2

Projective Geometry and Transformations of 2D

This chapter introduces the main geometric ideas and notation that are required to understand the material covered in this book. Some of these ideas are relatively familiar, such as vanishing point formation or representing conics, whilst others are more esoteric, such as using circular points to remove perspective distortion from an image. These ideas can be understood more easily in the planar (2D) case because they are more easily visualized here. The geometry of 3-space, which is the subject of the later parts of this book, is only a simple generalization of this planar case.

In particular, the chapter covers the geometry of projective transformations of the plane. These transformations model the geometric distortion which arises when a plane is imaged by a perspective camera. Under perspective imaging certain geometric properties are preserved, such as collinearity (a straight line is imaged as a straight line), whilst others are not, for example parallel lines are not imaged as parallel lines in general. Projective geometry models this imaging and also provides a mathematical representation appropriate for computations.

We begin by describing the representation of points, lines and conics in homogeneous notation, and how these entities map under projective transformations. The line at infinity and the circular points are introduced, and it is shown that these capture the affine and metric properties of the plane. Algorithms for rectifying planes are then given which enable affine and metric properties to be computed from images. We end with a description of fixed points under projective transformations.

2.1 Planar geometry

The basic concepts of planar geometry are familiar to anyone who has studied mathematics even at an elementary level. In fact, they are so much a part of our everyday experience that we take them for granted. At an elementary level, geometry is the study of points and lines and their relationships.

To the purist, the study of geometry ought properly to be carried out from a "geometric" or coordinate-free viewpoint. In this approach, theorems are stated and proved in terms of geometric primitives only, without the use of algebra. The classical approach of Euclid is an example of this method. Since Descartes, however, it has been seen that geometry may be algebraicized, and indeed the theory of geometry may be developed from an algebraic viewpoint. Our approach in this book will be a hybrid approach, sometimes using geometric, and sometimes algebraic methods. In the algebraic approach, geometric entities are described in terms of coordinates and algebraic entities. Thus, for instance a point is identified with a vector in terms of some coordinate basis. A line is also identified with a vector, and a conic section (more briefly, a conic) is represented by a symmetric matrix. In fact, we often carry this identification so far as to consider that the vector actually *is* a point, or the symmetric matrix *is* a conic, at least for convenience of language. A significant advantage of the algebraic approach to geometry is that results derived in this way may more easily be used to derive algorithms and practical computational methods. Computation and algorithms are a major concern in this book, which justifies the use of the algebraic method.

2.2 The 2D projective plane

As we all know, a point in the plane may be represented by the pair of coordinates (x, y) in $\mathbb{R}^2$. Thus, it is common to identify the plane with $\mathbb{R}^2$. Considering $\mathbb{R}^2$ as a vector space, the coordinate pair (x, y) is a vector – a point is identified as a vector. In this section we introduce the *homogeneous* notation for points and lines on a plane.

Row and column vectors. Later on, we will want to consider linear mappings between vector spaces, and represent such mappings as matrices. In the usual manner, the product of a matrix and a vector is another vector, the image under the mapping. This brings up the distinction between "column" and "row" vectors, since a matrix may be multiplied on the right by a column and on the left by a row vector. Geometric entities will by default be represented by column vectors. A bold-face symbol such as $\mathbf{x}$ always represents a column vector, and its transpose is the row vector $\mathbf{x}^\mathsf{T}$. In accordance with this convention, a point in the plane will be represented by the column vector $(x, y)^\mathsf{T}$, rather than its transpose, the row vector (x, y). We write $\mathbf{x} = (x, y)^\mathsf{T}$, both sides of this equation representing column vectors.

2.2.1 Points and lines

Homogeneous representation of lines. A line in the plane is represented by an equation such as $ax + by + c = 0$, different choices of a, b and c giving rise to different lines. Thus, a line may naturally be represented by the vector $(a, b, c)^\mathsf{T}$. The correspondence between lines and vectors $(a, b, c)^\mathsf{T}$ is not one-to-one, since the lines $ax + by + c = 0$ and $(ka)x + (kb)y + (kc) = 0$ are the same, for any non-zero constant k. Thus, the vectors $(a, b, c)^\mathsf{T}$ and $k(a, b, c)^\mathsf{T}$ represent the same line, for any non-zero k. In fact, two such vectors related by an overall scaling are considered as being equivalent. An equivalence class of vectors under this equivalence relationship is known as a *homogeneous* vector. Any particular vector $(a, b, c)^\mathsf{T}$ is a representative of the equivalence class. The set of equivalence classes of vectors in $\mathbb{R}^3 - (0, 0, 0)^\mathsf{T}$ forms the *projective space* $\mathbb{P}^2$. The notation $-(0, 0, 0)^\mathsf{T}$ indicates that the vector $(0, 0, 0)^\mathsf{T}$, which does not correspond to any line, is excluded.

Homogeneous representation of points. A point $\mathbf{x} = (x, y)^\top$ lies on the line $\mathbf{l} = (a, b, c)^\top$ if and only if $ax + by + c = 0$. This may be written in terms of an inner product of vectors representing the point as $(x, y, 1)(a, b, c)^\top = (x, y, 1)\mathbf{l} = 0$; that is the point $(x, y)^\top$ in $\mathbb{R}^2$ is represented as a 3-vector by adding a final coordinate of 1. Note that for any non-zero constant k and line $\mathbf{l}$ the equation $(kx, ky, k)\mathbf{l} = 0$ if and only if $(x, y, 1)\mathbf{l} = 0$. It is natural, therefore, to consider the set of vectors $(kx, ky, k)^\top$ for varying values of k to be a representation of the point $(x, y)^\top$ in $\mathbb{R}^2$. Thus, just as with lines, points are represented by homogeneous vectors. An arbitrary homogeneous vector representative of a point is of the form $\mathbf{x} = (x_1, x_2, x_3)^\top$, representing the point $(x_1/x_3, x_2/x_3)^\top$ in $\mathbb{R}^2$. Points, then, as homogeneous vectors are also elements of $\mathbb{P}^2$.

One has a simple equation to determine when a point lies on a line, namely

Result 2.1. *The point* $\mathbf{x}$ *lies on the line* $\mathbf{l}$ *if and only if* $\mathbf{x}^\top\mathbf{l} = 0$.

Note that the expression $\mathbf{x}^\top\mathbf{l}$ is just the inner or scalar product of the two vectors $\mathbf{l}$ and $\mathbf{x}$. The scalar product $\mathbf{x}^\top\mathbf{l} = \mathbf{l}^\top\mathbf{x} = \mathbf{x}.\mathbf{l}$. In general, the transpose notation $\mathbf{l}^\top\mathbf{x}$ will be preferred, but occasionally, we will use a . to denote the inner product. We distinguish between the *homogeneous coordinates* $\mathbf{x} = (x_1, x_2, x_3)^\top$ of a point, which is a 3-vector, and the *inhomogeneous* coordinates $(x, y)^\top$, which is a 2-vector.

Degrees of freedom (dof). It is clear that in order to specify a point two values must be provided, namely its x- and y-coordinates. In a similar manner a line is specified by two parameters (the two independent ratios $\{a : b : c\}$) and so has two degrees of freedom. For example, in an inhomogeneous representation, these two parameters could be chosen as the gradient and y intercept of the line.

Intersection of lines. Given two lines $\mathbf{l} = (a, b, c)^\top$ and $\mathbf{l}' = (a', b', c')^\top$, we wish to find their intersection. Define the vector $\mathbf{x} = \mathbf{l} \times \mathbf{l}'$, where $\times$ represents the vector or cross product. From the triple scalar product identity $\mathbf{l}.(\mathbf{l} \times \mathbf{l}') = \mathbf{l}'.(\mathbf{l} \times \mathbf{l}') = 0$, we see that $\mathbf{l}^\top\mathbf{x} = \mathbf{l}'^\top\mathbf{x} = 0$. Thus, if $\mathbf{x}$ is thought of as representing a point, then $\mathbf{x}$ lies on both lines $\mathbf{l}$ and $\mathbf{l}'$, and hence is the intersection of the two lines. This shows:

Result 2.2. *The intersection of two lines* $\mathbf{l}$ *and* $\mathbf{l}'$ *is the point* $\mathbf{x} = \mathbf{l} \times \mathbf{l}'$.

Note that the simplicity of this expression for the intersection of the two lines is a direct consequence of the use of homogeneous vector representations of lines and points.

Example 2.3. Consider the simple problem of determining the intersection of the lines $x = 1$ and $y = 1$. The line $x = 1$ is equivalent to $-1x + 1 = 0$, and thus has homogeneous representation $\mathbf{l} = (-1, 0, 1)^\top$. The line $y = 1$ is equivalent to $-1y + 1 = 0$, and thus has homogeneous representation $\mathbf{l}' = (0, -1, 1)^\top$. From result 2.2 the intersection point is

$$\mathbf{x} = \mathbf{l} \times \mathbf{l}' = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{vmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

which is the inhomogeneous point $(1, 1)^\top$ as required. △

Line joining points. An expression for the line passing through two points $\mathbf{x}$ and $\mathbf{x}'$ may be derived by an entirely analogous argument. Defining a line $\mathbf{l}$ by $\mathbf{l} = \mathbf{x} \times \mathbf{x}'$, it may be verified that both points $\mathbf{x}$ and $\mathbf{x}'$ lie on $\mathbf{l}$. Thus

Result 2.4. *The line through two points* $\mathbf{x}$ *and* $\mathbf{x}'$ *is* $\mathbf{l} = \mathbf{x} \times \mathbf{x}'$.

2.2.2 Ideal points and the line at infinity

Intersection of parallel lines. Consider two lines $ax+by+c = 0$ and $ax+by+c' = 0$. These are represented by vectors $\mathbf{l} = (a, b, c)^\mathsf{T}$ and $\mathbf{l}' = (a, b, c')^\mathsf{T}$ for which the first two coordinates are the same. Computing the intersection of these lines gives no difficulty, using result 2.2. The intersection is $\mathbf{l} \times \mathbf{l}' = (c' - c)(b, -a, 0)^\mathsf{T}$, and ignoring the scale factor $(c' - c)$, this is the point $(b, -a, 0)^\mathsf{T}$.

Now if we attempt to find the inhomogeneous representation of this point, we obtain $(b/0, -a/0)^\mathsf{T}$, which makes no sense, except to suggest that the point of intersection has infinitely large coordinates. In general, points with homogeneous coordinates $(x, y, 0)^\mathsf{T}$ do not correspond to any finite point in $\mathbb{R}^2$. This observation agrees with the usual idea that parallel lines meet at infinity.

Example 2.5. Consider the two lines $x = 1$ and $x = 2$. Here the two lines are parallel, and consequently intersect "at infinity". In homogeneous notation the lines are $\mathbf{l} = (-1, 0, 1)^\mathsf{T}$, $\mathbf{l}' = (-1, 0, 2)^\mathsf{T}$, and from result 2.2 their intersection point is

$$\mathbf{x} = \mathbf{l} \times \mathbf{l}' = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -1 & 0 & 1 \\ -1 & 0 & 2 \end{vmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

which is the point at infinity in the direction of the y-axis. △

Ideal points and the line at infinity. Homogeneous vectors $\mathbf{x} = (x_1, x_2, x_3)^\mathsf{T}$ such that $x_3 \neq 0$ correspond to finite points in $\mathbb{R}^2$. One may augment $\mathbb{R}^2$ by adding points with last coordinate $x_3 = 0$. The resulting space is the set of all homogeneous 3-vectors, namely the projective space $\mathbb{P}^2$. The points with last coordinate $x_3 = 0$ are known as *ideal* points, or points at infinity. The set of all ideal points may be written $(x_1, x_2, 0)^\mathsf{T}$, with a particular point specified by the ratio $x_1 : x_2$. Note that this set lies on a single line, the *line at infinity*, denoted by the vector $\mathbf{l}_\infty = (0, 0, 1)^\mathsf{T}$. Indeed, one verifies that $(0, 0, 1)(x_1, x_2, 0)^\mathsf{T} = 0$.

Using result 2.2 one finds that a line $\mathbf{l} = (a, b, c)^\mathsf{T}$ intersects $\mathbf{l}_\infty$ in the ideal point $(b, -a, 0)^\mathsf{T}$ (since $(b, -a, 0)\mathbf{l} = 0$). A line $\mathbf{l}' = (a, b, c')^\mathsf{T}$ *parallel* to $\mathbf{l}$ intersects $\mathbf{l}_\infty$ in the same ideal point $(b, -a, 0)^\mathsf{T}$ irrespective of the value of c'. In inhomogeneous notation $(b, -a)^\mathsf{T}$ is a vector tangent to the line, and orthogonal to the line normal (a, b), and so represents the line's *direction*. As the line's direction varies the ideal point $(b, -a, 0)^\mathsf{T}$ varies over $\mathbf{l}_\infty$. For these reasons the line at infinity can be thought of as the set of directions of lines in the plane.

Note how the introduction of the concept of points at infinity serves to simplify the intersection properties of points and lines. In the projective plane $\mathbb{P}^2$, one may state without qualification that two distinct lines meet in a single point and two distinct

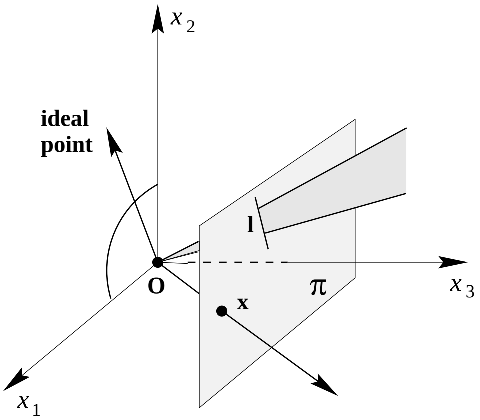


Fig. 2.1. **A model of the projective plane.** *Points and lines of $\mathbb{P}^2$ are represented by rays and planes, respectively, through the origin in $\mathbb{R}^3$. Lines lying in the x_1x_2-plane represent ideal points, and the x_1x_2-plane represents $\mathbf{l}_\infty$.*

points lie on a single line. This is not true in the standard Euclidean geometry of $\mathbb{R}^2$, in which parallel lines form a special case.

The study of the geometry of $\mathbb{P}^2$ is known as projective geometry. In a coordinate-free purely geometric study of projective geometry, one does not make any distinction between points at infinity (ideal points) and ordinary points. It will, however, serve our purposes in this book sometimes to distinguish between ideal points and non-ideal points. Thus, the line at infinity will at times be considered as a special line in projective space.

A model for the projective plane. A fruitful way of thinking of $\mathbb{P}^2$ is as a set of rays in $\mathbb{R}^3$. The set of all vectors $k(x_1, x_2, x_3)^\mathsf{T}$ as k varies forms a ray through the origin. Such a ray may be thought of as representing a single point in $\mathbb{P}^2$. In this model, the lines in $\mathbb{P}^2$ are planes passing through the origin. One verifies that two non-identical rays lie on exactly one plane, and any two planes intersect in one ray. This is the analogue of two distinct points uniquely defining a line, and two lines always intersecting in a point.

