8

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 $\mathbf{x}'^{\top}\mathtt{F}\mathbf{x} = 0$.

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 9. 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 10.

8.1 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 8.1a 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

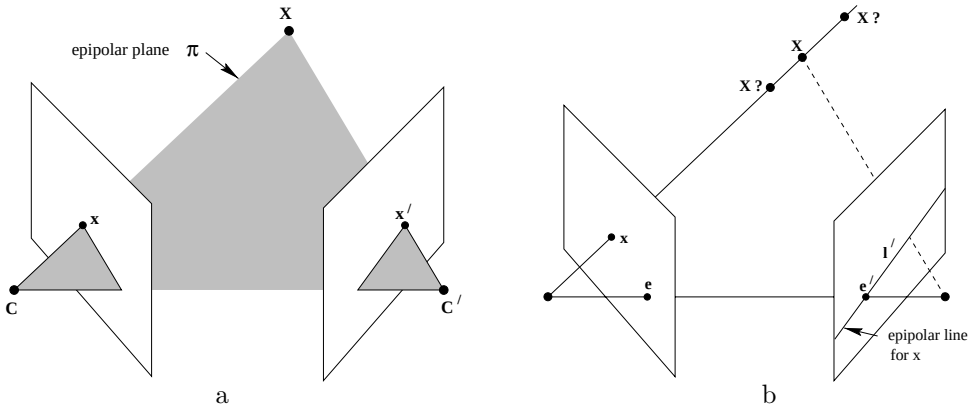


Fig. 8.1. **Point correspondence geometry.** *(a) The two cameras are indicated by their centres* $\mathbf{C}$ *and* $\mathbf{C}'$ *and image planes. The camera centres, 3-space point* $\mathbf{X}$*, and its images* $\mathbf{x}$ *and* $\mathbf{x}'$ *lie in a common plane* $\boldsymbol{\pi}$*. (b) An 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* $\mathbf{l}'$*.*

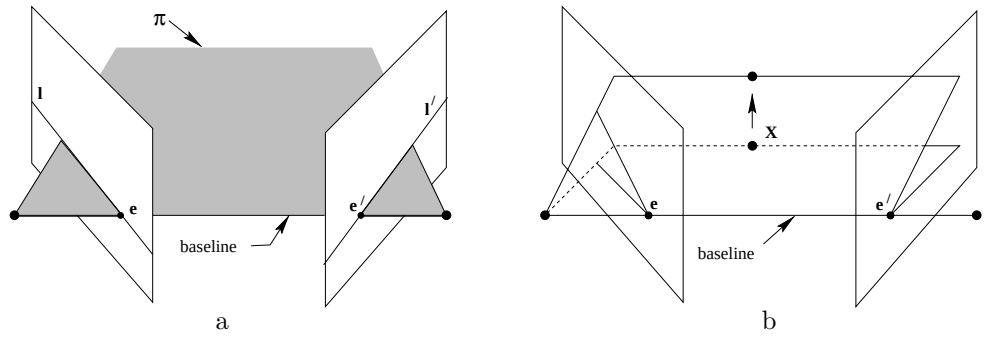


Fig. 8.2. **Epipolar geometry.** *(a) The camera baseline intersects each image plane at the epipoles* $\mathbf{e}$ *and* $\mathbf{e}'$*. Any plane* $\boldsymbol{\pi}$ *containing the baseline is an epipolar plane, and intersects the image planes in corresponding epipolar lines* $\mathbf{l}$ *and* $\mathbf{l}'$*. (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.*

$\mathbf{x}$ 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.

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

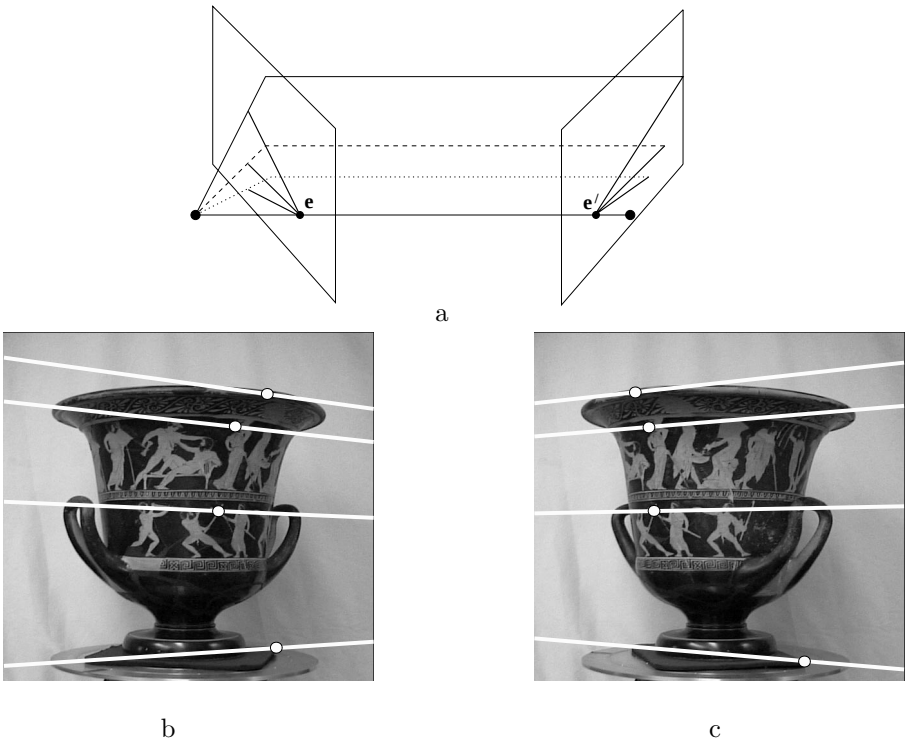


Fig. 8.3. **Converging cameras.** *(a) Epipolar geometry for converging cameras. (b) and (c) A pair of images with superimposed corresponding points and their epipolar lines (in white). The motion between the views is a translation and rotation. In each image, the direction of the other camera may be inferred from the intersection of the pencil of epipolar lines. In this case, both epipoles lie outside of the visible image.*

The geometric entities involved in epipolar geometry are illustrated in figure 8.2. 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 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 8.3 and figure 8.4. 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 10.6(*p*274).

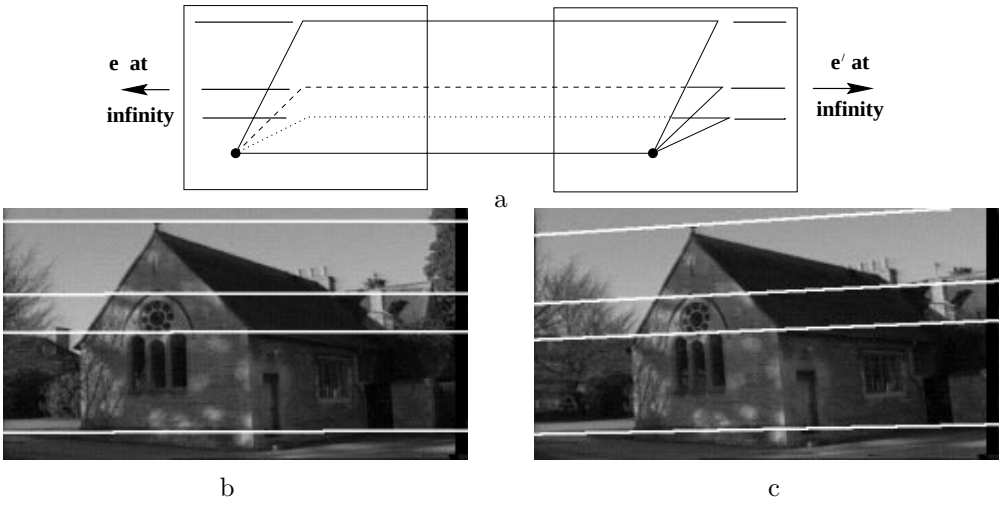


Fig. 8.4. **Motion parallel to the image plane.** *In the case of a special motion where the translation is parallel to the image plane, and the rotation axis is perpendicular to the image plane, the intersection of the baseline with the image plane is at infinity. Consequently the epipoles are at infinity, 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.*

8.2 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 8.1 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 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 F, the fundamental matrix.

8.2.1 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

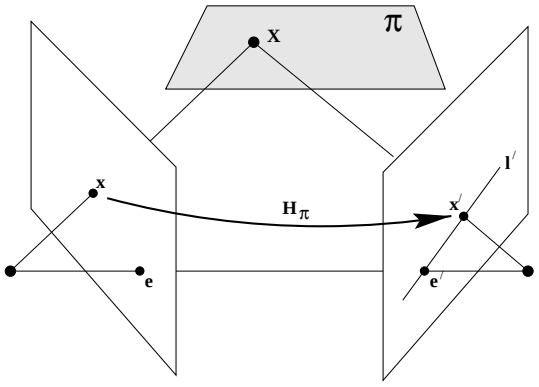


Fig. 8.5. *A point* $\mathbf{x}$ *in one image is transferred via the plane* $\boldsymbol{\pi}$ *to a matching point* $\mathbf{x}'$ *in the second image. The epipolar line through* $\mathbf{x}'$ *is obtained by joining* $\mathbf{x}'$ *to the epipole* $\mathbf{e}'$. *In symbols one may write* $\mathbf{x}' = \mathtt{H}_{\boldsymbol{\pi}}\mathbf{x}$ *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.*

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 8.5. 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 figure 8.1b. 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 (A3.4–p554)). 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 8.1. *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 12.

