17

N-Linearities and Multiple View Tensors

This chapter introduces the quadrifocal tensor Q^{ijkl} between four views, which is the analogue of the fundamental matrix for two and the trifocal tensor for three views. The quadrifocal tensor encapsulates the relationships between imaged points and lines seen in four views.

It is shown that multiple view relations may be derived directly and uniformly from the intersection properties of back-projected lines and points. From this analysis the fundamental matrix $\mathtt{F}$, trifocal tensor $\mathcal{T}_i^{jk}$, and quadrifocal tensor Q^{ijkl} appear in a common framework involving matrix determinants. Specific formulae are given for each of these tensors in terms of the camera matrices.

We also develop general counting arguments for the degrees of freedom of the tensors and the number of point and line correspondences required for tensor computation. These are given for configurations in general position and for the important special case where four or more of the elements are coplanar.

17.1 Bilinear relations

We consider first the relationship that holds between the coordinates of a point seen in two separate views. Thus, let $\mathbf{x} \leftrightarrow \mathbf{x}'$ be a pair of corresponding points which are the images of the same point $\mathbf{X}$ in space as seen in the two separate views. It will be convenient, for clarity of notation, to represent the two camera matrices by $\mathtt{A}$ and $\mathtt{B}$, instead of the usual notation, $\mathtt{P}$ and $\mathtt{P}'$. The projection from space to image can now be expressed as $k\mathbf{x} = \mathtt{A}\mathbf{X}$ and $k'\mathbf{x}' = \mathtt{B}\mathbf{X}$ where k and k' are two undetermined constants. This pair of equations may be written down as one equation

$$\begin{bmatrix} \mathtt{A} & \mathbf{x} & \mathbf{0} \\ \mathtt{B} & \mathbf{0} & \mathbf{x}' \end{bmatrix} \begin{pmatrix} \mathbf{X} \\ -k \\ -k' \end{pmatrix} = \mathbf{0}$$

and it may easily be verified that this is equivalent to the two equations. This can be written in a more detailed form by denoting the i-th row of the matrix $\mathtt{A}$ by $\mathbf{a}^i$, and similarly the i-th row of the matrix $\mathtt{B}$ by $\mathbf{b}^i$. We also write $\mathbf{x} = (x^1, x^2, x^3)^\mathsf{T}$ and

$\mathbf{x}' = (x'^1, x'^2, x'^3)^\top$. The set of equations is now

$$\left[\begin{array}{ccc} \mathbf{a}^1 & x^1 & \\ \mathbf{a}^2 & x^2 & \\ \mathbf{a}^3 & x^3 & \\ \hline \mathbf{b}^1 & & x'^1 \\ \mathbf{b}^2 & & x'^2 \\ \mathbf{b}^3 & & x'^3 \end{array}\right] \begin{pmatrix} \mathbf{X} \\ -k \\ -k' \end{pmatrix} = \mathbf{0}. \tag{17.1}$$

Now, this is a 6×6 set of equations which by hypothesis has a non-zero solution, the vector $(\mathbf{X}^\top, -k, -k')^\top$. It follows that the matrix of coefficients in (17.1) must have zero determinant. It will be seen that this condition leads to a bilinear relationship between the entries of the vectors $\mathbf{x}$ and $\mathbf{x}'$ expressed by the fundamental matrix $\mathtt{F}$. We will now look specifically at the form of this relationship.

Consider the matrix appearing in (17.1). Denote it by $\mathtt{X}$. The determinant of $\mathtt{X}$ may be written as an expression in terms of the quantities x^i and x'^i. Notice that the entries x^i and x'^i appear in only two columns of $\mathtt{X}$. This implies that the determinant of $\mathtt{X}$ may be expressed as a quadratic expression in terms of the x^i and x'^i. In fact, since all the entries x^i appear in the same column, there can be no terms of the form $x^i x^j$ or $x'^i x'^j$. Briefly, in terms of the x^i and x'^i, the determinant of $\mathtt{X}$ is a bilinear expression. The fact that the determinant is zero may be written as an equation

$$(x'^1, x'^2, x'^3)\mathtt{F}(x^1, x^2, x^3)^\top = x'^i x^j F_{ij} = 0 \tag{17.2}$$

where $\mathtt{F}$ is a 3×3 matrix, the fundamental matrix.

We may compute a specific formula for the entries of the matrix $\mathtt{F}$ as follows. The entry F_{ij} of $\mathtt{F}$ is the coefficient of the term $x'^i x^j$ in the expansion of the determinant of $\mathtt{X}$. In order to find this coefficient, we must eliminate the rows and columns of the matrix containing x'^i and x^j, take the determinant of the resulting matrix and multiply by ± 1 as appropriate. For instance, the coefficient of $x'^1 x^1$ is obtained by eliminating two rows and the last two columns of the matrix $\mathtt{X}$. The remaining matrix is

$$\begin{bmatrix} \mathbf{a}^2 \\ \mathbf{a}^3 \\ \mathbf{b}^2 \\ \mathbf{b}^3 \end{bmatrix}$$

and the coefficient of $x'^1 x^1$ is equal to the determinant of this 4×4 matrix. In general, we may write

$$F_{ji} = (-1)^{i+j} \det \begin{bmatrix} \sim\mathbf{a}^i \\ \sim\mathbf{b}^j \end{bmatrix}. \tag{17.3}$$

In this expression, the notation $\sim\mathbf{a}^i$ has been used to denote the matrix obtained from $\mathtt{A}$ by *omitting* the row $\mathbf{a}^i$. Thus the symbol $\sim$ may be read as *omit*, and $\sim\mathbf{a}^i$ represents two rows of $\mathtt{A}$. The determinant appearing on the right side of (17.3) is therefore a 4×4 determinant.

A different way of writing the expression for F_{ji} makes use of the tensor ϵ_{rst} (defined in section A1.1(p563)) as follows:[1]

$$F_{ji} = \left(\frac{1}{4}\right) \epsilon_{ipq}\epsilon_{jrs} \det \begin{bmatrix} \mathbf{a}^p \\ \mathbf{a}^q \\ \mathbf{b}^r \\ \mathbf{b}^s \end{bmatrix}. \tag{17.4}$$

To see this, note that F_{ji} is defined in (17.4) in terms of a sum of determinants over all values of p, q, r and s. However, for a given value of i, the tensor ϵ_{ipq} is zero unless p and q are different from i and from each other. This leaves only two remaining choices of p and q (for example if $i = 1$, then we may choose $p = 2,\ q = 3$ or $p = 3,\ q = 2$). Similarly, there are only two different choices of r and s giving rise to non-zero terms. Thus the sum consists of four non-zero terms only. Furthermore, the determinants appearing in these four terms consist of the same four rows of the matrices $\mathtt{A}$ and $\mathtt{B}$ and hence have equal values, except for sign. However, the value of $\epsilon_{ipq}\epsilon_{jrs}$ is such that the four terms all have the same sign and are equal. Thus, the sum (17.4) is equal to the single term appearing in (17.3).

17.1.1 Epipoles as tensors

The expression (17.3) for the fundamental matrix involves determinants of matrices containing two rows from each of $\mathtt{A}$ and $\mathtt{B}$. If we consider instead the determinants of matrices containing all three rows from one matrix and one row from the other matrix, the resulting determinants represent the epipoles. Specifically we have

$$e^i = \det \begin{bmatrix} \mathbf{a}^i \\ \mathtt{B} \end{bmatrix} \qquad e'^j = \det \begin{bmatrix} \mathtt{A} \\ \mathbf{b}^j \end{bmatrix} \tag{17.5}$$

where $\mathbf{e}$ and $\mathbf{e}'$ are the epipoles in the two images. To see this, note that the epipole is defined by $e^i = \mathbf{a}^i\mathbf{C}'$, where $\mathbf{C}'$ is the centre of the second camera, defined by $\mathtt{B}\mathbf{C}' = \mathbf{0}$. The formula (17.5) is now obtained by expanding the determinant by cofactors along the first row (in a similar manner to the derivation of (3.4–p67)).