Points and lines may be obtained by intersecting this set of rays and planes by the plane $x_3 = 1$. As illustrated in figure 2.1 the rays representing ideal points and the plane representing $\mathbf{l}_\infty$ are parallel to the plane $x_3 = 1$.

Duality. The reader has probably noticed how the role of points and lines may be interchanged in statements concerning the properties of lines and points. In particular, the basic incidence equation $\mathbf{l}^\mathsf{T}\mathbf{x} = 0$ for line and point is symmetric, since $\mathbf{l}^\mathsf{T}\mathbf{x} = 0$ implies $\mathbf{x}^\mathsf{T}\mathbf{l} = 0$, in which the positions of line and point are swapped. Similarly, result 2.2 and result 2.4 giving the intersection of two lines and the line through two points are essentially the same, with the roles of points and lines swapped. One may enunciate a general principle, the *duality principle* as follows:

Result 2.6. Duality principle. *To any theorem of 2-dimensional projective geometry there corresponds a dual theorem, which may be derived by interchanging the roles of points and lines in the original theorem.*

In applying this principle, concepts of incidence must be appropriately translated as well. For instance, the line through two points is dual to the point through (that is the point of intersection of) two lines.

Note that is it not necessary to prove the dual of a given theorem once the original theorem has been proved. The proof of the dual theorem will be the dual of the proof of the original theorem.

2.2.3 Conics and dual conics

A conic is a curve described by a second-degree equation in the plane. In Euclidean geometry conics are of three main types: hyperbola, ellipse, and parabola (apart from so-called degenerate conics, to be defined later). Classically these three types of conic arise as conic sections generated by planes of differing orientation (the degenerate conics arise from planes which contain the cone vertex). However, it will be seen that in 2D projective geometry all non-degenerate conics are equivalent under projective transformations.

The equation of a conic in inhomogeneous coordinates is

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

i.e. a polynomial of degree 2. "Homogenizing" this by the replacements:
$x \mapsto x_1/x_3,\ y \mapsto x_2/x_3$ gives

$$a{x_1}^2 + bx_1x_2 + c{x_2}^2 + dx_1x_3 + ex_2x_3 + f{x_3}^2 = 0 \tag{2.1}$$

or in matrix form

$$\mathbf{x}^{\mathsf{T}}\mathtt{C}\mathbf{x} = 0 \tag{2.2}$$

where the conic coefficient matrix $\mathtt{C}$ is given by

$$\mathtt{C} = \begin{bmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{bmatrix}. \tag{2.3}$$

Note that the conic coefficient matrix is symmetric. As in the case of the homogeneous representation of points and lines, only the ratios of the matrix elements are important, since multiplying $\mathtt{C}$ by a non-zero scalar does not affect the above equations. Thus $\mathtt{C}$ is a homogeneous representation of a conic. The conic has five degrees of freedom which can be thought of as the ratios $\{a : b : c : d : e : f\}$ or equivalently the six elements of a symmetric matrix less one for scale.

Five points define a conic. Suppose we wish to compute the conic which passes through a set of points, $\mathbf{x}_i$. How many points are we free to specify before the conic is determined uniquely? The question can be answered constructively by providing an algorithm to determine the conic. From (2.1) each point $\mathbf{x}_i$ places one constraint on the conic coefficients, since if the conic passes through (x_i, y_i) then

$$a{x_i}^2 + bx_iy_i + c{y_i}^2 + dx_i + ey_i + f = 0.$$

This constraint can be written as

$$\begin{pmatrix} x_i^2 & x_iy_i & y_i^2 & x_i & y_i & 1 \end{pmatrix} \mathbf{c} = 0$$

where $\mathbf{c} = (a, b, c, d, e, f)^\mathsf{T}$ is the conic $\mathtt{C}$ represented as a 6-vector.

Stacking the constraints from five points we obtain

$$\begin{bmatrix} x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2y_2 & y_2^2 & x_2 & y_2 & 1 \\ x_3^2 & x_3y_3 & y_3^2 & x_3 & y_3 & 1 \\ x_4^2 & x_4y_4 & y_4^2 & x_4 & y_4 & 1 \\ x_5^2 & x_5y_5 & y_5^2 & x_5 & y_5 & 1 \end{bmatrix} \mathbf{c} = \mathbf{0} \tag{2.4}$$

and the conic is the null vector of this 5×6 matrix. This shows that a conic is determined uniquely (up to scale) by five points in general position. The method of fitting a geometric entity (or relation) by determining a null space will be used frequently in the computation chapters throughout this book.

Tangent lines to conics. The line $\mathbf{l}$ tangent to a conic at a point $\mathbf{x}$ has a particularly simple form in homogeneous coordinates:

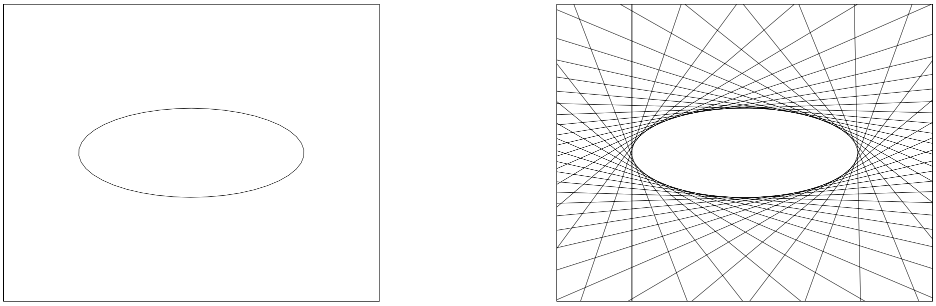
Result 2.7. *The line* $\mathbf{l}$ *tangent to* $\mathtt{C}$ *at a point* $\mathbf{x}$ *on* $\mathtt{C}$ *is given by* $\mathbf{l} = \mathtt{C}\mathbf{x}$.

Proof. The line $\mathbf{l} = \mathtt{C}\mathbf{x}$ passes through $\mathbf{x}$, since $\mathbf{l}^\mathsf{T}\mathbf{x} = \mathbf{x}^\mathsf{T}\mathtt{C}\mathbf{x} = 0$. If $\mathbf{l}$ has one-point contact with the conic, then it is a tangent, and we are done. Otherwise suppose that $\mathbf{l}$ meets the conic in another point $\mathbf{y}$. Then $\mathbf{y}^\mathsf{T}\mathtt{C}\mathbf{y} = 0$ and $\mathbf{x}^\mathsf{T}\mathtt{C}\mathbf{y} = \mathbf{l}^\mathsf{T}\mathbf{y} = 0$. From this it follows that $(\mathbf{x} + \alpha\mathbf{y})^\mathsf{T}\mathtt{C}(\mathbf{x} + \alpha\mathbf{y}) = 0$ for all α, which means that the whole line $\mathbf{l} = \mathtt{C}\mathbf{x}$ joining $\mathbf{x}$ and $\mathbf{y}$ lies on the conic $\mathtt{C}$, which is therefore degenerate (see below). □

Dual conics. The conic $\mathtt{C}$ defined above is more properly termed a *point* conic, as it defines an equation on points. Given the duality result 2.6 of $\mathbb{P}^2$ it is not surprising that there is also a conic which defines an equation on lines. This *dual* (or line) conic is also represented by a 3×3 matrix, which we denote as $\mathtt{C}^*$. A line $\mathbf{l}$ *tangent* to the conic $\mathtt{C}$ satisfies $\mathbf{l}^\mathsf{T}\mathtt{C}^*\mathbf{l} = 0$. The notation $\mathtt{C}^*$ indicates that $\mathtt{C}^*$ is the adjoint matrix of $\mathtt{C}$ (the adjoint is defined in section A4.2(*p*580) of appendix 4(*p*578)). For a non-singular symmetric matrix $\mathtt{C}^* = \mathtt{C}^{-1}$ (up to scale).

The equation for a dual conic is straightforward to derive in the case that $\mathtt{C}$ has full rank: From result 2.7, at a point $\mathbf{x}$ on $\mathtt{C}$ the tangent is $\mathbf{l} = \mathtt{C}\mathbf{x}$. Inverting, we find the point $\mathbf{x}$ at which the line $\mathbf{l}$ is tangent to $\mathtt{C}$ is $\mathbf{x} = \mathtt{C}^{-1}\mathbf{l}$. Since $\mathbf{x}$ satisfies $\mathbf{x}^\mathsf{T}\mathtt{C}\mathbf{x} = 0$ we obtain $(\mathtt{C}^{-1}\mathbf{l})^\mathsf{T}\mathtt{C}(\mathtt{C}^{-1}\mathbf{l}) = \mathbf{l}^\mathsf{T}\mathtt{C}^{-1}\mathbf{l} = 0$, the last step following from $\mathtt{C}^{-\mathsf{T}} = \mathtt{C}^{-1}$ because $\mathtt{C}$ is symmetric.

Dual conics are also known as conic envelopes, and the reason for this is illustrated



a b

Fig. 2.2. *(a) Points* $\mathbf{x}$ *satisfying* $\mathbf{x}^\mathsf{T}\mathtt{C}\mathbf{x} = 0$ *lie on a point conic. (b) Lines* $\mathbf{l}$ *satisfying* $\mathbf{l}^\mathsf{T}\mathtt{C}^*\mathbf{l} = 0$ *are tangent to the point conic* $\mathtt{C}$*. The conic* $\mathtt{C}$ *is the envelope of the lines* $\mathbf{l}$*.*

in figure 2.2. A dual conic has five degrees of freedom. In a similar manner to points defining a point conic, it follows that five lines in general position define a dual conic.

Degenerate conics. If the matrix $\mathtt{C}$ is not of full rank, then the conic is termed degenerate. Degenerate point conics include two lines (rank 2), and a repeated line (rank 1).

Example 2.8. The conic

$$\mathtt{C} = \mathbf{l}\,\mathbf{m}^\mathsf{T} + \mathbf{m}\,\mathbf{l}^\mathsf{T}$$

is composed of two lines $\mathbf{l}$ and $\mathbf{m}$. Points on $\mathbf{l}$ satisfy $\mathbf{l}^\mathsf{T}\mathbf{x} = 0$, and are on the conic since $\mathbf{x}^\mathsf{T}\mathtt{C}\mathbf{x} = (\mathbf{x}^\mathsf{T}\mathbf{l})(\mathbf{m}^\mathsf{T}\mathbf{x}) + (\mathbf{x}^\mathsf{T}\mathbf{m})(\mathbf{l}^\mathsf{T}\mathbf{x}) = 0$. Similarly, points satisfying $\mathbf{m}^\mathsf{T}\mathbf{x} = 0$ also satisfy $\mathbf{x}^\mathsf{T}\mathtt{C}\mathbf{x} = 0$. The matrix $\mathtt{C}$ is symmetric and has rank 2. The null vector is $\mathbf{x} = \mathbf{l} \times \mathbf{m}$ which is the intersection point of $\mathbf{l}$ and $\mathbf{m}$. △

Degenerate *line* conics include two points (rank 2), and a repeated point (rank 1). For example, the line conic $\mathtt{C}^* = \mathbf{x}\mathbf{y}^\mathsf{T} + \mathbf{y}\mathbf{x}^\mathsf{T}$ has rank 2 and consists of lines passing through either of the two points $\mathbf{x}$ and $\mathbf{y}$. Note that for matrices that are not invertible $(\mathtt{C}^*)^* \neq \mathtt{C}$.

2.3 Projective transformations

In the view of geometry set forth by Felix Klein in his famous "Erlangen Program", [Klein-39], geometry is the study of properties invariant under groups of transformations. From this point of view, 2D projective geometry is the study of properties of the projective plane $\mathbb{P}^2$ that are invariant under a group of transformations known as *projectivities*.

A projectivity is an invertible mapping from points in $\mathbb{P}^2$ (that is homogeneous 3-vectors) to points in $\mathbb{P}^2$ that maps lines to lines. More precisely,

Definition 2.9. A *projectivity* is an invertible mapping h from $\mathbb{P}^2$ to itself such that three points $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$ lie on the same line if and only if $h(\mathbf{x}_1)$, $h(\mathbf{x}_2)$ and $h(\mathbf{x}_3)$ do.

Projectivities form a group since the inverse of a projectivity is also a projectivity, and so is the composition of two projectivities. A projectivity is also called a *collineation* (a helpful name), a *projective transformation* or a *homography*: the terms are synonymous.

In definition 2.9, a projectivity is defined in terms of a coordinate-free geometric concept of point line incidence. An equivalent algebraic definition of a projectivity is possible, based on the following result.

Theorem 2.10. *A mapping $h : \mathbb{P}^2 \rightarrow \mathbb{P}^2$ is a projectivity if and only if there exists a non-singular 3×3 matrix $\mathtt{H}$ such that for any point in $\mathbb{P}^2$ represented by a vector $\mathbf{x}$ it is true that $h(\mathbf{x}) = \mathtt{H}\mathbf{x}$.*

To interpret this theorem, any point in $\mathbb{P}^2$ is represented as a homogeneous 3-vector, $\mathbf{x}$, and $\mathtt{H}\mathbf{x}$ is a linear mapping of homogeneous coordinates. The theorem asserts that any projectivity arises as such a linear transformation in homogeneous coordinates, and that conversely any such mapping is a projectivity. The theorem will not be proved in full here. It will only be shown that any invertible linear transformation of homogeneous coordinates is a projectivity.

Proof. Let $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$ lie on a line $\mathbf{l}$. Thus $\mathbf{l}^\mathsf{T}\mathbf{x}_i = 0$ for $i = 1, \ldots, 3$. Let $\mathtt{H}$ be a non-singular 3×3 matrix. One verifies that $\mathbf{l}^\mathsf{T}\mathtt{H}^{-1}\mathtt{H}\mathbf{x}_i = 0$. Thus, the points $\mathtt{H}\mathbf{x}_i$ all lie on the line $\mathtt{H}^{-\mathsf{T}}\mathbf{l}$, and collinearity is preserved by the transformation.

The converse is considerably harder to prove, namely that each projectivity arises in this way. □

As a result of this theorem, one may give an alternative definition of a projective transformation (or collineation) as follows.

Definition 2.11. Projective transformation. A planar projective transformation is a linear transformation on homogeneous 3-vectors represented by a non-singular 3×3 matrix:

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \tag{2.5}$$

or more briefly, $\mathbf{x}' = \mathtt{H}\mathbf{x}$.

Note that the matrix $\mathtt{H}$ occurring in this equation may be changed by multiplication by an arbitrary non-zero scale factor without altering the projective transformation. Consequently we say that $\mathtt{H}$ is a *homogeneous* matrix, since as in the homogeneous representation of a point, only the ratio of the matrix elements is significant. There are eight independent ratios amongst the nine elements of $\mathtt{H}$, and it follows that a projective transformation has eight degrees of freedom.

A projective transformation projects every figure into a projectively equivalent figure, leaving all its projective properties invariant. In the ray model of figure 2.1 a projective transformation is simply a linear transformation of $\mathbb{R}^3$.