8.2.2 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 (5.13–p148) 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

$$\mathtt{F} = [\mathbf{e}']_{\times}\mathtt{P}'\mathtt{P}^{+}. \tag{8.1}$$

This is essentially the same formula for the fundamental matrix as the one derived in the previous section, the homography $\mathtt{H}_{\boldsymbol{\pi}}$ having the explicit form $\mathtt{H}_{\boldsymbol{\pi}} = \mathtt{P}'\mathtt{P}^{+}$ 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 (8.1) is the zero matrix.

Example 8.2. 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

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

and

$$\begin{aligned} \mathtt{F} &= [\mathtt{P}'\mathbf{C}]_\times \mathtt{P}'\mathtt{P}^+ \\ &= [\mathtt{K}'\mathbf{t}]_\times \mathtt{K}'\mathtt{R}\mathtt{K}^{-1} = \mathtt{K}'^{-\top}[\mathbf{t}]_\times \mathtt{R}\mathtt{K}^{-1} = \mathtt{K}'^{-\top}\mathtt{R}[\mathtt{R}^\top\mathbf{t}]_\times \mathtt{K}^{-1} = \mathtt{K}'^{-\top}\mathtt{R}\mathtt{K}^\top[\mathtt{K}\mathtt{R}^\top\mathbf{t}]_\times \end{aligned} \quad (8.2)$$

where the various forms follow from result A3.3(p555). Note that the epipoles (defined as the image of the other camera centre) are

$$\mathbf{e} = \mathtt{P}\begin{pmatrix} -\mathtt{R}^\top\mathbf{t} \\ 1 \end{pmatrix} = \mathtt{K}\mathtt{R}^\top\mathbf{t} \qquad \mathbf{e}' = \mathtt{P}'\begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} = \mathtt{K}'\mathbf{t}. \quad (8.3)$$

Thus we may write (8.2) as

$$\mathtt{F} = [\mathbf{e}']_\times \mathtt{K}'\mathtt{R}\mathtt{K}^{-1} = \mathtt{K}'^{-\top}[\mathbf{t}]_\times \mathtt{R}\mathtt{K}^{-1} = \mathtt{K}'^{-\top}\mathtt{R}[\mathtt{R}^\top\mathbf{t}]_\times \mathtt{K}^{-1} = \mathtt{K}'^{-\top}\mathtt{R}\mathtt{K}^\top[\mathbf{e}]_\times. \quad (8.4)$$

△

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, (16.3–p400) expresses F in terms of 4×4 determinants composed from rows of the camera matrices for each view.

8.2.3 Correspondence condition

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

Result 8.3. *The fundamental matrix satisfies the condition that for any pair of corresponding points $\mathbf{x} \leftrightarrow \mathbf{x}'$ in the two images*

$$\mathbf{x}'^\top \mathtt{F}\mathbf{x} = 0.$$

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 $0 = \mathbf{x}'^\top\mathbf{l}' = \mathbf{x}'^\top\mathtt{F}\mathbf{x}$. Conversely, if image points satisfy the relation $\mathbf{x}'^\top\mathtt{F}\mathbf{x} = 0$ 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 8.3 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 F to be computed from image correspondences alone. We have seen from (8.1) that F may be computed from the two camera matrices, $\mathtt{P}, \mathtt{P}'$, and in particular that F is determined uniquely from the cameras, up to an overall scaling. However, we may now enquire how many correspondences are required to compute F from $\mathbf{x}'^\top\mathtt{F}\mathbf{x} = 0$, and the circumstances under which the matrix is uniquely defined by these correspondences. The details of this are postponed until chapter 10, where it will be seen that in general at least 7 correspondences are required to compute 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 $\mathbf{x}'^{\top}\mathtt{F}\mathbf{x} = 0$.
- **Epipolar lines**:
  - $\diamond$ $\mathbf{l}' = \mathtt{F}\mathbf{x}$ is the epipolar line corresponding to $\mathbf{x}$.
  - $\diamond$ $\mathbf{l} = \mathtt{F}^{\top}\mathbf{x}'$ is the epipolar line corresponding to $\mathbf{x}'$.
- **Epipoles**:
  - $\diamond$ $\mathtt{F}\mathbf{e} = \mathbf{0}$.
  - $\diamond$ $\mathtt{F}^{\top}\mathbf{e}' = \mathbf{0}$.
- **Computation from camera matrices** $\mathtt{P}, \mathtt{P}'$:
  - $\diamond$ 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 $\mathtt{P}\mathbf{C} = \mathbf{0}$.
  - $\diamond$ Canonical cameras, $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$, $\mathtt{P}' = [\mathtt{M} \mid \mathbf{m}]$, $\mathtt{F} = [\mathbf{e}']_{\times}\mathtt{M} = \mathtt{M}^{-\top}[\mathbf{e}]_{\times}$, where $\mathbf{e}' = \mathbf{m}$ and $\mathbf{e} = \mathtt{M}^{-1}\mathbf{m}$.
  - $\diamond$ Cameras not at infinity $\mathtt{P} = \mathtt{K}[\mathtt{I} \mid \mathbf{0}]$, $\mathtt{P}' = \mathtt{K}'[\mathtt{R} \mid \mathbf{t}]$, $\mathtt{F} = \mathtt{K}'^{-\top}[\mathbf{t}]_{\times}\mathtt{R}\mathtt{K}^{-1} = [\mathtt{K}'\mathbf{t}]_{\times}\mathtt{K}'\mathtt{R}\mathtt{K}^{-1} = \mathtt{K}'^{-\top}\mathtt{R}\mathtt{K}^{\top}[\mathtt{K}\mathtt{R}^{\top}\mathbf{t}]_{\times}$.

Table 8.1. *Summary of fundamental matrix properties.*

8.2.4 Properties of the fundamental matrix

Definition 8.4. 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

$$\mathbf{x}'^{\top}\mathtt{F}\mathbf{x} = 0 \tag{8.5}$$

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 8.1.

(i) **Transpose**: If $\mathtt{F}$ is the fundamental matrix of the pair of cameras $(\mathtt{P}, \mathtt{P}')$, then $\mathtt{F}^{\top}$ 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, $\mathbf{l} = \mathtt{F}^{\top}\mathbf{x}'$ 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 $\mathbf{e}'^{\top}(\mathtt{F}\mathbf{x}) = (\mathbf{e}'^{\top}\mathtt{F})\mathbf{x} = 0$ for all $\mathbf{x}$. It follows that $\mathbf{e}'^{\top}\mathtt{F} = \mathbf{0}$, i.e. $\mathbf{e}'$ is the left null-space of $\mathtt{F}$. Similarly $\mathtt{F}\mathbf{e} = \mathbf{0}$, i.e. $\mathbf{e}$ is the right null-space of $\mathtt{F}$.

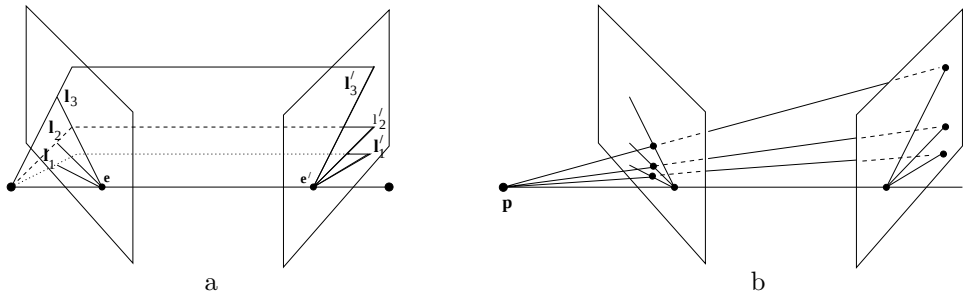


Fig. 8.6. **Epipolar line homography.** *(a) There is a pencil of epipolar lines in each image centred on the epipole. The correspondence between epipolar lines, $\mathbf{l}_i \leftrightarrow \mathbf{l}'_i$, is defined by the pencil of planes with axis the baseline. (b) The corresponding lines are related 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.*

(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 1.28(p39)). 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 8.6a) 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).

8.2.5 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 8.6b 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 8.4. Here we give an explicit formula for this mapping.

Result 8.5. *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,* $\mathbf{l} = \mathtt{F}^\top [\mathbf{k}']_\times \mathbf{l}'$*.*

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 $\mathbf{l}'$. □

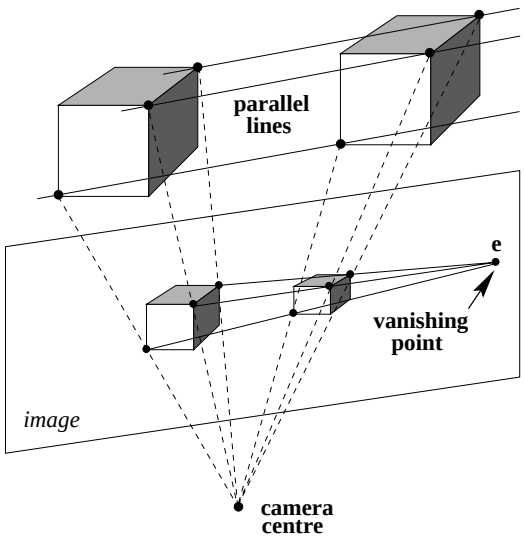


Fig. 8.7. *Under a pure translational camera motion, 3D points appear to slide along parallel rails. The images of these parallel lines intersect in a vanishing point corresponding to the translation direction. The epipole* $\mathbf{e}$ *is the vanishing point.*

Furthermore a convenient choice for $\mathbf{k}$ is the line $\mathbf{e}$, since $\mathbf{k}^\top\mathbf{e} = \mathbf{e}^\top\mathbf{e} \neq 0$, 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