17.1.2 Affine specialization

In the case where both the cameras are affine cameras, the fundamental matrix has a particularly simple form. Recall that an affine camera matrix is one for which the final row is $(0, 0, 0, 1)$. Now, note from (17.3) that if neither i nor j is equal to 3, then the third rows of both $\mathtt{A}$ and $\mathtt{B}$ are present in this expression for F_{ij}. The determinant has two equal rows, and hence equals zero. Thus, $\mathtt{F}$ is of the form

$$\mathtt{F}_{\mathrm{A}} = \begin{bmatrix} & & a \\ & & b \\ c & d & e \end{bmatrix}$$

[1] Of course the factor 1/4 is inessential since $\mathtt{F}$ is defined only up to scale. It is included here just to show the relationship to (17.3).

with all other entries being zero. Thus the affine fundamental matrix has just 5 non-zero entries, and hence 4 degrees of freedom. Its properties are described in section 14.2(p345).

Note that this argument relies solely on the fact that both cameras have the same third row. Since the third row of a camera matrix represents the principal plane of the camera (see section 6.2.1(p158)), it follows that the fundamental matrix for two cameras sharing the same principal plane is of the above form.

17.2 Trilinear relations

The determinant method of deriving the fundamental matrix can be used to derive relationships between the coordinates of points seen in three views. This analysis results in a formula for the trifocal tensor. Unlike the fundamental matrix, the trifocal tensor relates both lines and points in the three images. We begin by describing the relationships for corresponding points.

17.2.1 Trifocal point relations

Consider a point correspondence across three views: $\mathbf{x} \leftrightarrow \mathbf{x}' \leftrightarrow \mathbf{x}''$. Let the third camera matrix be $\mathtt{C}$ and let $\mathbf{c}^i$ be its i-th row. Analogous to (17.1) we can write an equation describing the projection of a point $\mathbf{X}$ into the three images as

$$\begin{bmatrix} \mathtt{A} & \mathbf{x} & & \\ \mathtt{B} & & \mathbf{x}' & \\ \mathtt{C} & & & \mathbf{x}'' \end{bmatrix} \begin{pmatrix} \mathbf{X} \\ -k \\ -k' \\ -k'' \end{pmatrix} = \mathbf{0}. \tag{17.6}$$

This matrix, which as before we will call $\mathtt{X}$, has 9 rows and 7 columns. From the existence of a solution to this set of equations, we deduce that its rank must be at most 6. Hence any 7×7 minor has zero determinant. This fact gives rise to the trilinear relationships that hold between the coordinates of the points $\mathbf{x}$, $\mathbf{x}'$ and $\mathbf{x}''$.

There are essentially two different types of 7×7 minors of $\mathtt{X}$. In choosing 7 rows of $\mathtt{X}$, we may choose either

(i) Three rows from each of two camera matrices and one row from the third, or
(ii) Three rows from one camera matrix and two rows from each of the two others.

Let us consider the first type. A typical such 7×7 minor of $\mathtt{X}$ is of the form

$$\begin{bmatrix} \mathtt{A} & \mathbf{x} & & \\ \mathtt{B} & & \mathbf{x}' & \\ \mathbf{c}^i & & & x''^i \end{bmatrix}. \tag{17.7}$$

Note that this matrix contains only one entry in the last column, namely x''^i. Expanding the determinant by cofactors down this last column reveals that the determinant is equal to

$$x''^i \det \begin{bmatrix} \mathtt{A} & \mathbf{x} & \\ \mathtt{B} & & \mathbf{x}' \end{bmatrix}.$$

Apart from the factor x''^i, this just leads to the bilinear relationship expressed by the fundamental matrix, as discussed in section 17.1.

The other sort of 7×7 minor is of more interest. An example of such a determinant is of the form

$$\det \begin{bmatrix} \mathtt{A} & \mathbf{x} & & \\ \mathbf{b}^j & & x'^j & \\ \mathbf{b}^l & & x'^l & \\ \mathbf{c}^k & & & x''^k \\ \mathbf{c}^m & & & x''^m \end{bmatrix} . \tag{17.8}$$

By the same sort of argument as with the bilinear relations it is seen that setting the determinant to zero leads to a trilinear relation of the form $f(\mathbf{x}, \mathbf{x}', \mathbf{x}'') = 0$. By expanding this determinant down the column containing x^i, we can find a specific formula as follows.

$$\det \mathtt{X}_{uv} = -\frac{1}{2} x^i x'^j x''^k \epsilon_{ilm} \epsilon_{jqu} \epsilon_{krv} \det \begin{bmatrix} \mathbf{a}^l \\ \mathbf{a}^m \\ \mathbf{b}^q \\ \mathbf{c}^r \end{bmatrix} = 0_{uv} \tag{17.9}$$

where u and v are free indices corresponding to the rows omitted from the matrices B and C to produce (17.8). We introduce the tensor

$$\mathcal{T}_i^{qr} = \frac{1}{2} \epsilon_{ilm} \det \begin{bmatrix} \mathbf{a}^l \\ \mathbf{a}^m \\ \mathbf{b}^q \\ \mathbf{c}^r \end{bmatrix} . \tag{17.10}$$

The trilinear relationship (17.9) may then be written

$$x^i x'^j x''^k \epsilon_{jqu} \epsilon_{krv} \mathcal{T}_i^{qr} = 0_{uv} . \tag{17.11}$$

The tensor $\mathcal{T}_i^{qr}$ is the trifocal tensor, and (17.11) is a trilinear relation such as those discussed in section 15.2.1(*p*378). The indices u and v are free indices, and each choice of u and v leads to a different trilinear relation.

Just as in the case of the fundamental matrix, one may write the formula for the tensor $\mathcal{T}_i^{qr}$ in a slightly different way

$$\mathcal{T}_i^{qr} = (-1)^{i+1} \det \begin{bmatrix} \sim\mathbf{a}^i \\ \mathbf{b}^q \\ \mathbf{c}^r \end{bmatrix} . \tag{17.12}$$

As in section 17.1, the expression $\sim\mathbf{a}^i$ means the matrix A with row i omitted. Note that we omit row i from the first camera matrix, but *include* rows q and r from the other two camera matrices.

In the often-considered case where the first camera matrix A has the canonical form $[\mathtt{I} \mid \mathbf{0}]$, the expression (17.12) for the trifocal tensor may be written simply as

$$\mathcal{T}_i^{qr} = b_i^q c_4^r - b_4^q c_i^r . \tag{17.13}$$

Note that there are in fact 27 possible trilinear relations that may be formed in this way (refer to (17.8)). Specifically, note that each relation arises from taking all three rows from one camera matrix along with two rows from each of the other two matrices. This gives the following computation.

- 3 ways to choose the first camera matrix from which to take all three rows.
- 3 ways to choose the row to omit from the second camera matrix.
- 3 ways to choose the row to omit from the third camera matrix.

This gives a total of 27 trilinear relations. However, among the 9 ways of choosing two rows from the second and third camera matrices, only 4 are linearly independent (we return to this in section 17.6). This means that there are a total of 12 linearly independent trilinear relations.

