judged orientation of the boat approximately corresponds to the mean rotation of the objects reconstructed using 4 sets of cues (Figure A.6).

Set 1: the fixed height of the man and the length of the boat.

Set 2: the rectangle from the stern width to the corresponding width in the front of the boat.

Set 3: the fixed length and width of the boat, except 14° and 166° viewing directions.

Set 4: oars are of equal length.

The best fit is given by averaging set 4, which results in no rotation, with any of the sets 1-3.

• The change of difference of the angle of orientation of the boat and the direction to the ship is small: 11° over 152° interval the judged angle of orientation of the boat.

Experiment 4

Subjects. 17.

Stimuli. Monochrome reproduction of the picture from Experiment 3.

Viewing conditions. Unrestricted viewing from viewing points from Experiment 3 with the viewing direction 150° (the viewing distance was scaled as the size of the picture was different.)

Task.

a) Subjects were asked to make separates judgement of

1) the plan-view orientation of the dory that is indicated by its long axis, disregarding all other cues;

2) the orientation with respect to the picture plane of a straight line connecting the blades of the oars;

3) the angles between the line connecting the blades of the oars and the boat axis.

To accomplish this, a pair of pointers pivoted on a common axis were used, to represent two sides of the angle to be estimated.

b) Subjects were asked to identify the "correct" boat among 5 plan-view outline drawings with length-to-width ratios of 1.5, 2.5, 3.33, 5.0, and 7.0.

Results.

- Mean length axis orientation was approx. 163°. For a similar viewing position in Experiment 3, the mean orientation was 136°, indicating some influence of other cues.
- The oar-crossing angle implied by separate judgements of the oar direction and the boat axis direction , was 125.7°, while the mean oar-crossing angle, judged directly, was 102°.
- There were strong individual differences.
- The results of this experiment suggest that cues can influence judgments in different ways, depending on the task, and that the ability to separate cues differs among subjects.

Experiment 5

Subjects. 6.

Stimuli. 3 stimuli similar to [Goldstein, 1987] (Figure A.12), except they were perspective line drawings: each depicted a 15.2 cm ball with one or two 1.9 cm square dots on it. One dot was located on the intersection of the central perpendicular to the picture plane and the sphere (the image of this dot was in the center of the image of the ball), the other one was on the big circle of the ball, which lies in the plane perpendicular to the picture plane, rotated 60° from the position of the first dot.

Viewing conditions. Unrestricted viewing, the viewing distance equals the projection distance (46 cm), viewing direction was 160°.

Task. For the two-dot picture, subjects were asked to indicate the angle between dots with reference to the center of the ball, for the single-dot pictures subjects were asked the dot's direction from the picture plane. Pointer arrangement from Experiment 2 was used.

Results. For single-dot pictures the average judgements were 147.5° for the center dot and 32.8° for the side dot, resulting in the relative angle of 114.7°. For the two-dot pictures, the mean judgement was 67.7°.

Experiment 6

Subjects. 10.

Stimuli. Same as in Experiment 4.

Viewing conditions. Same as in Experiment 4, and also from the viewing point on the principal ray 42 cm away from the picture.

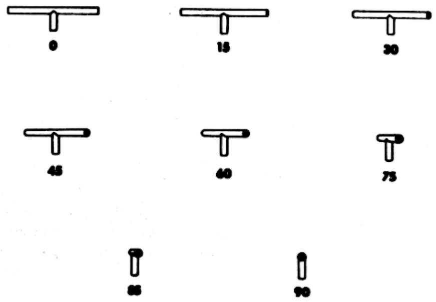


Figure A.7: (Left) Stimuli from Experiment 1 [Goldstein, 1979] □

Figure A.8: (Right) "The Village of Bequigny" by Theodore Rousseau, [Goldstein, 1979] □

Task. Subjects were asked to indicate spatial orientation of the dory using a pointer with two degrees of freedom.

Results.

• Mean judged about angles were 118.7° for 90° viewing direction, and 172.6° for 150° viewing direction. Compared to the Experiment 3, the rotation between positions was twice as much.

• Mean judged pitch (angle between the axis of the boat and the horizontal plane) was 22.8° and 28.8°; 5 observers indicated no or little increase, others indicated substantial increase.

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

Experiment 1.

Subjects. 5.

Stimuli. Line drawings of rods (Figure A.7) with different swings around a vertical axis. Depicted swings were 0, 15, 30, 45, 60, 75, 85, 90°. The 0° rod was 59 mm across.

Viewing conditions. Viewing was monocular; position of the right eye was fixed. The pictures were rotated so that the viewing direction was 15°, 45°, 70°, 90°, 110°, 135° and 165°. The viewing distance was 55 cm, corresponding to the viewing angle of approximately 6° for 90° viewing direction.

Task. The task was to set the direction of a pointer parallel to the direction of the horizontal rod in the picture.

Results. All the rods appear to be rotating as the viewing direction changes. The amount of rotation increases with depicted angle from 60° for the 0° rod, up to 149° for the 90° rod.

Experiment 2.

Subjects. Same.

Stimuli. Black-and-white photograph of the painting "The Village of Bequigny" by Theodore Rousseau (Figure A.8)

Viewing conditions. Same as in Experiment 1. The viewing angle was 17° for the viewing direction 90°.

Task. Observers were asked to judge the orientation of the road (A), the rut in the road (B), the house (C), and the line defined by two trees (D).

Results. Road rotates 150°, the house rotates 103°, the trees rotate 91°, the rut rotates 38°. Mostly confirms the results of the Experiment 1, although the house rotates more then the direction between trees, while appearing to be more perpendicular to the road with 90° viewing direction.

Experiment 3.

Subjects. Same.

Stimuli. Similar to those in the Experiment 1, but the length of the horizontal rods was only 1 mm (only disks are left).

Viewing conditions. Same as in Experiment 1.

Task. Same as in Experiment 1.

Results. The amount of rotation significantly increased for small angles (up to 115° for 15° rod), but remained unchanged for larger angles. The differential rotation decreased.

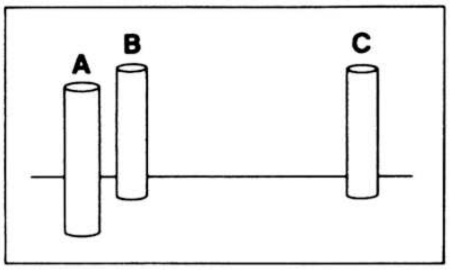
Experiment 4.

Subjects. Same.

Stimuli. Same as in Experiment 2.

Viewing conditions. Same as in Experiment 2.

Task. Observers were given a sheet of paper with two parallel lines on it indicating the direction of the road. The task was to indicate the position of the rut, of the house and of the trees relative to the road.



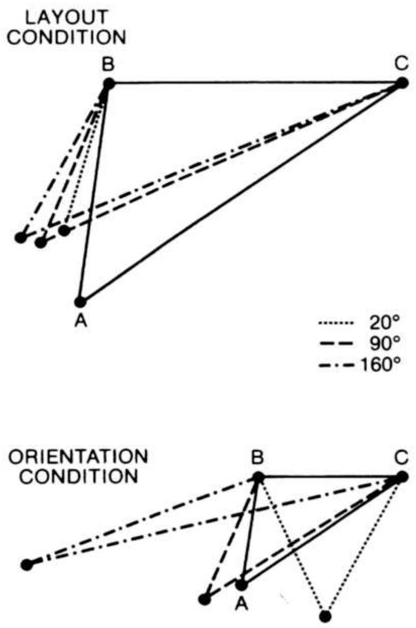


Figure A.9: (Left) Stimuli from Experiment 1, [Goldstein, 1987] □

Figure A.10: (Right) Results for Experiments 1 and 2, [Goldstein, 1987] □

Results. There was little dependence on the viewing direction. There was some change in the relative distances, but almost no change in the orientation.

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

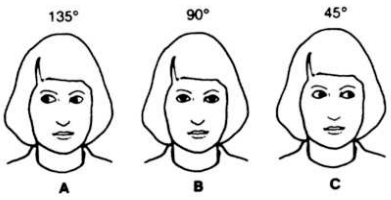
Experiment 1

Subjects. 5.

Stimuli. Black-and-white photographs of three vertical dowels with horizontal stripes on a homegeneous white surface against a homogeneous white background (Figure A.9). The size of the picture was 10.2 by 12.7 cm.

Viewing conditions. Monocular viewing from a fixed position, with viewing distance equal to the projection distance and viewing direction 20° , 45°, 70°, 90°, 110°, 135°, 160°. Viewing direction was changed by rotating the picture.

Task. Observers judged the spatial layout of the dowels by arranging three discs on a piece of paper.