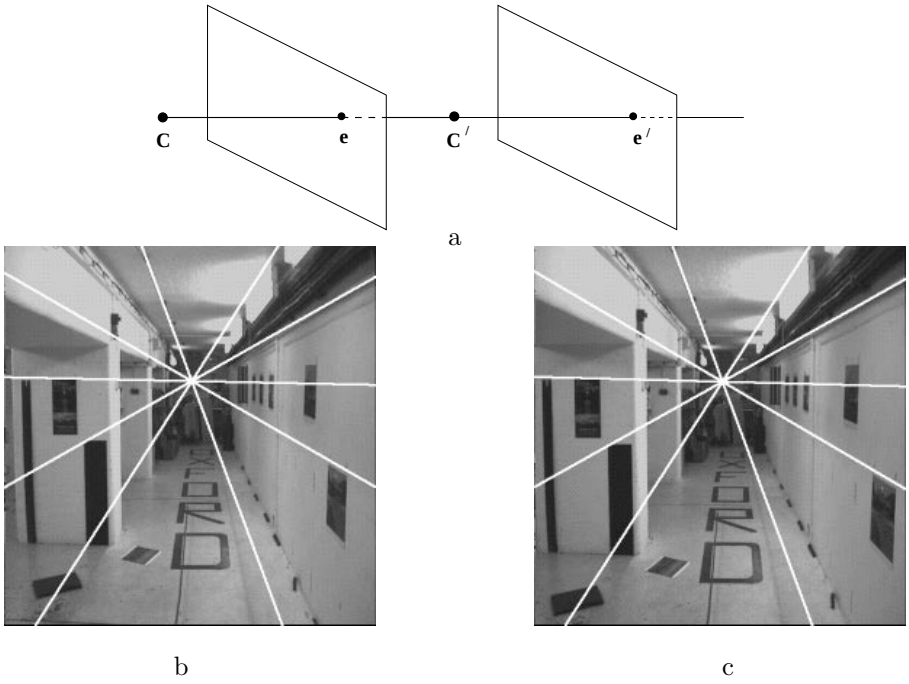
$$\mathbf{l}' = \mathtt{F}[\mathbf{e}]_\times\mathbf{l} \qquad \mathbf{l} = \mathtt{F}^\top[\mathbf{e}']_\times\mathbf{l}' \ .$$

8.3 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 2.4.1(*p*58)). 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.

8.3.1 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



a b c

Fig. 8.8. **Pure translational motion.** *(a) under the motion the epipole is a fixed point, i.e. has the same coordinates in both images, and points appear to move along lines radiating from the epipole. The epipole in this case is termed the* Focus of Expansion *(FOE). (b) and (c) the same epipolar lines are overlaid in both cases. Note the motion of the posters on the wall which slide along the epipolar line.*

of $\mathbf{t}$. This is illustrated in figure 8.7 and figure 8.8. 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.

Example 8.6. Suppose the motion of the cameras is a pure translation with no rotation and no change in the internal parameters. One may assume that the two cameras are $\mathtt{P} = \mathtt{K}[\mathtt{I} \mid \mathbf{0}]$ and $\mathtt{P}' = \mathtt{K}[\mathtt{I} \mid \mathbf{t}]$. Then from (8.4) (using $\mathtt{R} = \mathtt{I}$ and $\mathtt{K} = \mathtt{K}'$)

$$\mathtt{F} = [\mathbf{e}']_\times \mathtt{K}\mathtt{K}^{-1} = [\mathbf{e}']_\times.$$

If the camera translation is parallel to the x-axis, then $\mathbf{e}' = (1, 0, 0)^\top$, so

$$\mathtt{F} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

The relation between corresponding points, $\mathbf{x}'^\top \mathtt{F}\mathbf{x} = 0$, reduces to $y = y'$, i.e. the epipolar lines are corresponding rasters. This is the situation that is sought by image rectification described in section 10.12(p289). △

Indeed if the image point $\mathbf{x}$ is normalized as $\mathbf{x} = (x, y, 1)^\top$, then from $\mathbf{x} = \mathtt{P}\mathbf{X} = \mathtt{K}[\mathtt{I} \mid \mathbf{0}]\mathbf{X}$, the space point's (inhomogeneous) coordinates are $(\mathrm{X}, \mathrm{Y}, \mathrm{Z})^\top =$

$\mathtt{z}\mathtt{K}^{-1}\mathbf{x}$, where $\mathtt{z}$ is the depth of the point $\mathbf{X}$ (the distance of $\mathbf{X}$ from the camera centre measured along the principal axis of the first camera). It then follows from $\mathbf{x}' = \mathtt{P}'\mathbf{X} = \mathtt{K}[\mathtt{I} \mid \mathbf{t}]\mathbf{X}$ that the mapping from an image point $\mathbf{x}$ to an image point $\mathbf{x}'$ is

$$\mathbf{x}' = \mathbf{x} + \mathtt{K}\mathbf{t}/\mathtt{z}. \tag{8.6}$$

The motion $\mathbf{x}' = \mathbf{x} + \mathtt{K}\mathbf{t}/\mathtt{z}$ of (8.6) shows that the image point "starts" at $\mathbf{x}$ and then moves along the line defined by $\mathbf{x}$ and the epipole $\mathbf{e} = \mathbf{e}' = \mathbf{v}$. The extent of the motion depends on the magnitude of the translation $\mathbf{t}$ (which is not a homogeneous vector here) and the inverse depth $\mathtt{z}$, so that points closer to the camera appear to move faster than those further away – a common experience when looking out of a train window.

Note that in this case of pure translation $\mathtt{F} = [\mathbf{e}']_\times$ is skew-symmetric and has only 2 degrees of freedom, which correspond to the position of the epipole. The epipolar line of $\mathbf{x}$ is $\mathbf{l}' = \mathtt{F}\mathbf{x} = [\mathbf{e}]_\times\mathbf{x}$, and $\mathbf{x}$ lies on this line since $\mathbf{x}^\top[\mathbf{e}]_\times\mathbf{x} = 0$, i.e. $\mathbf{x}$, $\mathbf{x}'$ and $\mathbf{e} = \mathbf{e}'$ are collinear (assuming both images are overlaid on top of each other). This collinearity property is termed *auto-epipolar*, and does not hold for general motion.

General motion. The pure translation case gives additional insight into the general motion case. Given two arbitrary cameras, we may rotate the camera used for the first image so that it is aligned with the second camera. This rotation may be simulated by applying a projective transformation to the first image. A further correction may be applied to the first image to account for any difference in the calibration matrices of the two images. The result of these two corrections is a projective transformation $\mathtt{H}$ of the first image. If one assumes these corrections to have been made, then the effective relationship of the two cameras to each other is that of a pure translation. Consequently, the fundamental matrix corresponding to the corrected first image and the second image is of the form $\hat{\mathtt{F}} = [\mathbf{e}']_\times$, satisfying $\mathbf{x}'^\top\hat{\mathtt{F}}\hat{\mathbf{x}} = 0$, where $\hat{\mathbf{x}} = \mathtt{H}\mathbf{x}$ is the corrected point in the first image. From this one deduces that $\mathbf{x}'^\top[\mathbf{e}']_\times\mathtt{H}\mathbf{x} = 0$, and so the fundamental matrix corresponding to the initial point correspondences $\mathbf{x} \leftrightarrow \mathbf{x}'$ is $\mathtt{F} = [\mathbf{e}']_\times\mathtt{H}$. This is illustrated in figure 8.9.

Example 8.7. Continuing from example 8.2, assume again that the two cameras are $\mathtt{P} = \mathtt{K}[\mathtt{I} \mid \mathbf{0}]$ and $\mathtt{P}' = \mathtt{K}'[\mathtt{R} \mid \mathbf{t}]$. Then as described in section 7.4.2(p194) the requisite projective transformation is $\mathtt{H} = \mathtt{K}'\mathtt{R}\mathtt{K}^{-1} = \mathtt{H}_\infty$, where $\mathtt{H}_\infty$ is the infinite homography (see section 12.4(p327)), and $\mathtt{F} = [\mathbf{e}']_\times\mathtt{H}_\infty$.

If the image point $\mathbf{x}$ is normalized as $\mathbf{x} = (x, y, 1)^\top$, as in example 8.6, then $(\mathtt{X}, \mathtt{Y}, \mathtt{Z})^\top = \mathtt{z}\mathtt{K}^{-1}\mathbf{x}$, and from $\mathbf{x} = \mathtt{P}'\mathbf{X} = \mathtt{K}'[\mathtt{R} \mid \mathbf{t}]\mathbf{X}$ the mapping from an image point $\mathbf{x}$ to an image point $\mathbf{x}'$ is

$$\mathbf{x}' = \mathtt{K}'\mathtt{R}\mathtt{K}^{-1}\mathbf{x} + \mathtt{K}'\mathbf{t}/\mathtt{z}. \tag{8.7}$$

The mapping is in two parts: the first term depends on the image position alone,

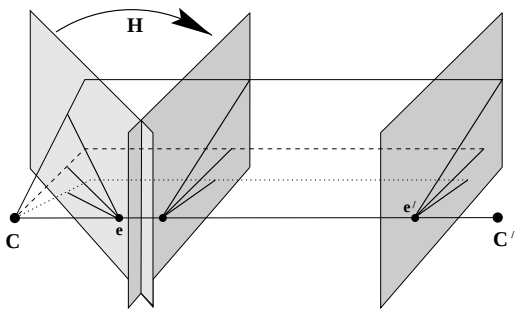


Fig. 8.9. **General camera motion.** *The first camera (on the left) may be rotated and corrected to simulate a pure translational motion. The fundamental matrix for the original pair is the product* $\mathtt{F} = [\mathbf{e}']_\times \mathtt{H}$*, where* $[\mathbf{e}']_\times$ *is the fundamental matrix of the translation, and* $\mathtt{H}$ *is the projective transformation corresponding to the correction of the first camera.*

i.e. $\mathbf{x}$, but not the point's depth $\mathtt{z}$, and takes account of the camera rotation and change of internal parameters; the second term depends on the depth, but not on the image position $\mathbf{x}$, and takes account of camera translation. In the case of pure translation ($\mathtt{R} = \mathtt{I}, \mathtt{K} = \mathtt{K}'$) (8.7) reduces to (8.6). △

8.3.2 Pure planar motion

In this case the rotation axis is orthogonal to the translation direction. Orthogonality imposes one constraint on the motion, and it is shown in the exercises at the end of this chapter that if $\mathtt{K}' = \mathtt{K}$ then $\mathtt{F_s}$, the symmetric part of $\mathtt{F}$, has rank 2 in this planar motion case (note, for a general motion the symmetric part of $\mathtt{F}$ has full rank). Thus, the condition that $\det \mathtt{F_s} = 0$ is an additional constraint on $\mathtt{F}$ and reduces the number of degrees of freedom from 7, for a general motion, to 6 degrees of freedom for a pure planar motion.

