Olivier Faugeras

Three-Dimensional Computer Vision

A Geometric Viewpoint

The MIT Press
Cambridge, Massachusetts
London, England

at infinity, for the affine case, and the choice of two special points on that line, called the *absolute points*, for the euclidean case. The last notion in this section, that of a quadratic transformation, is used primarily in chapter 7 in the proof of the five-point algorithm; this also can be skipped on the first reading.

Section 2.5 takes us to the projective space. Since we will often model the real world as embedded in a projective space, it is important to read and understand this section. Its plan is very similar to that of the previous one. We introduce four fundamental concepts. First is the principle of duality, which extends to this case and makes points and planes equivalent. Second is the concept of the plane at infinity, which plays the same role as the line at infinity of the projective plane and the point at infinity of the projective line in helping us to understand the relationship between the projective space and the more familiar affine space. Third is the cross-ratio of four planes intersecting along a line which, like those defined in the two previous sections, is a very useful invariant. Fourth is the idea of a pencil of planes, which also plays an important role in clarifying the epipolar geometry. The concept of the absolute conic, although essential for one to understand the relationship between the euclidean and projective space, is used only in chapter 3 to interpret the significance of the intrinsic parameters of a camera, and in chapter 7 in the proof of the five-point algorithm. It may therefore be skipped on the first reading.

Projective spaces

We will begin with a general study of projective spaces of any dimension. A point of an n dimensional projective space, $\mathcal{P}^n$, is represented by an $n+1$ vector of coordinates $\mathbf{x} = [x_1, \ldots x_{n+1}]^T$, where at least one of the x_i is nonzero. The numbers x_i are sometimes called the *homogeneous* or *projective coordinates* of the point, and the vector $\mathbf{x}$ is called a *coordinate vector*. Two $n+1$ vectors $[x_1, \ldots x_{n+1}]^T$ and $[y_1, \ldots y_{n+1}]^T$ represent the same point if and only if there exists a nonzero scalar λ such that $x_i = \lambda y_i$ for $1 \le i \le n+1$. Therefore, the correspondence between points and coordinate vectors is not one to one, and this makes the application of linear algebra to projective geometry a little more complicated.

2.2.1 Collineations

We will now look at the linear transformations of a projective space. An $(n+1) \times (n+1)$ matrix $\mathbf{A}$ such that $det(\mathbf{A})$ is different from 0 defines a linear transformation or *collineation* from $\mathcal{P}^n$ into itself. It is easy to see that the set of collineations is a group. This group is also known as the *projective group*. The matrix associated with a given collineation is defined up to a nonzero scale factor, which we usually denote by

$$\rho\mathbf{y} = \mathbf{A}\mathbf{x} \quad \text{and also} \quad \mathbf{x} \,\overline{\wedge}\, \mathbf{y}$$

2.2.2 Projective basis

A *projective basis* is a set of $n+2$ points of $\mathcal{P}^n$ such that no $n+1$ of them are linearly dependent. For example, the set $\mathbf{e}_i = [0, \ldots, 1, \ldots, 0]^T$, $i = 1, \ldots, n+1$, where 1 is in the ith position, and $\mathbf{e}_{n+2} = [1, 1, \ldots, 1]^T$, is a projective basis, called the *standard projective basis*. Any point $\mathbf{x}$ of $\mathcal{P}^n$ can be described as a linear combination of any $n+1$ points of the standard basis. For example:

$$\mathbf{x} = \sum_{i=1}^{n+1} x_i \mathbf{e}_i$$

Let us now prove a very important proposition that we borrow from the book by Semple and Kneebone [SK52].

Proposition 2.1
Let $\mathbf{x}_1, \ldots, \mathbf{x}_{n+2}$ be $n+2$ coordinate vectors of points in $\mathcal{P}^n$, no $n+1$ of which are linearly dependent, i.e., a projective basis. If $\mathbf{e}_1, \ldots, \mathbf{e}_{n+1}, \mathbf{e}_{n+2}$ is the standard projective basis, there exist nonsingular matrices $\mathbf{A}$ such that $\mathbf{A}\mathbf{e}_i = \lambda_i \mathbf{x}_i$, $i = 1, \ldots, n+2$, where the λ_i are nonzero scalars; any two matrices with this property differ at most by a scalar factor.

Proof The matrix $\mathbf{A}$ satisfies the $n+1$ conditions

$$\mathbf{A}\mathbf{e}_i = \lambda_i \mathbf{x}_i \qquad i = 1, \ldots, n+1$$

if and only if it can be written $\mathbf{A} = [\lambda_1 \mathbf{x}_1, \ldots, \lambda_{n+1} \mathbf{x}_{n+1}]$. We must show that we can choose the values of $\lambda_1, \ldots, \lambda_{n+1}$, and λ_{n+2} in such a way that the equation

$$\mathbf{A}\mathbf{e}_{n+2} = \lambda_{n+2}\mathbf{x}_{n+2}$$

is also satisfied. But this is equivalent to

$$[\mathbf{x}_1, \ldots, \mathbf{x}_{n+1}] \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_{n+1} \end{bmatrix} = \lambda_{n+2}\mathbf{x}_{n+2}$$

By the hypothesis concerning the linear independence of the vectors $\mathbf{x}_i$, the matrix on the left-hand side of the previous equation is of rank $n + 1$. Thus the ratios of the λ_i are uniquely determined and, furthermore, none of the λ_i is zero. The matrix $\mathbf{A}$ is thus uniquely determined up to a scalar factor and is clearly nonsingular. ■

This proposition will help us characterize the set of collineations.