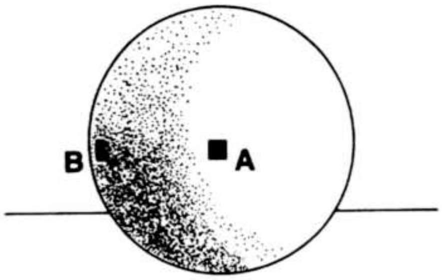


Figure A.11: (Left) Drawings of the face stimuli from Experiment 3, [Goldstein, 1987] □

Figure A.12: (Right) A line drawing of the sphere from Experiment 6, [Goldstein, 1987] □

Results. Averaged results are shown in Figure A.10 for viewing directions 20°, 90°, 160°. Size BC was scaled to the same size for all observers. Judged layout changed only slightly with the viewing direction.

Experiment 2

Subjects. Same.

Stimuli. Same.

Viewing conditions. Same.

Task. Observers were asked to set a pointer parallel to the directions BA, BC, CA.

Results. Total changes of the directions were: BA – 123°, CA – 79°, BC – 29°. The layouts, reconstructed from the directions are shown in (Figure A.10). These layout differ greatly from each other.

Experiment 3

Subjects. 7.

Stimuli. Seven photographs of a frontally oriented human face with different gaze directions resulting from looking at seven different targets, positioned at 15° intervals in the range of directions from 45° to 135°. (Figure A.11).

Viewing conditions. The viewing was monocular; position of the right eye was fixed. The pictures were rotated so that the viewing direction was 20°, 45°, 70°, 90°, 110°, 135°, 160°.

Task. The task was to set the direction of a pointer parallel to the direction of the gaze in the picture.

Results. Total rotation of the judged direction of the gaze was 137° for the depicted gaze of 90°, and only 38° for the 45° gaze, and about 40° for 135° gaze. Total amount of rotation increased towards 90° gaze.

Experiment 4

Subjects. 4.

Stimuli. Similar to the photographs from Experiment 3. in addition to the frontal position of the face, three additional head orientations were used (75°, 60° and 45°).

Viewing conditions. Same.

Task. Same.

Results.

• The results were qualitatively similar to the results of Experiment 3. The maximum of total rotation is shifted to smaller angles for 60° and 45° head orientations.

• Judged gaze direction is relatively independent from the head orientation.

Experiment 5

Subjects. 4.

Stimuli. Line drawings of two vertical cylinders similar to those used in Experiments 1 and 2. Black-on-white pictures were used for viewing in the light and negatives were used for viewing in the dark.

Viewing conditions. Similar to Experiments 3 and 4. Two modes of viewing were used: in the light, regular line drawings, and in the dark, back-lit negatives.

Task. Subjects were asked to orient a pointer in the direction parallel to the direction from cylinder A to cylinder B.

Results. Viewing the picture in the dark, when the picture frame and surface are invisible, increases total amount of rotation of the judged direction for oblique orientation.

Experiment 6 (not numbered in the paper).

Subjects. 3.

Stimuli. Photograph of a sphere with a marked point either at 90° horizontal direction from the center of the sphere or at 120° (counted from the direction parallel to the picture plane). (Figure A.12).

Viewing conditions. Exact conditions unknown. Viewing direction was 20°.

Task. Subjects were asked to judge the directions from the center of the sphere to the dots (presumably, using a pointer).

Results. The average judgements of the orientations of the points A and B were 25° and 153°, which results in calculated angular separation of 128°.

A.5 Rectangular Corners

[Perkins, 1972] Perkins, D. N. "Visual discrimination between rectangular and nonrectangular parallelepipeds," *Perception & Psychophysics*, 12(5):396–400. 1972.

Experiment

Subjects. 27.

Stimuli. 128 drawings of boxes, some of which could have been and some of which could not have been orthogonal parallel projections of rectangular boxes according to the first Perkins' law (Section 2.3.1.) There were 16 different drawings, the rest were produced by rotating the original drawings by 0, 45°, ...315°. The original drawings all had angle C 120° (Figure A.13, and angle A varied from 70 to 170°. Three edges radiating from the central vertex were 3 cm, size of the pictures was 4.5×6 in

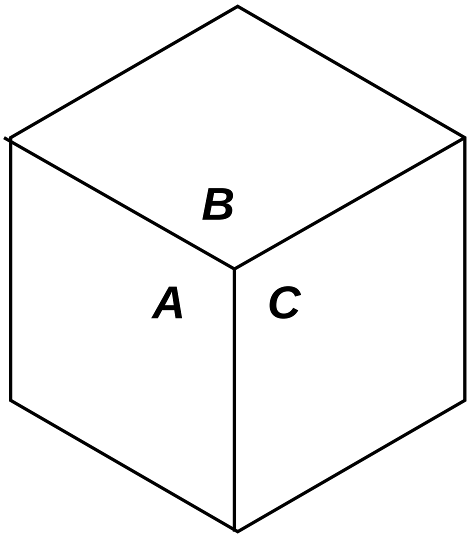
Viewing conditions. Unrestricted viewing from a position 6 feet away from the pictures. Three methods of presentation were used: –"straight", when pictures were presented one by one;

– "samples", when two pictures with A equal to 70° (nonrectangular case) and 120° (rectangular case) were permanently displayed next to the judged picture;

– "pairs", when pictures were sorted into two piles and the "rectangularity status" (rectangular, nonrectangular, borderline) was different for corresponding cards in two piles.

Task. In the first two cases, the task was to say if the box in the picture appeared to be rectangular. In the last case, the task was to say which box appears to be more rectangular.

Results. For each procedure the percentage of correct judgements for rectangular pictures was close to 90%, for nonrectangular it varied from 81% for the "straight" procedure, to 9% for "pairs". Any box was shown in 8 orientations. Distribution of the number of rectangular judgements (maximum 8) is shown in Figure A.14.



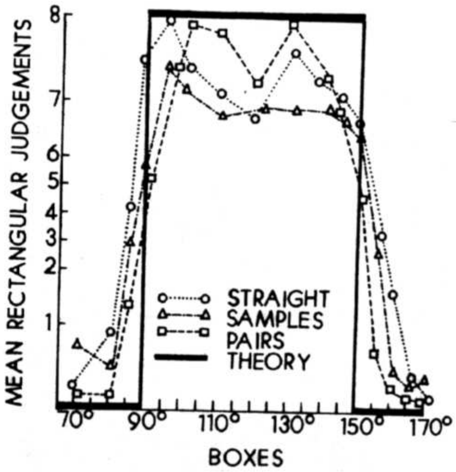


Figure A.13: (Left) One of the stimuli for [Perkins, 1972] □

Figure A.14: (Right) Results of [Perkins, 1972] □

[Perkins, 1973] Perkins, D. N. "Compensating for distortion in viewing pictures obliquely," *Perception & Psychophysics*, 14(1):13–18. 1973.

Experiment

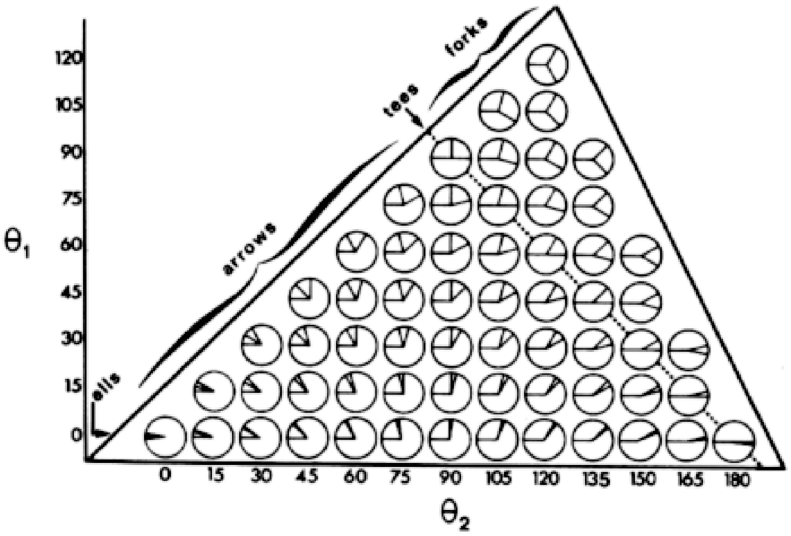
Subjects. 16.

Stimuli. Same as in [Perkins, 1972].

Viewing conditions. First, subject judged rectangularity of the pictures from 90° viewing direction, the viewing distance 6 ft. (32 cards only). Next, two viewing directions were used: 40° and 26°. A suspended wire frame ensured correct viewing direction, otherwise viewing conditions were unrestricted. The choice of the viewing directions was motivated by the following observation: classification of the reprojection of the drawings into an oblique plane according to the Perkins' laws change for a large number of pictures around 30°. In the neighborhood of 40 and 26° classification of these projections were stable (cf. Section 2.5.1).

Task. Subjects were asked to judge each picture as rectangular or nonrectangular box.

Results. All boxes for each of the two viewing directions could be classified into one of four categories, using four possible combinations of two criteria: whether the picture satisfies the first Perkins' law, and whether the projection of the picture onto the plane perpendicular to the viewing direction satisfies this law. The borderline case was excluded.