8.4 Geometric representation of the fundamental matrix

This section is not essential for a first reading and the reader may optionally skip to section 8.5.

In this section the fundamental matrix is decomposed into its symmetric and asymmetric parts, and each part is given a geometric representation. The symmetric and asymmetric parts of the fundamental matrix are

$$\mathtt{F_s} = \left(\mathtt{F} + \mathtt{F}^\top\right)/2 \qquad \mathtt{F_a} = \left(\mathtt{F} - \mathtt{F}^\top\right)/2$$

so that $\mathtt{F} = \mathtt{F_s} + \mathtt{F_a}$.

To motivate the decomposition, consider the points $\mathbf{X}$ in 3-space that map to the same point in two images. These image points are fixed under the camera motion so that $\mathbf{x} = \mathbf{x}'$. Clearly such points are corresponding and thus satisfy $\mathbf{x}^\top \mathtt{F} \mathbf{x} = 0$, which is a necessary condition on corresponding points. Now, for any

skew-symmetric matrix $\mathtt{A}$ the form $\mathbf{x}^\top\mathtt{A}\mathbf{x}$ is identically zero. Consequently only the symmetric part of $\mathtt{F}$ contributes to $\mathbf{x}^\top\mathtt{F}\mathbf{x} = 0$, which then reduces to $\mathbf{x}^\top\mathtt{F}_\mathtt{S}\mathbf{x} = 0$. As will be seen below the matrix $\mathtt{F}_\mathtt{S}$ may be thought of as a conic in the image plane.

Geometrically the conic arises as follows. The locus of all points in 3-space for which $\mathbf{x} = \mathbf{x}'$ is known as the *horopter* curve. Generally this is a twisted cubic curve in 3-space (see section 2.3(p57)) passing through the two camera centres [Maybank-93]. The image of the horopter is the conic defined by $\mathtt{F}_\mathtt{S}$. We return to the horopter in chapter 21.

Symmetric part. The matrix $\mathtt{F}_\mathtt{S}$ is symmetric and is of rank 3 in general. It has 5 degrees of freedom and is identified with a point conic, called the *Steiner conic* (the name is explained below). The epipoles $\mathbf{e}$ and $\mathbf{e}'$ lie on the conic $\mathtt{F}_\mathtt{S}$. To see that the epipoles lie on the conic, i.e. that $\mathbf{e}^\top\mathtt{F}_\mathtt{S}\mathbf{e} = 0$, start from $\mathtt{F}\mathbf{e} = \mathbf{0}$. Then $\mathbf{e}^\top\mathtt{F}\mathbf{e} = \mathbf{0}$ and so $\mathbf{e}^\top\mathtt{F}_\mathtt{S}\mathbf{e} + \mathbf{e}^\top\mathtt{F}_\mathtt{A}\mathbf{e} = 0$. However, $\mathbf{e}^\top\mathtt{F}_\mathtt{A}\mathbf{e} = 0$, since for any anti-symmetric matrix $\mathtt{S}$, $\mathbf{x}^\top\mathtt{S}\mathbf{x} = 0$. Thus $\mathbf{e}^\top\mathtt{F}_\mathtt{S}\mathbf{e} = 0$. The derivation for $\mathbf{e}'$ follows in a similar manner.

Anti-symmetric part. The matrix $\mathtt{F}_\mathtt{A}$ is skew-symmetric and may be written as $\mathtt{F}_\mathtt{A} = [\mathbf{x}_\mathtt{A}]_\times$, where $\mathbf{x}_\mathtt{A}$ is the null-vector of $\mathtt{F}_\mathtt{A}$. The skew-symmetric part has 2 degrees of freedom and is identified with the point $\mathbf{x}_\mathtt{A}$.

The relation between the point $\mathbf{x}_\mathtt{A}$ and conic $\mathtt{F}_\mathtt{S}$ is shown in figure 8.10a. The polar of $\mathbf{x}_\mathtt{A}$ intersects the Steiner conic $\mathtt{F}_\mathtt{S}$ at the epipoles $\mathbf{e}$ and $\mathbf{e}'$ (the pole–polar relation is described in section 1.2.3(p8)). The proof of this result is left as an exercise.

Epipolar line correspondence. It is a classical theorem of projective geometry due to Steiner [Semple-79] that for two line pencils related by a homography, the locus of intersections of corresponding lines is a conic. This is precisely the situation here. The pencils are the epipolar pencils, one through $\mathbf{e}$ and the other through $\mathbf{e}'$. The epipolar lines are related by a 1D homography as described in section 8.2.5. The locus of intersection is the conic $\mathtt{F}_\mathtt{S}$.

The conic and epipoles enable epipolar lines to be determined by a geometric construction as illustrated in figure 8.10b. This construction is based on the fixed point property of the Steiner conic $\mathtt{F}_\mathtt{S}$. The epipolar line $\mathbf{l} = \mathbf{x} \times \mathbf{e}$ in the first view defines an epipolar plane in 3-space which intersects the horopter in a point, which we will call $\mathbf{X}_\mathrm{c}$. The point $\mathbf{X}_\mathrm{c}$ is imaged in the first view at $\mathbf{x}_\mathrm{c}$, which is the point at which $\mathbf{l}$ intersects the conic $\mathtt{F}_\mathtt{S}$ (since $\mathtt{F}_\mathtt{S}$ is the image of the horopter). Now the image of $\mathbf{X}_\mathrm{c}$ is also $\mathbf{x}_\mathrm{c}$ in the second view due to the fixed-point property of the horopter. So $\mathbf{x}_\mathrm{c}$ is the image in the second view of a point on the epipolar plane of $\mathbf{x}$. It follows that $\mathbf{x}_\mathrm{c}$ lies on the epipolar line $\mathbf{l}'$ of $\mathbf{x}$, and consequently $\mathbf{l}'$ may be computed as $\mathbf{l}' = \mathbf{x}_\mathrm{c} \times \mathbf{e}'$.

The conic together with two points on the conic account for the 7 degrees of

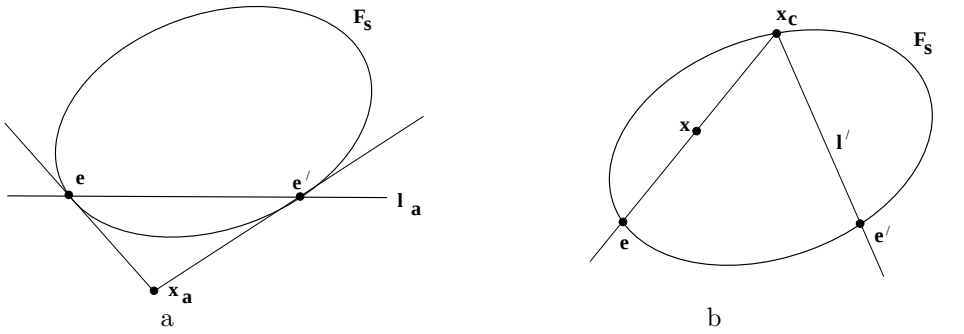


Fig. 8.10. **Geometric representation of $\mathtt{F}$.** *(a) The conic $\mathtt{F_S}$ represents the symmetric part of $\mathtt{F}$, and the point $\mathbf{x_a}$ the skew-symmetric part. The conic $\mathtt{F_S}$ is the locus of intersection of corresponding epipolar lines, assuming both images are overlaid on top of each other. It is the image of the horopter curve. The line $\mathbf{l_a}$ is the polar of $\mathbf{x_a}$ with respect to the conic $\mathtt{F_S}$. It intersects the conic at the epipoles $\mathbf{e}$ and $\mathbf{e}'$. (b) The epipolar line $\mathbf{l}'$ corresponding to a point $\mathbf{x}$ is constructed as follows: intersect the line defined by the points $\mathbf{e}$ and $\mathbf{x}$ with the conic. This intersection point is $\mathbf{x}_c$. Then $\mathbf{l}'$ is the line defined by the points $\mathbf{x}_c$ and $\mathbf{e}'$.*

freedom of $\mathtt{F}$: 5 degrees of freedom for the conic and one each to specify the two epipoles on the conic. Given $\mathtt{F}$, then the conic $\mathtt{F_S}$, epipoles $\mathbf{e}, \mathbf{e}'$ and skew-symmetric point $\mathbf{x_a}$ are defined uniquely. However, $\mathtt{F_S}$ and $\mathbf{x_a}$ do not uniquely determine $\mathtt{F}$ since the identity of the epipoles is not recovered, i.e. the polar of $\mathbf{x_a}$ determines the epipoles but does not determine which one is $\mathbf{e}$ and which one $\mathbf{e}'$.