It is important to distinguish between the number of trilinear relations, however, and the number of different trifocal tensors. As is shown by (17.11), several different trilinear relations may be expressed in terms of just one trifocal tensor. In (17.11) each distinct choice of the free indices u and v gives rise to a different trilinear relation, all of which are expressible in terms of the same trifocal tensor $\mathcal{T}_i^{qr}$. On the other hand, in the definition of the trifocal tensor given in (17.10), the camera matrix $\mathtt{A}$ is treated differently from the other two, in that $\mathtt{A}$ contributes two rows (after omitting row i) to the determinant defining any given entry of $\mathcal{T}_i^{qr}$, whereas the other two camera matrices contribute just one row. This means that there are in fact three different trifocal tensors corresponding to the choice of which of the three camera matrices contributes two rows.

17.2.2 Trifocal line relations

A line in an image is represented by a covariant vector l_i, and the condition for a point $\mathbf{x}$ to lie on the line is that $l_i x^i = 0$. Let X^j represent a point $\mathbf{X}$ in space, and a_j^i represent a camera matrix $\mathtt{A}$. The 3D point X^j is mapped to the image point as $x^i = a_j^i \mathrm{X}^j$. It follows that the condition for the point X^j to project to a point on the line l_i is that $l_i a_j^i \mathrm{X}^j = 0$. Another way of looking at this is that $l_i a_j^i$ represents a plane consisting of all points that project onto the line l_i.

Consider the situation where a point X^j maps to a point x^i in one image and to some point on lines l'_q and l''_r in two other images. This may be expressed by equations

$$x^i = k a_j^i \mathrm{X}^j \quad l'_q b_j^q \mathrm{X}^j = 0 \quad l''_r c_j^r \mathrm{X}^j = 0.$$

These may be written as a single matrix equation of the form

$$\begin{bmatrix} \mathtt{A} & \mathbf{x} \\ l'_q \mathbf{b}^q & 0 \\ l''_r \mathbf{c}^r & 0 \end{bmatrix} \begin{pmatrix} \mathbf{X} \\ -k \end{pmatrix} = \mathbf{0}. \tag{17.14}$$

Since this set of equations has a solution, it follows that $\det \mathtt{X} = 0$, where $\mathtt{X}$ is the matrix on the left of the equation. Expanding this determinant down the last column gives

$$
\begin{aligned}
0 = -\det \mathtt{X} &= \frac{1}{2} x^i \epsilon_{ilm} \det \begin{bmatrix} \mathbf{a}^l \\ \mathbf{a}^m \\ l'_q \mathbf{b}^q \\ l''_r \mathbf{c}^r \end{bmatrix} = \frac{1}{2} x^i l'_q l''_r \epsilon_{ilm} \det \begin{bmatrix} \mathbf{a}^l \\ \mathbf{a}^m \\ \mathbf{b}^q \\ \mathbf{c}^r \end{bmatrix} \\
&= x^i l'_q l''_r \mathcal{T}_i^{qr}. \qquad (17.15)
\end{aligned}
$$

This shows the connection of the trifocal tensor with sets of lines. The two lines l'_q and l''_r back-project to planes meeting in a line in space. The image of this line in the first image is a line, which may be represented by l_i. For any point x^i on that line the relation (17.15) holds. It follows that $l'_q l''_r \mathcal{T}_i^{qr}$ is the representation of the line l_i. Thus, we see that for three corresponding lines in the three images

$$l_p = l'_q l''_r \mathcal{T}_p^{qr} \qquad (17.16)$$

where, of course, the two sides are equal only up to a scale factor. Since the two sides of the relation (17.16) are vectors, this may be interpreted as meaning that the vector product of the two sides vanishes. Expressing this vector product using the tensor ϵ^{ijk}, we arrive at an equation

$$l_p l'_q l''_r \epsilon^{ipw} \mathcal{T}_i^{qr} = 0^w. \qquad (17.17)$$

In an analogous manner to the derivation of (17.11) and (17.15) we may derive a relationship between corresponding points in two images and a line in a third image. In particular, if a point $\mathtt{X}^j$ in space maps to points x^i and x'^j in the first two images, and to some point on a line l''_r in the third image, then the relation is

$$x^i x'^j l''_r \epsilon_{jqu} \mathcal{T}_i^{qr} = 0_u. \qquad (17.18)$$

In this relation, the index u is free, and there is one such relation for each choice of $u = 1, \ldots, 3$, of which two are linearly independent.

We summarize the results of this section in table 17.1, where the final column denotes the number of linearly independent equations.

Correspondence	Relation	Number of equations
three points	$x^i x'^j x''^k \epsilon_{jqu} \epsilon_{krv} \mathcal{T}_i^{qr} = 0_{uv}$	4
two points, one line	$x^i x'^j l''_r \epsilon_{jqu} \mathcal{T}_i^{qr} = 0_u$	2
one point, two lines	$x^i l'_q l''_r \mathcal{T}_i^{qr} = 0$	1
three lines	$l_p l'_q l''_r \epsilon^{piw} \mathcal{T}_i^{qr} = 0^w$	2

Table 17.1. *Trilinear relations (see also table 16.1(*p*391)).*

Note how the different equation sets are related to each other. For instance, the second line of the table is derived from the first by replacing $x''^k \epsilon_{krv}$ by the line l''_r and deleting the free index v.

17.2.3 Relations between two views and the trifocal tensor

To this point we have considered correspondences across three views and the trifocal tensor. Here we describe the constraints that arise if the correspondence is only across two views. From the two view case, where point correspondences constrain the fundamental matrix, we would expect some constraint on $\mathcal{T}$.

Consider the case of corresponding points x'^j and x''^k in the second and third images. This means that there is a point $\mathbf{X}$ in space mapping to the points x'^j and x''^k. The point $\mathbf{X}$ also maps to some point x^i in the first image, but x^i is not known. Nevertheless, there exists a relationship $x^i x'^j x''^k \epsilon_{jqu}\epsilon_{krv}\mathcal{T}_i^{qr} = 0_{uv}$ between these points. For each choice of u and v, denote $\mathtt{A}_{i,uv} = x'^j x''^k \epsilon_{jqu}\epsilon_{krv}\mathcal{T}_i^{qr}$. The entries of $\mathtt{A}_{i,uv}$ are linear expressions in the entries of $\mathcal{T}_i^{qr}$ that may be determined explicitly in terms of the known points x'^j and x''^k. There exists a point $\mathbf{x}$ such that $x^i\mathtt{A}_{i,uv} = 0$. For each choice of u, v we may consider $\mathtt{A}_{i,uv}$ as being a 3-vector indexed by i, and for the different choices of u and v, there are 4 linearly independent such expressions. Thus, $\mathtt{A}$ may be considered as a 3×4 matrix, and the condition that $x^i\mathtt{A}_{i,uv} = 0$ means that $\mathtt{A}_{i,uv}$ has rank 2. This means that every 3×3 subdeterminant of $\mathtt{A}$ is zero, which leads to cubic constraints on the elements of the trifocal tensor. For geometric reasons, it appears that the equations $x^i\mathtt{A}_{i,uv}$ are not algebraically independent for the four choices of u and v. Consequently we obtain a single cubic constraint on $\mathcal{T}_i^{jk}$ from a two-view point correspondence. Details are left to the reader.

In the case where the point correspondence is between the first and second (or third) views the analysis is slightly different. However, the result in each case is that although a point correspondence across two views leads to a constraint on the trifocal tensor, this constraint is not a linear constraint as it is in the case of correspondences across three views.

17.2.4 Affine trifocal tensor

In the case where all three cameras are affine, the trifocal tensor satisfies certain constraints. A camera matrix is affine if the last row is $(0, 0, 0, 1)$. It follows that if two of the rows in the matrix in (17.12) are of this form, then the corresponding element of $\mathcal{T}_i^{jk}$ is zero. This is the case for elements $\mathcal{T}_1^{j3}, \mathcal{T}_2^{j3}, \mathcal{T}_1^{3k}, \mathcal{T}_2^{3k}$ and $\mathcal{T}_3^{33}$ – a total of 11 elements. Thus the trifocal tensor contains 16 non-zero entries, defined up to scale. As in the case of the affine fundamental matrix, this analysis is equally valid for the case of cameras sharing the same principal plane.