2.2.3 Change of projective basis

Let us consider two sets of $n + 2$ points represented by the coordinate vectors $\mathbf{x}_1, \ldots, \mathbf{x}_{n+2}$ and $\mathbf{y}_1, \ldots, \mathbf{y}_{n+2}$. We will prove that if the points in these two sets are in general position, there exists a unique collineation that maps the first set of points onto the second.

Proposition 2.2
If $\mathbf{x}_1, \ldots, \mathbf{x}_{n+2}$ and $\mathbf{y}_1, \ldots, \mathbf{y}_{n+2}$ are two sets of $n + 2$ coordinate vectors such that in either set no $n + 1$ vectors are linearly dependent, i.e., form two projective bases, then there exists a nonsingular $(n + 1) \times (n + 1)$ matrix $\mathbf{P}$ such that $\mathbf{P}\mathbf{x}_i = \rho_i\mathbf{y}_i, i = 1, \ldots, n + 2$, where the ρ_i are scalars, and the matrix $\mathbf{P}$ is uniquely determined apart from a scalar factor.

Proof By the previous proposition, we can choose a nonsingular matrix $\mathbf{A}$ and a set of nonzero scalars $\lambda_1, \ldots, \lambda_{n+2}$ such that

$$\mathbf{A}\mathbf{e}_i = \lambda_i\mathbf{x}_i \qquad i = 1, \ldots, n + 2$$

Similarly, we can choose $\mathbf{B}$ and $\mu_1, \ldots, \mu_{n+2}$ so that

$$\mathbf{B}\mathbf{e}_i = \mu_i\mathbf{y}_i \qquad i = 1, \ldots, n + 2$$

Then

$$\mathbf{B}\mathbf{A}^{-1}\mathbf{x}_i = \frac{\mu_i}{\lambda_i}\mathbf{y}_i \qquad i = 1, \ldots, n + 2$$

and we can take $\mathbf{P} = \mathbf{BA}^{-1}$ and $\rho_i = \frac{\mu_i}{\lambda_i}$. Furthermore, if $\mathbf{Px}_i = \rho_i \mathbf{y}_i$ and $\mathbf{Qx}_i = \sigma_i \mathbf{y}_i$, then $\mathbf{PAe}_i = \lambda_i \rho_i \mathbf{y}_i$ and $\mathbf{QAe}_i = \mu_i \rho_i \mathbf{y}_i$ and hence, by the previous proposition, $\mathbf{PA} = \tau\mathbf{QA}$, i.e., $\mathbf{P} = \tau\mathbf{Q}$ for some scalar τ. ∎

This proposition shows that a collineation is defined by $n + 2$ pairs of corresponding points. We will use this property many times in the next chapters.

2.2.4 The relationship between $\mathcal{P}^m$ and the unit sphere $\mathcal{S}^m$ of R^{m+1}

Here we state a theorem that will turn out to be extremely useful in chapter 5.

Theorem 2.1
The space $\mathcal{P}^m$ is topologically equivalent to the unit sphere S^m of R^{m+1} in which we have identified antipodal points.

Proof This proof can be found in all books on algebraic topology, for example in the book by Greenberg and Harper [GH81]. We can develop an intuition about what is going on as follows. A point x of S^m is represented by a vector $\mathbf{x} = [x_1, \ldots, x_{n+1}]^T$ such that $\sum_{i=1}^{n+1} x_i^2 = 1$. This also represents a point of $\mathcal{P}^m$. The vector $-\mathbf{x}$ represents the antipodal point of x, which is also on S^m but represents the *same* point of $\mathcal{P}^m$. ∎

In particular, this theorem says that all projective spaces are compact spaces. This comes as a bit of a surprise since we all know, at least vaguely, that projective spaces are about points at infinity and that compact subsets of R^n are bounded. But we should not follow our intuition. Projective spaces are indeed compact, although the attentive reader will have no doubt realized that the kind of compact space we are talking about, a sphere that has been folded upon itself by identifying antipodal points, is by no means easy to picture in the mind's eye.

In chapter 5 we will also study the differential structure of the space $\mathcal{P}^n$ and find it very simple, almost as simple as that of R^{m+1}. The importance of this theorem is due to the fact that the folded unit spheres of R^2 and R^3 very naturally appear in the problems of representing directions; the folded unit sphere of R^4 appears in the problem of using quaternions for representing three-dimensional rotations. Since $\mathcal{P}^n$ is simpler to use than the folded sphere S^n, we will use this theorem to construct simple representations of directions, lines, and rotations.