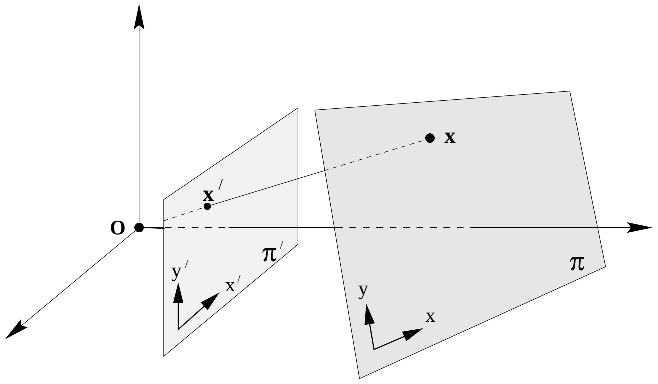


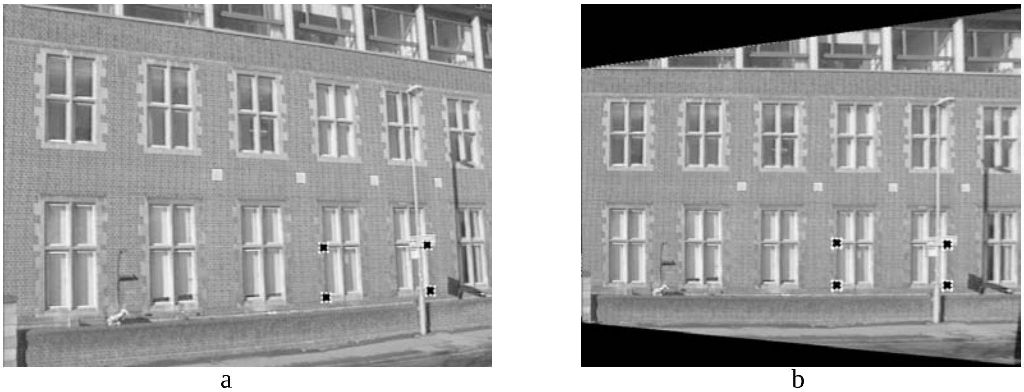
Fig. 2.3. **Central projection maps points on one plane to points on another plane.** *The projection also maps lines to lines as may be seen by considering a plane through the projection centre which intersects with the two planes π and π'. Since lines are mapped to lines, central projection is a projectivity and may be represented by a linear mapping of homogeneous coordinates* $\mathbf{x}' = \mathtt{H}\mathbf{x}$.

Mappings between planes. As an example of how theorem 2.10 may be applied, consider figure 2.3. Projection along rays through a common point (the centre of projection) defines a mapping from one plane to another. It is evident that this point-to-point mapping preserves lines in that a line in one plane is mapped to a line in the other. If a coordinate system is defined in each plane and points are represented in homogeneous coordinates, then the *central projection* mapping may be expressed by $\mathbf{x}' = \mathtt{H}\mathbf{x}$ where $\mathtt{H}$ is a non-singular 3×3 matrix. Actually, if the two coordinate systems defined in the two planes are both Euclidean (rectilinear) coordinate systems then the mapping defined by central projection is more restricted than an arbitrary projective transformation. It is called a *perspectivity* rather than a full projectivity, and may be represented by a transformation with six degrees of freedom. We return to perspectivities in section A7.4(*p*632).

Example 2.12. Removing the projective distortion from a perspective image of a plane.

Shape is distorted under perspective imaging. For instance, in figure 2.4a the windows are not rectangular in the image, although the originals are. In general parallel lines on a scene plane are not parallel in the image but instead converge to a finite point. We have seen that a central projection image of a plane (or section of a plane) is related to the original plane via a projective transformation, and so the image is a projective distortion of the original. It is possible to "undo" this projective transformation by computing the inverse transformation and applying it to the image. The result will be a new synthesized image in which the objects in the plane are shown with their correct geometric shape. This will be illustrated here for the front of the building of figure 2.4a. Note that since the ground and the front are not in the same plane, the projective transformation that must be applied to rectify the front is not the same as the one used for the ground.

Computation of a projective transformation from point-to-point correspondences will be considered in great detail in chapter 4. For now, a method for computing the trans-



a b

Fig. 2.4. **Removing perspective distortion.** *(a) The original image with perspective distortion – the lines of the windows clearly converge at a finite point. (b) Synthesized frontal orthogonal view of the front wall. The image (a) of the wall is related via a projective transformation to the true geometry of the wall. The inverse transformation is computed by mapping the four imaged window corners to corners of an appropriately sized rectangle. The four point correspondences determine the transformation. The transformation is then applied to the whole image. Note that sections of the image of the ground are subject to a further projective distortion. This can also be removed by a projective transformation.*

formation is briefly indicated. One begins by selecting a section of the image corresponding to a planar section of the world. Local 2D image and world coordinates are selected as shown in figure 2.3. Let the inhomogeneous coordinates of a pair of matching points $\mathbf{x}$ and $\mathbf{x}'$ in the world and image plane be (x, y) and (x', y') respectively. We use inhomogeneous coordinates here instead of the homogeneous coordinates of the points, because it is these inhomogeneous coordinates that are measured directly from the image and from the world plane. The projective transformation of (2.5) can be written in inhomogeneous form as

$$x' = \frac{x'_1}{x'_3} = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}, \qquad y' = \frac{x'_2}{x'_3} = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}}.$$

Each point correspondence generates two equations for the elements of $\mathtt{H}$, which after multiplying out are

$$\begin{aligned} x'\left(h_{31}x + h_{32}y + h_{33}\right) &= h_{11}x + h_{12}y + h_{13} \\ y'\left(h_{31}x + h_{32}y + h_{33}\right) &= h_{21}x + h_{22}y + h_{23}. \end{aligned}$$

These equations are *linear* in the elements of $\mathtt{H}$. Four point correspondences lead to eight such linear equations in the entries of $\mathtt{H}$, which are sufficient to solve for $\mathtt{H}$ up to an insignificant multiplicative factor. The only restriction is that the four points must be in “general position”, which means that no three points are collinear. The inverse of the transformation $\mathtt{H}$ computed in this way is then applied to the whole image to undo the effect of perspective distortion on the selected plane. The results are shown in figure 2.4b. △

Three remarks concerning this example are appropriate: first, the computation of the rectifying transformation $\mathtt{H}$ in this way does not require knowledge of *any* of the camera’s parameters or the pose of the plane; second, it is not always necessary to

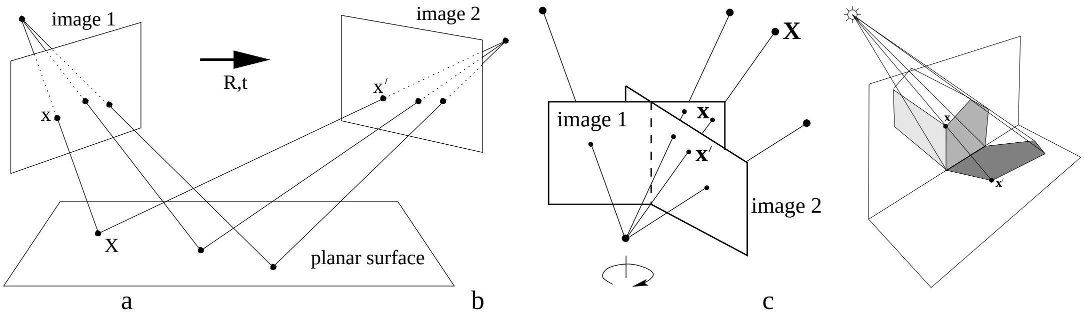


Fig. 2.5. **Examples of a projective transformation, $\mathbf{x}' = \mathtt{H}\mathbf{x}$, arising in perspective images.** *(a) The projective transformation between two images induced by a world plane (the concatenation of two projective transformations is a projective transformation); (b) The projective transformation between two images with the same camera centre (e.g. a camera rotating about its centre or a camera varying its focal length); (c) The projective transformation between the image of a plane (the end of the building) and the image of its shadow onto another plane (the ground plane). Figure (c) courtesy of Luc Van Gool.*

know coordinates for four points in order to remove projective distortion: alternative approaches, which are described in section 2.7, require less, and different types of, information; third, superior (and preferred) methods for computing projective transformations are described in chapter 4.

Projective transformations are important mappings representing many more situations than the perspective imaging of a world plane. A number of other examples are illustrated in figure 2.5. Each of these situations is covered in more detail later in the book.

2.3.1 Transformations of lines and conics

Transformation of lines. It was shown in the proof of theorem 2.10 that if points $\mathbf{x}_i$ lie on a line $\mathbf{l}$, then the transformed points $\mathbf{x}'_i = \mathtt{H}\mathbf{x}_i$ under a projective transformation lie on the line $\mathbf{l}' = \mathtt{H}^{-\mathsf{T}}\mathbf{l}$. In this way, incidence of points on lines is preserved, since $\mathbf{l}'^{\mathsf{T}}\mathbf{x}'_i = \mathbf{l}^{\mathsf{T}}\mathtt{H}^{-1}\mathtt{H}\mathbf{x}_i = 0$. This gives the transformation rule for lines:
Under the point transformation $\mathbf{x}' = \mathtt{H}\mathbf{x}$, a line transforms as

$$\mathbf{l}' = \mathtt{H}^{-\mathsf{T}}\mathbf{l}. \tag{2.6}$$

One may alternatively write $\mathbf{l}'^{\mathsf{T}} = \mathbf{l}^{\mathsf{T}}\mathtt{H}^{-1}$. Note the fundamentally different way in which lines and points transform. Points transform according to $\mathtt{H}$, whereas lines (as rows) transform according to $\mathtt{H}^{-1}$. This may be explained in terms of "covariant" or "contravariant" behaviour. One says that points transform *contravariantly* and lines transform *covariantly*. This distinction will be taken up again, when we discuss tensors in chapter 15 and is fully explained in appendix 1(*p*562).

Transformation of conics. Under a point transformation $\mathbf{x}' = \mathtt{H}\mathbf{x}$, (2.2) becomes

$$\begin{aligned}\mathbf{x}^{\mathsf{T}}\mathtt{C}\mathbf{x} &= \mathbf{x}'^{\mathsf{T}}[\mathtt{H}^{-1}]^{\mathsf{T}}\mathtt{C}\mathtt{H}^{-1}\mathbf{x}' \\ &= \mathbf{x}'^{\mathsf{T}}\mathtt{H}^{-\mathsf{T}}\mathtt{C}\mathtt{H}^{-1}\mathbf{x}'\end{aligned}$$

a b c

Fig. 2.6. **Distortions arising under central projection.** *Images of a tiled floor. (a)* **Similarity***: the circular pattern is imaged as a circle. A square tile is imaged as a square. Lines which are parallel or perpendicular have the same relative orientation in the image. (b)* **Affine***: The circle is imaged as an ellipse. Orthogonal world lines are not imaged as orthogonal lines. However, the sides of the square tiles, which are parallel in the world are parallel in the image. (c)* **Projective***: Parallel world lines are imaged as converging lines. Tiles closer to the camera have a larger image than those further away.*

which is a quadratic form $\mathbf{x}'^{\mathsf{T}}\mathtt{C}'\mathbf{x}'$ with $\mathtt{C}' = \mathtt{H}^{-\mathsf{T}}\mathtt{C}\mathtt{H}^{-1}$. This gives the transformation rule for a conic:

Result 2.13. *Under a point transformation* $\mathbf{x}' = \mathtt{H}\mathbf{x}$*, a conic* $\mathtt{C}$ *transforms to* $\mathtt{C}' = \mathtt{H}^{-\mathsf{T}}\mathtt{C}\mathtt{H}^{-1}$.

The presence of $\mathtt{H}^{-1}$ in this equation may be expressed by saying that a conic transforms *covariantly*. The transformation rule for a dual conic is derived in a similar manner. This gives:

Result 2.14. *Under a point transformation* $\mathbf{x}' = \mathtt{H}\mathbf{x}$*, a dual conic* $\mathtt{C}^*$ *transforms to* $\mathtt{C}^{*\prime} = \mathtt{H}\mathtt{C}^*\mathtt{H}^{\mathsf{T}}$.

2.4 A hierarchy of transformations

In this section we describe the important specializations of a projective transformation and their geometric properties. It was shown in section 2.3 that projective transformations form a group. This group is called the *projective linear group*, and it will be seen that these specializations are *subgroups* of this group.

The group of invertible $n \times n$ matrices with real elements is the (real) general linear group on n dimensions, or $GL(n)$. To obtain the projective linear group the matrices related by a scalar multiplier are identified, giving $PL(n)$ (this is a quotient group of $GL(n)$). In the case of projective transformations of the plane $n = 3$.

The important subgroups of $PL(3)$ include the *affine group*, which is the subgroup of $PL(3)$ consisting of matrices for which the last row is $(0, 0, 1)$, and the *Euclidean group*, which is a subgroup of the affine group for which in addition the upper left hand 2×2 matrix is orthogonal. One may also identify the *oriented Euclidean group* in which the upper left hand 2×2 matrix has determinant 1.

We will introduce these transformations starting from the most specialized, the isometries, and progressively generalizing until projective transformations are reached.

This defines a *hierarchy* of transformations. The distortion effects of various transformations in this hierarchy are shown in figure 2.6.

Some transformations of interest are not groups, for example, perspectivities (because the composition of two perspectivities is a projectivity, not a perspectivity). This point is covered in section A7.4(*p*632).

Invariants. An alternative to describing the transformation *algebraically*, i.e. as a matrix acting on coordinates of a point or curve, is to describe the transformation in terms of those elements or quantities that are preserved or *invariant*. A (scalar) invariant of a geometric configuration is a function of the configuration whose value is unchanged by a particular transformation. For example, the separation of two points is unchanged by a Euclidean transformation (translation and rotation), but not by a similarity (e.g. translation, rotation and isotropic scaling). Distance is thus a Euclidean, but not similarity invariant. The angle between two lines is both a Euclidean and a similarity invariant.

2.4.1 Class I: Isometries

Isometries are transformations of the plane $\mathbb{R}^2$ that preserve Euclidean distance (from *iso* = same, *metric* = measure). An isometry is represented as

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} \epsilon\cos\theta & -\sin\theta & t_x \\ \epsilon\sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

where $\epsilon = \pm 1$. If $\epsilon = 1$ then the isometry is *orientation-preserving* and is a *Euclidean* transformation (a composition of a translation and rotation). If $\epsilon = -1$ then the isometry reverses orientation. An example is the composition of a reflection, represented by the matrix $\text{diag}(-1, 1, 1)$, with a Euclidean transformation.

Euclidean transformations model the motion of a rigid object. They are by far the most important isometries in practice, and we will concentrate on these. However, the orientation reversing isometries often arise as ambiguities in structure recovery.

A planar Euclidean transformation can be written more concisely in block form as

$$\mathbf{x}' = \mathtt{H}_{\mathrm{E}}\mathbf{x} = \begin{bmatrix} \mathtt{R} & \mathbf{t} \\ \mathbf{0}^{\mathsf{T}} & 1 \end{bmatrix} \mathbf{x} \tag{2.7}$$

where $\mathtt{R}$ is a 2×2 rotation matrix (an orthogonal matrix such that $\mathtt{R}^{\mathsf{T}}\mathtt{R} = \mathtt{R}\mathtt{R}^{\mathsf{T}} = \mathtt{I}$), $\mathbf{t}$ a translation 2-vector, and $\mathbf{0}$ a null 2-vector. Special cases are a pure rotation (when $\mathbf{t} = \mathbf{0}$) and a pure translation (when $\mathtt{R} = \mathtt{I}$). A Euclidean transformation is also known as a *displacement*.