Pure planar motion. We return to the case of planar motion discussed above in section 8.3.2, where $\mathtt{F_S}$ has rank 2. It is evident that in this case the Steiner conic is degenerate and from section 1.2.3(p8) is equivalent to two non-coincident lines:

$$\mathtt{F_S} = \mathbf{l}_h \mathbf{l}_s^\top + \mathbf{l}_s \mathbf{l}_h^\top$$

as depicted in figure 8.11a. The geometric construction of the epipolar line $\mathbf{l}'$ corresponding to a point $\mathbf{x}$ of section 8.4 has a simple algebraic representation in this case. As in the general motion case, there are three steps, illustrated in figure 8.11b: first the line $\mathbf{l} = \mathbf{e} \times \mathbf{x}$ joining $\mathbf{e}$ and $\mathbf{x}$ is computed; second, its intersection point with the "conic" $\mathbf{x}_c = \mathbf{l}_s \times \mathbf{l}$ is determined; third the epipolar line $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}_c$ is the join of $\mathbf{x}_c$ and $\mathbf{e}'$. Putting these steps together we find

$$\mathbf{l}' = \mathbf{e}' \times [\mathbf{l}_s \times (\mathbf{e} \times \mathbf{x})] = [\mathbf{e}']_\times [\mathbf{l}_s]_\times [\mathbf{e}]_\times \mathbf{x}.$$

It follows that $\mathtt{F}$ may be written as

$$\mathtt{F} = [\mathbf{e}']_\times [\mathbf{l}_s]_\times [\mathbf{e}]_\times. \tag{8.8}$$

The 6 degrees of freedom of $\mathtt{F}$ are accounted for as 2 degrees of freedom for each of the two epipoles and 2 degrees of freedom for the line.

The geometry of this situation can be easily visualized: the horopter for this motion is a degenerate twisted cubic consisting of a circle in the plane of the motion

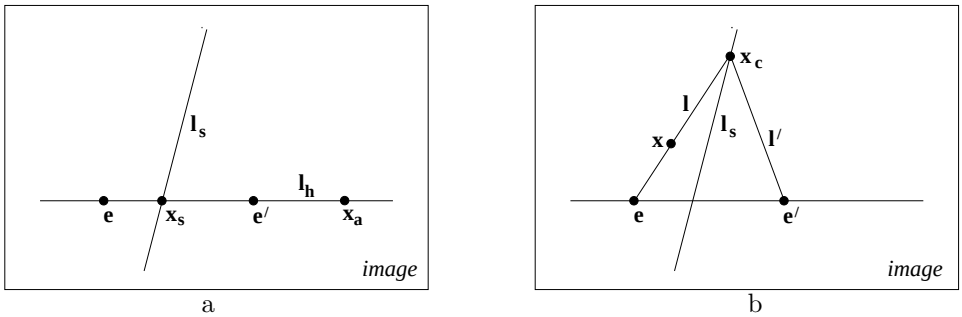


Fig. 8.11. **Geometric representation of $\mathtt{F}$ for planar motion.** *(a) The lines $\mathbf{l}_s$ and $\mathbf{l}_h$ constitute the Steiner conic for this motion, which is degenerate. Compare this figure with the conic for general motion shown in figure 8.10. (b) The epipolar line $\mathbf{l}'$ corresponding to a point $\mathbf{x}$ is constructed as follows: intersect the line defined by the points $\mathbf{e}$ and $\mathbf{x}$ with the (conic) line $\mathbf{l}_s$. This intersection point is $\mathbf{x}_c$. Then $\mathbf{l}'$ is the line defined by the points $\mathbf{x}_c$ and $\mathbf{e}'$.*

(the plane orthogonal to the rotation axis and containing the camera centres), and a line parallel to the rotation axis and intersecting the circle. The line is the screw axis (see section 2.4.1(p58)). The motion is equivalent to a rotation about the screw axis with zero translation. Under this motion points on the screw axis are fixed, and consequently their images are fixed. The line $\mathbf{l}_s$ is the image of the screw axis. The line $\mathbf{l}_h$ is the intersection of the image with the plane of the motion. This geometry is used for auto-calibration in chapter 18.

8.5 Retrieving the camera matrices

To this point we have examined the properties of $\mathtt{F}$ and of image relations for a point correspondence $\mathbf{x} \leftrightarrow \mathbf{x}'$. We now turn to one of the most significant properties of $\mathtt{F}$, that the matrix may be used to determine the camera matrices of the two views.

8.5.1 Projective invariance and canonical cameras

It is evident from the derivations of section 8.2 that the map $\mathbf{l}' = \mathtt{F}\mathbf{x}$ and the correspondence condition $\mathbf{x}'^{\top}\mathtt{F}\mathbf{x} = 0$ are *projective* relationships: the derivations have involved only projective geometric relationships, such as the intersection of lines and planes, and in the algebraic development only the linear mapping of the projective camera between world and image points. Consequently, the relationships depend only on projective coordinates in the image, and not, for example on Euclidean measurements such as the angle between rays. In other words the image relationships are projectively invariant: under a projective transformation of the image coordinates $\hat{\mathbf{x}} = \mathtt{H}\mathbf{x}, \hat{\mathbf{x}}' = \mathtt{H}'\mathbf{x}'$, there is a corresponding map $\hat{\mathbf{l}}' = \widehat{\mathtt{F}}\hat{\mathbf{x}}$ with $\widehat{\mathtt{F}} = \mathtt{H}'^{-\top}\mathtt{F}\mathtt{H}^{-1}$ the corresponding rank 2 fundamental matrix.

Similarly, $\mathtt{F}$ only depends on projective properties of the cameras $\mathtt{P}, \mathtt{P}'$. The

camera matrix relates 3-space measurements to image measurements and so depends on both the image coordinate frame and the choice of world coordinate frame. $\mathtt{F}$ does not depend on the choice of world frame, for example a rotation of world coordinates changes $\mathtt{P}, \mathtt{P}'$, but not $\mathtt{F}$. In fact, the fundamental matrix is unchanged by a projective transformation of 3-space. More precisely,

Result 8.8. *If* $\mathtt{H}$ *is a* 4×4 *matrix representing a projective transformation of 3-space, then the fundamental matrices corresponding to the pairs of camera matrices* $(\mathtt{P}, \mathtt{P}')$ *and* $(\mathtt{PH}, \mathtt{P}'\mathtt{H})$ *are the same.*

Proof Observe that $\mathtt{P}\mathbf{X} = (\mathtt{PH})(\mathtt{H}^{-1}\mathbf{X})$, and similarly for $\mathtt{P}'$. Thus if $\mathbf{x} \leftrightarrow \mathbf{x}'$ are matched points with respect to the pair of cameras $(\mathtt{P}, \mathtt{P}')$, corresponding to a 3D point $\mathbf{X}$, then they are also matched points with respect to the pair of cameras $(\mathtt{PH}, \mathtt{P}'\mathtt{H})$, corresponding to the point $\mathtt{H}^{-1}\mathbf{X}$. □

Thus, although from (8.1–*p*224) a pair of camera matrices $(\mathtt{P}, \mathtt{P}')$ uniquely determine a fundamental matrix $\mathtt{F}$, the converse is not true. The fundamental matrix determines the pair of camera matrices at best up to right-multiplication by a 3D projective transformation. It will be seen below that this is the full extent of the ambiguity, and indeed the camera matrices are determined up to a projective transformation by the fundamental matrix.

Canonical form of camera matrices. Given this ambiguity, it is common to define a specific *canonical form* for the pair of camera matrices corresponding to a given fundamental matrix in which the first matrix is of the simple form $[\mathtt{I} \mid \mathbf{0}]$, where $\mathtt{I}$ is the 3×3 identity matrix and $\mathbf{0}$ a null 3-vector. To see that this is always possible, let $\mathtt{P}$ be augmented by one row to make a 4×4 non-singular matrix, denoted $\mathtt{P}^*$. Now letting $\mathtt{H} = \mathtt{P}^{*-1}$, one verifies that $\mathtt{PH} = [\mathtt{I} \mid \mathbf{0}]$ as desired.

The following result is very frequently used

Result 8.9. *The fundamental matrix corresponding to a pair of camera matrices* $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$ *and* $\mathtt{P}' = [\mathtt{M} \mid \mathbf{m}]$ *is equal to* $[\mathbf{m}]_\times \mathtt{M}$.

This is easily derived as a special case of (8.1–*p*224).

8.5.2 Projective ambiguity of cameras given F

It has been seen that a pair of camera matrices determines a unique fundamental matrix. This mapping is not injective (one-to-one) however, since pairs of camera matrices that differ by a projective transformation give rise to the same fundamental matrix. It will now be shown that this is the only ambiguity. We will show that a given fundamental matrix determines the pair of camera matrices up to right multiplication by a projective transformation. Thus, the fundamental matrix captures the projective relationship of the two cameras.

Theorem 8.10. *Let* $\mathtt{F}$ *be a fundamental matrix and let* $(\mathtt{P}, \mathtt{P}')$ *and* $(\tilde{\mathtt{P}}, \tilde{\mathtt{P}}')$ *be two pairs of camera matrices such that* $\mathtt{F}$ *is the fundamental matrix corresponding to each of these pairs. Then there exists a non-singular* 4×4 *matrix* $\mathtt{H}$ *such that* $\tilde{\mathtt{P}} = \mathtt{PH}$ *and* $\tilde{\mathtt{P}}' = \mathtt{P}'\mathtt{H}$.