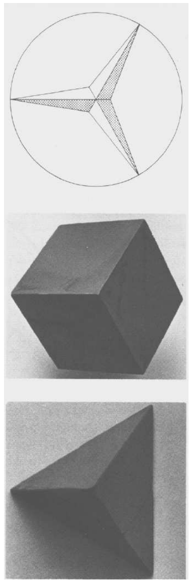


Figure A.15: (Left) Stimuli from [Shepard, 1981] □

Figure A.16: (Right) Three-dimensional objects from [Shepard, 1981] □

Judgements were made mostly according to the classification of the picture, not of the projection. In the cases of a conflict, the percent of judgments coinciding with the picture classification was lower, but still high (above 68%) in all cases except orthogonally nonrectangular, projectively rectangular (43%). There was very little projective influence at 41°.

[Shepard, 1981] Shepard, R. N. "Psychophysical complimentarity," In *Perceptual Organization*, Kubovy, M. and Pomerantz, J. R., editors, pages 279–343. Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey. 1981.

Experiment

Subjects. Unknown.

Stimuli. 61 line drawings consisting of three radial lines meeting at different angles at one point (Figure A.15.)

Viewing conditions. Circular aperture in front of the drawing, otherwise unrestricted viewing.

Task. Subjects were asked if each pattern could be a projection of a a corner of a cube (all three angles 90°), of a tetrahedron (all three angle 60°) or of a planar "Mercedes-Benz" sign (all three angles 120°) (Figure A.16.) Subject could rotate each drawing in its plane around the center with a remote control.

Results. Dashed lines in Figure A.17 indicate the boundaries of the areas of orthogonal parallel projections of corresponding spatial corners. For the corners there is no difference between orthogonal parallel and central projections assuming that the principal ray of the central projection goes through the point where the lines meet. The size of the circles is proportional to the fraction of positive answers.

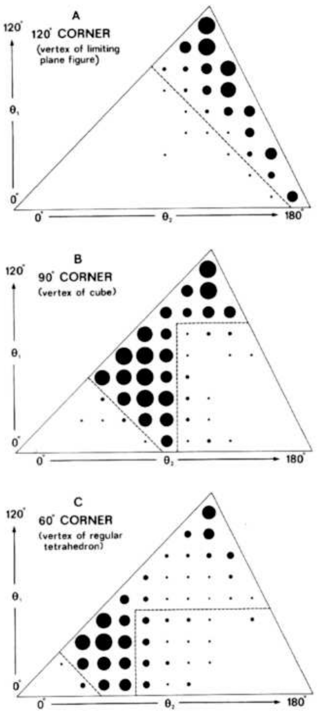


Figure A.17: Results of [Shepard, 1981] □

• For the cubic corner there is close agreement of 50% threshold of positive answers with orthogonal projection boundaries. This confirms Perkins' first law.

• For the "Mersedes-Benz" figure the agreement was less than for the cubic corner, and for 60° corner even less. However, most of the change was in the direction of decreasing the ranges of acceptable configurations.

A.6 Foreshortening

[Nicholls and Kennedy, 1993a] Nicholls, A. L. and Kennedy, J. M. "Angular subtense effects on perception of polar and parallel projections of cubes.," *Perception & Psychophysics*, 54(6):763–772. 1993.

Experiment 1

Subjects. 14.

Stimuli. Line drawings of a cube, orthogonal parallel projection and central projection with 15° and 33° projection angles.

Viewing conditions. Projections were viewed monocularly, viewing angles were 35°, 15°, 5°, 4°, 2°.

Task. The subjects were asked to compare the pictures with a standard (the picture with the projection angle 35° viewed at a distance corresponding to the viewing angle 15°.)

Results.

• Moderate parallel foreshortening (15°) was always the best; parallel projection was second best and strong parallel foreshortening (35°) was the worst.

• Parallel projection is perceived as distorted only at close distances; only for 35° it was worse than strongly convergent.

Experiment 2

Subjects. 45.

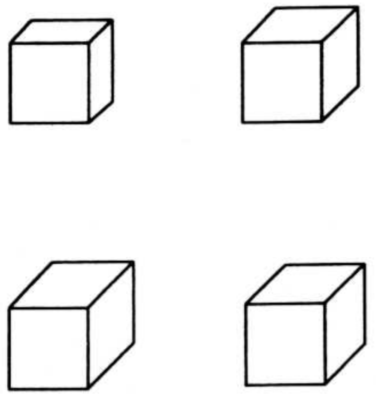
Stimuli. Same as 1, but the image with moderate degree of parallel foreshortening (15° projection angle) positioned at the distance corresponding to the viewing angle 5° was used as standard.

Viewing conditions. Same.

Results. Order of preference was preserved, except for the closest distance, when the pictures with the strongest parallel foreshortening was preferred over the parallel projection (no parallel foreshortening,) but they were still judged to be slightly worse than the pictures with moderate parallel foreshortening.

Experiment 3.

Subjects. 15.



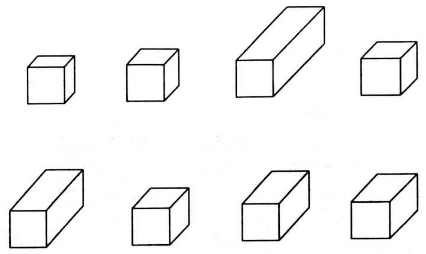


Figure A.18: (Left) Stimuli from [Nicholls and Kennedy, 1993b], Experiment 1 □

Figure A.19: (Right) Stimuli from [Nicholls and Kennedy, 1993b], Experiment 2 □

Stimuli. Same.

Viewing conditions. The viewing angles were 50°, 45°, 40°, 35° and 30°. Standard image (15° projection, moderate degree of parallel foreshortening)set at a position corresponding to the viewing angle 45°.

Results. Moderate parallel foreshortening was still rated the best at all distances, but now parallel projection had lower ratings than strongly divergent.

[Nicholls and Kennedy, 1993b] Nicholls, A. L. and Kennedy, J. M. "Foreshortening and the perception of parallel projections.," *Perception & Psychophysics*, 54(5):665–674. 1993.

Experiment 1.

Subjects. 90.

Stimuli. Line drawings of a cube in oblique parallel projection. One face was always depicted as a square, oblique lines representing the lateral edges of the cube met the horizontal at 45, 30, 60°, the degree of perpendicular foreshortening (the ratio of the length of the edges of the front to the length of the oblique lines) was 1:0.38, 1:0.65, 1:0.73 and 1:0.81 (Figure A.18).

Viewing conditions. Unrestricted viewing.

Task. Subjects were asked to choose best drawings among the set with a fixed angle.

Results. 1:0.65 ratio was consistently preferred.

Experiment 2.

Subjects. 144.

Stimuli. Three wooden blocks with square cross-sections with length-to-width ratios of 0.65, 1 and 2; line drawings with receding lines equal to 0.4, 0.5, 0.65, 0.8, 1.1 , 1.2, 1.5, 2 of the front edges (Figure A.19).

Viewing conditions. Unrestricted viewing.

Task. The subjects examined one out of three wooden blocks. After that, they were asked to choose the best drawing.

Results. .

For 1:0.65 block 1:0.42 picture was preferred, for 1:1 block 1:0.65 was preferred, and for 1:2 block 1:1.6 was preferred. In first two cases the preferred perpendicular foreshortening was 1:0.65, in the third case 1:0.8. About 40% of the subjects made the dominating choice.

Experiment 3.

Subjects. 72.

Stimuli. Same blocks as before, but the pictures were either frontal orthogonal parallel projections or frontal orthogonal central projections. (Figure A.20).

Viewing conditions. Unrestricted viewing.

Task. Same.

Results. For parallel projection, 1:0.65 and 1:1 blocks, the preferred perpendicular foreshortening was 0.8, for 1:2 block - 1:1. For central projections it was 0.5 and 0.65 respectively. About 30% of the subjects made the dominating choice.

[Hagen and Elliot, 1976] Hagen, M. A. and Elliot, H. B. "An investigation of the relationship between viewing condition and preference for true and modified linear perspective with adults," *Journal of Experimental Psychology: Human Perception & Performance*, 2(4):479–490. 1976.

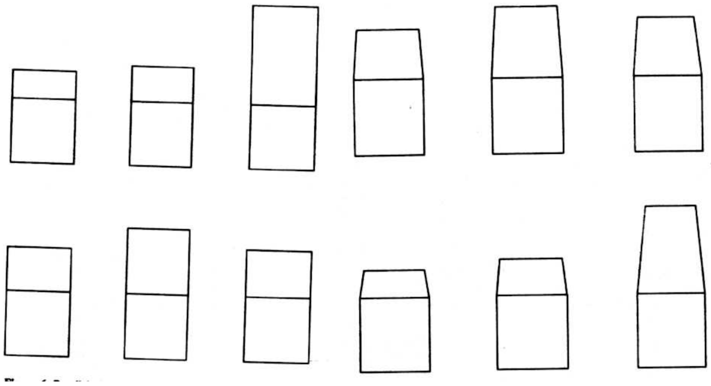


Figure A.20: Stimuli from [Nicholls and Kennedy, 1993b], Experiment 3 □

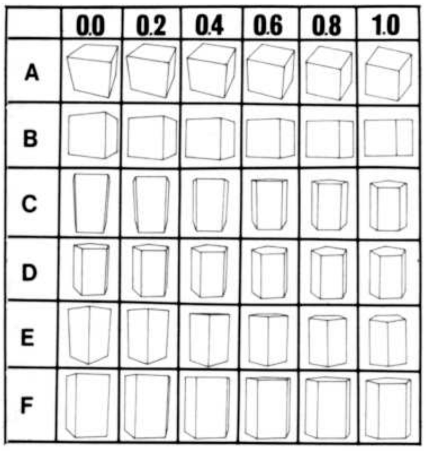


Figure A.21: Stimulus from [Hagen and Elliot, 1976] □

Experiment 1-2.