We now turn to the more detailed study of the cases $n = 1, 2$, and 3, which are the cases encountered in practice in computer vision.

The projective line

The space $\mathcal{P}^1$ is known as the projective line. It is the simplest of all projective spaces, which is the first reason why we start with it. The second reason is that many structures embedded in higher-dimensional projective spaces have the same structure as $\mathcal{P}^1$.

The standard projective basis of the projective line is $\mathbf{e}_1 = [1, 0]^T$, $\mathbf{e}_2 = [0, 1]^T$, and $\mathbf{e}_3 = [1, 1]^T$. A point on the line can be written as

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 \tag{2.1}$$

with x_1 and x_2 not both equal to 0. Let us consider a subset of $\mathcal{P}^1$ of the points such that $x_2 \neq 0$. This is the same as excluding the point represented by $\mathbf{e}_1$. Now since the homogeneous coordinates are defined up to a scalar, these points are described by a parameter α, $-\infty \leq \alpha \leq +\infty$ so that

$$\mathbf{x} = \alpha\mathbf{e}_1 + \mathbf{e}_2$$

where $\alpha = \frac{x_1}{x_2}$. The parameter α is often called the *projective parameter* of the point. Note that the point represented by $\mathbf{e}_2$ has a projective parameter equal to 0.

2.3.1 The point at infinity

The point represented by $\mathbf{e}_1$ is called the *point at infinity* of the line $\mathcal{P}^1$. It is defined by the linear equation $x_2 = 0$. The reason for this terminology is that, if we think of the projective line as containing the usual affine line under the correspondence $\alpha \to \alpha\mathbf{e}_1 + \mathbf{e}_2$, then the projective parameter α of the point gives us a one-to-one correspondence between the projective and affine lines for all values of α that are different from ∞. The values $\alpha = \pm\infty$ correspond to the point $\mathbf{e}_1$, which is outside the affine line but is the limit of points of the affine line with large values of α. This is an extremely useful interpretation of the relationship between the affine and projective lines and, as we show later, can be generalized to higher dimensions.

2.3.2 Collineations of $\mathcal{P}^1$

A collineation of $\mathcal{P}^1$ is defined by a 2×2 matrix $\mathbf{P} = \begin{bmatrix} r & s \\ t & u \end{bmatrix}$ of rank 2. This matrix is defined up to a scale factor. If a point has projective parameter α, the transformed point has projective parameter β given by

$$\beta = \frac{r\alpha + s}{t\alpha + u}$$

Note that the condition that the rank of $\mathbf{P}$ equals 2 is equivalent to $ru - st \neq 0$. According to proposition 2.1, a collineation of $\mathcal{P}^1$ is defined by three pairs of corresponding points.

2.3.3 The cross-ratio of four points

We now define the important concept of the cross-ratio, which is a quantity that remains invariant under the group of collineations. Let a, b, c, and d be four points of $\mathcal{P}^1$ with their respective projective parameters α_a, α_b, α_c, and α_d. Then the cross-ratio $\{a, b; c, d\}$ is defined to be

$$\{a, b; c, d\} = \frac{\alpha_a - \alpha_c}{\alpha_a - \alpha_d} : \frac{\alpha_b - \alpha_c}{\alpha_b - \alpha_d} \tag{2.2}$$

The significance of the cross-ratio is that it is invariant under collineations of $\mathcal{P}^1$. In particular, $\{a, b; c, d\}$ is independent of the choice of coordinates in $\mathcal{P}^1$ (see problem 1).

2.4 The projective plane

The space $\mathcal{P}^2$ is known as the *projective plane*. The main reason for its importance is the fact that it is useful to model the image plane as a projective plane (see chapters 3, 6, and 7).

2.4.1 Points and lines

A point in $\mathcal{P}^2$ is defined by three numbers, not all zero (x_1, x_2, and x_3). They form a coordinate vector $\mathbf{x}$ defined up to a scale factor. In $\mathcal{P}^2$ there are objects other than points, such as lines. A line is also defined by a triplet of numbers (u_1, u_2, and u_3), not all zero. They form a coordinate vector $\mathbf{u}$ defined up to a scale factor. The equation of the line is

$$\sum_{i=1}^{3} u_i x_i = 0 \tag{2.3}$$

in the standard projective basis ($\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, and $\mathbf{e}_4$) of $\mathcal{P}^2$.

Formally, there is no difference between points and lines in $\mathcal{P}^2$. This is known as the *principle of duality*.[1] A point represented by $\mathbf{x}$ can be thought of as the set of lines through it. These lines are represented by the coordinate vector $\mathbf{u}$, satisfying $\mathbf{u}^T\mathbf{x} = 0$. This is sometimes referred to as the *line equation* of the point. Inversely, a line represented by $\mathbf{u}$ can be thought of as the set of points represented by $\mathbf{x}$ and satisfying the same equation, called the *point equation* of the line. We can now prove the following result, which will be useful later:

Proposition 2.3
A line going through two points m_1 and m_2, represented by $\mathbf{m}_1$ and $\mathbf{m}_2$, is represented by the cross-product $\mathbf{m}_1 \wedge \mathbf{m}_2$.