A planar Euclidean transformation has three degrees of freedom, one for the rotation and two for the translation. Thus three parameters must be specified in order to define the transformation. The transformation can be computed from two point correspondences.

Invariants. The invariants are very familiar, for instance: length (the distance between two points), angle (the angle between two lines), and area.

Groups and orientation. An isometry is orientation-preserving if the upper left hand 2×2 matrix has determinant 1. Orientation-*preserving* isometries form a group, orientation-*reversing* ones do not. This distinction applies also in the case of similarity and affine transformations which now follow.

2.4.2 Class II: Similarity transformations

A similarity transformation (or more simply a *similarity*) is an isometry composed with an isotropic scaling. In the case of a Euclidean transformation composed with a scaling (i.e. no reflection) the similarity has matrix representation

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} s\cos\theta & -s\sin\theta & t_x \\ s\sin\theta & s\cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \tag{2.8}$$

This can be written more concisely in block form as

$$\mathbf{x}' = \mathtt{H}_{\mathrm{S}}\mathbf{x} = \begin{bmatrix} s\mathtt{R} & \mathbf{t} \\ \mathbf{0}^{\mathsf{T}} & 1 \end{bmatrix} \mathbf{x} \tag{2.9}$$

where the scalar s represents the isotropic scaling. A similarity transformation is also known as an *equi-form* transformation, because it preserves "shape" (form). A planar similarity transformation has four degrees of freedom, the scaling accounting for one more degree of freedom than a Euclidean transformation. A similarity can be computed from two point correspondences.

Invariants. The invariants can be constructed from Euclidean invariants with suitable provision being made for the additional scaling degree of freedom. Angles between lines are not affected by rotation, translation or isotropic scaling, and so are similarity invariants. In particular parallel lines are mapped to parallel lines. The length between two points is not a similarity invariant, but the *ratio* of two lengths is an invariant, because the scaling of the lengths cancels out. Similarly a ratio of areas is an invariant because the scaling (squared) cancels out.

Metric structure. A term that will be used frequently in the discussion on reconstruction (chapter 10) is *metric*. The description *metric structure* implies that the structure is defined up to a similarity.

2.4.3 Class III: Affine transformations

An affine transformation (or more simply an *affinity*) is a non-singular linear transformation followed by a translation. It has the matrix representation

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{2.10}$$

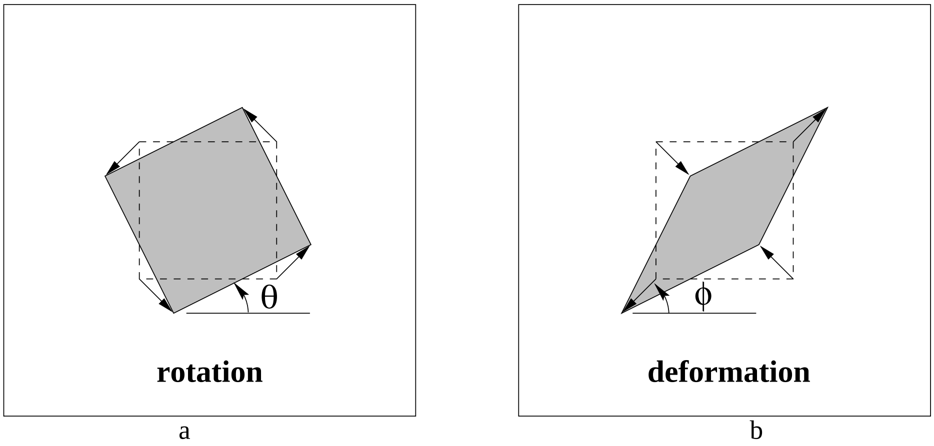


Fig. 2.7. **Distortions arising from a planar affine transformation.** *(a) Rotation by* $\mathtt{R}(\theta)$*. (b) A deformation* $\mathtt{R}(-\phi)\,\mathtt{D}\,\mathtt{R}(\phi)$*. Note, the scaling directions in the deformation are orthogonal.*

or in block form

$$\mathbf{x}' = \mathtt{H}_{\mathtt{A}}\mathbf{x} = \begin{bmatrix} \mathtt{A} & \mathbf{t} \\ \mathbf{0}^{\mathsf{T}} & 1 \end{bmatrix}\mathbf{x} \tag{2.11}$$

with $\mathtt{A}$ a 2×2 non-singular matrix. A planar affine transformation has six degrees of freedom corresponding to the six matrix elements. The transformation can be computed from three point correspondences.

A helpful way to understand the geometric effects of the linear component $\mathtt{A}$ of an affine transformation is as the composition of two fundamental transformations, namely rotations and non-isotropic scalings. The affine matrix $\mathtt{A}$ can always be decomposed as

$$\mathtt{A} \;=\; \mathtt{R}(\theta)\,\mathtt{R}(-\phi)\,\mathtt{D}\,\mathtt{R}(\phi) \tag{2.12}$$

where $\mathtt{R}(\theta)$ and $\mathtt{R}(\phi)$ are rotations by θ and ϕ respectively, and $\mathtt{D}$ is a diagonal matrix:

$$\mathtt{D} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

This decomposition follows directly from the SVD (section A4.4(*p*585)): writing $\mathtt{A} = \mathtt{UDV}^{\mathsf{T}} = (\mathtt{UV}^{\mathsf{T}})(\mathtt{VDV}^{\mathsf{T}}) = \mathtt{R}(\theta)\,(\mathtt{R}(-\phi)\,\mathtt{D}\,\mathtt{R}(\phi))$, since $\mathtt{U}$ and $\mathtt{V}$ are orthogonal matrices.

The affine matrix $\mathtt{A}$ is hence seen to be the concatenation of a rotation (by ϕ); a scaling by λ_1 and λ_2 respectively in the (rotated) x and y directions; a rotation back (by $-\phi$); and finally another rotation (by θ). The only "new" geometry, compared to a similarity, is the non-isotropic scaling. This accounts for the two extra degrees of freedom possessed by an affinity over a similarity. They are the angle ϕ specifying the scaling direction, and the ratio of the scaling parameters $\lambda_1 : \lambda_2$. The essence of an affinity is this scaling in orthogonal directions, oriented at a particular angle. Schematic examples are given in figure 2.7.

Invariants. Because an affine transformation includes non-isotropic scaling, the similarity invariants of length ratios and angles between lines are not preserved under an affinity. Three important invariants are:

(i) **Parallel lines.** Consider two parallel lines. These intersect at a point $(x_1, x_2, 0)^\mathsf{T}$ at infinity. Under an affine transformation this point is mapped to another point at infinity. Consequently, the parallel lines are mapped to lines which still intersect at infinity, and so are parallel after the transformation.

(ii) **Ratio of lengths of parallel line segments.** The length scaling of a line segment depends only on the angle between the line direction and scaling directions. Suppose the line is at angle α to the x-axis of the orthogonal scaling direction, then the scaling magnitude is $\sqrt{\lambda_1^2 \cos^2 \alpha + \lambda_2^2 \sin^2 \alpha}$. This scaling is common to all lines with the same direction, and so cancels out in a ratio of parallel segment lengths.

(iii) **Ratio of areas.** This invariance can be deduced directly from the decomposition (2.12). Rotations and translations do not affect area, so only the scalings by λ_1 and λ_2 matter here. The effect is that area is scaled by $\lambda_1 \lambda_2$ which is equal to $\det \mathtt{A}$. Thus the area of any shape is scaled by $\det \mathtt{A}$, and so the scaling cancels out for a ratio of areas. It will be seen that this does not hold for a projective transformation.

An affinity is orientation-preserving or -reversing according to whether $\det \mathtt{A}$ is positive or negative respectively. Since $\det \mathtt{A} = \lambda_1 \lambda_2$ the property depends only on the sign of the scalings.

2.4.4 Class IV: Projective transformations

A projective transformation was defined in (2.5). It is a general non-singular linear transformation of *homogeneous* coordinates. This generalizes an affine transformation, which is the composition of a general non-singular linear transformation of *inhomogeneous* coordinates and a translation. We have earlier seen the action of a projective transformation (in section 2.3). Here we examine its block form

$$\mathbf{x}' = \mathtt{H}_\mathrm{P}\mathbf{x} = \begin{bmatrix} \mathtt{A} & \mathbf{t} \\ \mathbf{v}^\mathsf{T} & v \end{bmatrix} \mathbf{x} \tag{2.13}$$

where the vector $\mathbf{v} = (v_1, v_2)^\mathsf{T}$. The matrix has nine elements with only their ratio significant, so the transformation is specified by eight parameters. Note, it is not always possible to scale the matrix such that v is unity since v might be zero. A projective transformation between two planes can be computed from four point correspondences, with no three collinear on either plane. See figure 2.4.

Unlike the case of affinities, it is not possible to distinguish between orientation preserving and orientation reversing projectivities in $\mathbb{P}^2$. We will return to this point in section 2.6.

Invariants. The most fundamental projective invariant is the cross ratio of four collinear points: a ratio of lengths on a line is invariant under affinities, but not under projectivities. However, a ratio of ratios or *cross ratio* of lengths on a line is a projective invariant. We return to properties of this invariant in section 2.5.

2.4.5 Summary and comparison

Affinities (6 dof) occupy the middle ground between similarities (4 dof) and projectivities (8 dof). They generalize similarities in that angles are not preserved, so that shapes are skewed under the transformation. On the other hand their action is homogeneous over the plane: for a given affinity the $\det \mathtt{A}$ scaling in area of an object (e.g. a square) is the same anywhere on the plane; and the orientation of a transformed line depends only on its initial orientation, not on its position on the plane. In contrast, for a given projective transformation, area scaling varies with position (e.g. under perspective a more distant square on the plane has a smaller image than one that is nearer, as in figure 2.6); and the orientation of a transformed line depends on both the orientation and position of the source line (however, it will be seen later in section 8.6(*p*213) that a line's vanishing point depends only on line orientation, not position).

The key difference between a projective and affine transformation is that the vector $\mathbf{v}$ is not null for a projectivity. This is responsible for the non-linear effects of the projectivity. Compare the mapping of an ideal point $(x_1, x_2, 0)^\mathsf{T}$ under an affinity and projectivity: First the affine transformation

$$\begin{bmatrix} \mathtt{A} & \mathbf{t} \\ \mathbf{0}^\mathsf{T} & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} = \begin{pmatrix} \mathtt{A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ 0 \end{pmatrix}. \tag{2.14}$$

Second the projective transformation

$$\begin{bmatrix} \mathtt{A} & \mathbf{t} \\ \mathbf{v}^\mathsf{T} & v \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} = \begin{pmatrix} \mathtt{A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ v_1 x_1 + v_2 x_2 \end{pmatrix}. \tag{2.15}$$

In the first case the ideal point remains ideal (i.e. at infinity). In the second it is mapped to a finite point. It is this ability which allows a projective transformation to model vanishing points.

2.4.6 Decomposition of a projective transformation

A projective transformation can be decomposed into a chain of transformations, where each matrix in the chain represents a transformation higher in the hierarchy than the previous one.

$$\mathtt{H} = \mathtt{H}_\mathrm{S}\, \mathtt{H}_\mathrm{A}\, \mathtt{H}_\mathrm{P} = \begin{bmatrix} s\mathtt{R} & \mathbf{t} \\ \mathbf{0}^\mathsf{T} & 1 \end{bmatrix} \begin{bmatrix} \mathtt{K} & \mathbf{0} \\ \mathbf{0}^\mathsf{T} & 1 \end{bmatrix} \begin{bmatrix} \mathtt{I} & \mathbf{0} \\ \mathbf{v}^\mathsf{T} & v \end{bmatrix} = \begin{bmatrix} \mathtt{A} & \mathbf{t} \\ \mathbf{v}^\mathsf{T} & v \end{bmatrix} \tag{2.16}$$

with $\mathtt{A}$ a non-singular matrix given by $\mathtt{A} = s\mathtt{R}\mathtt{K} + \mathbf{t}\mathbf{v}^\mathsf{T}$, and $\mathtt{K}$ an upper-triangular matrix normalized as $\det \mathtt{K} = 1$. This decomposition is valid provided $v \neq 0$, and is unique if s is chosen positive.

Each of the matrices $\mathtt{H}_\mathrm{S}, \mathtt{H}_\mathrm{A}, \mathtt{H}_\mathrm{P}$ is the "essence" of a transformation of that type (as indicated by the subscripts S, A, P). Consider the process of rectifying the perspective image of a plane as in example 2.12: $\mathtt{H}_\mathrm{P}$ (2 dof) moves the line at infinity; $\mathtt{H}_\mathrm{A}$ (2 dof) affects the affine properties, but does not move the line at infinity; and finally, $\mathtt{H}_\mathrm{S}$ is a general similarity transformation (4 dof) which does not affect the affine or projective properties. The transformation $\mathtt{H}_\mathrm{P}$ is an *elation,* described in section A7.3(*p*631).

Example 2.15. The projective transformation

$$\mathtt{H} = \begin{bmatrix} 1.707 & 0.586 & 1.0 \\ 2.707 & 8.242 & 2.0 \\ 1.0 & 2.0 & 1.0 \end{bmatrix}$$

may be decomposed as

$$\mathtt{H} = \begin{bmatrix} 2\cos 45^\circ & -2\sin 45^\circ & 1 \\ 2\sin 45^\circ & 2\cos 45^\circ & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0.5 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}.$$

△

This decomposition can be employed when the objective is to only partially determine the transformation. For example, if one wants to measure length ratios from the perspective image of a plane, then it is only necessary to determine (rectify) the transformation up to a similarity. We return to this approach in section 2.7.

Taking the inverse of $\mathtt{H}$ in (2.16) gives $\mathtt{H}^{-1} = \mathtt{H}_\mathrm{P}^{-1}\,\mathtt{H}_\mathrm{A}^{-1}\,\mathtt{H}_\mathrm{S}^{-1}$. Since $\mathtt{H}_\mathrm{P}^{-1}, \mathtt{H}_\mathrm{A}^{-1}$ and $\mathtt{H}_\mathrm{S}^{-1}$ are still projective, affine and similarity transformations respectively, a general projective transformation may also be decomposed in the form

$$\mathtt{H} = \mathtt{H}_\mathrm{P}\,\mathtt{H}_\mathrm{A}\,\mathtt{H}_\mathrm{S} = \begin{bmatrix} \mathtt{I} & \mathbf{0} \\ \mathbf{v}^\mathsf{T} & 1 \end{bmatrix} \begin{bmatrix} \mathtt{K} & \mathbf{0} \\ \mathbf{0}^\mathsf{T} & 1 \end{bmatrix} \begin{bmatrix} s\mathtt{R} & \mathbf{t} \\ \mathbf{0}^\mathsf{T} & 1 \end{bmatrix} \tag{2.17}$$

Note that the actual values of $\mathtt{K}$, $\mathtt{R}$, $\mathbf{t}$ and $\mathbf{v}$ will be different from those of (2.16).