17.3 Quadrilinear relations

Similar arguments work in the case of four views. Once more, consider a point correspondence across 4 views: $\mathbf{x} \leftrightarrow \mathbf{x}' \leftrightarrow \mathbf{x}'' \leftrightarrow \mathbf{x}'''$. With camera matrices $\mathtt{A}$, $\mathtt{B}$, $\mathtt{C}$ and $\mathtt{D}$, the projection equations may be written as

$$\begin{bmatrix} \mathtt{A} & \mathbf{x} & & & \\ \mathtt{B} & & \mathbf{x}' & & \\ \mathtt{C} & & & \mathbf{x}'' & \\ \mathtt{D} & & & & \mathbf{x}''' \end{bmatrix} \begin{pmatrix} \mathbf{X} \\ -k \\ -k' \\ -k'' \\ -k''' \end{pmatrix} = \mathbf{0}. \tag{17.19}$$

Since this equation has a solution, the matrix $\mathtt{X}$ on the left has rank at most 7, and so all 8×8 determinants are zero. As in the trilinear case, any determinant containing only one row from one of the camera matrices gives rise to a trilinear or bilinear relation between the remaining views. A different case occurs when we consider 8×8 determinants containing two rows from each of the camera matrices. Such a determinant leads to a new quadrilinear relationship of the form

$$x^i x'^j x''^k x'''^l \epsilon_{ipw}\epsilon_{jqx}\epsilon_{kry}\epsilon_{lsz} Q^{pqrs} = 0_{wxyz} \tag{17.20}$$

where each choice of the free variables w, x, y and z gives a different equation, and the 4-dimensional *quadrifocal tensor* Q^{pqrs} is defined by

$$Q^{pqrs} = \det \begin{bmatrix} \mathbf{a}^p \\ \mathbf{b}^q \\ \mathbf{c}^r \\ \mathbf{d}^s \end{bmatrix}. \tag{17.21}$$

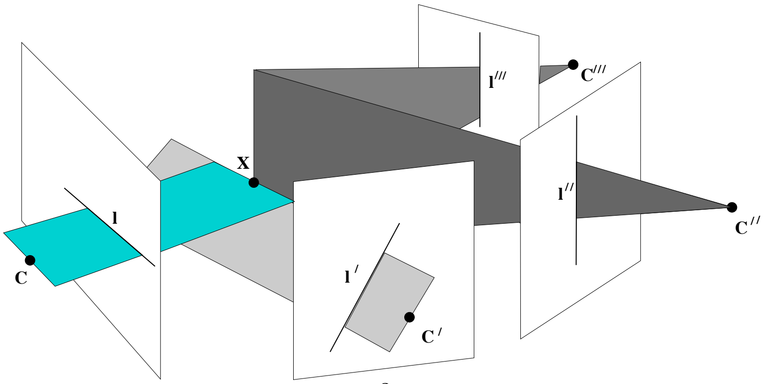
Note that the four indices of the four-view tensor are contravariant, and there is no distinguished view as there is in the case of the trifocal tensor. There is only one four-view tensor corresponding to four given views, and this one tensor gives rise to 81 different quadrilinear relationships, of which 16 are linearly independent (see section 17.6).

As in the case of the trifocal tensor, there are also relations between lines and points in the case of the four-view tensor. Equations relating points are really just special cases of the relationship for lines. In the case of a 4-line correspondence, however, something different happens, as will now be explained. The relationship between a set of four lines and the quadrifocal tensor is given by the formula

$$l_p l'_q l''_r l'''_s Q^{pqrs} = 0 \tag{17.22}$$

for any set of corresponding lines l_p, l'_q, l''_r and l'''_s. However, the derivation shows that this condition will hold as long as there is a single point in space that projects onto the four image lines. It is not necessary that the four image lines correspond (in the sense that they are the image of a common line in space). This configuration is illustrated in figure 17.1a.

Now, consider the case where three of the lines (for instance l'_q, l''_r and l'''_s) correspond by deriving from a single 3D line (figure 17.1b). Now let l_p be any *arbitrary* line in the first image. The back-projection of this line is a plane, which will meet the 3D line in a single point, $\mathbf{X}$, and the conditions are present for (17.22) to hold. Since this is true for any arbitrary line l_p, it must follow that $l'_q l''_r l'''_s Q^{pqrs} = 0^p$. This gives three linearly independent equations involving l'_q, l''_r and l'''_s. However, given a set of corresponding



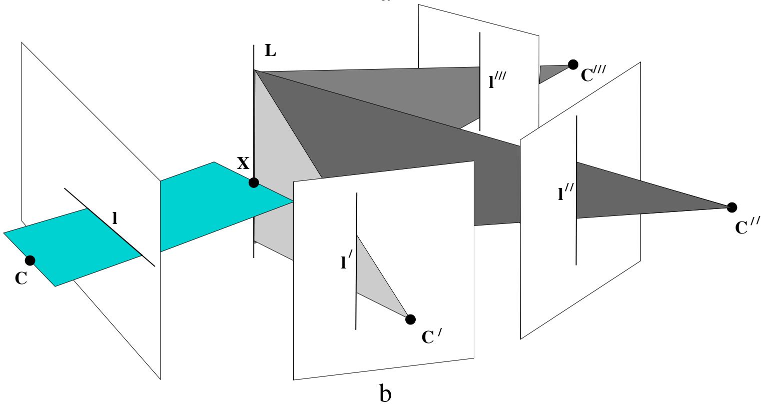


Fig. 17.1. **Four line "correspondence".** *The four lines* $\mathbf{l}, \mathbf{l}', \mathbf{l}'', \mathbf{l}'''$ *satisfy the quadrilinear relation (17.22) since their back-projections intersect in a common point* $\mathbf{X}$. *(a) No three planes intersect in a common line. (b) The three lines* $\mathbf{l}' \leftrightarrow \mathbf{l}'' \leftrightarrow \mathbf{l}'''$ *are the image of the same line* $\mathbf{L}$ *in 3-space.*

lines in four images, as above, we may choose a subset of three lines, and for each line-triplet obtain three equations in this way. Since there are four choices of line-triplets, the total number of equations is thus 12.

However only 9 of these equations are independent, and this may be seen as follows. Suppose that $\mathbf{l} = \mathbf{l}' = \mathbf{l}'' = \mathbf{l}''' = (1, 0, 0)^{\mathsf{T}}$. This is equivalent to the general situation since by applying projective transformations to each of the four images we can transform an arbitrary line correspondence to this case. Now the equation $l'_q l''_r l'''_s Q^{pqrs} = 0^p$ means that any element $Q^{p111} = 0$. Applying this argument to all four triplets of three views, we find that $Q^{pqrs} = 0$ whenever at least three of the indices are 1. There are a total of 9 such elements. Since the set of equations generated by the line correspondence is equivalent to setting each of these elements to zero, among the total of 12 equations there are just 9 independent ones.

The four-view relations are summarized in table 17.2. No equation is given here for the case of three lines and one point, since this gives no more restrictions on the tensor than just the three-line correspondence.