Proof Suppose that a given fundamental matrix $\mathtt{F}$ corresponds to two different pairs of camera matrices $(\mathtt{P}, \mathtt{P}')$ and $(\tilde{\mathtt{P}}, \tilde{\mathtt{P}}')$. As a first step, we may simplify the problem by assuming that each of the two pair of camera matrices is in canonical form with $\mathtt{P} = \tilde{\mathtt{P}} = [\mathtt{I} \mid \mathbf{0}]$, since this may be done by applying projective transformations to each pair as necessary. Thus, suppose that $\mathtt{P} = \tilde{\mathtt{P}} = [\mathtt{I} \mid \mathbf{0}]$ and that $\mathtt{P}' = [\mathtt{A} \mid \mathbf{a}]$ and $\tilde{\mathtt{P}}' = [\tilde{\mathtt{A}} \mid \tilde{\mathbf{a}}]$. According to result 8.9 the fundamental matrix may then be written $\mathtt{F} = [\mathbf{a}]_\times \mathtt{A} = [\tilde{\mathbf{a}}]_\times \tilde{\mathtt{A}}$.
We will need the following lemma:

Lemma 8.11. *Suppose the rank 2 matrix* $\mathtt{F}$ *can be decomposed in two different ways as* $\mathtt{F} = [\mathbf{a}]_\times \mathtt{A}$ *and* $\mathtt{F} = [\tilde{\mathbf{a}}]_\times \tilde{\mathtt{A}}$*; then* $\tilde{\mathbf{a}} = k\mathbf{a}$ *and* $\tilde{\mathtt{A}} = k^{-1}(\mathtt{A} + \mathbf{a}\mathbf{v}^\top)$ *for some non-zero constant* k *and 3-vector* $\mathbf{v}$.

Proof First, note that $\mathbf{a}^\top \mathtt{F} = \mathbf{a}^\top [\mathbf{a}]_\times \mathtt{A} = \mathbf{0}$, and similarly, $\tilde{\mathbf{a}}^\top \mathtt{F} = \mathbf{0}$. Since $\mathtt{F}$ has rank 2, it follows that $\tilde{\mathbf{a}} = k\mathbf{a}$ as required. Next, from $[\mathbf{a}]_\times \mathtt{A} = [\tilde{\mathbf{a}}]_\times \tilde{\mathtt{A}}$ it follows that $[\mathbf{a}]_\times (k\tilde{\mathtt{A}} - \mathtt{A}) = \mathtt{0}$, and so $k\tilde{\mathtt{A}} - \mathtt{A} = \mathbf{a}\mathbf{v}^\top$ for some $\mathbf{v}$. Hence, $\tilde{\mathtt{A}} = k^{-1}(\mathtt{A} + \mathbf{a}\mathbf{v}^\top)$ as required. □

Applying this result to the two camera matrices $\mathtt{P}'$ and $\tilde{\mathtt{P}}'$ shows that $\mathtt{P}' = [\mathtt{A} \mid \mathbf{a}]$ and $\tilde{\mathtt{P}}' = [k^{-1}(\mathtt{A} + \mathbf{a}\mathbf{v}^\top) \mid k\mathbf{a}]$ if they are to generate the same $\mathtt{F}$. It only remains now to show that these camera pairs are projectively related. Let $\mathtt{H}$ be the matrix $\mathtt{H} = \begin{bmatrix} k^{-1}\mathtt{I} & \mathbf{0} \\ k^{-1}\mathbf{v}^\top & k \end{bmatrix}$. Then one verifies that $\mathtt{PH} = k^{-1}[\mathtt{I} \mid \mathbf{0}] = k^{-1}\tilde{\mathtt{P}}$, and furthermore,

$$\mathtt{P}'\mathtt{H} = [\mathtt{A} \mid \mathbf{a}]\mathtt{H} = [k^{-1}(\mathtt{A} + \mathbf{a}\mathbf{v}^\top) \mid k\mathbf{a}] = [\tilde{\mathtt{A}} \mid \tilde{\mathbf{a}}] = \tilde{\mathtt{P}}'$$

so that the pairs $\mathtt{P}, \mathtt{P}'$ and $\tilde{\mathtt{P}}, \tilde{\mathtt{P}}'$ are indeed projectively related. □

This can be tied precisely to a counting argument: the two cameras $\mathtt{P}$ and $\mathtt{P}'$ each have 11 degrees of freedom, making a total of 22 degrees of freedom. To specify a projective world frame requires 15 degrees of freedom (section 2.1($p45$)), so once the degrees of freedom of the world frame are removed from the two cameras $22 - 15 = 7$ degrees of freedom remain – which corresponds to the 7 degrees of freedom of the fundamental matrix.

8.5.3 Canonical cameras given F

We have shown that $\mathtt{F}$ determines the camera pair up to a projective transformation of 3-space. We will now derive a specific formula for a pair of cameras with canonical

form given $\mathtt{F}$. We will make use of the following characterization of the fundamental matrix $\mathtt{F}$ corresponding to a pair of camera matrices:

Result 8.12. *A non-zero matrix* $\mathtt{F}$ *is the fundamental matrix corresponding to a pair of camera matrices* $\mathtt{P}$ *and* $\mathtt{P}'$ *if and only if* ${\mathtt{P}'}^\top\mathtt{F}\mathtt{P}$ *is skew-symmetric.*

Proof The condition that ${\mathtt{P}'}^\top\mathtt{F}\mathtt{P}$ is skew-symmetric is equivalent to $\mathbf{X}^\top{\mathtt{P}'}^\top\mathtt{F}\mathtt{P}\mathbf{X} = 0$ for all $\mathbf{X}$. Setting $\mathbf{x}' = \mathtt{P}'\mathbf{X}$ and $\mathbf{x} = \mathtt{P}\mathbf{X}$, this is equivalent to ${\mathbf{x}'}^\top\mathtt{F}\mathbf{x} = 0$, which is the defining equation for the fundamental matrix. □

One may write down a particular solution for the pairs of camera matrices in canonical form that correspond to a fundamental matrix as follows:

Result 8.13. *Let* $\mathtt{F}$ *be a fundamental matrix and* $\mathtt{S}$ *any skew-symmetric matrix. Define the pair of camera matrices*

$$\mathtt{P} = [\mathtt{I} \mid \mathbf{0}] \quad \textit{and} \quad \mathtt{P}' = [\mathtt{S}\mathtt{F} \mid \mathbf{e}'],$$

where $\mathbf{e}'$ *is the epipole such that* ${\mathbf{e}'}^\top\mathtt{F} = \mathbf{0}$*, and assume that* $\mathtt{P}'$ *so defined is a valid camera matrix (has rank 3). Then* $\mathtt{F}$ *is the fundamental matrix corresponding to the pair* $(\mathtt{P}, \mathtt{P}')$*.*

To demonstrate this, we invoke result 8.12 and simply verify that

$$[\mathtt{S}\mathtt{F} \mid \mathbf{e}']^\top\mathtt{F}[\mathtt{I} \mid \mathbf{0}] = \begin{bmatrix} \mathtt{F}^\top\mathtt{S}^\top\mathtt{F} & \mathbf{0} \\ {\mathbf{e}'}^\top\mathtt{F} & 0 \end{bmatrix} = \begin{bmatrix} \mathtt{F}^\top\mathtt{S}^\top\mathtt{F} & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{bmatrix} \tag{8.9}$$

which is skew-symmetric.

The skew-symmetric matrix $\mathtt{S}$ may be written in terms of its null-vector as $\mathtt{S} = [\mathbf{s}]_\times$. Then $[[\mathbf{s}]_\times\mathtt{F} \mid \mathbf{e}']$ has rank 3 provided $\mathbf{s}^\top\mathbf{e}' \neq 0$, according to the following argument. Since $\mathbf{e}'\mathtt{F} = \mathbf{0}$, the column space (span of the columns) of $\mathtt{F}$ is perpendicular to $\mathbf{e}'$. But if $\mathbf{s}^\top\mathbf{e}' \neq 0$, then $\mathbf{s}$ is not perpendicular to $\mathbf{e}'$, and hence not in the column space of $\mathtt{F}$. Now, the column space of $[\mathbf{s}']_\times\mathtt{F}$ is spanned by the cross-products of $\mathbf{s}$ with the columns of $\mathtt{F}$, and therefore equals the plane perpendicular to $\mathbf{s}$. So $[\mathbf{s}]_\times\mathtt{F}$ has rank 2. Since $\mathbf{e}'$ is not perpendicular to $\mathbf{s}$, it does not lie in this plane, and so $[[\mathbf{s}]_\times\mathtt{F} \mid \mathbf{e}']$ has rank 3, as required.

As suggested by Luong and Viéville [Luong-96] a good choice for $\mathtt{S}$ is $\mathtt{S} = [\mathbf{e}']_\times$, for in this case ${\mathbf{e}'}^\top\mathbf{e}' \neq 0$, which leads to the following useful result.

Result 8.14. *The camera matrices corresponding to a fundamental matrix* $\mathtt{F}$ *may be chosen as* $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$ *and* $\mathtt{P}' = [[\mathbf{e}']_\times\mathtt{F} \mid \mathbf{e}']$.

Note that the camera matrix $\mathtt{P}'$ has left 3×3 submatrix $[\mathbf{e}']_\times\mathtt{F}$ which has rank 2. This corresponds to a camera with centre on $\boldsymbol{\pi}_\infty$. However, there is no particular reason to avoid this situation.