2.4.7 The number of invariants

The question naturally arises as to how many invariants there are for a given geometric configuration under a particular transformation. First the term "number" needs to be made more precise, for if a quantity is invariant, such as length under Euclidean transformations, then any function of that quantity is invariant. Consequently, we seek a counting argument for the number of functionally independent invariants. By considering the number of transformation parameters that must be eliminated in order to form an invariant, it can be seen that:

Result 2.16. *The number of functionally independent invariants is equal to, or greater than, the number of degrees of freedom of the configuration less the number of degrees of freedom of the transformation.*

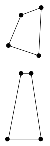
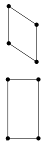
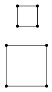
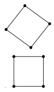
Group	Matrix	Distortion	Invariant properties
Projective 8 dof	$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$		Concurrency, collinearity, **order of contact**: intersection (1 pt contact); tangency (2 pt contact); inflections (3 pt contact with line); tangent discontinuities and cusps. cross ratio (ratio of ratio of lengths).
Affine 6 dof	$\begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Parallelism, ratio of areas, ratio of lengths on collinear or parallel lines (e.g. midpoints), linear combinations of vectors (e.g. centroids). The line at infinity, $\mathbf{l}_\infty$.
Similarity 4 dof	$\begin{bmatrix} sr_{11} & sr_{12} & t_x \\ sr_{21} & sr_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Ratio of lengths, angle. The circular points, **I**, **J** (see section 2.7.3).
Euclidean 3 dof	$\begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Length, area

Table 2.1. **Geometric properties invariant to commonly occurring *planar* transformations.** *The matrix $\mathtt{A} = [a_{ij}]$ is an invertible 2×2 matrix, $\mathtt{R} = [r_{ij}]$ is a 2D rotation matrix, and (t_x, t_y) a 2D translation. The distortion column shows typical effects of the transformations on a square. Transformations higher in the table can produce all the actions of the ones below. These range from Euclidean, where only translations and rotations occur, to projective where the square can be transformed to any arbitrary quadrilateral (provided no three points are collinear).*

For example, a configuration of four points in general position has 8 degrees of freedom (2 for each point), and so 4 similarity, 2 affinity and zero projective invariants since these transformations have respectively 4, 6 and 8 degrees of freedom.

Table 2.1 summarizes the 2D transformation groups and their invariant properties. Transformations lower in the table are specializations of those above. A transformation lower in the table inherits the invariants of those above.

2.5 The projective geometry of 1D

The development of the projective geometry of a line, $\mathbb{P}^1$, proceeds in much the same way as that of the plane. A point x on the line is represented by homogeneous coordinates $(x_1, x_2)^\mathsf{T}$, and a point for which $x_2 = 0$ is an ideal point of the line. We will use the notation $\bar{\mathbf{x}}$ to represent the 2-vector $(x_1, x_2)^\mathsf{T}$. A projective transformation of a line is represented by a 2×2 homogeneous matrix,

$$\bar{\mathbf{x}}' = \mathtt{H}_{2\times 2}\bar{\mathbf{x}}$$

and has 3 degrees of freedom corresponding to the four elements of the matrix less one for overall scaling. A projective transformation of a line may be determined from three corresponding points.

6

Camera Models

A camera is a mapping between the 3D world (object space) and a 2D image. The principal camera of interest in this book is *central projection*. This chapter develops a number of camera *models* which are matrices with particular properties that represent the camera mapping.

It will be seen that all cameras modelling central projection are specializations of the *general projective camera*. The anatomy of this most general camera model is examined using the tools of projective geometry. It will be seen that geometric entities of the camera, such as the projection centre and image plane, can be computed quite simply from its matrix representation. Specializations of the general projective camera inherit its properties, for example their geometry is computed using the same algebraic expressions.

The specialized models fall into two major classes – those that model cameras with a finite centre, and those that model cameras with centre "at infinity". Of the cameras at infinity the *affine camera* is of particular importance because it is the natural generalization of parallel projection.

This chapter is principally concerned with the projection of points. The action of a camera on other geometric entities, such as lines, is deferred until chapter 8.

6.1 Finite cameras

In this section we start with the most specialized and simplest camera model, which is the basic pinhole camera, and then progressively generalize this model through a series of gradations.

The models we develop are principally designed for CCD like sensors, but are also applicable to other cameras, for example X-ray images, scanned photographic negatives, scanned photographs from enlarged negatives, etc.

The basic pinhole model. We consider the central projection of points in space onto a plane. Let the centre of projection be the origin of a Euclidean coordinate system, and consider the plane $\mathrm{Z} = f$, which is called the *image plane* or *focal plane*. Under the pinhole camera model, a point in space with coordinates $\mathbf{X} = (\mathrm{X}, \mathrm{Y}, \mathrm{Z})^\mathsf{T}$ is mapped to the point on the image plane where a line joining the point $\mathbf{X}$ to the centre of projection meets the image plane. This is shown in figure 6.1. By similar triangles, one quickly

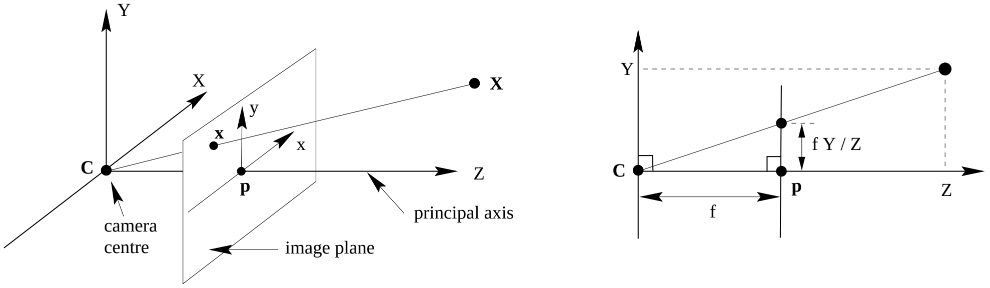


Fig. 6.1. **Pinhole camera geometry.** $\mathbf{C}$ *is the camera centre and* $\mathbf{p}$ *the principal point. The camera centre is here placed at the coordinate origin. Note the image plane is placed in front of the camera centre.*

computes that the point $(\mathrm{X}, \mathrm{Y}, \mathrm{Z})^\mathsf{T}$ is mapped to the point $(f\mathrm{X}/\mathrm{Z}, f\mathrm{Y}/\mathrm{Z}, f)^\mathsf{T}$ on the image plane. Ignoring the final image coordinate, we see that

$$(\mathrm{X}, \mathrm{Y}, \mathrm{Z})^\mathsf{T} \mapsto (f\mathrm{X}/\mathrm{Z}, f\mathrm{Y}/\mathrm{Z})^\mathsf{T} \tag{6.1}$$

describes the central projection mapping from world to image coordinates. This is a mapping from Euclidean 3-space $\mathbb{R}^3$ to Euclidean 2-space $\mathbb{R}^2$.

The centre of projection is called the *camera centre*. It is also known as the *optical centre*. The line from the camera centre perpendicular to the image plane is called the *principal axis* or *principal ray* of the camera, and the point where the principal axis meets the image plane is called the *principal point*. The plane through the camera centre parallel to the image plane is called the *principal plane* of the camera.

Central projection using homogeneous coordinates. If the world and image points are represented by homogeneous vectors, then central projection is very simply expressed as a linear mapping between their homogeneous coordinates. In particular, (6.1) may be written in terms of matrix multiplication as

$$\begin{pmatrix} \mathrm{X} \\ \mathrm{Y} \\ \mathrm{Z} \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} f\mathrm{X} \\ f\mathrm{Y} \\ \mathrm{Z} \end{pmatrix} = \begin{bmatrix} f & & & 0 \\ & f & & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} \mathrm{X} \\ \mathrm{Y} \\ \mathrm{Z} \\ 1 \end{pmatrix}. \tag{6.2}$$

The matrix in this expression may be written as $\mathrm{diag}(f, f, 1)[\mathtt{I} \mid \mathbf{0}]$ where $\mathrm{diag}(f, f, 1)$ is a diagonal matrix and $[\mathtt{I} \mid \mathbf{0}]$ represents a matrix divided up into a 3×3 block (the identity matrix) plus a column vector, here the zero vector.

We now introduce the notation $\mathbf{X}$ for the world point represented by the homogeneous 4-vector $(\mathrm{X}, \mathrm{Y}, \mathrm{Z}, 1)^\mathsf{T}$, $\mathbf{x}$ for the image point represented by a homogeneous 3-vector, and $\mathtt{P}$ for the 3×4 homogeneous *camera projection matrix*. Then (6.2) is written compactly as

$$\mathbf{x} = \mathtt{P}\mathbf{X}$$

which defines the camera matrix for the pinhole model of central projection as

$$\mathtt{P} = \mathrm{diag}(f, f, 1)\, [\mathtt{I} \mid \mathbf{0}].$$

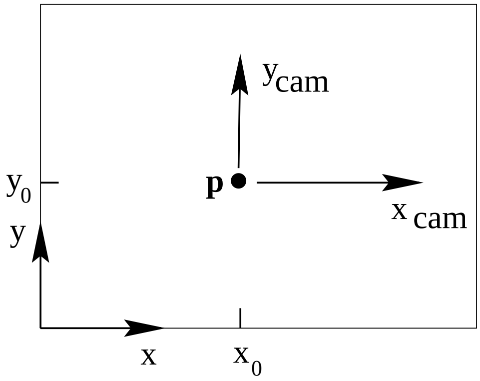


Fig. 6.2. *Image* (x, y) *and camera* $(x_{\rm cam}, y_{\rm cam})$ *coordinate systems.*

Principal point offset. The expression (6.1) assumed that the origin of coordinates in the image plane is at the principal point. In practice, it may not be, so that in general there is a mapping

$$(\mathtt{X}, \mathtt{Y}, \mathtt{Z})^{\mathsf{T}} \mapsto (f\mathtt{X}/\mathtt{Z} + p_x, f\mathtt{Y}/\mathtt{Z} + p_y)^{\mathsf{T}}$$

where $(p_x, p_y)^{\mathsf{T}}$ are the coordinates of the principal point. See figure 6.2. This equation may be expressed conveniently in homogeneous coordinates as

$$\begin{pmatrix} \mathtt{X} \\ \mathtt{Y} \\ \mathtt{Z} \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} f\mathtt{X} + \mathtt{Z}p_x \\ f\mathtt{Y} + \mathtt{Z}p_y \\ \mathtt{Z} \end{pmatrix} = \begin{bmatrix} f & & p_x & 0 \\ & f & p_y & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} \mathtt{X} \\ \mathtt{Y} \\ \mathtt{Z} \\ 1 \end{pmatrix}. \tag{6.3}$$

Now, writing

$$\mathtt{K} = \begin{bmatrix} f & & p_x \\ & f & p_y \\ & & 1 \end{bmatrix} \tag{6.4}$$

then (6.3) has the concise form

$$\mathbf{x} = \mathtt{K}[\mathtt{I} \mid \mathbf{0}]\mathbf{X}_{\rm cam}. \tag{6.5}$$

The matrix $\mathtt{K}$ is called the *camera calibration matrix*. In (6.5) we have written $(\mathtt{X}, \mathtt{Y}, \mathtt{Z}, 1)^{\mathsf{T}}$ as $\mathbf{X}_{\rm cam}$ to emphasize that the camera is assumed to be located at the origin of a Euclidean coordinate system with the principal axis of the camera pointing straight down the $\mathtt{Z}$-axis, and the point $\mathbf{X}_{\rm cam}$ is expressed in this coordinate system. Such a coordinate system may be called the *camera coordinate frame*.

Camera rotation and translation. In general, points in space will be expressed in terms of a different Euclidean coordinate frame, known as the *world coordinate frame*. The two coordinate frames are related via a rotation and a translation. See figure 6.3. If $\widetilde{\mathbf{X}}$ is an inhomogeneous 3-vector representing the coordinates of a point in the world coordinate frame, and $\widetilde{\mathbf{X}}_{\rm cam}$ represents the same point in the camera coordinate frame, then we may write $\widetilde{\mathbf{X}}_{\rm cam} = \mathtt{R}(\widetilde{\mathbf{X}} - \widetilde{\mathbf{C}})$, where $\widetilde{\mathbf{C}}$ represents the coordinates of the camera

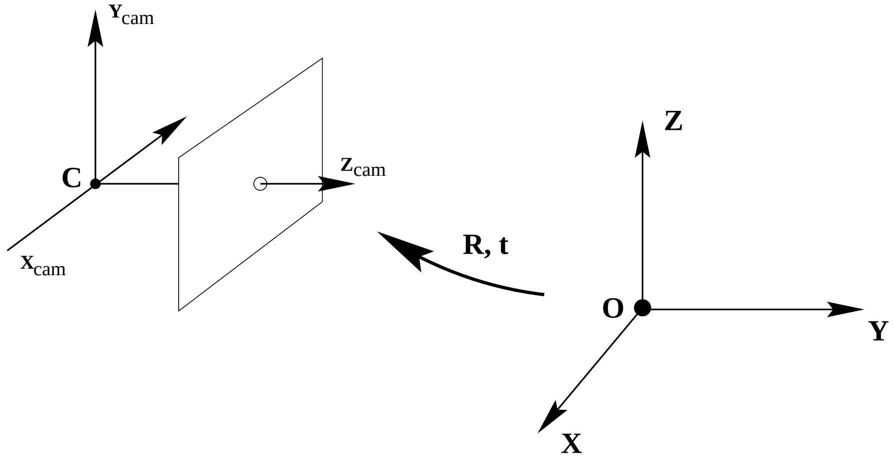


Fig. 6.3. *The Euclidean transformation between the world and camera coordinate frames.*

centre in the world coordinate frame, and $\mathtt{R}$ is a 3×3 rotation matrix representing the orientation of the camera coordinate frame. This equation may be written in homogeneous coordinates as

$$\mathbf{X}_{\text{cam}} = \begin{bmatrix} \mathtt{R} & -\mathtt{R}\widetilde{\mathbf{C}} \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \mathrm{X} \\ \mathrm{Y} \\ \mathrm{Z} \\ 1 \end{pmatrix} = \begin{bmatrix} \mathtt{R} & -\mathtt{R}\widetilde{\mathbf{C}} \\ 0 & 1 \end{bmatrix} \mathbf{X}. \tag{6.6}$$

Putting this together with (6.5) leads to the formula

$$\mathbf{x} = \mathtt{KR}[\mathtt{I} \mid -\widetilde{\mathbf{C}}]\mathbf{X} \tag{6.7}$$

where $\mathbf{X}$ is now in a world coordinate frame. This is the general mapping given by a pinhole camera. One sees that a general pinhole camera, $\mathtt{P} = \mathtt{KR}[\mathtt{I} \mid -\widetilde{\mathbf{C}}]$, has 9 degrees of freedom: 3 for $\mathtt{K}$ (the elements f, p_x, p_y), 3 for $\mathtt{R}$, and 3 for $\widetilde{\mathbf{C}}$. The parameters contained in $\mathtt{K}$ are called the *internal* camera parameters, or the *internal orientation* of the camera. The parameters of $\mathtt{R}$ and $\widetilde{\mathbf{C}}$ which relate the camera orientation and position to a world coordinate system are called the *external* parameters or the *exterior orientation*.