Proof A coordinate vector of a point on the line is given by

$$\mathbf{x} = \alpha \mathbf{m}_1 + \beta \mathbf{m}_2$$

for arbitrary values of the scalars α and β. This is equivalent to writing that the determinant $(\mathbf{x}, \mathbf{m}_1, \mathbf{m}_2) = 0$. But this determinant can also be written as

$$\mathbf{x}^T(\mathbf{m}_1 \wedge \mathbf{m}_2) = 0$$

from which the proposition follows. ■

---

1. Note that this idea is at the origin of the so called Hough transform (see chapter 11), which has been patented [Hou62].

2.4.2 The line at infinity

Among all possible lines, the one whose equation is $x_3 = 0$ is called the *line at infinity* of $\mathcal{P}^2$, denoted by l_∞. The reason for this terminology is that we think of the projective plane as containing the usual affine plane under the correspondence $[X_1, X_2]^T \to [X_1, X_2, 1]^T$ or $X_1\mathbf{e}_1 + X_2\mathbf{e}_2 + \mathbf{e}_3$. This is a one-to-one correspondence between the affine plane and the projective plane minus the line of equation $x_3 = 0$. For each projective point of coordinates (x_1, x_2, x_3) that is not on that line, we have

$$X_1 = \frac{x_1}{x_3} \quad X_2 = \frac{x_2}{x_3} \tag{2.4}$$

If $X_1 \to \infty$ while X_2 does not, we obtain $\mathbf{e}_1$, which is on l_∞. Similarly, when $X_2 \to \infty$ while X_1 does not, we obtain $\mathbf{e}_2$.

Each line in the projective plane of the form of equation (2.3) intersects l_∞ at the point $(-u_2, u_1, 0)$, which is that line's point at infinity. Note that the vector $[-u_2, u_1]^T$ gives the direction of the affine line of equation $u_1X_1 + u_2X_2 + u_3 = 0$. This gives us a neat interpretation of the line at infinity: Each point on that line, with coordinates $(x_1, x_2, 0)$, can be thought of as a direction in the underlying affine plane, the direction parallel to the vector $[x_1, x_2]^T$. Indeed, it does not matter if x_1 and x_2 are defined only up to a scale factor since the direction does not change. We will use this observation in chapter 5 when we discuss the problem of representing two-dimensional directions.

Another useful property is the following:

Proposition 2.4
The representation of the point of intersection of two distinct projective lines is the cross-product of their representations.

Proof To see this, simply apply the principle of duality. We have seen that the representation of the line going through two points is the cross-product of the representation of those points, and this implies that the representation of the point of intersection of two lines is the cross-product of their representations. For an alternate proof see problem 4. ∎

Note that this implies that, in projective geometry, two distinct lines always intersect.

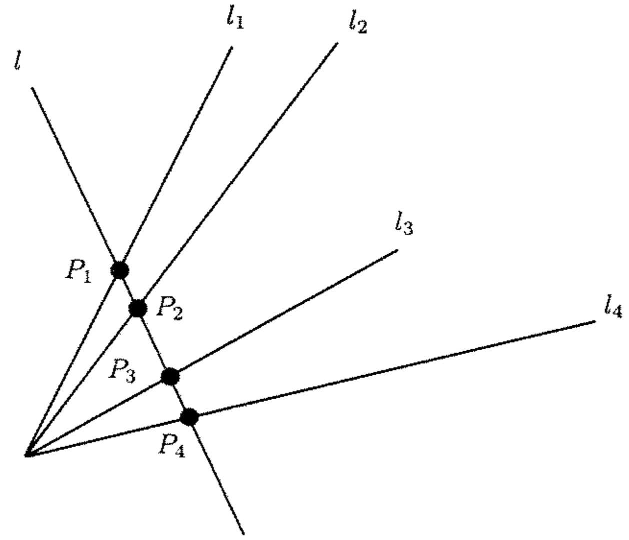

Figure 2.1 Cross-ratio of four lines: $\{l_1, l_2; l_3, l_4\} = \{P_1, P_2; P_3, P_4\}$.

2.4.3 The cross-ratio of four lines intersecting at a point

Let us now generalize the notion of cross-ratio, which was introduced in section 2.3.3 for four points of $\mathcal{P}^1$, to four lines of $\mathcal{P}^2$ intersecting at a point. Given four lines l_1, l_2, l_3, l_4 of $\mathcal{P}^2$ that intersect at a point, their cross-ratio $\{l_1, l_2; l_3, l_4\}$ is defined as the cross-ratio $\{P_1, P_2; P_3, P_4\}$ of their four points of intersection with any line l (see figure 2.1). This value is of course independent of the choice of l (see problem 5).