Correspondence	Relation	Number of equations
four points	$x^i x'^j x''^k x'''^l \epsilon_{ipw}\epsilon_{jqx}\epsilon_{kry}\epsilon_{lsz} Q^{pqrs} = 0_{wxyz}$	16
three points, one line	$x^i x'^j x''^k l'''_s \epsilon_{ipw}\epsilon_{jqx}\epsilon_{kry} Q^{pqrs} = 0_{wxy}$	8
two points, two lines	$x^i x'^j l''_r l'''_s \epsilon_{ipw}\epsilon_{jqx} Q^{pqrs} = 0_{wx}$	4
three lines	$l_p l'_q l''_r Q^{pqrs} = 0^s$	3
four lines	$l_p l'_q l''_r Q^{pqrs} = 0^s, \;\; l_p l'_q l'''_s Q^{pqrs} = 0^r, \;\; \ldots$	9

Table 17.2. *Quadrilinear relations.*

17.4 Intersections of four planes

The multi-view tensors may be given a different derivation, which sheds a little more light on their meaning. In this interpretation, the basic geometric property is the intersection of four planes. Four planes in space will generally not meet in a common point. A necessary and sufficient condition for them to do so is that the determinant of the 4×4 matrix formed from the vectors representing the planes should vanish.

Notation. In this section only we shall represent the determinant of a 4×4 matrix with rows $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ by $\mathbf{a} \wedge \mathbf{b} \wedge \mathbf{c} \wedge \mathbf{d}$. In a more general context, the symbol $\wedge$ represents the meet (or intersection) operator in the double algebra (see literature section of this chapter). However, for the present purposes the reader need only consider it as a shorthand for the determinant.

We start with the quadrifocal tensor for which the derivation is easiest. Consider four lines $\mathbf{l}$, $\mathbf{l}'$, $\mathbf{l}''$ and $\mathbf{l}'''$ in images formed from four cameras with camera matrices $\mathtt{A}$, $\mathtt{B}$, $\mathtt{C}$ and $\mathtt{D}$. The back projection of a line $\mathbf{l}$ through camera $\mathtt{A}$ is written as the plane $l_i \mathbf{a}^i$, with notation as in (17.14). The condition that these four planes are coincident may be written as

$$(l_p \mathbf{a}^p) \wedge (l'_q \mathbf{b}^q) \wedge (l''_r \mathbf{c}^r) \wedge (l'''_s \mathbf{d}^s) = 0.$$

However, since the determinant is linear in each row, this may be written as

$$0 = l_p l'_q l''_r l'''_s (\mathbf{a}^p \wedge \mathbf{b}^q \wedge \mathbf{c}^r \wedge \mathbf{d}^s) \stackrel{\text{def}}{=} l_p l'_q l''_r l'''_s Q^{pqrs}. \tag{17.23}$$

This corresponds to the definition (17.21) and line relation (17.22) for the quadrifocal tensor. The basic geometric property is the intersection of the four planes in space.

Trifocal tensor derivation. Consider now a point–line–line relationship $x^i \leftrightarrow l'_j \leftrightarrow l''_k$ for three views and let l^1_p and l^2_q be two lines in the first image that pass through the image point $\mathbf{x}$. The planes back-projected from the four lines meet in a point (see figure 17.2). So we can write:

$$l^1_l l^2_m l'_q l''_r (\mathbf{a}^l \wedge \mathbf{a}^m \wedge \mathbf{b}^q \wedge \mathbf{c}^r) = 0\,.$$

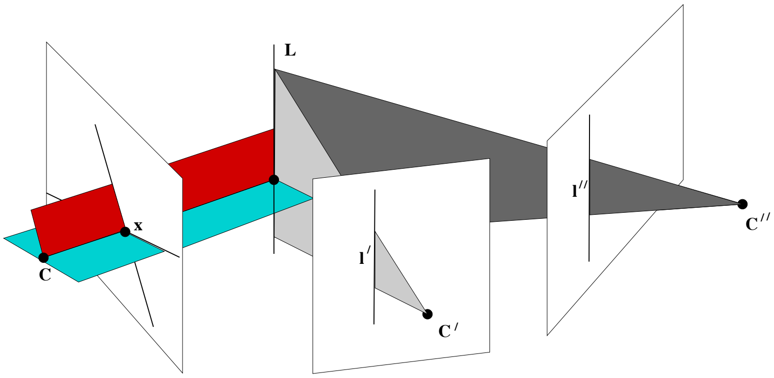


Fig. 17.2. *A point–line–line correspondence* $\mathbf{x} \leftrightarrow \mathbf{l}' \leftrightarrow \mathbf{l}''$ *involving three images may be interpreted as follows. Two arbitrary lines are chosen to pass through the point* $\mathbf{x}$ *in the first image. The four lines then back-project to planes that meet in a point in space.*

The next step is an algebraic trick – to multiply this equation by the $\epsilon^{ilm}\epsilon_{ilm}$. This is a scalar value (in fact equal to 6, the number of permutations of (ilm)). The result after regrouping is

$$\left(l^1_l l^2_m \epsilon^{ilm}\right) l'_q l''_r \epsilon_{ilm} \left(\mathbf{a}^l \wedge \mathbf{a}^m \wedge \mathbf{b}^q \wedge \mathbf{c}^r\right) = 0\,.$$

Now the expression $l^1_l l^2_m \epsilon^{ilm}$ is simply the cross-product of the two lines l_l and l_m, in other words their intersection point, x^i. Thus finally we can write

$$0 = x^i l'_q l''_r \left(\epsilon_{ilm}(\mathbf{a}^l \wedge \mathbf{a}^m \wedge \mathbf{b}^q \wedge \mathbf{c}^r)\right) \stackrel{\text{def}}{=} x^i l'_q l''_r \mathcal{T}_i^{qr} \tag{17.24}$$

which are the definition (17.10) and basic incidence relation (17.15) for the trifocal tensor.

Fundamental matrix. We can derive the fundamental matrix in the same manner. Given a correspondence $\mathbf{x} \leftrightarrow \mathbf{x}'$, select pairs of lines l^1_p and l^2_q passing through $\mathbf{x}$, and l'^1_r and l'^2_s passing through $\mathbf{x}'$. The back-projected planes all meet in a point, so we write

$$l^1_p\, l^2_q\, l'^1_r\, l'^2_s (\mathbf{a}^p \wedge \mathbf{a}^q \wedge \mathbf{b}^r \wedge \mathbf{b}^s) = 0\,.$$

Multiplying by $(\epsilon_{ipq}\epsilon^{ipq})(\epsilon_{jrs}\epsilon^{jrs})$ and proceeding as before leads to the coplanarity constraint

$$0 = x^i x'^j \left(\epsilon_{ipq}\epsilon_{jrs}(\mathbf{a}^p \wedge \mathbf{a}^q \wedge \mathbf{b}^r \wedge \mathbf{b}^s)\right) \stackrel{\text{def}}{=} x^i x'^j F_{ji} \tag{17.25}$$

which can be compared with (17.4).

17.5 Counting arguments

In this section we specify the number of points or lines required to carry out reconstruction from several views. This analysis is related to counting the number of degrees of freedom of the associated tensors. However, in doing this it is necessary to distinguish between the number of degrees of freedom of the tensor viewed as an unconstrained algebraic object, and the number of degrees of freedom that arises from a configuration of cameras and their camera matrices.

For instance, consider the fundamental matrix. At one level, the fundamental matrix may be considered as a homogeneous 3×3 matrix, and hence has 8 degrees of freedom (9 minus 1 for the indeterminate scale). On the other hand, the fundamental matrix arising from a pair of camera matrices according to (17.3–*p*412) must satisfy the additional constraint $\det \mathtt{F} = 0$. Hence, such a fundamental matrix has just 7 degrees of freedom. Since the camera matrices may be determined up to a 3D projectivity from the fundamental matrix (and vice versa), we may count the number of degrees of freedom of a fundamental matrix by counting the degrees of freedom of the camera matrices. Two camera matrices have a total of 22 degrees of freedom (two homogeneous 3×4 matrices). A 3D homography is represented by a 4×4 homogeneous matrix, and so has 15 degrees of freedom. This gives a total of $7 = 22 - 15$ degrees of freedom for the configuration of the two cameras, modulo a projective transformation. This agrees with the 7 degrees of freedom of the fundamental matrix, providing a check on the previous calculation.