It is often convenient not to make the camera centre explicit, and instead to represent the world to image transformation as $\widetilde{\mathbf{X}}_{\text{cam}} = \mathtt{R}\widetilde{\mathbf{X}} + \mathbf{t}$. In this case the camera matrix is simply

$$\mathtt{P} = \mathtt{K}[\mathtt{R} \mid \mathbf{t}] \tag{6.8}$$

where from (6.7) $\mathbf{t} = -\mathtt{R}\widetilde{\mathbf{C}}$.

CCD cameras. The pinhole camera model just derived assumes that the image coordinates are Euclidean coordinates having equal scales in both axial directions. In the case of CCD cameras, there is the additional possibility of having non-square pixels. If image coordinates are measured in pixels, then this has the extra effect of introducing unequal scale factors in each direction. In particular if the number of pixels per unit distance in image coordinates are m_x and m_y in the x and y directions, then the transformation from world coordinates to pixel coordinates is obtained by multiplying (6.4) on the left by an extra factor diag$(m_x, m_y, 1)$. Thus, the general form of the calibration matrix of a CCD camera is

$$\mathtt{K} = \begin{bmatrix} \alpha_x & & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix} \tag{6.9}$$

where $\alpha_x = f m_x$ and $\alpha_y = f m_y$ represent the focal length of the camera in terms of pixel dimensions in the x and y direction respectively. Similarly, $\tilde{\mathbf{x}}_0 = (x_0, y_0)$ is the principal point in terms of pixel dimensions, with coordinates $x_0 = m_x p_x$ and $y_0 = m_y p_y$. A CCD camera thus has 10 degrees of freedom.

Finite projective camera. For added generality, we can consider a calibration matrix of the form

$$\mathtt{K} = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}. \tag{6.10}$$

The added parameter s is referred to as the *skew* parameter. The skew parameter will be zero for most normal cameras. However, in certain unusual instances which are described in section 6.2.4, it can take non-zero values.

A camera

$$\mathtt{P} = \mathtt{KR}[\mathtt{I} \mid -\widetilde{\mathbf{C}}] \tag{6.11}$$

for which the calibration matrix $\mathtt{K}$ is of the form (6.10) will be called a *finite projective* camera. A finite projective camera has 11 degrees of freedom. This is the same number of degrees of freedom as a 3×4 matrix, defined up to an arbitrary scale.

Note that the left hand 3×3 submatrix of $\mathtt{P}$, equal to $\mathtt{KR}$, is non-singular. Conversely, any 3×4 matrix $\mathtt{P}$ for which the left hand 3×3 submatrix is non-singular is the camera matrix of some finite projective camera, because $\mathtt{P}$ can be decomposed as $\mathtt{P} = \mathtt{KR}[\mathtt{I} \mid -\widetilde{\mathbf{C}}]$. Indeed, letting $\mathtt{M}$ be the left 3×3 submatrix of $\mathtt{P}$, one decomposes $\mathtt{M}$ as a product $\mathtt{M} = \mathtt{KR}$ where $\mathtt{K}$ is upper-triangular of the form (6.10) and $\mathtt{R}$ is a rotation matrix. This decomposition is essentially the RQ matrix decomposition, described in section A4.1.1(*p*579), of which more will be said in section 6.2.4. The matrix $\mathtt{P}$ can therefore be written $\mathtt{P} = \mathtt{M}[\mathtt{I} \mid \mathtt{M}^{-1}\mathbf{p}_4] = \mathtt{KR}[\mathtt{I} \mid -\widetilde{\mathbf{C}}]$ where $\mathbf{p}_4$ is the last column of $\mathtt{P}$. In short

- *The set of camera matrices of finite projective cameras is identical with the set of homogeneous* 3×4 *matrices for which the left hand* 3×3 *submatrix is non-singular.*

General projective cameras. The final step in our hierarchy of projective cameras is to remove the non-singularity restriction on the left hand 3×3 submatrix. A *general projective* camera is one represented by an arbitrary homogeneous 3×4 matrix of rank 3. It has 11 degrees of freedom. The rank 3 requirement arises because if the rank is

Camera centre. The camera centre is the 1-dimensional right null-space $\mathbf{C}$ of $\mathtt{P}$, i.e. $\mathtt{P}\mathbf{C} = \mathbf{0}$.

- ◇ **Finite camera** ($\mathtt{M}$ is not singular) $\mathbf{C} = \begin{pmatrix} -\mathtt{M}^{-1}\mathbf{p}_4 \\ 1 \end{pmatrix}$
- ◇ **Camera at infinity** ($\mathtt{M}$ is singular) $\mathbf{C} = \begin{pmatrix} \mathbf{d} \\ 0 \end{pmatrix}$ where $\mathbf{d}$ is the null 3-vector of $\mathtt{M}$, i.e. $\mathtt{M}\mathbf{d} = \mathbf{0}$.

Column points. For $i = 1, \ldots, 3$, the column vectors $\mathbf{p}_i$ are vanishing points in the image corresponding to the $\mathtt{X}$, $\mathtt{Y}$ and $\mathtt{Z}$ axes respectively. Column $\mathbf{p}_4$ is the image of the coordinate origin.

Principal plane. The principal plane of the camera is $\mathbf{P}^3$, the last row of $\mathtt{P}$.

Axis planes. The planes $\mathbf{P}^1$ and $\mathbf{P}^2$ (the first and second rows of $\mathtt{P}$) represent planes in space through the camera centre, corresponding to points that map to the image lines $x = 0$ and $y = 0$ respectively.

Principal point. The image point $\mathbf{x}_0 = \mathtt{M}\mathbf{m}^3$ is the principal point of the camera, where $\mathbf{m}^{3\mathsf{T}}$ is the third row of $\mathtt{M}$.

Principal ray. The principal ray (axis) of the camera is the ray passing through the camera centre $\mathbf{C}$ with direction vector $\mathbf{m}^{3\mathsf{T}}$. The principal axis vector $\mathbf{v} = \det(\mathtt{M})\mathbf{m}^3$ is directed towards the front of the camera.

Table 6.1. *Summary of the properties of a projective camera* $\mathtt{P}$. *The matrix is represented by the block form* $\mathtt{P} = [\mathtt{M} \mid \mathbf{p}_4]$.

less than this then the range of the matrix mapping will be a line or point and not the whole plane; in other words not a 2D image.

6.2 The projective camera

A general projective camera $\mathtt{P}$ maps world points $\mathbf{X}$ to image points $\mathbf{x}$ according to $\mathbf{x} = \mathtt{P}\mathbf{X}$. Building on this mapping we will now dissect the camera model to reveal how geometric entities, such as the camera centre, are encoded. Some of the properties that we consider will apply only to finite projective cameras and their specializations, whilst others will apply to general cameras. The distinction will be evident from the context. The derived properties of the camera are summarized in table 6.1.

6.2.1 Camera anatomy

A general projective camera may be decomposed into blocks according to $\mathtt{P} = [\mathtt{M} \mid \mathbf{p}_4]$, where $\mathtt{M}$ is a 3×3 matrix. It will be seen that if $\mathtt{M}$ is non-singular, then this is a finite camera, otherwise it is not.

Camera centre. The matrix $\mathtt{P}$ has a 1-dimensional right null-space because its rank is 3, whereas it has 4 columns. Suppose the null-space is generated by the 4-vector $\mathbf{C}$, that is $\mathtt{P}\mathbf{C} = \mathbf{0}$. It will now be shown that $\mathbf{C}$ is the camera centre, represented as a homogeneous 4-vector.

Consider the line containing $\mathbf{C}$ and any other point $\mathbf{A}$ in 3-space. Points on this line may be represented by the join

$$\mathbf{X}(\lambda) = \lambda\mathbf{A} + (1 - \lambda)\mathbf{C} \ .$$

9

Epipolar Geometry and the Fundamental Matrix

The epipolar geometry is the intrinsic projective geometry between two views. It is independent of scene structure, and only depends on the cameras' internal parameters and relative pose.

The fundamental matrix $\mathtt{F}$ encapsulates this intrinsic geometry. It is a 3×3 matrix of rank 2. If a point in 3-space $\mathbf{X}$ is imaged as $\mathbf{x}$ in the first view, and $\mathbf{x}'$ in the second, then the image points satisfy the relation $\mathbf{x}'^{\mathsf{T}}\mathtt{F}\mathbf{x} = 0$.

We will first describe epipolar geometry, and derive the fundamental matrix. The properties of the fundamental matrix are then elucidated, both for general motion of the camera between the views, and for several commonly occurring special motions. It is next shown that the cameras can be retrieved from $\mathtt{F}$ up to a projective transformation of 3-space. This result is the basis for the projective reconstruction theorem given in chapter 10. Finally, if the camera internal calibration is known, it is shown that the Euclidean motion of the cameras between views may be computed from the fundamental matrix up to a finite number of ambiguities.

The fundamental matrix is independent of scene structure. However, it can be computed from correspondences of imaged scene points alone, without requiring knowledge of the cameras' internal parameters or relative pose. This computation is described in chapter 11.

9.1 Epipolar geometry

The epipolar geometry between two views is essentially the geometry of the intersection of the image planes with the pencil of planes having the baseline as axis (the baseline is the line joining the camera centres). This geometry is usually motivated by considering the search for corresponding points in stereo matching, and we will start from that objective here.

Suppose a point $\mathbf{X}$ in 3-space is imaged in two views, at $\mathbf{x}$ in the first, and $\mathbf{x}'$ in the second. What is the relation between the corresponding image points $\mathbf{x}$ and $\mathbf{x}'$? As shown in figure 9.1a the image points $\mathbf{x}$ and $\mathbf{x}'$, space point $\mathbf{X}$, and camera centres are coplanar. Denote this plane as $\boldsymbol{\pi}$. Clearly, the rays back-projected from $\mathbf{x}$ and $\mathbf{x}'$ intersect at $\mathbf{X}$, and the rays are coplanar, lying in $\boldsymbol{\pi}$. It is this latter property that is of most significance in searching for a correspondence.

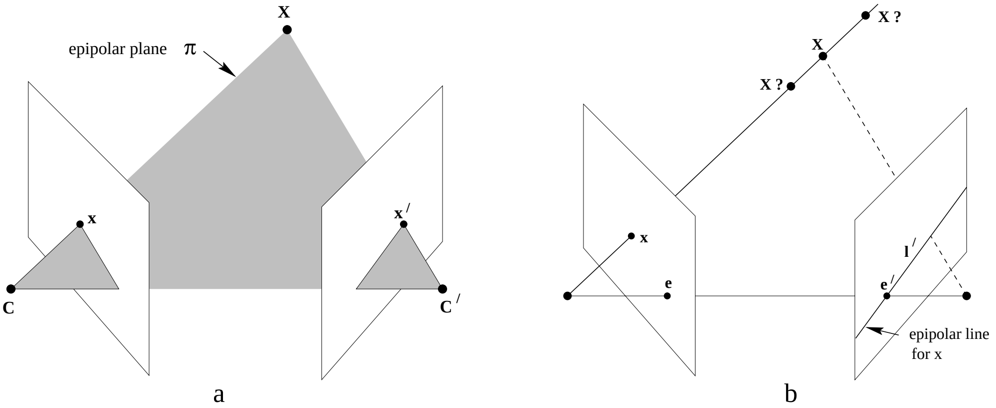


Fig. 9.1. **Point correspondence geometry.** *(a) The two cameras are indicated by their centres* $\mathbf{C}$ *and* $\mathbf{C}'$ *and image planes. The camera centres, 3-space point* $\mathbf{X}$*, and its images* $\mathbf{x}$ *and* $\mathbf{x}'$ *lie in a common plane* $\boldsymbol{\pi}$*. (b) An image point* $\mathbf{x}$ *back-projects to a ray in 3-space defined by the first camera centre,* $\mathbf{C}$*, and* $\mathbf{x}$*. This ray is imaged as a line* $\mathbf{l}'$ *in the second view. The 3-space point* $\mathbf{X}$ *which projects to* $\mathbf{x}$ *must lie on this ray, so the image of* $\mathbf{X}$ *in the second view must lie on* $\mathbf{l}'$*.*

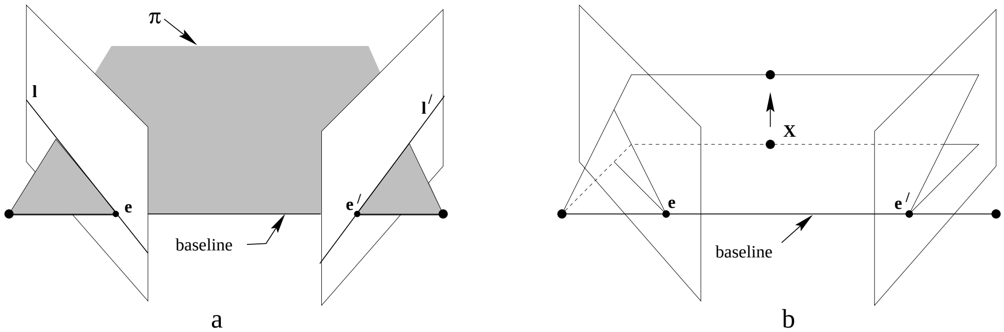


Fig. 9.2. **Epipolar geometry.** *(a) The camera baseline intersects each image plane at the epipoles* $\mathbf{e}$ *and* $\mathbf{e}'$*. Any plane* $\boldsymbol{\pi}$ *containing the baseline is an epipolar plane, and intersects the image planes in corresponding epipolar lines* $\mathbf{l}$ *and* $\mathbf{l}'$*. (b) As the position of the 3D point* $\mathbf{X}$ *varies, the epipolar planes "rotate" about the baseline. This family of planes is known as an epipolar pencil. All epipolar lines intersect at the epipole.*

Supposing now that we know only $\mathbf{x}$, we may ask how the corresponding point $\mathbf{x}'$ is constrained. The plane $\boldsymbol{\pi}$ is determined by the baseline and the ray defined by $\mathbf{x}$. From above we know that the ray corresponding to the (unknown) point $\mathbf{x}'$ lies in $\boldsymbol{\pi}$, hence the point $\mathbf{x}'$ lies on the line of intersection $\mathbf{l}'$ of $\boldsymbol{\pi}$ with the second image plane. This line $\mathbf{l}'$ is the image in the second view of the ray back-projected from $\mathbf{x}$. It is the *epipolar line* corresponding to $\mathbf{x}$. In terms of a stereo correspondence algorithm the benefit is that the search for the point corresponding to $\mathbf{x}$ need not cover the entire image plane but can be restricted to the line $\mathbf{l}'$.