Subjects. 24 college students in Experiment 1, 33 college students in Experiment 2.

Stimuli. 7 sets of images of polyhedra (Figure A.21) generated by the method described in [Reggini, 1975] (see also Section 4.1), convergence degree from 0.0 to 1.0, Angular size from 19° to 58°.

Viewing conditions. a) monocular peephole, from fixed distance, from 90° and 45° viewing directions. b) free viewing. 3D solids were shown prior to the experiment.

Task. a) assign rank in the range 1 to 6. b) Sort the pictures from the most to the least natural looking.

Results. Parallel projection is consistently rated the best from both viewing points. For the free viewing the result is same; the only exception are the pictures of cubes (0.8 is the most frequently chosen).

Experiment 3

Subjects. 20 college students.

Stimuli. Same.

Viewing conditions. Viewing point coincided with the projection point for the degree of convergence 0.0 (regular perspective.) Monocular viewing, peephole.

Task. Subjects were asked to rank the pictures from most to least natural looking or from most to least accurate drawing.

Results. Parallel projection has the highest rating. No dependence on the form of instructions was observed.

Experiment 4.

Subjects. 28 college students.

Stimuli. Sets D,F from Experiments 1-3.

Viewing conditions. Viewing point coincided with the projection point for the degree of convergence 0.0. monocular viewing, peephole. Stimuli were presented three at a time, each picture was oriented perpendicular to the viewing direction.

Task. Subjects were asked to choose the most and the least natural drawing from each triplet.

Results. Parallel projection was chosen as the most realistic more frequently than any other degree of convergence.

Note. This study was critisized by other authors. The images in sets C, E, F differed not only in the degree of convergence, but also in visibility of faces of the polyhedra. The use of Reggini's perspective seems to be unnecessary, because the degree perspective convergence (parallel foreshortening) could be varied simply by changing the viewing point.

[Hagen et al., 1978a] Hagen, M. A., Elliott, H. B., and Jones, R. K. "A distinctive characteristic of pictorial perception: The zoom effect," *Perception*, 7(6):625–633. 1978.

Experiment 1.

Subjects. 60 college students.

Stimuli. Picture sets D,F from [Hagen and Elliot, 1976]. 3 types of stimuli: a) black and white line drawings, normal light; b) Dayglo (fluorescent paper), under normal light; c) Dayglo stimuli under ultraviolet light.

Viewing conditions. Monocular viewing, peephole, the viewing point coincides with the projection point for the degree of convergence 0.0. Three pictures from the same set were presented simultaneously. The viewing direction was 90°. 3D objects were shown to the subjects before the experiment.

Task. Subjects were instructed to choose the most realistic and natural looking drawing out of three.

Results.

- Under the normal light parallel projection was preferred for all types of stimuli.
- For Dayglo stimuli under ultraviolet light there was no particular preference.

Experiment 2.

Subjects. 20.

Stimuli. Dayglo stimuli from Experiment 1 under ultraviolet light.

Viewing conditions. Monocular viewing with head motion allowed.

Task. Same as in Experiment 1.

Results. Results are similar to those obtained under normal light - parallel projection is preferred.

Experiment 3.

Subjects. 20.

Stimuli. Same.

Viewing conditions. Binocular viewing with fixed position of the head.

Task. Same.

Results. Slight change in the direction of normal light preferences, but not nearly as radical.

A.7 Visual field

[Finke and Kurtzman, 1981] Finke, R. A. and Kurtzman, H. S. “Mapping the visual field in mental imagery," *Journal of Experimental Psychology: General*, 110(4):501–517. 1981.

Experiment 1.

Subjects. 30.

Stimuli. A circle 76 cm in diameter, with eight radial lines. A disk 2.5 cm in diameter was placed in the middle; each half of the disk contained gratings 1, 3 or 9 cycle-per-degree (cpd). Gratings had the same number of cpd and were perpendicular to each other (Figure A.22). The disk in the middle was oriented so that the gratings were parallel or perpendicular to one of the drawn diameters of the circle.

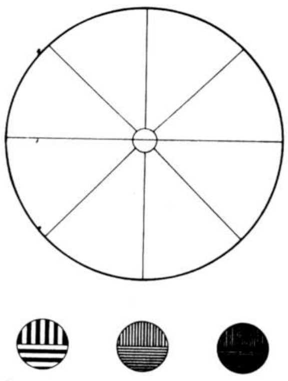


Figure A.22: Stimuli from [Finke and Kurtzman, 1981] □

Viewing conditions. The viewing point was fixed 28 cm away from the disk, resulting in the angular size of the small disk approximately 5°. and the size of the circle approximately 104°. Viewing was binocular.

Task.

a) Subjects were asked to move their eyes along each diameter, until they could not see that the disk in the middle consists of two distinct parts. A small red dot on the end of a rod was used as a fixation point.

b) The subjects were trained to *imagine* the pattern in the middle (the actual pattern was removed) and the same experiment was performed.

Note. The details of instructions are important for part b); see the paper for details.

Results.

• The average field size both for real and imagined stimuli decreased with increase in the spatial frequency of the gratings: 31.2 for 1 cpd, 27.9 for 3 cpd, 14.3 for 9 cpd.

• The measured visual field was longer horizontally than vertically (aspect ratio 1.27, 1,24, 1.17 for 1,3,9 cpd respectively.)

• Imagery fields were in close correspondence with the perceptual fields.

Experiment 2.

Subjects. 10.

This experiment was designed to check the influence of knowledge or expectations on the results of part b) of the previous experiment. Subjects were asked to guess the outcome of the experiment that was described to them (actual stimuli were demonstrated).

Results. Resulting field sizes for 3 and 9 cpd gratings were considerably lower than in Experiment 1; the field size estimate for 3 cpd was closer to the 9 cpd estimate contrary to the results of the Experiment 1. Average predicted visual fields were circular (no eccentricity). 3 subjects, however, did predict considerable eccentricity in the correct direction.

Note. Experiment 3 was mostly irrelevant for our purposes and is omitted.

A.8 Straightness and Curvature

[Watt et al., 1987] Watt, R. J., Ward, R. M., and Casco, C. "The detection of deviation from straightness in lines," *Vision Research*, 27(9):1659–1678. 1987.

Experiment 1.

Subjects. 3 (authors).

Stimuli. Various perturbations of a straight line segment presented on the screen of an oscilloscope. The segments were 60 arc min. in length, the size of the perturbations along the line was in the range 5–40 arc min. There were 3 types of perturbations: vernier-like, bump-like, S-like. Each type of perturbation had various types of smoothing: none (discontinuous constant bumps), Gaussian, linear (Figure A.23, Figure A.24, Figure A.25).

The maximal amplitude of the perturbation in the direction perpendicular¡ to the bump varied. The

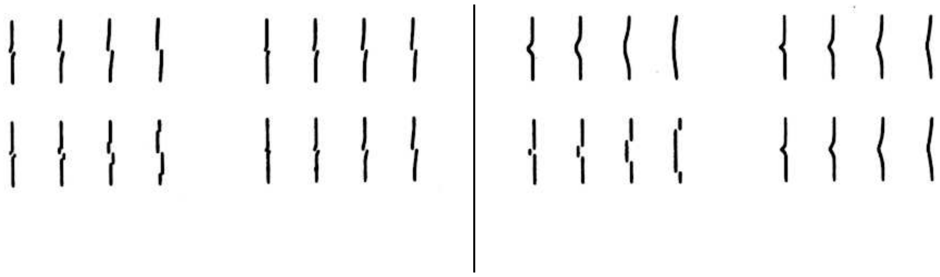


Figure A.23: (Left) Vernier-like stimuli from [Watt et al., 1987] □

Figure A.24: (Right) Bump-like stimuli from [Watt et al., 1987] □

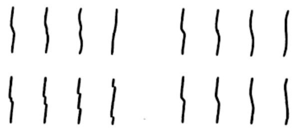


Figure A.25: S-like stimuli from [Watt et al., 1987] □

resolution of the display was 0.3 arc min. The direction (sign) of the perturbation varied randomly. The orientation of the line varied randomly in a 4° range around the vertical axis.