2.4.4 Pencils of lines

There is a structure of the projective plane that has numerous applications, especially in stereo and motion. The name of this structure is a *pencil of lines*. The set of lines in $\mathcal{P}^2$ passing through a fixed point is a one-dimensional projective space known as a pencil of lines. Let us consider two lines l_1 and l_2 of the pencil represented by their coordinate vectors $\mathbf{u}_1$ and $\mathbf{u}_2$. Any line l of the pencil goes through the point of intersection of l_1 and l_2 represented by $\mathbf{u}_1 \wedge \mathbf{u}_2$ (proposition 2.4). Thus, its coordinate vector $\mathbf{u}$ satisfies $\mathbf{u}^T(\mathbf{u}_1 \wedge \mathbf{u}_2) = 0$, or equivalently

$$\mathbf{u} = \alpha \mathbf{u}_1 + \beta \mathbf{u}_2$$

Modeling and Calibrating Cameras

A guide to this chapter

This chapter contains three sections. The first section, which should be read thoroughly, establishes the fundamental properties of the pinhole camera considered as a (projective) geometric engine. The second section explores the effect of changing coordinate systems in the world or in the retina. In particular, it gives a detailed interpretation of the intrinsic parameters of a camera in relation to the somewhat mysterious entity that we called the absolute conic in chapter 2. The practical reader may skip this section almost entirely and look only at equations (3.20) and (3.21).

The third section studies the practical problem of calibrating a real camera with a theoretical as well as a concrete eye. It contains a characterization of the configuration of degenerate reference points that do not yield a unique solution to the calibration problem. Since these configurations, which are described in proposition 3.1, are extremely unlikely to occur, the proposition can be skipped the first time; however, proposition 3.2 is important, and an understanding of it is necessary in order to avoid gross errors.

Modeling cameras

Since one of the goals of this book is to develop methods for performing metric measurements from images, and since very often these images

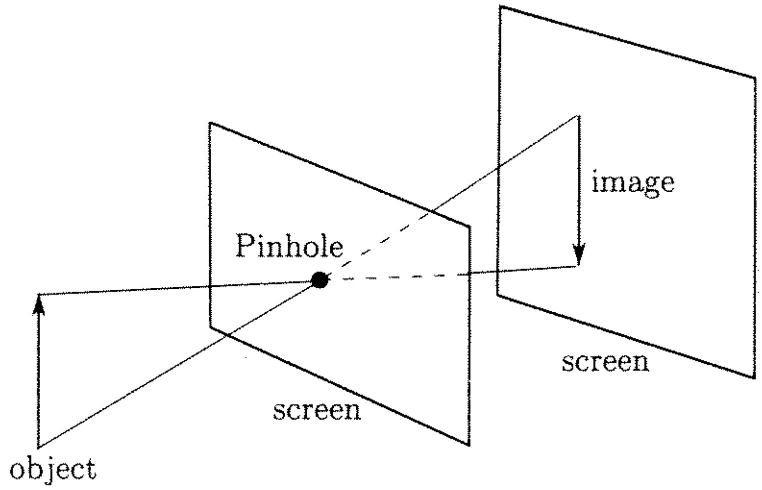

Figure 3.1 Image formation in a pinhole camera.

have been acquired using television or photographic cameras, we want to spend some time defining accurate quantitative models of these devices. There is a deep relationship between these models and projective geometry, and many of the ideas developed in the previous chapter will be put to use. The reader should be familiar with the contents of sections 2.4.1, 2.4.2, and 2.5.1.

In this chapter and in the remainder of the book, we will be using the $\tilde{}$ notation to indicate projective quantities when there is a possibility of confusing them with affine or metric quantities. For example, $\tilde{\mathbf{x}}$ denotes a projective coordinate vector, which is defined up to a multiplicative nonzero scalar, and $\mathbf{x}$ denotes a vector of R^n.

3.2.1 A simple camera model

We will look at a camera model first from the geometric standpoint, then from the physical standpoint.

3.2.1.1 A geometric model

Let us consider the system depicted in figure 3.1. It consists of two screens. A small hole has been punched in the first screen, and through this hole some of the rays of light emitted or reflected by the object pass,

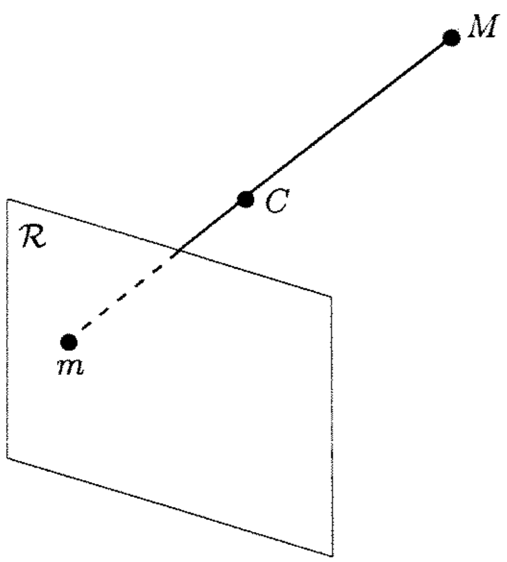

Figure 3.2 The pinhole camera model.

forming an inverted image of that object on the second screen. A beautiful picture taken with such a device can be found in the book by Jenkins and White [JW76].