The proof of theorem 8.10 shows that the four parameter family of camera pairs in canonical form $\tilde{\mathtt{P}} = [\mathtt{I} \mid \mathbf{0}]$, $\tilde{\mathtt{P}}' = [\mathtt{A} + \mathbf{a}\mathbf{v}^\top \mid k\mathbf{a}]$ have the same fundamental matrix

as the canonical pair, $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$, $\mathtt{P}' = [\mathtt{A} \mid \mathbf{a}]$; and that this is the most general solution. To summarize:

Result 8.15. *The general formula for a pair of canonic camera matrices corresponding to a fundamental matrix* $\mathtt{F}$ *is given by*

$$\mathtt{P} = [\mathtt{I} \mid \mathbf{0}] \qquad \mathtt{P}' = [[\mathbf{e}']_\times \mathtt{F} + \mathbf{e}'\mathbf{v}^\top \mid \lambda \mathbf{e}'] \tag{8.10}$$

where $\mathbf{v}$ *is any 3-vector, and* λ *a non-zero scalar.*

8.6 The essential matrix

The essential matrix is the specialization of the fundamental matrix to the case of normalized image coordinates (see below). Historically, the essential matrix was introduced (by Longuet-Higgins) before the fundamental matrix, and the fundamental matrix may be thought of as the generalization of the essential matrix in which the (inessential) assumption of calibrated cameras is removed. The essential matrix has fewer degrees of freedom, and additional properties, compared to the fundamental matrix. These properties are described below.

Normalized coordinates. Consider a camera matrix decomposed as $\mathtt{P} = \mathtt{K}[\mathtt{R} \mid \mathbf{t}]$, and let $\mathbf{x} = \mathtt{P}\mathbf{X}$ be a point in the image. If the calibration matrix $\mathtt{K}$ is known, then we may apply its inverse to the point $\mathbf{x}$ to obtain the point $\hat{\mathbf{x}} = \mathtt{K}^{-1}\mathbf{x}$. Then $\hat{\mathbf{x}} = [\mathtt{R} \mid \mathbf{t}]\mathbf{X}$, where $\hat{\mathbf{x}}$ is the image point expressed in *normalized coordinates*. It may be thought of as the image of the point $\mathbf{X}$ with respect to a camera $[\mathtt{R} \mid \mathbf{t}]$ having the identity matrix $\mathtt{I}$ as calibration matrix. The camera matrix $\mathtt{K}^{-1}\mathtt{P} = [\mathtt{R} \mid \mathbf{t}]$ is called a *normalized camera matrix*, the effect of the known calibration matrix having been removed.

Now, consider a pair of normalized camera matrices $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$ and $\mathtt{P}' = [\mathtt{R} \mid \mathbf{t}]$. The fundamental matrix corresponding to the pair of normalized cameras is customarily called the *essential matrix*, and according to (8.2–*p*225) it has the form

$$\mathtt{E} = [\mathbf{t}]_\times \mathtt{R} = \mathtt{R}\,[\mathtt{R}^\top \mathbf{t}]_\times .$$

Definition 8.16. The defining equation for the essential matrix is

$$\hat{\mathbf{x}}'^\top \mathtt{E} \hat{\mathbf{x}} = 0 \tag{8.11}$$

in terms of the normalized image coordinates for corresponding points $\mathbf{x} \leftrightarrow \mathbf{x}'$.

Substituting for $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}'$ gives $\mathbf{x}'^\top \mathtt{K}'^{-\top} \mathtt{E} \mathtt{K}^{-1} \mathbf{x} = 0$. Comparing this with the relation $\mathbf{x}'^\top \mathtt{F} \mathbf{x} = 0$ for the fundamental matrix, it follows that the relationship between the fundamental and essential matrices is

$$\mathtt{E} = \mathtt{K}'^\top \mathtt{F} \mathtt{K}. \tag{8.12}$$

8.6.1 Properties of the essential matrix

The essential matrix, $\mathtt{E} = [\mathbf{t}]_\times \mathtt{R}$, has only five degrees of freedom: both the rotation matrix $\mathtt{R}$ and the translation $\mathbf{t}$ have three degrees of freedom, but there is an overall scale ambiguity – like the fundamental matrix, the essential matrix is a homogeneous quantity.

The reduced number of degrees of freedom translates into extra constraints that are satisfied by an essential matrix, compared with a fundamental matrix. We investigate what these constraints are.

Result 8.17. *A 3×3 matrix is an essential matrix if and only if two of its singular values are equal, and the third is zero.*

Proof This is easily deduced from the decomposition of $\mathtt{E}$ as $[\mathbf{t}]_\times \mathtt{R} = \mathtt{SR}$, where $\mathtt{S}$ is skew-symmetric. We will use the matrices

$$\mathtt{W} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathtt{Z} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{8.13}$$

It may be verified that $\mathtt{W}$ is orthogonal and $\mathtt{Z}$ is skew-symmetric. From Result A3.1-($p554$), which gives a block decomposition of a general skew-symmetric matrix, the 3×3 skew-symmetric matrix $\mathtt{S}$ may be written as $\mathtt{S} = k\mathtt{U}\mathtt{Z}\mathtt{U}^\top$ where $\mathtt{U}$ is orthogonal. Noting that, up to sign, $\mathtt{Z} = \mathrm{diag}(1, 1, 0)\mathtt{W}$, then up to scale, $\mathtt{S} = \mathtt{U}\,\mathrm{diag}(1, 1, 0)\mathtt{W}\mathtt{U}^\top$, and $\mathtt{E} = \mathtt{SR} = \mathtt{U}\,\mathrm{diag}(1, 1, 0)(\mathtt{W}\mathtt{U}^\top\mathtt{R})$. This is a singular value decomposition of $\mathtt{E}$ with two equal singular values, as required. Conversely, a matrix with two equal singular values may be factored as $\mathtt{SR}$ in this way. □

Since $\mathtt{E} = \mathtt{U}\,\mathrm{diag}(1, 1, 0)\mathtt{V}^\top$, it may seem that $\mathtt{E}$ has six degrees of freedom and not five, since both $\mathtt{U}$ and $\mathtt{V}$ have three degrees of freedom. However, because the two singular values are equal, the SVD is not unique – in fact there is a one-parameter family of SVDs for $\mathtt{E}$. Indeed, an alternative SVD is given by $\mathtt{E} = (\mathtt{U}\,\mathrm{diag}(\mathtt{R}_{2\times2}, 1))\,\mathrm{diag}(1, 1, 0)(\mathrm{diag}(\mathtt{R}_{2\times2}{}^\top, 1))\mathtt{V}^\top$ for any 2×2 rotation matrix $\mathtt{R}$.

8.6.2 Extraction of cameras from the essential matrix

The essential matrix may be computed directly from (8.11) using normalized image coordinates, or else computed from the fundamental matrix using (8.12). (Methods of computing the fundamental matrix are deferred to chapter 10). Once the essential matrix is known, the camera matrices may be retrieved from $\mathtt{E}$ as will be described next. In contrast with the fundamental matrix case, where there is a projective ambiguity, the camera matrices may be retrieved from the essential matrix up to scale and a four-fold ambiguity. That is there are four possible solutions, except for overall scale, which cannot be determined.

We may assume that the first camera matrix is $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$. In order to compute

the second camera matrix, $\mathtt{P}'$, it is necessary to factor $\mathtt{E}$ into the product $\mathtt{SR}$ of a skew-symmetric matrix and a rotation matrix.

Result 8.18. *Suppose that the SVD of* $\mathtt{E}$ *is* $\mathtt{U}\,\mathrm{diag}(1,1,0)\mathtt{V}^\top$. *Using the notation of (8.13), there are (ignoring signs) two possible factorizations* $\mathtt{E} = \mathtt{SR}$ *as follows:*

$$\mathtt{S} = \mathtt{U}\mathtt{Z}\mathtt{U}^\top \quad \mathtt{R} = \mathtt{U}\mathtt{W}\mathtt{V}^\top \quad \textit{or} \quad \mathtt{U}\mathtt{W}^\top\mathtt{V}^\top \ . \tag{8.14}$$

Proof That the given factorization is valid is true by inspection. That there are no other factorizations is shown as follows. Suppose $\mathtt{E} = \mathtt{SR}$. The form of $\mathtt{S}$ is determined by the fact that its left null-space is the same as that of $\mathtt{E}$. Hence $\mathtt{S} = \mathtt{U}\mathtt{Z}\mathtt{U}^\top$. The rotation $\mathtt{R}$ may be written as $\mathtt{U}\mathtt{X}\mathtt{V}^\top$, where $\mathtt{X}$ is some rotation matrix. Then

$$\mathtt{U}\,\mathrm{diag}(1,1,0)\mathtt{V}^\top = \mathtt{E} = \mathtt{SR} = (\mathtt{U}\mathtt{Z}\mathtt{U}^\top)(\mathtt{U}\mathtt{X}\mathtt{V}^\top) = \mathtt{U}(\mathtt{Z}\mathtt{X})\mathtt{V}^\top$$

from which one deduces that $\mathtt{Z}\mathtt{X} = \mathrm{diag}(1,1,0)$. Since $\mathtt{X}$ is a rotation matrix, it follows that $\mathtt{X} = \mathtt{W}$ or $\mathtt{X} = \mathtt{W}^\top$, as required. □

The factorization (8.14) determines the $\mathbf{t}$ part of the camera matrix $\mathtt{P}'$, up to scale, from $\mathtt{S} = [\mathbf{t}]_\times$. However, the Frobenius norm of $\mathtt{S} = \mathtt{U}\mathtt{Z}\mathtt{U}^\top$ is 2, which means that if $\mathtt{S} = [\mathbf{t}]_\times$ *including scale* then $||\mathbf{t}|| = 1$, which is a convenient normalization for the baseline of the two camera matrices. Since $\mathtt{S}\mathbf{t} = \mathbf{0}$, it follows that $\mathbf{t} = \mathtt{U}\,(0,0,1)^\top = \mathbf{u}_3$, the last column of $\mathtt{U}$. However, the sign of $\mathtt{E}$, and consequently $\mathbf{t}$, cannot be determined. Thus, corresponding to a given essential matrix, there are four possible choices of the camera matrix $\mathtt{P}'$, based on the two possible choices of $\mathtt{R}$ and two possible signs of $\mathtt{t}$. To summarize:

Result 8.19. *For a given essential matrix* $\mathtt{E} = \mathtt{U}\,\mathrm{diag}(1,1,0)\mathtt{V}^\top$, *and first camera matrix* $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$, *there are four possible choices for the second camera matrix* $\mathtt{P}'$, *namely*

$$\mathtt{P}' = [\mathtt{U}\mathtt{W}\mathtt{V}^\top \mid +\mathbf{u}_3] \ \text{ or } \ [\mathtt{U}\mathtt{W}\mathtt{V}^\top \mid -\mathbf{u}_3] \ \text{ or } \ [\mathtt{U}\mathtt{W}^\top\mathtt{V}^\top \mid +\mathbf{u}_3] \ \text{ or } \ [\mathtt{U}\mathtt{W}^\top\mathtt{V}^\top \mid -\mathbf{u}_3].$$

8.6.3 Geometrical interpretation of the four solutions

It is clear that the difference between the first two solutions is simply that the direction of the translation vector from the first to the second camera is reversed.