Viewing conditions. Unrestricted viewing from the distance 2.8 m.

Task. The subjects indicated the sign of the presented stimulus, if they could not decide, they guessed.

Results. On the basis of the data, threshold amplitudes for various stimulus types and size were computed.

• The threshold amplitude was never below 3 arc sec.

• For vernier perturbations the threshold decreased with size of the perturbations; for the other two types it decreased for smaller sizes of the bump, but increased for larger sizes.

• The best fit to the data was provided by the hypothesis that the threshold is determined by the area of the maximal bump with respect to the least-square-root straight line through the data (threshold value 0.3 arc min.^{2}), if the amplitude is greater than 3 arc min.

Experiment 2a.

Subjects. Unknown.

Stimuli. Pairs of lines on the screen of an oscilloscope. One of the lines was straight, the dots of the other line were perturbed according to a Gaussian distribution with varying means. The length of the lines varied.

Viewing conditions. Same.

Task. Subjects were asked to choose the perturbed line.

Results. According to a computer simulation, the largest bump rule would produce threshold vs. length dependence with exponent -0.37. Probability summation (bumps are detected with some probability) rule gives exponents between -0.37 and -1.0. Complete summation (all bumps are used) corresponds to the exponent -1.0. The exponents found in the experiment were significantly less than -0.37 and greater than -1.0 (weak probability summation).

Experiment 2b.

Subjects. Unknown

Stimuli. Bump-like perturbation with no smoothing from the Experiment 1 was used. Line length was 15, 45 , 30 arc min. A wide range of perturbation lengths was considered.

Viewing conditions. Same.

Task. Same.

Results. It appears that for larger perturbation sizes (more than have of the line) the bumps on the ends of the line are detected rather than the large bump in the middle. This results in increase of the threshold for larger perturbation sizes (cf. Experiment 1).

Experiment 3a.

Subjects. Unknown

Stimuli. Lines with two Gauss-like bumps on one side, with length 60 arc min. displayed on the screen of a oscilloscope. A random Gaussian perturbation with variable standard deviation and zero mean was added.

Viewing conditions. Same.

Task. Same as in Experiment 1.

Results. Amplitude threshold was plotted against the standard deviation of the added noise. Assuming the existence of internal noise in determination of position of the points it was estimated to be approximately 3 arc sec.

Experiment 3b.

Subjects. Unknown.

Stimuli. Lines with three types of perturbations: the first, second and third derivatives of a Gaussian with variable spatial extent.

Viewing conditions. Same.

Task. Same as in Experiment 1.

Results. Threshold decreases up to certain size of the perturbation, than slightly increases. This critical size value is relatively independent of the type of the perturbation approximately 5 arc min.) This might be an indication that there is a limit for spatial integration.

[Ogilvie and Daicar, 1967] Ogilvie, J. and Daicar, E. "The perception of curvature," *Canadian Journal of Psychology*, 21(6):521–525. 1967.

Experiment

Subjects. 5.

Stimuli. Black curved lines 1.5 mm wide on a 5×7 in photographic plate. The radii of curvature were 2.62, 3.15, 4.06, 5.55 and 10.1 m. Two chord length were used: 122 and 51 m. Define normalized curvature as the chord length divided by the radius of curvature. Normalized curvature approximates the displacement of the end point of the line with respect to the tangent at the starting point expressed as a fraction of the chord length. For 122 mm chord length it was equal to 0.047, 0.039, 0.030, 0.022, 0.012.

Viewing conditions. The viewing point was 10.1 m away from the display, resulting in angular subtens of the lines of 412 and 17 arc min. A 90 cm aperture was used to restrict the field of view to 2.75°. Viewing was monocular.

Task. Lines were presented to the subjects in different orientations (vertical, horizontal, oblique 45°). Subjects were asked to indicate if the line is curved up or down (left or right). Each line was presented 20 times, 6 lines were presented in random order.

Results. Calculated average thresholds were best correlated with the segment angular area which was calculated using an approximate formula (2 ×chord length ×sagitta / 3, where sagitta is the distance from the middle of the arc to the chord.) Threshold was in the range 3.1-3.9 log sq. sec. (0.4 - 2.2 sq. min.)

A.9 Verticality

[DiLorenzo and Rock, 1982] DiLorenzo, J. R. and Rock, I. "The rod-and-frame effect as a function of the righting of the frame.," *Journal of Experimental Psychology: Human Perception & Performance*, 8(4):536–546. 1982.

Experiment 1.

Subjects. 12.

Stimuli. A luminous vertical rod inside a luminous frame. The tilt of the frame could be adjusted.

Viewing conditions. Unrestricted viewing from a fixed position, viewing angle 54°, in a dark room (only the frame and the rod were visible.)

Task. The subjects were shown the frame tilted at 20°. After 4 min., the subjects were asked to adjust the frame to the vertical position. After that, the rod tilted at 14° and the frame tilted at 20° in the same direction were shown. The subjects were asked to adjust the rod to the vertical position (the frame remained tilted). Same experiment was done with the head of the observer tilted at 45°.

Results. The average righting effect (the tilt of the single frame after adjustment) was 7.7°; the average rod-and-frame (RFE) effect was 11° and 8.7° for counterclockwise and clockwise directions. When the head was tilted, the effects increased.

Experiment 2.

Subjects. 10.

Stimuli. Similar to the Experiment 1, with an internal frame added (angular subtense 32°).

Viewing conditions. Same.

Task. The outer frame was presented in an upright position for 4 min. Than adjusted if the observer perceived it as tilted. After that it was returned to the objective vertical. Then the inner frame was added at a tilt of 20°. and Experiment 1 was performed with the inner frame as the frame. Same experiment was performed with tilted head.

Results. No righting effect or RFE were observed in both cases.

Experiment 3.

Subjects. 10.

Stimuli. Same as Experiment 2.

Viewing conditions. Same.

Task. Similar to Experiment 2, but with outer frame tilted and inner frame vertical.

Results. Approximately 8-9° righting effect for both frames and 5.8° RFE were observed. When the head was tilted, the righting effect increased for both frames, RFE effect increased mostly insignificantly.

Appendix B

Reconstruction of the center of projection from a projection of a rectangle

Suppose we know that a picture contains a perspective projection of a rectangle, which is a nondegenerate quadrangle. Let l be the line going through the center of projection O perpendicular to the projection plane p (Figure B.1.)

Let's assume in addition that we know the position of the intersection point O' of the line l and the projection plane p. If the picture was made with a regular camera, then O' is the center of the picture.

Introduce a coordinate system with the origin at point O, z-axis directed along the normal to the plane and x and y-axes parallel to the projection plane.

Let $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, $\mathbf{x}_4$ be the known projections of the vertices of a rectangle with unknown vertices $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, $\mathbf{v}_4$. Let $\mathbf{d} = \vec{OO'} = (0, 0, d)$. Let $\mathbf{e}$ be the unit vector in the direction of $\mathbf{d}$. Then for all j $x_j^3 = d$.

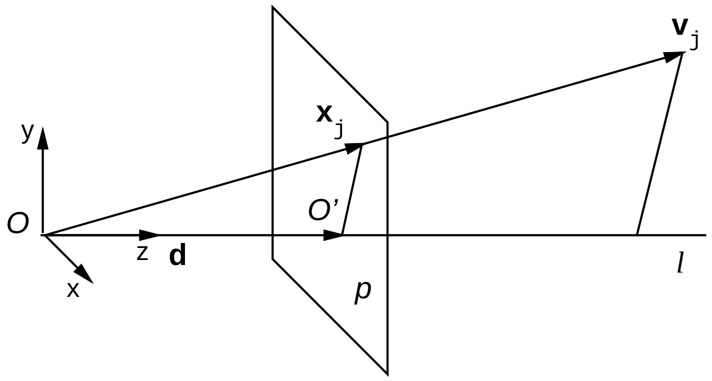


Figure B.1: □

We will use parentheses ( ) for the scalar product of a pair of vectors and for the triple product of a triple of vectors. We will use brackets [ ] for the vector product.