Similarly, the trifocal tensor encodes the projective structure of three camera matrices, and hence has $18 = 3{*}11{-}15$ degrees of freedom. In the same way the quadrifocal tensor has $29 = 4 * 11 - 15$ degrees of freedom. Generally for m cameras, we have

$$\#\,\text{dof} = 11m - 15.$$

Since the trifocal and quadrifocal tensors, considered just as homogeneous algebraic arrays have 26 and 80 degrees of freedom respectively, they must satisfy an additional set of constraints, dictated by the camera geometry – 8 constraints for the trifocal tensor and 51 constraints for the quadrifocal tensor.

In computing geometric structure, we may use linear algebraic methods based on estimating the multifocal tensor by solving the constraints from the multi-linearities. The number of correspondences required is determined from the number of equations generated by each point or line correspondence. This *linear method* takes no account of the constraints imposed on the tensors by their geometry.

On the other hand, the number of correspondences required may be determined by counting the number of geometric constraints given by each correspondence and comparing with the total number of degrees of freedom of the system. Consider a configuration of n points in m views. The total number of degrees of freedom of this system is $11m - 15 + 3n$, since each of the n 3D points has 3 degrees of freedom. To estimate the projective structure the available data are the images of the n points in m images, a total of $2mn$ measurements (each 2D point having two coordinates). Thus for reconstruction to be possible, we require $2mn \geq 11m - 15 + 3n$, or $(2m-3)n \geq 11m - 15$. Thus the required number of points is

$$n \geq \frac{11m - 15}{2m - 3} = 5 + \frac{m}{2m - 3}.$$

One may also think of this by observing that each point correspondence contributes

# views	tensor	# elems	# dof	linear		non-linear	
				# points	# lines	# points	# lines
2	$\mathtt{F}$	9	7	8	—	7^*	—
3	$\mathcal{T}$	27	18	7	13	6^*	9^*?
4	$\mathtt{Q}$	81	29	6	9	6	8?

Table 17.3. **Projective degrees of freedom and constraints.** *The linear column indicates the minimum number of correspondences across all views required to solve linearly for the tensor (up to scale). The non-linear is the minimum number of correspondences required. A star indicates multiple solutions, and a question-mark indicates that no practical reconstruction method is known.*

$2m - 3$ constraints on the cameras, that is $2m$ for the coordinates of the points in each view, less 3 for the added degree of freedom of the 3D point.

An analogous argument applies for line correspondences. A line has four degrees of freedom in 3-space, and its image line is described by 2 degrees of freedom, so that each line correspondence provides $2m - 4$ constraints and the number of lines l required is

$$l \geq \frac{11m - 15}{2m - 4}.$$

In either of these cases, if the number of constraints (equations) is equal to the number of degrees of freedom (unknowns) of the cameras and points or lines configuration, then we can in general expect multiple solutions, except in the linear case. However, if there are more equations than unknowns, the system is over-determined, and in the generic case there will exist a single solution.

Table 17.3 summarizes the number of correspondences required for reconstruction. A star (*) indicates that multiple solutions are expected. For the case of non-linear solutions, actual methods of solution are not always known, other than by brute-force generation and solution of a set of simultaneous polynomial equations. Situations where no simple method is known are marked with a question-mark. Note that not much is known about non-linear solutions for lines. Specific non-linear algorithms are:

(i) 7 points in 2 views: see section 11.1.2(p281). Three solutions are possible.

(ii) 6 points in 3 views: see section 20.2.4(p510). Three solutions are possible. This is the dual (see chapter 20) of the previous problem.

(iii) 6 points in 4 views. Solve using 6 points in 3 views, then use the camera resection DLT algorithm (section 7.1(p178)) to find the fourth view and eliminate all but one solution. However, unlike the previous two cases, the case of 6 points in 4 views is overdetermined, and this solution is only valid for perfect data. The estimation for noisy data is discussed in chapter 20.

# views	tensor	# non-zero elements	# dof	linear		non-linear	
				# points	# lines	# points	# lines
2	$\mathtt{F}_{\mathrm{A}}$	5	4	4	—	4	—
3	$\mathcal{T}_{\mathrm{A}}$	16	12	4	8	4	6*?
4	$\mathtt{Q}_{\mathrm{A}}$	48	20	4	6	4	5*?

Table 17.4. **Affine degrees of freedom and constraints.** *The camera for each view is affine. See caption of table 17.3 for details.*

17.5.1 Affine cameras

For affine cameras, the reconstruction problem requires fewer correspondences. The plane at infinity may be identified in a reconstruction, being the principal plane of all of the cameras, and the reconstruction ambiguity is affine, not projective. The number of degrees of freedom for m affine cameras is

$$\#\,\text{dof} \;= 8m - 12.$$

Each view adds 8 degrees of freedom for the new 3×4 affine camera matrix, and we subtract 12 for the degrees of freedom of a 3D affine transformation.

As in the projective case a point correspondence imposes $2m - 3$ constraints, and a line correspondence $2m - 4$ constraints. As before the number of required points may be computed as $n(2m - 3) \geq 8m - 12$, i.e.

$$n \geq \frac{8m - 12}{2m - 3} = 4.$$

For lines the result is

$$l \geq \frac{8m - 12}{2m - 4} = 4 + \frac{2}{m - 2}.$$

As for linear methods, the number of elements in the tensor is 3^m. In the affine case the tensor is, as always, only defined up to scale, but additionally many of the elements are zero, as seen in section 17.1.2 and section 17.2.4. This cuts down on the required number of correspondences. Counting results are given in table 17.4. Note that for point correspondences, the linear algorithms apply with the minimum number of correspondences given by the above equation. Hence, the non-linear algorithms are the same as the linear ones.

17.5.2 Knowing four coplanar points – plane plus parallax

The previous counting arguments were for the case of points and lines in general position. Here we consider the important case that four or more of the 3D points are coplanar. It will be see that the computation of the tensors, and consequently projective structure, is significantly simplified. This discussion is given in terms of knowing

# views	tensor	# dof	points		lines	
			# points	# constraints	# lines	# constraints
2	F	2	2	2	—	
3	$\mathcal{T}$	5	2	5	5	5
4	Q	8	2	8	4	8

Table 17.5. *Number of additional correspondences required to compute projective structure given 2D homographies between the views, induced by a plane. The homographies may be computed from four or more point matches derived from coplanar points, or by any other means. Points used to compute the homographies are not counted in this table.*

four coplanar points in the images. However, all that is really important here is that inter-image homographies induced by a plane should be known.

Computation of the fundamental matrix knowing four coplanar points was considered in section 13.3(*p*334) (see algorithm 13.2(*p*336)), and for the trifocal tensor in section 16.7.1. Now, we consider the four-view case. From the correspondence of 4 (or more) points derived from 3D points in a plane, we can compute the homographies $\mathtt{H}'$, $\mathtt{H}''$ and $\mathtt{H}'''$ from the first view to the second, third and fourth views respectively, induced by the plane. In a projective reconstruction, we choose the plane containing the points as the plane at infinity, in which case $\mathtt{H}'$, $\mathtt{H}''$ and $\mathtt{H}'''$ are the infinite homographies. Assuming further that the first image is at the origin, the four camera matrices may now be written as

$$\mathtt{A} = [\mathtt{I} \mid \mathbf{0}] \quad \mathtt{B} = [\mathtt{H}'|\mathbf{t}'] \quad \mathtt{C} = [\mathtt{H}''|\mathbf{t}''] \quad \mathtt{D} = [\mathtt{H}'''|\mathbf{t}''']$$

The vectors $\mathbf{t}', \mathbf{t}'', \mathbf{t}'''$ may be determined only up to a common scale.