We can directly build a geometric model of the pinhole camera as indicated in figure 3.2. It consists of a plane $\mathcal{R}$ called the *retinal plane* in which the image is formed through an operation called a *perspective projection*: a point C, the *optical center*, located at a distance f, the *focal length* of the optical system, is used to form the image m in the retinal plane of the 3-D point M as the intersection of the line $\langle C, M \rangle$ with the plane $\mathcal{R}$.

The *optical axis* is the line going through the optical center C and perpendicular to $\mathcal{R}$, which it pierces at a point c. Another plane of interest (see figure 3.3) is the plane $\mathcal{F}$ going through C and parallel to $\mathcal{R}$. It is called the *focal plane*. Points M situated in the focal plane do not have an image in the retinal plane since the line $\langle C, M \rangle$ is parallel to this plane and thus does not intersect it. To speak projectively, it intersects it at infinity. We return to this phenomenon in a later section. It is remarkable that such a simple system can accurately model the geometry and optics of most of the modern Vidicon, CCD, and CID cameras.

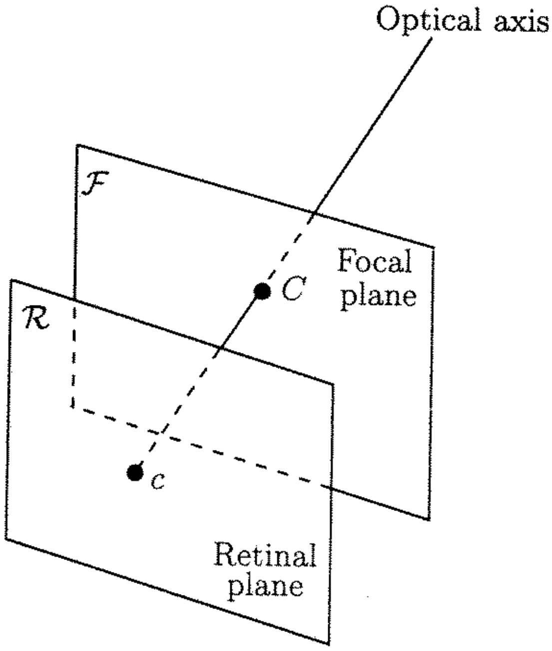

Figure 3.3 The optical axis, focal plane, and retinal plane.

3.2.1.2 A physical model

Next we will relate the amount of light that is reflected and emitted at a point on an object to the brightness of the image of that point in the retinal plane. The amount of light falling on a surface is called the *irradiance*, and it is measured in $\mathrm{W} \times \mathrm{m}^{-2}$, watts per square meter. The amount of light radiated from a surface is called the *radiance*, and it is measured in $\mathrm{W} \times \mathrm{m}^{-2} \times \mathrm{sr}^{-1}$, watts per square meter per steradian.[1] A simple computation that can be found, for example in the book by Horn [Hor86], shows that the relationship between the image irradiance E and the scene radiance L is a very simple linear relationship:

$$E = L\frac{\pi}{4}(\frac{d}{f})^2 \cos^4 \alpha$$

The parameters involved in this equation are defined in figure 3.4.

---

1. The steradian is the unit used to measure solid angles.

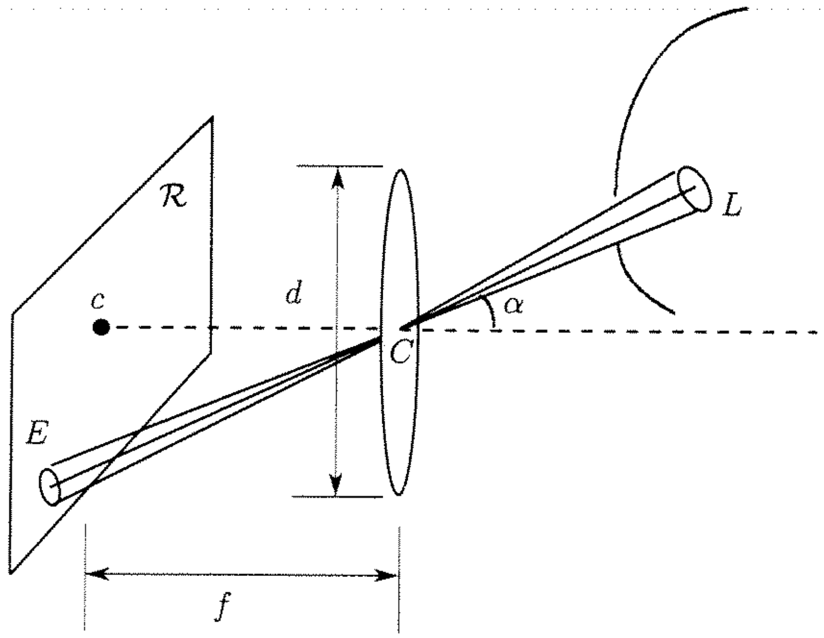

Figure 3.4 The relationship between image irradiance and scene radiance, which is linear.