The fact that $\mathbf{x}_\mathrm{j}$ is the projection of $\mathbf{v}_\mathrm{j}$ can be expressed as

$$\frac{\mathbf{x}_\mathrm{j}(\mathbf{v}_\mathrm{j}, \mathbf{d})}{d^2} = \mathbf{v}_\mathrm{j} \tag{B.1}$$

The sides of the rectangle should be equal:

$$\mathbf{v}_2 - \mathbf{v}_1 = \mathbf{v}_3 - \mathbf{v}_4 \tag{B.2}$$

The lengths of the diagonals should be equal

$$|\mathbf{v}_3 - \mathbf{v}_1| = |\mathbf{v}_4 - \mathbf{v}_2|$$

which is equivalent to

$$(\mathbf{v}_3 - \mathbf{v}_1)^2 = (\mathbf{v}_4 - \mathbf{v}_2)^2 \tag{B.3}$$

Subtracting $\mathbf{v}_1 + \mathbf{v}_3 = \mathbf{v}_2 + \mathbf{v}_4$, which follows from (B.2), from (B.3) and simplifying we get

$$(\mathbf{v}_3\ \mathbf{v}_1) = (\mathbf{v}_4\ \mathbf{v}_2) \tag{B.4}$$

Substituting the expression for $\mathbf{v}_\mathrm{j}$ from (B.1) into (B.2) and (B.3) results in

$$\frac{1}{d}\left((\mathbf{v}_1\ \mathbf{e})\mathbf{x}_1 - (\mathbf{v}_2\ \mathbf{e})\mathbf{x}_2 + (\mathbf{v}_3\ \mathbf{e})\mathbf{x}_3 - (\mathbf{v}_4\ \mathbf{e})\mathbf{x}_4\right) = 0 \tag{B.5}$$

$$\frac{1}{d}(\mathbf{x}_1\ \mathbf{x}_3)(\mathbf{v}_1\ \mathbf{e})(\mathbf{v}_3\ \mathbf{e}) = \frac{1}{d}(\mathbf{x}_2\ \mathbf{x}_4)(\mathbf{v}_2\ \mathbf{e})(\mathbf{v}_4\ \mathbf{e}) \tag{B.6}$$

Multiplying the vector equation (B.5) by $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$ we get the following system of equations in variables $(\mathbf{v}_1\ \mathbf{e})$, -$(\mathbf{v}_2\ \mathbf{e})$, $(\mathbf{v}_3\ \mathbf{e})$ (we leave $(\mathbf{v}_4\ \mathbf{e})$ as a parameter):

$$\begin{pmatrix} (\mathbf{x}_1\ \mathbf{x}_1) & (\mathbf{x}_1\ \mathbf{x}_2) & (\mathbf{x}_1\ \mathbf{x}_3) \\ (\mathbf{x}_2\ \mathbf{x}_1) & (\mathbf{x}_2\ \mathbf{x}_2) & (\mathbf{x}_2\ \mathbf{x}_3) \\ (\mathbf{x}_3\ \mathbf{x}_1) & (\mathbf{x}_3\ \mathbf{x}_2) & (\mathbf{x}_3\ \mathbf{x}_3) \end{pmatrix} \begin{pmatrix} (\mathbf{v}_1\ \mathbf{e}) \\ -(\mathbf{v}_2\ \mathbf{e}) \\ (\mathbf{v}_3\ \mathbf{e}) \end{pmatrix} = \begin{pmatrix} (\mathbf{x}_1\ \mathbf{x}_4) \\ (\mathbf{x}_2\ \mathbf{x}_4) \\ (\mathbf{x}_3\ \mathbf{x}_4) \end{pmatrix} (\mathbf{v}_4\ \mathbf{e}) \tag{B.7}$$

The matrix of the left-hand side of this system is the Gram's matrix of the vectors $\mathbf{x}_1$,$\mathbf{x}_2$, $\mathbf{x}_3$. By assumption, the projection of the rectangle is not a line segment. Therefore, the rank of this matrix must be 3, and the system has a unique solution.

We can also note that we can express the determinant of the matrix in the system (B.7) as $D = (\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3)^2$, where $(\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3)$ is the triple product $([\mathbf{x}_1\ \mathbf{x}_2]\,\mathbf{x}_3)$. This is a consequence of the following general identity:

$$(\mathbf{abc})\,(\mathbf{efg}) = \begin{vmatrix} (\mathbf{a}\ \mathbf{e}) & (\mathbf{a}\ \mathbf{f}) & (\mathbf{a}\ \mathbf{g}) \\ (\mathbf{b}\ \mathbf{e}) & (\mathbf{b}\ \mathbf{f}) & (\mathbf{b}\ \mathbf{g}) \\ (\mathbf{c}\ \mathbf{e}) & (\mathbf{c}\ \mathbf{f}) & (\mathbf{c}\ \mathbf{g}) \end{vmatrix} \tag{B.8}$$

The solutions of the system are

$$(\mathbf{v}_1\ \mathbf{e}) = \frac{(\mathbf{v}_4\ \mathbf{e})}{D} \begin{vmatrix} (\mathbf{x}_1\ \mathbf{x}_4) & (\mathbf{x}_1\ \mathbf{x}_2) & (\mathbf{x}_1\ \mathbf{x}_3) \\ (\mathbf{x}_2\ \mathbf{x}_4) & (\mathbf{x}_2\ \mathbf{x}_2) & (\mathbf{x}_2\ \mathbf{x}_3) \\ (\mathbf{x}_3\ \mathbf{x}_4) & (\mathbf{x}_3\ \mathbf{x}_2) & (\mathbf{x}_3\ \mathbf{x}_3) \end{vmatrix}$$

$$(\mathbf{v}_2\ \mathbf{e}) = \frac{(\mathbf{v}_4\ \mathbf{e})}{D} \begin{vmatrix} (\mathbf{x}_1\ \mathbf{x}_1) & (\mathbf{x}_1\ \mathbf{x}_4) & (\mathbf{x}_1\ \mathbf{x}_3) \\ (\mathbf{x}_2\ \mathbf{x}_1) & (\mathbf{x}_2\ \mathbf{x}_4) & (\mathbf{x}_2\ \mathbf{x}_3) \\ (\mathbf{x}_3\ \mathbf{x}_1) & (\mathbf{x}_3\ \mathbf{x}_4) & (\mathbf{x}_3\ \mathbf{x}_3) \end{vmatrix}$$

$$(\mathbf{v}_3\ \mathbf{e}) = \frac{(\mathbf{v}_4\ \mathbf{e})}{D}\begin{vmatrix} (\mathbf{x}_1\ \mathbf{x}_1) & (\mathbf{x}_1\ \mathbf{x}_2) & (\mathbf{x}_1\ \mathbf{x}_4) \\ (\mathbf{x}_2\ \mathbf{x}_1) & (\mathbf{x}_2\ \mathbf{x}_2) & (\mathbf{x}_2\ \mathbf{x}_4) \\ (\mathbf{x}_3\ \mathbf{x}_1) & (\mathbf{x}_3\ \mathbf{x}_2) & (\mathbf{x}_3\ \mathbf{x}_4) \end{vmatrix}$$

Using identity (B.8) the solutions can be simplified:

$$(\mathbf{v}_1\ \mathbf{e}) = \frac{(\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3)(\mathbf{x}_4\ \mathbf{x}_2\ \mathbf{x}_3)}{(\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3)^2}(\mathbf{v}_4\ \mathbf{e}) = \frac{(\mathbf{x}_4\ \mathbf{x}_2\ \mathbf{x}_3)}{(\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3)}(\mathbf{v}_4\ \mathbf{e})$$

$$-(\mathbf{v}_2\ \mathbf{e}) = \frac{(\mathbf{x}_1\ \mathbf{x}_4\ \mathbf{x}_3)}{(\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3)}(\mathbf{v}_4\ \mathbf{e}), \quad (\mathbf{v}_3\ \mathbf{e}) = \frac{(\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_4)}{(\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3)}(\mathbf{v}_4\ \mathbf{e})$$

Substituting these expressions for $(\mathbf{v}_1\ \mathbf{e})$, $(\mathbf{v}_2\ \mathbf{e})$, $(\mathbf{v}_3\ \mathbf{e})$ into (B.5) and dividing by $(\mathbf{v}_4\ \mathbf{e})^2/(\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3)^2$, we get the following equation:

$$(\mathbf{x}_1\ \mathbf{x}_3)(\mathbf{x}_4\ \mathbf{x}_2\ \mathbf{x}_3)(\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_4) = -(\mathbf{x}_2\ \mathbf{x}_4)(\mathbf{x}_1\ \mathbf{x}_4\ \mathbf{x}_3)(\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3) \tag{B.9}$$

The z-component of the vectors $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$ is equal to d. Manipulating the determinants, we get

$$(\mathbf{x}_\mathrm{i}\ \mathbf{x}_\mathrm{j}\ \mathbf{x}_\mathrm{k}) = \begin{vmatrix} x_\mathrm{i}^1 & x_\mathrm{j}^1 & x_\mathrm{k}^1 \\ x_\mathrm{i}^2 & x_\mathrm{j}^2 & x_\mathrm{k}^2 \\ d & d & d \end{vmatrix} = d\begin{vmatrix} x_\mathrm{i}^1 & x_\mathrm{j}^1 & x_\mathrm{k}^1 \\ x_\mathrm{i}^2 & x_\mathrm{j}^2 & x_\mathrm{k}^2 \\ 1 & 1 & 1 \end{vmatrix} =$$

$$d\left(\begin{vmatrix} x_\mathrm{i}^1 & x_\mathrm{j}^1 \\ x_\mathrm{i}^2 & x_\mathrm{j}^2 \end{vmatrix} - \begin{vmatrix} x_\mathrm{i}^1 & x_\mathrm{k}^1 \\ x_\mathrm{i}^2 & x_\mathrm{k}^2 \end{vmatrix} + \begin{vmatrix} x_\mathrm{j}^1 & x_\mathrm{k}^1 \\ x_\mathrm{j}^2 & x_\mathrm{k}^2 \end{vmatrix}\right) =$$

$$d\left(\begin{vmatrix} x_\mathrm{i}^1 & x_\mathrm{j}^1 - x_\mathrm{k}^1 \\ x_\mathrm{i}^2 & x_\mathrm{j}^2 - x_\mathrm{k}^2 \end{vmatrix} - \begin{vmatrix} x_\mathrm{j}^1 & x_\mathrm{j}^1 - x_\mathrm{k}^1 \\ x_\mathrm{j}^2 & x_\mathrm{j}^2 - x_\mathrm{k}^2 \end{vmatrix}\right) = d\begin{vmatrix} x_\mathrm{i}^1 - x_\mathrm{j}^1 & x_\mathrm{j}^1 - x_\mathrm{k}^1 \\ x_\mathrm{i}^2 - x_\mathrm{j}^2 & x_\mathrm{j}^2 - x_\mathrm{k}^2 \end{vmatrix}$$