Since the left hand 3×3 blocks of the camera matrices are now known, it is easily seen from (17.21) that the entries of Q^{pqrs} are linear in the remaining entries $\mathbf{t}'$, $\mathbf{t}''$ and $\mathbf{t}'''$ of the camera matrices. Using (17.21) we can write down an expression for the entries of Q^{pqrs} in terms of the entries of $\mathbf{t}'$, $\mathbf{t}''$ and $\mathbf{t}'''$. In fact, we may write this explicitly as $\mathbf{q} = \mathtt{M}\mathbf{t}$, where $\mathtt{M}$ is an 81×9 matrix and $\mathbf{q}$ and $\mathbf{t}$ are vectors representing the entries of Q and $\mathbf{t}', \ldots, \mathbf{t}'''$. Thus, the quadrifocal tensor may be linearly parametrized in terms of 9 homogeneous coordinates and hence has 8 degrees of freedom.

Now, given a set of equations $\mathtt{E}\mathbf{q} = \mathbf{0}$ derived from correspondences as in table 17.2, they may be rewritten in terms of the minimal parameter set $\mathbf{t}$ by substituting $\mathbf{q} = \mathtt{M}\mathbf{t}$ arriving at $\mathtt{EM}\mathbf{t} = \mathbf{0}$. This allows a linear solution for $\mathbf{t}$ and hence the camera matrices, and (if desired) the tensor, $\mathbf{q} = \mathtt{M}\mathbf{t}$. Note the important advantage here that the tensor so obtained automatically corresponds to a set of camera matrices, and so satisfies all geometric constraints – the troublesome 51 constraints that must be satisfied by the quadrifocal tensor disappear into thin air.

The above analysis was for the computation of the quadrifocal tensor, but it applies equally well to the fundamental matrix and trifocal tensor as well.

Counting argument. We return to the general case of m views and consider the situation from a geometric point of view. The number of degrees of freedom parametrizing the tensor is equal to the number of geometric degrees of freedom remaining in the camera matrices, namely $3m - 4$ where m is the number of views. This arises from $3(m - 1)$ for the last column of each camera matrix apart from the first, less one for the common scaling.

Counting the number of constraints imposed by a point or a line correspondence is a little tricky, however, and shows that one must always be careful with counting arguments not to neglect hidden dependencies. First, consider a line correspondence $\mathbf{l} \leftrightarrow \mathbf{l}' \leftrightarrow \mathbf{l}''$ across three views – the argument will hold in general for $m \geq 3$ views. The question to be resolved is just how much information can be derived from the measurement of the image lines. Surprisingly, knowledge of the plane homographies between the images reduces the amount of information that a line correspondence provides. For simplicity of argument, we may assume that planar homographies have been applied to the images so that the four coplanar reference points map to the same point in each image. As a result any further point in the reference plane will map to the same point in each image. Now the 3D line $\mathbf{L}$ from which the line correspondence is derived must meet the reference plane in a point $\mathbf{X}$. Since it lies on the reference plane, $\mathbf{X}$ projects to the same point $\mathbf{x}$ in all images, and $\mathbf{x}$ must lie on all three image lines $\mathbf{l}, \mathbf{l}', \mathbf{l}''$. Thus, corresponding lines in the three views cannot be arbitrary – they must all pass through a common image point. In the general case the number of degrees of freedom of the measurements of the lines in m views is $m + 2$. To specify the point $\mathbf{x}$ requires 2 degrees of freedom, and each line through the point then has one remaining degree of freedom (its orientation). Subtracting 4 for the four added degrees of freedom of the line in space, we see

- *Each line correspondence across m views generates $m - 2$ constraints on the remaining degrees of freedom of the cameras.*

Note how the condition that the image lines must meet in a point restricts the number of equations generated by the line correspondence. Without observing this condition, one would expect $2m - 4$ constraints from each line. However, with perfect data the set of $2m - 4$ equations will have rank only $m - 2$. With noisy data the image lines will not be exactly coincident, and the system may have full rank, but this will be entirely due to noise, and the smallest singular values of the system will be essentially random. One should, however, include all available equations in solving the system, since this will diminish the noise effects.

For point correspondences the argument is similar. The line through any two image points is the image of a line in space. The projections of this 3D line are constrained as discussed in the preceding discussion, and this imposes constraints on the matching points. The measurements have $3m + 2$ degrees of freedom: 2 for the intersection point of the line with the plane and, in each view, one for the line orientation and two corresponding to the position of each of the two points on the line. Subtracting 6 for the degrees of freedom of the two points in 3-space, the result is

- *Two point correspondences across m views generate $3m-4$ constraints on the remaining degrees of freedom of the cameras.*

Since this is the same as the number of degrees of freedom of the cameras, 2 points are sufficient to compute structure.

Since the number of geometric constraints on the cameras is the same as the number of degrees of freedom of the tensor, there is no distinction between constraints on the tensor and constraints on the geometry. Thus, point and line correspondences may be used to generate linear constraints on the tensor. There is no need for non-linear methods. The number of required correspondences is summarized in table 17.5.

17.6 Number of independent equations

It was asserted in table 17.2 that each four-view point correspondence gives rise to 16 linearly independent equations in the entries of the quadrifocal tensor. We now examine this point more closely.

Given sufficiently many point matches across four views, one may solve for the tensor Q^{pqrs}. Once Q is known, it is possible to solve for the camera matrices and hence compute projective structure. This was shown in [Heyden-95b, Heyden-97c, Hartley-98c] but is not discussed further in this book. A curious phenomenon occurs however when one counts the number of point matches necessary to do this. As indicated above, it appears to be the case that each point match gives 16 linearly independent equations in the entries of the tensor Q^{pqrs}, and it seems unlikely that the equations derived from two totally unrelated sets of point correspondences could have any dependencies. It would therefore appear that from five point correspondences one obtains 80 equations, which is enough to solve for the 81 entries of Q^{pqrs} up to scale. From this argument it would seem that it is possible to solve for the tensor from only five point matches across four views, and thence one may solve for the camera matrices, up to the usual projective ambiguity. This conclusion however is contradicted by the following remark.

- *It is not possible to determine the positions of four (or any number of) cameras from the images of five points.*

This follows from the counting arguments of section 17.5. Obviously there is some error in our counting of equations. The truth is contained in the following.

Result 17.1. *The full set of 81 linear equations (17.20) derived from a single point correspondence $\mathbf{x} \leftrightarrow \mathbf{x}' \leftrightarrow \mathbf{x}'' \leftrightarrow \mathbf{x}'''$ across four views contains 16 independent constraints on Q^{pqrs}. Furthermore, let the equations be written as $\mathtt{A}\mathbf{q} = \mathbf{0}$ where $\mathtt{A}$ is an 81×81 matrix and $\mathbf{q}$ is a vector containing the entries of Q^{pqrs}. Then the 16 non-zero singular values of $\mathtt{A}$ are all equal.*

What this result is saying is that indeed as expected one obtains 16 linearly independent equations from one point correspondence, and in fact it is possible to reduce this set of equations by an orthogonal transformation (multiplication of the equation matrix $\mathtt{A}$ on the left by an orthogonal matrix $\mathtt{U}$) to a set of 16 orthogonal equations. The proof is postponed until the end of this section. The surprising fact however is that the equation sets corresponding to two unrelated point correspondences have a dependency, as stated in the following result.