The geometric entities involved in epipolar geometry are illustrated in figure 9.2. The terminology is

- The **epipole** is the *point* of intersection of the line joining the camera centres (the baseline) with the image plane. Equivalently, the epipole is the image in one view

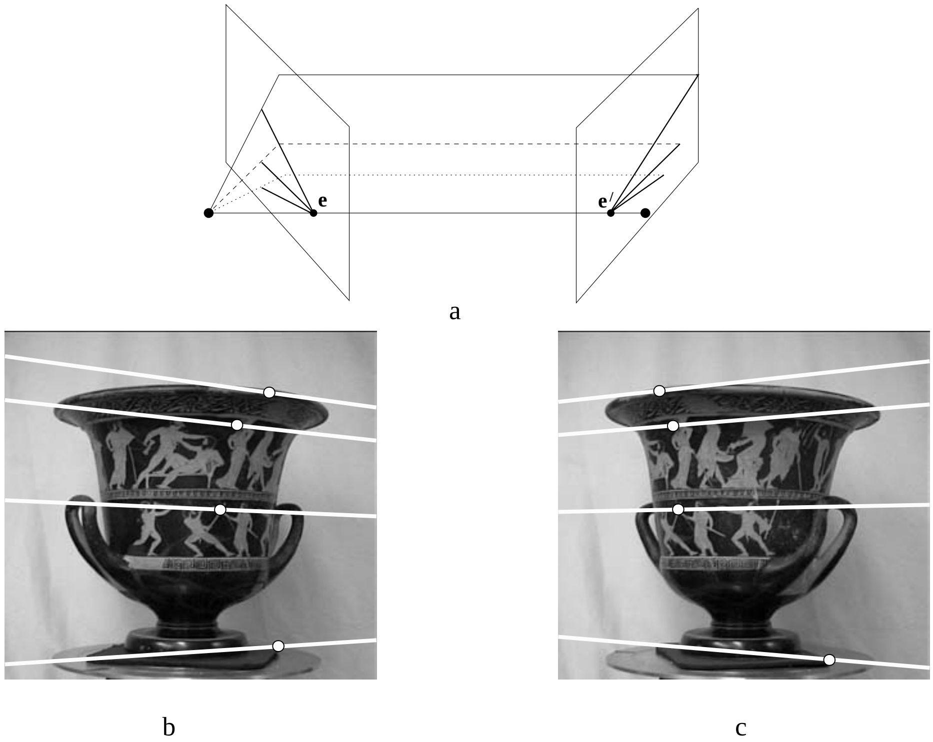


Fig. 9.3. **Converging cameras.** *(a) Epipolar geometry for converging cameras. (b) and (c) A pair of images with superimposed corresponding points and their epipolar lines (in white). The motion between the views is a translation and rotation. In each image, the direction of the other camera may be inferred from the intersection of the pencil of epipolar lines. In this case, both epipoles lie outside of the visible image.*

of the camera centre of the other view. It is also the vanishing point of the baseline (translation) direction.

- An **epipolar plane** is a plane containing the baseline. There is a one-parameter family (a pencil) of epipolar planes.
- An **epipolar line** is the intersection of an epipolar plane with the image plane. All epipolar lines intersect at the epipole. An epipolar plane intersects the left and right image planes in epipolar lines, and defines the correspondence between the lines.

Examples of epipolar geometry are given in figure 9.3 and figure 9.4. The epipolar geometry of these image pairs, and indeed all the examples of this chapter, is computed directly from the images as described in section 11.6(*p*290).

9.2 The fundamental matrix F

The fundamental matrix is the algebraic representation of epipolar geometry. In the following we derive the fundamental matrix from the mapping between a point and its epipolar line, and then specify the properties of the matrix.

Given a pair of images, it was seen in figure 9.1 that to each point $\mathbf{x}$ in one image, there exists a corresponding epipolar line $\mathbf{l}'$ in the other image. Any point $\mathbf{x}'$ in the second image matching the point $\mathbf{x}$ must lie on the epipolar line $\mathbf{l}'$. The epipolar line

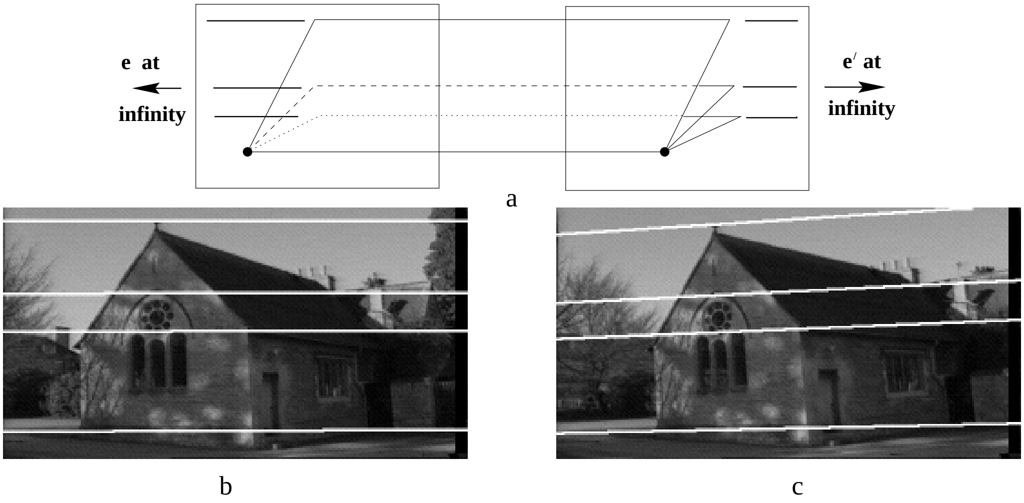


Fig. 9.4. **Motion parallel to the image plane.** *In the case of a special motion where the translation is parallel to the image plane, and the rotation axis is perpendicular to the image plane, the intersection of the baseline with the image plane is at infinity. Consequently the epipoles are at infinity, and epipolar lines are parallel. (a) Epipolar geometry for motion parallel to the image plane. (b) and (c) a pair of images for which the motion between views is (approximately) a translation parallel to the* x*-axis, with no rotation. Four corresponding epipolar lines are superimposed in white. Note that corresponding points lie on corresponding epipolar lines.*

is the projection in the second image of the ray from the point $\mathbf{x}$ through the camera centre $\mathbf{C}$ of the first camera. Thus, there is a map

$$\mathbf{x} \mapsto \mathbf{l}'$$

from a point in one image to its corresponding epipolar line in the other image. It is the nature of this map that will now be explored. It will turn out that this mapping is a (singular) *correlation*, that is a projective mapping from points to lines, which is represented by a matrix $\mathtt{F}$, the fundamental matrix.

9.2.1 Geometric derivation

We begin with a geometric derivation of the fundamental matrix. The mapping from a point in one image to a corresponding epipolar line in the other image may be decomposed into two steps. In the first step, the point $\mathbf{x}$ is mapped to some point $\mathbf{x}'$ in the other image lying on the epipolar line $\mathbf{l}'$. This point $\mathbf{x}'$ is a potential match for the point $\mathbf{x}$. In the second step, the epipolar line $\mathbf{l}'$ is obtained as the line joining $\mathbf{x}'$ to the epipole $\mathbf{e}'$.

Step 1: Point transfer via a plane. Refer to figure 9.5. Consider a plane $\boldsymbol{\pi}$ in space not passing through either of the two camera centres. The ray through the first camera centre corresponding to the point $\mathbf{x}$ meets the plane $\boldsymbol{\pi}$ in a point $\mathbf{X}$. This point $\mathbf{X}$ is then projected to a point $\mathbf{x}'$ in the second image. This procedure is known as transfer via the plane $\boldsymbol{\pi}$. Since $\mathbf{X}$ lies on the ray corresponding to $\mathbf{x}$, the projected point $\mathbf{x}'$ must lie on the epipolar line $\mathbf{l}'$ corresponding to the image of this ray, as illustrated in

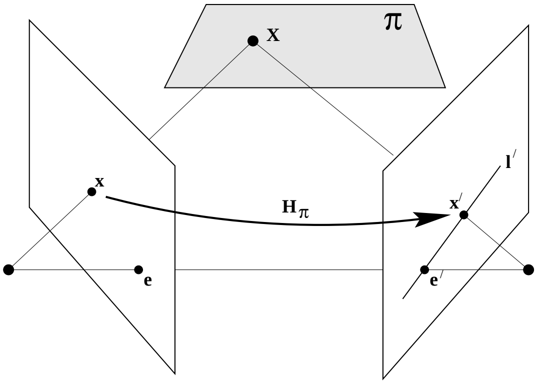


Fig. 9.5. *A point $\mathbf{x}$ in one image is transferred via the plane $\boldsymbol{\pi}$ to a matching point $\mathbf{x}'$ in the second image. The epipolar line through $\mathbf{x}'$ is obtained by joining $\mathbf{x}'$ to the epipole $\mathbf{e}'$. In symbols one may write $\mathbf{x}' = \mathtt{H}_{\boldsymbol{\pi}}\mathbf{x}$ and $\mathbf{l}' = [\mathbf{e}']_\times \mathbf{x}' = [\mathbf{e}']_\times \mathtt{H}_{\boldsymbol{\pi}}\mathbf{x} = \mathtt{F}\mathbf{x}$ where $\mathtt{F} = [\mathbf{e}']_\times \mathtt{H}_{\boldsymbol{\pi}}$ is the fundamental matrix.*

figure 9.1b. The points $\mathbf{x}$ and $\mathbf{x}'$ are both images of the 3D point $\mathbf{X}$ lying on a plane. The set of all such points $\mathbf{x}_i$ in the first image and the corresponding points $\mathbf{x}'_i$ in the second image are projectively equivalent, since they are each projectively equivalent to the planar point set $\mathbf{X}_i$. Thus there is a 2D homography $\mathtt{H}_{\boldsymbol{\pi}}$ mapping each $\mathbf{x}_i$ to $\mathbf{x}'_i$.

Step 2: Constructing the epipolar line. Given the point $\mathbf{x}'$ the epipolar line $\mathbf{l}'$ passing through $\mathbf{x}'$ and the epipole $\mathbf{e}'$ can be written as $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}' = [\mathbf{e}']_\times \mathbf{x}'$ (the notation $[\mathbf{e}']_\times$ is defined in (A4.5–*p*581)). Since $\mathbf{x}'$ may be written as $\mathbf{x}' = \mathtt{H}_{\boldsymbol{\pi}}\mathbf{x}$, we have

$$\mathbf{l}' = [\mathbf{e}']_\times \mathtt{H}_{\boldsymbol{\pi}}\mathbf{x} = \mathtt{F}\mathbf{x}$$

where we define $\mathtt{F} = [\mathbf{e}']_\times \mathtt{H}_{\boldsymbol{\pi}}$, the fundamental matrix. This shows

Result 9.1. *The fundamental matrix $\mathtt{F}$ may be written as $\mathtt{F} = [\mathbf{e}']_\times \mathtt{H}_{\boldsymbol{\pi}}$, where $\mathtt{H}_{\boldsymbol{\pi}}$ is the transfer mapping from one image to another via any plane $\boldsymbol{\pi}$. Furthermore, since $[\mathbf{e}']_\times$ has rank 2 and $\mathtt{H}_{\boldsymbol{\pi}}$ rank 3, $\mathtt{F}$ is a matrix of rank 2.*

Geometrically, $\mathtt{F}$ represents a mapping from the 2-dimensional projective plane $\mathbb{P}^2$ of the first image to the pencil of epipolar lines through the epipole $\mathbf{e}'$. Thus, it represents a mapping from a 2-dimensional onto a 1-dimensional projective space, and hence must have rank 2.

Note, the geometric derivation above involves a scene plane $\boldsymbol{\pi}$, but a plane is *not* required in order for $\mathtt{F}$ to exist. The plane is simply used here as a means of defining a point map from one image to another. The connection between the fundamental matrix and transfer of points from one image to another via a plane is dealt with in some depth in chapter 13.

9.2.2 Algebraic derivation

The form of the fundamental matrix in terms of the two camera projection matrices, $\mathtt{P}, \mathtt{P}'$, may be derived algebraically. The following formulation is due to Xu and Zhang [Xu-96].

The ray back-projected from $\mathbf{x}$ by $\mathtt{P}$ is obtained by solving $\mathtt{P}\mathbf{X} = \mathbf{x}$. The one-parameter family of solutions is of the form given by (6.13–*p*162) as

$$\mathbf{X}(\lambda) = \mathtt{P}^{+}\mathbf{x} + \lambda\mathbf{C}$$

where $\mathtt{P}^{+}$ is the pseudo-inverse of $\mathtt{P}$, i.e. $\mathtt{P}\mathtt{P}^{+} = \mathtt{I}$, and $\mathbf{C}$ its null-vector, namely the camera centre, defined by $\mathtt{P}\mathbf{C} = \mathbf{0}$. The ray is parametrized by the scalar λ. In particular two points on the ray are $\mathtt{P}^{+}\mathbf{x}$ (at $\lambda = 0$), and the first camera centre $\mathbf{C}$ (at $\lambda = \infty$). These two points are imaged by the second camera $\mathtt{P}'$ at $\mathtt{P}'\mathtt{P}^{+}\mathbf{x}$ and $\mathtt{P}'\mathbf{C}$ respectively in the second view. The epipolar line is the line joining these two projected points, namely $\mathbf{l}' = (\mathtt{P}'\mathbf{C}) \times (\mathtt{P}'\mathtt{P}^{+}\mathbf{x})$. The point $\mathtt{P}'\mathbf{C}$ is the epipole in the second image, namely the projection of the first camera centre, and may be denoted by $\mathbf{e}'$. Thus, $\mathbf{l}' = [\mathbf{e}']_{\times}(\mathtt{P}'\mathtt{P}^{+})\mathbf{x} = \mathtt{F}\mathbf{x}$, where $\mathtt{F}$ is the matrix

$$\mathtt{F} = [\mathbf{e}']_{\times}\mathtt{P}'\mathtt{P}^{+}. \tag{9.1}$$

This is essentially the same formula for the fundamental matrix as the one derived in the previous section, the homography $\mathtt{H}_{\pi}$ having the explicit form $\mathtt{H}_{\pi} = \mathtt{P}'\mathtt{P}^{+}$ in terms of the two camera matrices. Note that this derivation breaks down in the case where the two camera centres are the same for, in this case, $\mathbf{C}$ is the common camera centre of both $\mathtt{P}$ and $\mathtt{P}'$, and so $\mathtt{P}'\mathbf{C} = \mathbf{0}$. It follows that $\mathtt{F}$ defined in (9.1) is the zero matrix.