We denote this determinant D_{ijk}. Note that

$$|D_{ijk}| = |[x_\mathrm{i} - x_\mathrm{j}\ x_\mathrm{j} - x_\mathrm{k}]| \tag{B.10}$$

$x_\mathrm{i} = (x_\mathrm{i}^1, x_\mathrm{i}^2)$ are two-dimensional vectors in the projection plane.

Further, $(\mathbf{x}_3\ \mathbf{x}_1) = (x_3\ x_1) + d^2$, $(\mathbf{x}_2\ \mathbf{x}_4) = (x_4\ x_1) + d^2$. Equation (B.9) takes the form

$$D_{423}D_{124}((x_1\ x_3)+d^2)=-D_{123}D_{143}((x_2\ x_4)+d^2),$$

which is equivalent to

$$D_{423}D_{124}(\mathbf{x}_1\ \mathbf{x}_3)+D_{123}D_{143}(\mathbf{x}_2\ \mathbf{x}_4)=-d^2(D_{423}D_{124}+D_{123}D_{143}$$

.

This equation has a unique solution if

$$D_{423}D_{124}+D_{123}D_{143}\neq 0 \tag{B.11}$$

Consider the geometric interpretation of the quantities $(\mathbf{x}_\mathrm{i}\ \mathbf{x}_\mathrm{j}\ \mathbf{x}_\mathrm{k})$. It is the volume of the parallelepiped spanned by $\mathbf{x}_\mathrm{i}$, $\mathbf{x}_\mathrm{j}$, $\mathbf{x}_\mathrm{k}$ taken with the positive sign if the vectors form a right-hand triple, and with the negative sign if the vectors form a left-hand triple.

Assume counterclockwise ordering of the vertices of the quadrangle $\mathbf{x}_1\mathbf{x}_2\mathbf{x}_3\mathbf{x}_4$ if we look down from O (Figure B.2.) This assumption can be made without loss of generality, because if we change the ordering, the signs of all the volumes that we consider will change simultaneously, which doesn't affect the form of (B.9).

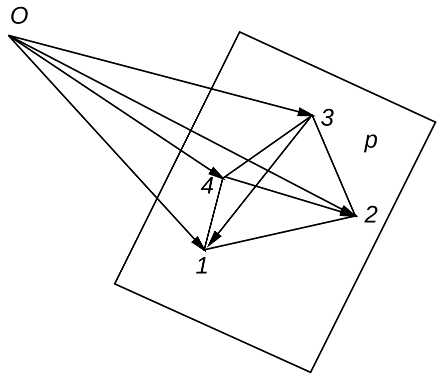


Figure B.2: □

From Figure B.2 we see that

$$(\mathbf{x}_4\ \mathbf{x}_2\ \mathbf{x}_3)=-Volume(O423),\qquad (\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_4)=-Volume(O124),$$

$$(\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3) = Volume(O123), \qquad (\mathbf{x}_1\ \mathbf{x}_4\ \mathbf{x}_3) = -Volume(O143),$$

where $Oijk$ denotes the pyramid with apex O and base ijk. The heights of all pyramids are equal to d. As $D_{ijk} = (\mathbf{x}_\mathrm{i}\ \mathbf{x}_\mathrm{j}\ \mathbf{x}_\mathrm{k})/d$, $|D_{ijk}| = S_{ijk}$, where S_{ijk} is the area of the triangle ijk. This can be also seen from (B.10).

We can use the information about the signs of $(\mathbf{x}_\mathrm{i}\ \mathbf{x}_\mathrm{j}\ \mathbf{x}_\mathrm{k})$ to write (B.9) in the following form:

$$S_{423}S_{124}(\mathbf{x}_1\ \mathbf{x}_3) - S_{123}S_{143}(\mathbf{x}_2\ \mathbf{x}_4) = -d^2\,(S_{423}S_{124} - S_{123}S_{143})$$

.

Condition (B.11) is

$$S_{423}S_{124} \neq S_{123}S_{143}. \tag{B.12}$$

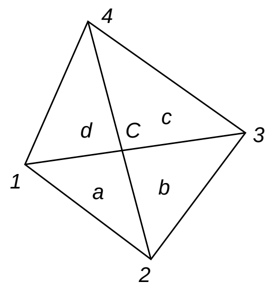


Figure B.3: □

The diagonals of the quadrangle 1234 divide it into four parts. Denote their areas by a, b, c, d (Figure B.3.) Then condition (B.12) means that

$$(c+b)(d+a) \neq (d+c)(a+b)$$

After rearrangement we get

$$(c-a)(b-d) \neq 0$$

which means that nonadjacent pieces must have unequal areas. This is a necessary condition for d to be determined from (B.9).

Denote $y_{ij} = x_\mathrm{i} - x_\mathrm{j}$. Then the formula for d^2 is

$$d^2 = \frac{(x_2\ x_4)\,|[y_{14}\ y_{43}]|\,|[y_{12}\ y_{23}]| - (x_1\ x_3)\,|[y_{42}\ y_{23}]|\,|[y_{12}\ y_{24}]|}{|[y_{14}\ y_{43}]|\,|[y_{12}\ y_{23}]| - |[y_{42}\ y_{23}]|\,|[y_{12}\ y_{24}]|} \tag{B.13}$$

The distance from the center of projection to the projection plane is given by (B.13) whenever the right-hand side is defined and positive. For the right-hand side of (B.13) to be defined it is neccessary and sufficient that the area of the triangle $C12$ *is not equal to the area of the triangle* $C34$ *and the area of the triangle* $C14$ *is not equal to the area of the triangle* $C23$ *(Figure B.3.)*

Bibliography

[Alberti, 1976] Alberti, L. B. *On painting.* Greenwood Press, Westport, Connecticut. 1976.

[Arnheim, 1954] Arnheim, R. *Art and visual perception; a psychology of the creative eye.* University of California Press, Berkeley. 1954.

[Carterette and Friedman, 1975] Carterette, E. C. and Friedman, M. P., editors. *Handbook of Perception: Seeing*, volume 5. Academic Press, New York. 1975.

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

[da Vinci, 1970] da Vinci, L. *Notebooks I, II.* Dover, New York. 1970.

[Deregowski, 1989] Deregowski, J. B. "Geometric restitution of perspective: Bartel's method.," *Perception*, 18(5):595–600. 1989.

[Deregowski, 1991] Deregowski, J. B. "Intercultural search for the origins of perspective," In *Contemporary issues in cross-cultural psychology*, Bleichrodt, N. and Drenth, P. J. D., editors, pages 334–346. Swets & Zeitlinger, Amsterdam, Netherlands. 1991.

[Deregowski and Parker, 1992] Deregowski, J. B. and Parker, D. M. "Convergent and divergent perspective," *Perception*, 21(4):441–447. 1992.

[Deregowski et al., 1994] Deregowski, J. B., Parker, D. M., and Massironi, M. "The perception of spatial structure with oblique viewing: An explanation for byzantine perspective?," *Perception*, 23(1):5–13. 1994.

[DiLorenzo and Rock, 1982] DiLorenzo, J. R. and Rock, I. "The rod-and-frame effect as a function of the righting of the frame.," *Journal of Experimental Psychology: Human Perception & Performance*, 8(4):536–546. 1982.

[Ernst, 1976] Ernst, b. *The Magic Mirror of M C. Escher.* Random House, New York. 1976.

[Feininger, 1953] Feininger, A. *Feininger on photography.* Crown Publishers, Inc., New York, revised edition. 1953.

[Finke and Kurtzman, 1981] Finke, R. A. and Kurtzman, H. S. "Mapping the visual field in mental imagery," *Journal of Experimental Psychology: General*, 110(4):501–517. 1981.

[Flocon and Taton, 1963] Flocon, A. and Taton, R. *La perspective.* Que sais-je? Le point des connaissances actuelle. Presses universitaires de France, Paris. 1963.

[Glaeser, 1994] Glaeser, G. *Fast Algorithms for 3-D Graphics.* Springer-Verlag, New York. 1994.