3.2.2 The perspective projection matrix

We will now study our camera model in further detail. We can choose the coordinate system (C, x, y, z) for the three-dimensional space and (c, u, v) for the retinal plane as indicated in figure 3.5. The coordinate system (C, x, y, z) is called the *standard coordinate system* of the camera. From this figure it should be clear that the relationship between image coordinates and 3-D space coordinates can be written as

$$-\frac{f}{z} = \frac{u}{x} = \frac{v}{y} \tag{3.1}$$

which can be rewritten linearly as

$$\begin{bmatrix} U \\ V \\ S \end{bmatrix} = \begin{bmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{3.2}$$

where

$$u = U/S \quad v = V/S \qquad \text{if } S \neq 0 \tag{3.3}$$

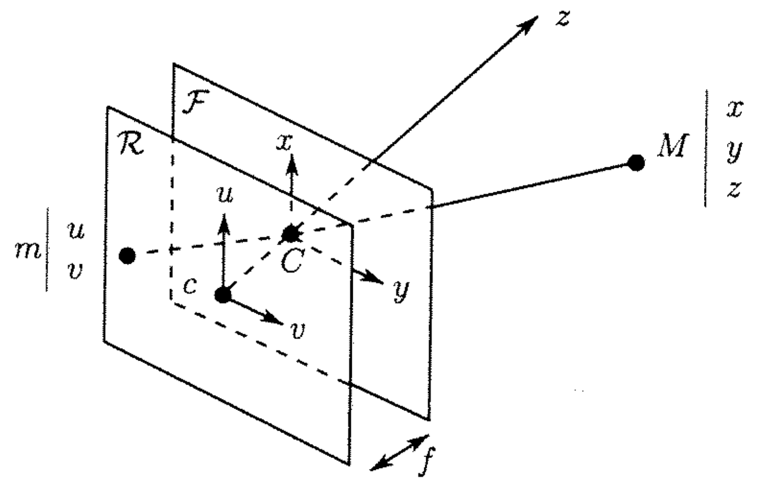

Figure 3.5 The focal plane (x, y) is parallel to the retinal plane (u, v) and at a distance f from it; f is the focal length.

Equations (3.3) should remind the reader of the equations (2.4). These equations allow us to interpret U, V, and S as the projective coordinates of a point in the retina. If $S = 0$, i.e., if $z = 0$, the 3-D point is in the focal plane of the camera. Thus the coordinates u and v are not defined, and the corresponding point is at infinity. We can convince ourselves of this by taking the point M out of the focal plane but placing it arbitrarily close to this plane (see figure 3.6). The points such that $S = 0$ are called *points at infinity* of the retinal plane. $S = 0$ is the equation of the line at infinity of the retinal plane, and this line is the "image" of the focal plane.

Note that equation (3.2) is projective, i.e., it is defined up to a scale factor, and we can rewrite it by using the projective coordinates $(X, Y, Z$, and $T)$ of M:

$$\begin{bmatrix} U \\ V \\ S \end{bmatrix} = \begin{bmatrix} -f & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ T \end{bmatrix} \tag{3.4}$$

The above formula expresses the fact that the relationship between image and space coordinates is linear in projective coordinates and can be written in matrix form as

$$\tilde{\mathbf{m}} = \tilde{\mathbf{P}}\tilde{\mathbf{M}} \tag{3.5}$$

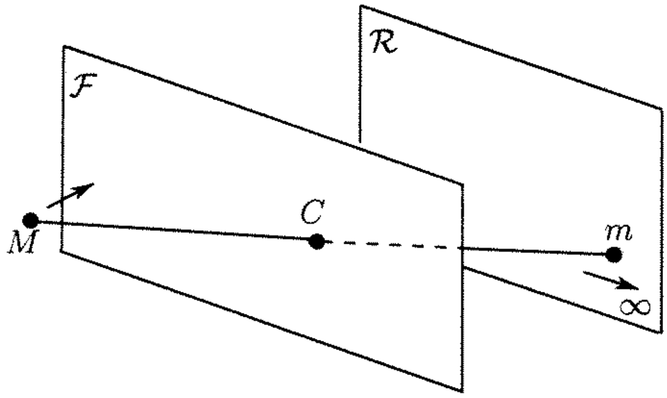

Figure 3.6 When M gets closer to the focal plane, its perspective projection m in the retinal plane goes to ∞.

where $\tilde{\mathbf{m}} = [U, V, S]^T$ and $\tilde{\mathbf{M}} = [X, Y, Z, T]^T$. A camera can be considered as a system that performs a linear projective transformation from the projective space $\mathcal{P}^3$ into the projective plane $\mathcal{P}^2$. This is one of the many examples in which the use of projective geometry makes things simpler: Instead of dealing with the nonlinear equations (3.1), we can use the linear relation (3.5) and the power of linear algebra.

