# 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

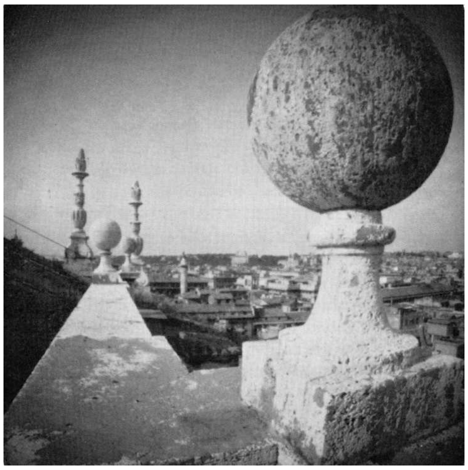


Figure 1.1: (Left) Wide-angle pinhole photograph taken on the roof of the Church of St. Ignatzio in Rome, classical example of perspective distortions from [Pirenne, 1970]. □

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;

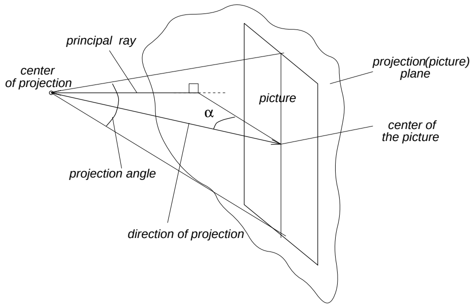
– to construct sample families of transformations in this framework.

## 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.



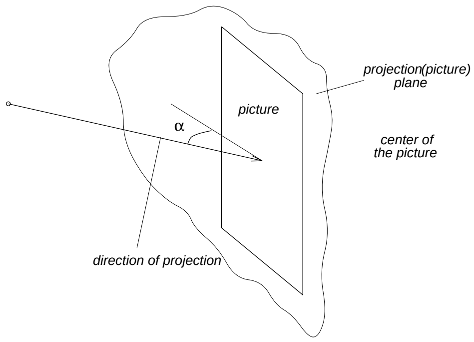




Figure 1.3: (Left) Central projection. □

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 $\alpha$ 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.*

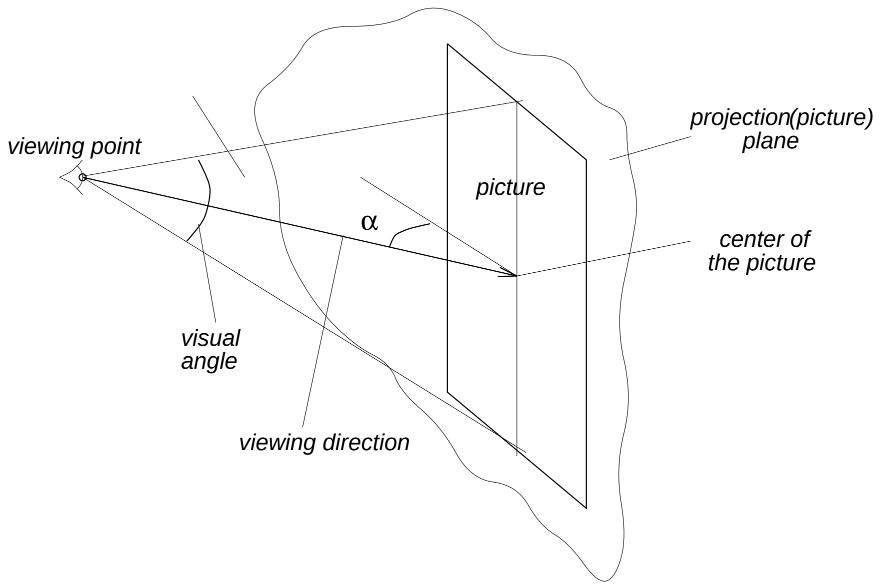



Figure 1.5: □

**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.