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

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

[Goncharenko et al., 1979] Goncharenko, E. N., Osipova, L. P., and Tomilin, M. G. "Lenses for wide-angle and panoramic photography," *Soviet Journal of Optical Technology*, 46(1):27–35. 1979.

[Goodman, 1976] Goodman, N. *Languages of art.* Hackett, Indianapolis, IN, 2nd edition edition. 1976. constructivist approach to picture perception.

[Hagen, 1976a] Hagen, M. A. "Development of ability to perceive and produce pictorial depth cue of overlapping," *Perceptual and Motor Skills*, 42(3, part 1):1007–1014. 1976.

[Hagen, 1976b] Hagen, M. A. "Influence of picture surface and station point on the ability to compensate for oblique view in pictorial perception.," *Developmental Psychology*, 12(1):57–63. 1976.

[Hagen, 1980] Hagen, M. A., editor. Academic Press series in cognition and perception. Academic Press, New York. 1980.

[Hagen, 1986] Hagen, M. A. *Varieties of realism : geometries of representational art.* Cambridge University Press, Cambridge ; New York. 1986.

[Hagen and Elliot, 1976] Hagen, M. A. and Elliot, H. B. “An investigation of the relationship between viewing condition and preference for true and modified linear perspective with adults,” *Journal of Experimental Psychology: Human Perception and Performance*, 2(4):479–490. 1976.

[Hagen et al., 1978a] Hagen, M. A., Elliott, H. B., and Jones, R. K. “A distinctive characteristic of pictorial perception: The zoom effect,” *Perception*, 7(6):625–633. 1978.

[Hagen and Jones, 1978] Hagen, M. A. and Jones, R. K. “Differential patterns of preference for modified linear perspective in children and adults,” *Journal of Experimental Child Psychology*, 26(2):205–215. 1978.

[Hagen et al., 1978b] Hagen, M. A., Jones, R. K., and Reed, E. S. “On a neglected variable in theories of pictorial perception: Truncation of the visual field,” *Perception and Psychophysics*, 23(4):326–330. 1978.

[Halloran, 1989] Halloran, T. O. “Picture perception is array-specific: Viewing angle versus apparent orientation,” *Perception and Psychophysics*, 45(5):467–482. 1989.

[Hill, 1990] Hill, F. S. *Computer Graphics.* Macmillan Publishing Co., Collier Macmillan Publishers, New York, London. 1990.

[Hockney, 1988] Hockney, D. *David Hockney : a retrospective / organized by Maurice Tuchman and Stephanie Barron.* Los Angeles County Museum of Art, Abrams, Los Angeles, New York. 1988.

[Inakage, 1991] Inakage, M. “Non-linear perspective projections,” In *Modeling in Computer Graphics: IFIP WG 5.10 Working Conference, Tokyo, Japan*, pages 203–215, Tokyo. Springer-Verlag. 1991.

[Klein, 1939] Klein, F. *Elementary mathematics from an advanced standpoint:geometry.* Macmillan company, New York. 1939.

[Kubovy, 1986] Kubovy, M. *The psychology of perspective and Renaissance art.* Cambridge University Press, Cambridge ¡Cambridgeshire¿ ; New York. 1986.

[La Gournerie, 1859] La Gournerie, J. d. *Traité de perspective linéare contenant les tracés pour les tableaux plans et courbes, les bas-reliefs et les décorations théatrales, avec une théorie des effets de perspective.* Dalmont et Dunod, Mallet-Bachelier, Paris. 1859. 1 volume and 1 atlas of plates.

[Lumsden, 1980] Lumsden, E. A. "Problems of magnification and minification: An explanation of the distortions of distance, slant, shape, and velocity," In [Hagen, 1980], pages 91–135. 1980.

[Mason, 1993] Mason, P. *History of Japanese Art.* Harry N. Abrams, Inc., New York. 1993.

[Milman, 1986] Milman, M. *Trompe-l'oeil, painted architecture.* Skira/Rizzoli, New York. 1986.

[Nicholls and Kennedy, 1993a] Nicholls, A. L. and Kennedy, J. M. "Angular subtense effects on perception of polar and parallel projections of cubes.," *Perception & Psychophysics*, 54(6):763–772. 1993.

[Nicholls and Kennedy, 1993b] Nicholls, A. L. and Kennedy, J. M. "Foreshortening and the perception of parallel projections.," *Perception & Psychophysics*, 54(5):665–674. 1993.

[Ogilvie and Daicar, 1967] Ogilvie, J. and Daicar, E. "The perception of curvature," *Canadian Journal of Psychology*, 21(6):521–525. 1967.

[Olmer, 1949] Olmer, P. *Perspective artistique.* Plon, Paris. 1943,1949.

[Ong and Kessinger, 1971] Ong, J. and Kessinger, D. J. "Perception of verticality with a rod and frame apparatus," *American Journal of Optometry & Archives of American Academy of Optometry*, 48(8):662–666. 1971.

[Perkins, 1968] Perkins, D. N. "Cubic corners," Quarterly Progress Report 89, MIT, Research Laboratory of Electronics. 1968.

[Perkins, 1972] Perkins, D. N. "Visual discrimination between rectangular and nonrectangular parallelopipeds," *Perception & Psychophysics*, 12(5):396–400. 1972.

[Perkins, 1973] Perkins, D. N. "Compensating for distortion in viewing pictures obliquely," *Perception and Psychophysics*, 14(1):13–18. 1973.

[Pirenne, 1970] Pirenne, M. H. *Optics, painting and photography.* Cambridge University Press, New York. 1970.

[Pontriagin, 1962] Pontriagin, L. S. *The mathematical theory of optimal processes.* Interscience Publishers, New York. 1962.

[Press et al., 1988] Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge University Press, Cambridge, UK. 1988.

[Raushenbakh, 1980] Raushenbakh, B. V. *Prostranstvennye postroenya v zhivopisi.* Nauka, Moskva. 1980. in Russian.

[Raushenbakh, 1986] Raushenbakh, B. V. *Sistemy perspektivy v izobrazitel'nom iskusstve : obshchaia teoriia perspektivy.* Nauka, Moskva. 1986. in Russian.

[Reggini, 1975] Reggini, H. C. "Perspective using curved projection rays and its computer application," *Leonardo*, 8:307–312. 1975.

[Rosinski and Faber, 1980] Rosinski, R. R. and Faber, J. "Compensation for viewing point in the perception of pictured space," In [Hagen, 1980], pages 137–176. 1980.

[Rosinski et al., 1980] Rosinski, R. R., Mulholland, T., Degelman, D., and Farber, J. "Picture perception: An analysis of visual compensation," *Perception & Psychophysics*, 28(6):521–526. 1980.

[Shepard, 1981] Shepard, R. N. "Psychophysical complimentarity," In *Perceptual Organization*, Kubovy, M. and Pomerantz, J. R., editors, pages 279–343. Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey. 1981.

[Smith, 1958a] Smith, O. W. "Comparison of apparent depth in a photograph viewed from two distances," *Perceptual and Motor Skills*, 8:79–81. 1958.

[Smith, 1958b] Smith, O. W. "Judgements of size and distance in photographs," *The American Journal of Psychology*, 71:529–538. 1958.

[Smith, 1961] Smith, Patricia C., S. O. W. "Ball throwing responses to photographically portrayed targets," *Journal of Experimental Psychology*, 62(3):223–233. 1961.

[Stark, 1928] Stark, F. *Das Netzhautbild: Verfahren zur Herstellung des wahren Sehbildes nach dem Grundprinzip des menslichen Sehens angewandt auf die zeihnerische Konstruktion der perspective.* by the author, Neuss am Rhein. 1928.

[Sullivan, 1989] Sullivan, M. *The Meeting of Eastern and Western Art.* University of California Press, Berkeley and Los Angeles. 1989.

[Taylor, 1963] Taylor, A. J. P. *The First World War: an illustrated history.* Hamish Hamilton, London. 1963.

[Thouless, 1931a] Thouless, R. H. "Phenomenal regression to the real object i," *British Journal of Psychology*, 21:339. 1931.

[Thouless, 1931b] Thouless, R. H. "Phenomenal regression to the real object ii," *British Journal of Psychology*, 22:1. 1931.

[Thouless, 1932] Thouless, R. H. "Individual differences in phenomenal regression to the real object," *British Journal of Psychology*, 22:216. 1932.

[Warner, 1983] Warner, F. W. *Foundations of differentiable manifolds and Lie groups*, volume 94 of *Graduate Texts in Mathematics.* Springer, New York, reprint edition. 1983.

[Watt et al., 1987] Watt, R. J., Ward, R. M., and Casco, C. "The detection of deviation from straightness in lines," *Vision Research*, 27(9):1659–1678. 1987.

[Zorin and Barr, 1995] Zorin, D. and Barr, A. H. "Corection of geometric perceptual distortions in pictures," In *Computer Graphics (SIGGRAPH '95 Proceedings).* 1995.