Example 9.2. Suppose the camera matrices are those of a calibrated stereo rig with the world origin at the first camera

$$\mathtt{P} = \mathtt{K}[\mathtt{I} \mid \mathbf{0}] \qquad \mathtt{P}' = \mathtt{K}'[\mathtt{R} \mid \mathbf{t}].$$

Then

$$\mathtt{P}^{+} = \begin{bmatrix} \mathtt{K}^{-1} \\ \mathbf{0}^{\mathsf{T}} \end{bmatrix} \qquad \mathbf{C} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}$$

and

$$\begin{aligned} \mathtt{F} &= [\mathtt{P}'\mathbf{C}]_{\times}\mathtt{P}'\mathtt{P}^{+} \\ &= [\mathtt{K}'\mathbf{t}]_{\times}\mathtt{K}'\mathtt{R}\mathtt{K}^{-1} = \mathtt{K}'^{-\mathsf{T}}[\mathbf{t}]_{\times}\mathtt{R}\mathtt{K}^{-1} = \mathtt{K}'^{-\mathsf{T}}\mathtt{R}[\mathtt{R}^{\mathsf{T}}\mathbf{t}]_{\times}\mathtt{K}^{-1} = \mathtt{K}'^{-\mathsf{T}}\mathtt{R}\mathtt{K}^{\mathsf{T}}[\mathtt{K}\mathtt{R}^{\mathsf{T}}\mathbf{t}]_{\times} \end{aligned} \tag{9.2}$$

where the various forms follow from result A4.3(*p*582). Note that the epipoles (defined as the image of the other camera centre) are

$$\mathbf{e} = \mathtt{P}\begin{pmatrix} -\mathtt{R}^{\mathsf{T}}\mathbf{t} \\ 1 \end{pmatrix} = \mathtt{K}\mathtt{R}^{\mathsf{T}}\mathbf{t} \qquad \mathbf{e}' = \mathtt{P}'\begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} = \mathtt{K}'\mathbf{t}. \tag{9.3}$$

Thus we may write (9.2) as

$$\mathtt{F} = [\mathbf{e}']_{\times}\mathtt{K}'\mathtt{R}\mathtt{K}^{-1} = \mathtt{K}'^{-\mathsf{T}}[\mathbf{t}]_{\times}\mathtt{R}\mathtt{K}^{-1} = \mathtt{K}'^{-\mathsf{T}}\mathtt{R}[\mathtt{R}^{\mathsf{T}}\mathbf{t}]_{\times}\mathtt{K}^{-1} = \mathtt{K}'^{-\mathsf{T}}\mathtt{R}\mathtt{K}^{\mathsf{T}}[\mathbf{e}]_{\times}. \tag{9.4}$$

$\triangle$

The expression for the fundamental matrix can be derived in many ways, and indeed will be derived again several times in this book. In particular, (17.3–*p*412) expresses $\mathtt{F}$ in terms of 4×4 determinants composed from rows of the camera matrices for each view.

9.2.3 Correspondence condition

Up to this point we have considered the map $\mathbf{x} \rightarrow \mathbf{l}'$ defined by $\mathtt{F}$. We may now state the most basic properties of the fundamental matrix.

Result 9.3. *The fundamental matrix satisfies the condition that for any pair of corresponding points $\mathbf{x} \leftrightarrow \mathbf{x}'$ in the two images*

$$\mathbf{x}'^{\mathsf{T}}\mathtt{F}\mathbf{x} = 0.$$

This is true, because if points $\mathbf{x}$ and $\mathbf{x}'$ correspond, then $\mathbf{x}'$ lies on the epipolar line $\mathbf{l}' = \mathtt{F}\mathbf{x}$ corresponding to the point $\mathbf{x}$. In other words $0 = \mathbf{x}'^{\mathsf{T}}\mathbf{l}' = \mathbf{x}'^{\mathsf{T}}\mathtt{F}\mathbf{x}$. Conversely, if image points satisfy the relation $\mathbf{x}'^{\mathsf{T}}\mathtt{F}\mathbf{x} = 0$ then the rays defined by these points are coplanar. This is a necessary condition for points to correspond.

The importance of the relation of result 9.3 is that it gives a way of characterizing the fundamental matrix without reference to the camera matrices, i.e. only in terms of corresponding image points. This enables $\mathtt{F}$ to be computed from image correspondences alone. We have seen from (9.1) that $\mathtt{F}$ may be computed from the two camera matrices, $\mathtt{P}, \mathtt{P}'$, and in particular that $\mathtt{F}$ is determined uniquely from the cameras, up to an overall scaling. However, we may now enquire how many correspondences are required to compute $\mathtt{F}$ from $\mathbf{x}'^{\mathsf{T}}\mathtt{F}\mathbf{x} = 0$, and the circumstances under which the matrix is uniquely defined by these correspondences. The details of this are postponed until chapter 11, where it will be seen that in general at least 7 correspondences are required to compute $\mathtt{F}$.

9.2.4 Properties of the fundamental matrix

Definition 9.4. Suppose we have two images acquired by cameras with non-coincident centres, then the **fundamental matrix** $\mathtt{F}$ is the unique 3×3 rank 2 homogeneous matrix which satisfies

$$\mathbf{x}'^{\mathsf{T}}\mathtt{F}\mathbf{x} = 0 \tag{9.5}$$

for all corresponding points $\mathbf{x} \leftrightarrow \mathbf{x}'$.

We now briefly list a number of properties of the fundamental matrix. The most important properties are also summarized in table 9.1.

(i) **Transpose**: If $\mathtt{F}$ is the fundamental matrix of the pair of cameras $(\mathtt{P}, \mathtt{P}')$, then $\mathtt{F}^{\mathsf{T}}$ is the fundamental matrix of the pair in the opposite order: $(\mathtt{P}', \mathtt{P})$.

(ii) **Epipolar lines**: For any point $\mathbf{x}$ in the first image, the corresponding epipolar line is $\mathbf{l}' = \mathtt{F}\mathbf{x}$. Similarly, $\mathbf{l} = \mathtt{F}^{\mathsf{T}}\mathbf{x}'$ represents the epipolar line corresponding to $\mathbf{x}'$ in the second image.

(iii) The **epipole**: for any point $\mathbf{x}$ (other than $\mathbf{e}$) the epipolar line $\mathbf{l}' = \mathtt{F}\mathbf{x}$ contains the epipole $\mathbf{e}'$. Thus $\mathbf{e}'$ satisfies $\mathbf{e}'^{\mathsf{T}}(\mathtt{F}\mathbf{x}) = (\mathbf{e}'^{\mathsf{T}}\mathtt{F})\mathbf{x} = 0$ for all $\mathbf{x}$. It follows that $\mathbf{e}'^{\mathsf{T}}\mathtt{F} = \mathbf{0}$, i.e. $\mathbf{e}'$ is the left null-vector of $\mathtt{F}$. Similarly $\mathtt{F}\mathbf{e} = \mathbf{0}$, i.e. $\mathbf{e}$ is the right null-vector of $\mathtt{F}$.

- $\mathtt{F}$ is a rank 2 homogeneous matrix with 7 degrees of freedom.
- **Point correspondence**: If $\mathbf{x}$ and $\mathbf{x}'$ are corresponding image points, then $\mathbf{x}'^{\mathsf{T}}\mathtt{F}\mathbf{x} = 0$.
- **Epipolar lines**:
  - ◇ $\mathbf{l}' = \mathtt{F}\mathbf{x}$ is the epipolar line corresponding to $\mathbf{x}$.
  - ◇ $\mathbf{l} = \mathtt{F}^{\mathsf{T}}\mathbf{x}'$ is the epipolar line corresponding to $\mathbf{x}'$.
- **Epipoles**:
  - ◇ $\mathtt{F}\mathbf{e} = \mathbf{0}$.
  - ◇ $\mathtt{F}^{\mathsf{T}}\mathbf{e}' = \mathbf{0}$.
- **Computation from camera matrices** $\mathtt{P}, \mathtt{P}'$:
  - ◇ General cameras, $\mathtt{F} = [\mathbf{e}']_\times \mathtt{P}'\mathtt{P}^+$, where $\mathtt{P}^+$ is the pseudo-inverse of $\mathtt{P}$, and $\mathbf{e}' = \mathtt{P}'\mathbf{C}$, with $\mathtt{P}\mathbf{C} = \mathbf{0}$.
  - ◇ Canonical cameras, $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$, $\mathtt{P}' = [\mathtt{M} \mid \mathbf{m}]$, $\mathtt{F} = [\mathbf{e}']_\times \mathtt{M} = \mathtt{M}^{-\mathsf{T}}[\mathbf{e}]_\times$, where $\mathbf{e}' = \mathbf{m}$ and $\mathbf{e} = \mathtt{M}^{-1}\mathbf{m}$.
  - ◇ Cameras not at infinity $\mathtt{P} = \mathtt{K}[\mathtt{I} \mid \mathbf{0}]$, $\mathtt{P}' = \mathtt{K}'[\mathtt{R} \mid \mathbf{t}]$, $\mathtt{F} = \mathtt{K}'^{-\mathsf{T}}[\mathbf{t}]_\times \mathtt{R}\mathtt{K}^{-1} = [\mathtt{K}'\mathbf{t}]_\times \mathtt{K}'\mathtt{R}\mathtt{K}^{-1} = \mathtt{K}'^{-\mathsf{T}}\mathtt{R}\mathtt{K}^{\mathsf{T}}[\mathtt{K}\mathtt{R}^{\mathsf{T}}\mathbf{t}]_\times$.

Table 9.1. *Summary of fundamental matrix properties.*

(iv) $\mathtt{F}$ has seven degrees of freedom: a 3×3 homogeneous matrix has eight independent ratios (there are nine elements, and the common scaling is not significant); however, $\mathtt{F}$ also satisfies the constraint $\det \mathtt{F} = 0$ which removes one degree of freedom.

(v) $\mathtt{F}$ is a *correlation*, a projective map taking a point to a line (see definition 2.29-(*p*59)). In this case a point in the first image $\mathbf{x}$ defines a line in the second $\mathbf{l}' = \mathtt{F}\mathbf{x}$, which is the epipolar line of $\mathbf{x}$. If $\mathbf{l}$ and $\mathbf{l}'$ are corresponding epipolar lines (see figure 9.6a) then any point $\mathbf{x}$ on $\mathbf{l}$ is mapped to the same line $\mathbf{l}'$. This means there is no inverse mapping, and $\mathtt{F}$ is not of full rank. For this reason, $\mathtt{F}$ is not a proper correlation (which would be invertible).

9.2.5 The epipolar line homography

The set of epipolar lines in each of the images forms a pencil of lines passing through the epipole. Such a pencil of lines may be considered as a 1-dimensional projective space. It is clear from figure 9.6b that corresponding epipolar lines are perspectively related, so that there is a homography between the pencil of epipolar lines centred at $\mathbf{e}$ in the first view, and the pencil centred at $\mathbf{e}'$ in the second. A homography between two such 1-dimensional projective spaces has 3 degrees of freedom.

The degrees of freedom of the fundamental matrix can thus be counted as follows: 2 for $\mathbf{e}$, 2 for $\mathbf{e}'$, and 3 for the epipolar line homography which maps a line through $\mathbf{e}$ to a line through $\mathbf{e}'$. A geometric representation of this homography is given in section 9.4. Here we give an explicit formula for this mapping.

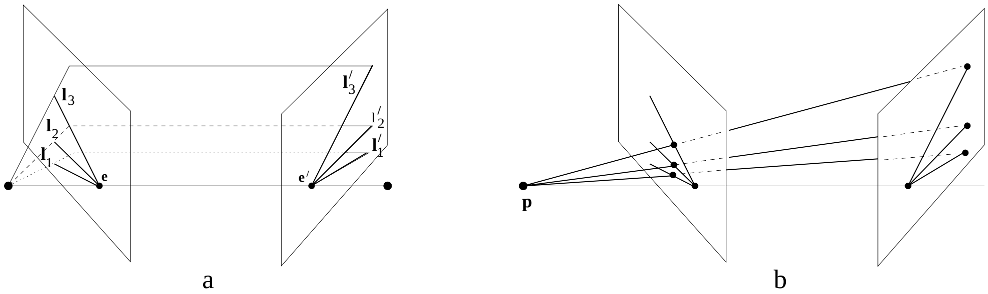


Fig. 9.6. **Epipolar line homography.** *(a) There is a pencil of epipolar lines in each image centred on the epipole. The correspondence between epipolar lines, $\mathbf{l}_i \leftrightarrow \mathbf{l}'_i$, is defined by the pencil of planes with axis the baseline. (b) The corresponding lines are related by a perspectivity with centre any point $\mathbf{p}$ on the baseline. It follows that the correspondence between epipolar lines in the pencils is a 1D homography.*

Result 9.5. *Suppose $\mathbf{l}$ and $\mathbf{l}'$ are corresponding epipolar lines, and $\mathbf{k}$ is any line not passing through the epipole $\mathbf{e}$, then $\mathbf{l}$ and $\mathbf{l}'$ are related by $\mathbf{l}' = \mathtt{F}[\mathbf{k}]_\times \mathbf{l}$. Symmetrically, $\mathbf{l} = \mathtt{F}^\mathsf{T}[\mathbf{k}']_\times \mathbf{l}'$.*

Proof. The expression $[\mathbf{k}]_\times \mathbf{l} = \mathbf{k} \times \mathbf{l}$ is the point of intersection of the two lines $\mathbf{k}$ and $\mathbf{l}$, and hence a point on the epipolar line $\mathbf{l}$ – call it $\mathbf{x}$. Hence, $\mathtt{F}[\mathbf{k}]_\times \mathbf{l} = \mathtt{F}\mathbf{x}$ is the epipolar line corresponding to the point $\mathbf{x}$, namely the line $\mathbf{l}'$. □

Furthermore a convenient choice for $\mathbf{k}$ is the line $\mathbf{e}$, since $\mathbf{k}^\mathsf{T}\mathbf{e} = \mathbf{e}^\mathsf{T}\mathbf{e} \neq 0$, so that the line $\mathbf{e}$ does not pass through the point $\mathbf{e}$ as is required. A similar argument holds for the choice of $\mathbf{k}' = \mathbf{e}'$. Thus the epipolar line homography may be written as

$$\mathbf{l}' = \mathtt{F}[\mathbf{e}]_\times \mathbf{l} \qquad \mathbf{l} = \mathtt{F}^\mathsf{T}[\mathbf{e}']_\times \mathbf{l}' \ .$$

9.3 Fundamental matrices arising from special motions

A special motion arises from a particular relationship between the translation direction, $\mathbf{t}$, and the direction of the rotation axis, $\mathbf{a}$. We will discuss two cases: *pure translation*, where there is no rotation; and *pure planar motion*, where $\mathbf{t}$ is orthogonal to $\mathbf{a}$ (the significance of the planar motion case is described in section 3.4.1(*p*77)). The 'pure' indicates that there is no change in the internal parameters. Such cases are important, firstly because they occur in practice, for example a camera viewing an object rotating on a turntable is equivalent to planar motion for pairs of views; and secondly because the fundamental matrix has a special form and thus additional properties.

9.3.1 Pure translation

In considering pure translations of the camera, one may consider the equivalent situation in which the camera is stationary, and the world undergoes a translation $-\mathbf{t}$. In this situation points in 3-space move on straight lines parallel to $\mathbf{t}$, and the imaged intersection of these parallel lines is the vanishing point $\mathbf{v}$ in the direction of $\mathbf{t}$. This is illustrated in figure 9.7 and figure 9.8. It is evident that $\mathbf{v}$ is the epipole for both views, and the imaged parallel lines are the epipolar lines. The algebraic details are given in the following example.