It is easy to convince ourselves that relation (3.5) still holds true with different matrices $\tilde{\mathbf{P}}$ and for any choice of the 3-D and retinal plane coordinate systems. We sometimes refer to the 3-D coordinate system as the *world coordinate system.*

We can give a geometric interpretation of the row vectors of matrix $\tilde{\mathbf{P}}$. We write this matrix in two different ways:

$$\tilde{\mathbf{P}} = \begin{bmatrix} \tilde{\mathbf{Q}}_1^T \\ \tilde{\mathbf{Q}}_2^T \\ \tilde{\mathbf{Q}}_3^T \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1^T & q_{14} \\ \mathbf{q}_2^T & q_{24} \\ \mathbf{q}_3^T & q_{34} \end{bmatrix} \tag{3.6}$$

where $\tilde{\mathbf{Q}}_i, i = 1, 2, 3$ are 4×1 vectors and $\tilde{\mathbf{Q}}_i^T = [\mathbf{q}_i^T, q_{i4}]$. Each such vector represents a projective plane with point equation $\tilde{\mathbf{Q}}_i^T\tilde{\mathbf{M}} = 0$. According to equation (3.4) the plane of equation $\tilde{\mathbf{Q}}_3^T\tilde{\mathbf{M}} = 0$ corresponds to points in the retinal plane such that $S = 0$, i.e., points at ∞. Therefore this is the focal plane. The plane of equation $\tilde{\mathbf{Q}}_1^T\tilde{\mathbf{M}} = 0$ and equation $\tilde{\mathbf{Q}}_2^T\tilde{\mathbf{M}} = 0$ corresponds to points in the retinal plane such that $U = 0$ and $V = 0$, respectively. The intersection of these two planes is the line going through the optical center C of the camera and the origin o in the retinal

Representing Geometric Primitives and Their Uncertainty

In this chapter we will study in detail how to represent some fundamental geometric primitives. A *representation* is a mapping from the set of geometric primitives under study to a set of numerical parameters, a subset of R^n. The variable n is the dimension of the representation. Given the problem of choosing among several representations of a set of geometric primitives, it is important to ask the following questions [MK78, BR78]:

1. **Is the representation unique?** Does every representable geometric primitive have a unique representation? This is equivalent to saying that the previous mapping is one-to-one.
2. **Is the representation complete?** In other words, does every geometric primitive admit a representation?
3. **Is the representation minimal?** This means that the number of parameters used in the representation is minimal. This number is characteristic of the set of geometric primitives and is called its *dimension*. It is important to note that that there exist nonminimal representations whose dimensions are larger than the dimension of the set of geometric primitives.[1]
4. **Is the representation smooth?** This means that, if a geometric primitive varies smoothly, then its representation also varies smoothly.

---

1. There may also exist representations whose dimensions are smaller than the dimension of the set of geometric primitives. We exclude those from our consideration because they do not allow for the reconstruction of the primitives.

ometric primitives such as points, lines, planes, orientations, directions, and displacements, which are routinely manipulated in computer vision problems. The second need is to give us tools for chapters 6, 8, and 11. The reader who is interested only in the tools can read the definition of a manifold in section 5.2, section 5.3.3 on the representation of 2-D lines, section 5.4.3 on the representation of planes, section 5.4.4 on the representation of 3-D lines, section 5.5.2 on quaternions, and section 5.5.4 on exponentials of antisymmetric matrices.

anifolds

Suppose we are given a set X of primitives (see figure 5.1). It is convenient to think of those primitives as points in R^n forming a subset of that space. We assume that X is such that there exists a family U_i of open sets of R^n covering X and such that, for each U_i, there exists a one-to-one mapping φ_i from $U_i \cap X$ into R^d. The maps φ_i should be thought of as the representations of the set of primitives X. They must satisfy the following *coherence condition:*

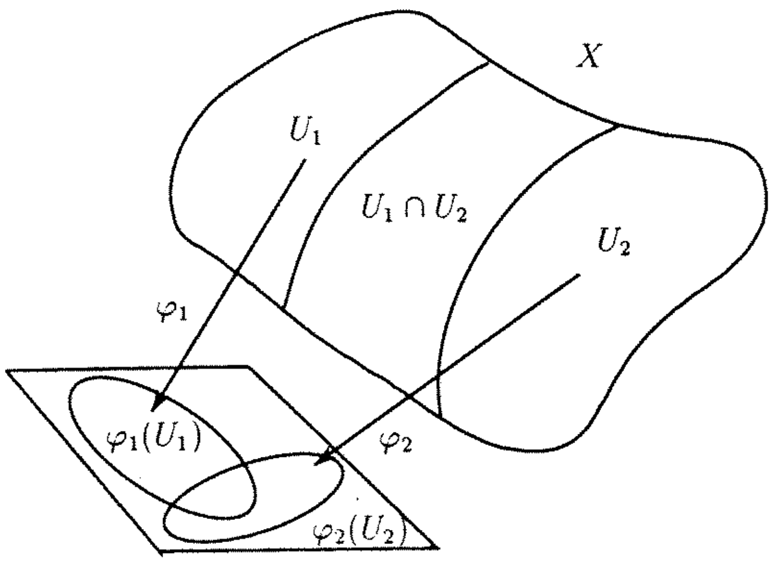

Figure 5.1 A manifold of dimension 2.