The relationship of the first and third solutions in result 8.19 is a little more complicated. However, it may be verified that

$$[\mathtt{U}\mathtt{W}^\top\mathtt{V}^\top \mid \mathbf{u}_3] = [\mathtt{U}\mathtt{W}\mathtt{V}^\top \mid \mathbf{u}_3] \begin{bmatrix} \mathtt{V}\mathtt{W}^\top\mathtt{W}^\top\mathtt{V}^\top & \\ & 1 \end{bmatrix}$$

and $\mathtt{V}\mathtt{W}^\top\mathtt{W}^\top\mathtt{V}^\top = \mathtt{V}\,\mathrm{diag}(-1,-1,1)\mathtt{V}^\top$ is a rotation through 180° about the line joining the two camera centres. Two solutions related in this way are known as a "twisted pair".

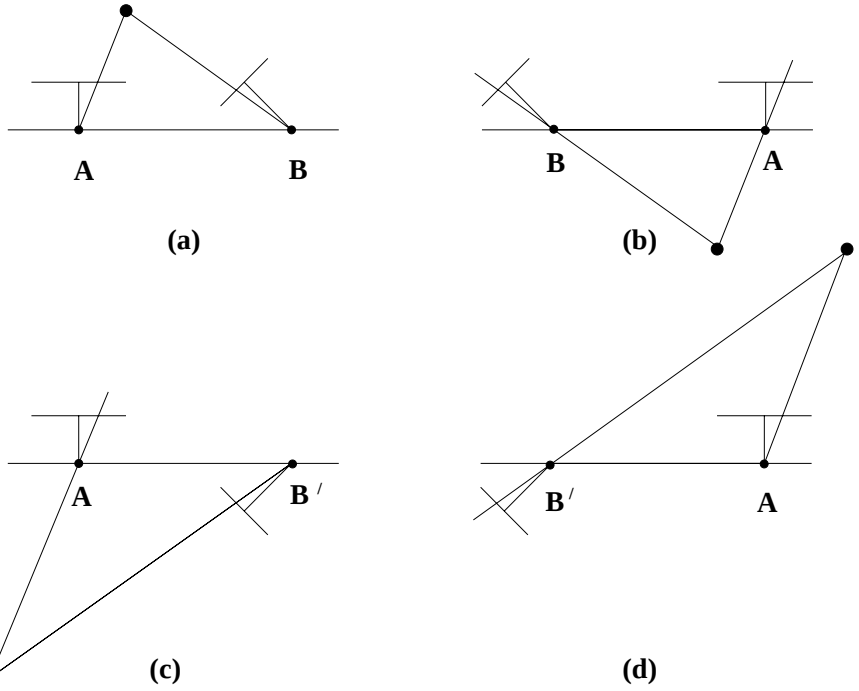


Fig. 8.12. **The four possible solutions for calibrated reconstruction from E.** *Between the left and right sides there is a baseline reversal. Between the top and bottom rows camera B rotates* 180° *about the baseline. Note, only in (a) is the reconstructed point in front of both cameras.*

The four solutions are illustrated in figure 8.12, where it is shown that a reconstructed point $\mathbf{X}$ will be in front of both cameras in one of these four solutions only. Thus, testing with a single point to determine if it is in front of both cameras is sufficient to decide between the four different solutions for the camera matrix $\mathtt{P}'$.

Note. The point of view has been taken here that the essential matrix is a homogeneous quantity. An alternative point of view is that the essential matrix is defined exactly by the equation $\mathtt{E} = [\mathbf{t}]_\times \mathtt{R}$, (i.e. including scale), and is determined only up to indeterminate scale by the equation $\mathbf{x}'^\top \mathtt{E}\mathbf{x} = 0$. The choice of point of view depends on which of these two equations one regards as the defining property of the essential matrix.

8.7 Closure

8.7.1 The literature

The essential matrix was introduced to the computer vision community by Longuet-Higgins [LonguetHiggins-81], with a matrix analogous to E appearing in the photogrammetry literature, e.g. [VonSanden-08]. Many properties of the essential matrix have been elucidated particularly by Huang and Faugeras [Huang-89], [Maybank-93], and [Horn-90].

The realization that the essential matrix could also be applied in uncalibrated situations, as it represented a projective relation, developed in the early part of the 1990s, and was published simultaneously by Faugeras [Faugeras-92a, Faugeras-92b], and Hartley *et al.* [Hartley-92a, Hartley-92c].

The special case of pure planar motion was examined by [Maybank-93] for the essential matrix. The corresponding case for the fundamental matrix is investigated by Beardsley and Zisserman [Beardsley-95b] and Viéville and Lingrand [Viéville-95], where further properties are given.

8.7.2 Notes and exercises

(i) **Fixating cameras.** Suppose two cameras fixate on a point in space such that their principal axes intersect at that point. Show that if the image coordinates are normalized so that the coordinate origin coincides with the principal point then the $\mathtt{F}_{33}$ element of the fundamental matrix is zero.

(ii) **Mirror images.** Suppose that a camera views an object and its reflection in a plane mirror. Show that this situation is equivalent to two views of the object, and that the fundamental matrix is skew-symmetric. Compare the fundamental matrix for this configuration with that of: (a) a pure translation, and (b) a pure planar motion. Show that the fundamental matrix is auto-epipolar (as is (a)).

(iii) Show that if the vanishing line of a plane contains the epipole then the plane is parallel to the baseline.

(iv) **Steiner conic.** Show that the polar of $\mathbf{x}_\mathtt{a}$ intersects the Steiner conic $\mathtt{F}_\mathtt{S}$ at the epipoles (figure 8.10a). Hint, start from $\mathtt{F}\mathbf{e} = \mathtt{F}_\mathtt{S}\mathbf{e} + \mathtt{F}_\mathtt{a}\mathbf{e} = \mathbf{0}$. Since $\mathbf{e}$ lies on the conic $\mathtt{F}_\mathtt{S}$, then $\mathbf{l}_1 = \mathtt{F}_\mathtt{S}\mathbf{e}$ is the tangent line at $\mathbf{e}$, and $\mathbf{l}_2 = \mathtt{F}_\mathtt{a}\mathbf{e} = [\mathbf{x}_\mathtt{a}]_\times\mathbf{e} = \mathbf{x}_\mathtt{a} \times \mathbf{e}$ is a line through $\mathbf{x}_\mathtt{a}$ and $\mathbf{e}$.

(v) **Planar motion.** It is shown by [Maybank-93] that if the rotation axis direction is orthogonal or parallel to the translation direction then the symmetric part of the essential matrix has rank 2. We assume here that $\mathtt{K} = \mathtt{K}'$. Then from (8.12), $\mathtt{F} = \mathtt{K}^{-\top}\mathtt{E}\mathtt{K}^{-1}$, and so

$$\mathtt{F}_\mathtt{S} = (\mathtt{F} + \mathtt{F}^\top)/2 = \mathtt{K}^{-\top}(\mathtt{E} + \mathtt{E}^\top)\mathtt{K}^{-1}/2 = \mathtt{K}^{-\top}\mathtt{E}_\mathtt{S}\mathtt{K}^{-1}.$$

It follows from $\det(\mathtt{F}_\mathtt{S}) = \det(\mathtt{K}^{-1})^2 \det(\mathtt{E}_\mathtt{S})$ that the symmetric part of $\mathtt{F}$ is also singular. Does this result hold if $\mathtt{K} \neq \mathtt{K}'$?

(vi) Any matrix $\mathtt{F}$ of rank 2 is the fundamental matrix corresponding to some pair of camera matrices $(\mathtt{P}, \mathtt{P}')$ This follows directly from result 8.14 since the solution for the canonical cameras depends only on the rank 2 property of $\mathtt{F}$.

(vii) Show that the 3D points determined from one of the ambiguous reconstructions obtained from $\mathtt{E}$ are related to the corresponding 3D points determined from another reconstruction by either (i) an inversion through the second

camera centre; or (ii) a harmonic homology of 3-space (see section A5.2-(p584)), where the homology plane is perpendicular to the baseline and through the second camera centre, and the vertex is the first camera centre.

(viii) Following a similar development to section 8.2.2, derive the form of the fundamental matrix for two linear pushbroom cameras. Details of this matrix are given in [Gupta-97] where it is shown that affine reconstruction is possible from a pair of images.