Result 17.2. *The set of equations (17.20–p419) derived from a set of n general point correspondences across four views has rank $16n - \binom{n}{2}$, for $n \leq 5$.*

The notation $\binom{n}{2}$ means the number of choices of 2 among n, specifically, $\binom{n}{2} = n(n-1)/2$. Thus for 5 points there are only 70 independent equations, not enough to solve for Q^{pqrs}. For $n = 6$ points, $16n - \binom{n}{2} = 81$, and we have enough equations to solve for the 81 entries of Q^{pqrs}.

We now prove the two results above. The key point in the proof of result 17.1 concerns the singular values of a skew-symmetric matrix (see result A4.1(p581)).

Result 17.3. *A 3×3 skew-symmetric matrix has two equal non-zero singular values.*

The rest of the proof of result 17.1 is quite straightforward as long as one does not get lost in notation.

Proof. (Result 17.1) The full set of 81 equations derived from a single point correspondence is of the form $x^i \epsilon_{ipw} x'^j \epsilon_{jqx} x''^k \epsilon_{kryv} x'''^l \epsilon_{lsz} Q^{pqrs} = 0_{wxyz}$. A total of 81 equations are generated by varying w, x, y, z over the range $1, \ldots, 3$. Thus, the equation matrix $\mathtt{A}$ may be written as

$$\mathtt{A}_{(wxyz)(pqrs)} = x^i \epsilon_{ipw} x'^j \epsilon_{jqx} x''^k \epsilon_{kry} x'''^l \epsilon_{lsz} \tag{17.26}$$

where the indices $(wxyz)$ index the row and $(pqrs)$ index the column of $\mathtt{A}$. We will consider a set of indices, such as $(wxyz)$ in this case, as a single index for the row or column of a matrix. This situation will be indicated by enclosing the indices in parentheses as here, and referring to them as a *combined index*.

We consider now the expression $x^i \epsilon_{ipw}$. This may be considered as a matrix indexed by the free indices p and w. Furthermore, since $x^i \epsilon_{ipw} = -x^i \epsilon_{iwp}$ we see that it is a skew-symmetric matrix, and hence has equal singular values. We denote this matrix by $\mathtt{S}_{wp}$. Writing result 17.3 using tensor notation, we have

$$\mathtt{U}^w_a \mathtt{S}_{wp} \mathtt{V}^p_e = k \mathtt{D}_{ae} \tag{17.27}$$

where the diagonal matrix $\mathtt{D}$ is as in result 17.3. The matrix $\mathtt{A}$ in (17.26) may be written as $\mathtt{A}_{(wxyz)(pqrs)} = \mathtt{S}_{wp} \mathtt{S}'_{xq} \mathtt{S}''_{yr} \mathtt{S}'''_{zs}$. Consequently, applying (17.27) we may write

$$\mathtt{U}^w_a \mathtt{U}'^x_b \mathtt{U}''^y_c \mathtt{U}'''^z_d \mathtt{A}_{(wxyz)(pqrs)} \mathtt{V}^p_e \mathtt{V}'^q_f \mathtt{V}''^r_g \mathtt{V}'''^s_h = k k' k'' k''' \mathtt{D}_{ae} \mathtt{D}_{bf} \mathtt{D}_{cg} \mathtt{D}_{dh}. \tag{17.28}$$

Now, writing

$$\hat{\mathtt{U}}^{(wxyz)}_{(abcd)} = \mathtt{U}^w_a \mathtt{U}'^x_b \mathtt{U}''^y_c \mathtt{U}'''^z_d \qquad \hat{\mathtt{V}}^{(pqrs)}_{(efgh)} = \mathtt{V}^p_e \mathtt{V}'^q_f \mathtt{V}''^r_g \mathtt{V}'''^s_h \qquad \hat{\mathtt{D}}_{(abcd)(efgh)} = \mathtt{D}_{ae} \mathtt{D}_{bf} \mathtt{D}_{cg} \mathtt{D}_{dh}$$

and $\hat{k} = k k' k'' k'''$ we see that (17.28) may be written as

$$\hat{\mathtt{U}}^{(wxyz)}_{(abcd)} \mathtt{A}_{(wxyz)(pqrs)} \hat{\mathtt{V}}^{(pqrs)}_{(efgh)} = \hat{k} \hat{\mathtt{D}}_{(abcd)(efgh)}. \tag{17.29}$$

As a matrix, $\mathtt{D}_{(abcd)(efgh)}$ is diagonal with 16 non-zero diagonal entries, all equal to unity. To show that (17.29) is the SVD of the matrix $\mathtt{A}_{(wxyz)(pqrs)}$, and hence to complete the proof, it remains only to show that $\mathtt{U}^{(wxyz)}_{(abcd)}$ and $\mathtt{V}^{(pqrs)}_{(efgh)}$ are orthogonal matrices. It is necessary only to show that the columns are orthonormal. This is straightforward and is left as an exercise. □

Proof. (Result 17.2) We consider two point correspondences across four views, namely $x^i \leftrightarrow x'^i \leftrightarrow x''^i \leftrightarrow x'''^i$ and $y^i \leftrightarrow y'^i \leftrightarrow y''^i \leftrightarrow y'''^i$. These give rise to two sets of equations $\mathtt{A}^{\mathbf{x}}\mathbf{q} = \mathbf{0}$ and $\mathtt{A}^{\mathbf{y}}\mathbf{q} = \mathbf{0}$ of the form (17.26). Each of these matrices has rank 16, and if the rows of $\mathtt{A}^{\mathbf{x}}$ are independent of the rows of $\mathtt{A}^{\mathbf{y}}$, then the rank of the combined set of equations is 32. However, if there is a linear dependence between the rows of $\mathtt{A}^{\mathbf{x}}$ and those of $\mathtt{A}^{\mathbf{y}}$ then their combined rank is at most 31.

We define a vector $\mathbf{s}_{\mathbf{x}}$ with combined index $(pqrs)$ by $\mathbf{s}_{\mathbf{x}}^{(pqrs)} = x^p x'^q x''^r x'''^s$. A vector $\mathbf{s}_{\mathbf{y}}$ is similarly defined. We will demonstrate that $\mathbf{s}_{\mathbf{y}}^{\mathsf{T}}\mathtt{A}^{\mathbf{x}} = \mathbf{s}_{\mathbf{x}}^{\mathsf{T}}\mathtt{A}^{\mathbf{y}}$, which means that the rows of $\mathtt{A}^{\mathbf{x}}$ and $\mathtt{A}^{\mathbf{y}}$ are linearly dependent.

Expanding $\mathbf{s}_{\mathbf{y}}^{\mathsf{T}}\mathtt{A}^{\mathbf{x}}$ gives

$$\begin{aligned}
\mathbf{s}_{\mathbf{y}}^{\mathsf{T}}\mathtt{A}^{\mathbf{x}} &= \mathbf{s}_{\mathbf{y}}^{(wxyz)}\mathtt{A}^{\mathbf{x}}_{(wxyz)(pqrs)} \\
&= (y^w y''^x y'''^y y''''^z) x^i \epsilon_{ipw} x'^j \epsilon_{jqx} x''^k \epsilon_{kry} x'''^l \epsilon_{lsz} \\
&= (x^i x''^j x'''^k x''''^l) y^w \epsilon_{wpi} y'^x \epsilon_{xqj} y''^y \epsilon_{yrk} y'''^z \epsilon_{zsl} \\
&= \mathbf{s}_{\mathbf{x}}^{\mathsf{T}}\mathtt{A}^{\mathbf{y}}.
\end{aligned}$$

This demonstrates that the rows of $\mathtt{A}^{\mathbf{x}}$ and $\mathtt{A}^{\mathbf{y}}$ are dependent, and their combined rank is at most 31. We now consider the possibility that the combined rank is less than 31. In such a case, all 31×31 subdeterminants of the matrix $[\mathtt{A}^{\mathbf{x}\mathsf{T}}, \mathtt{A}^{\mathbf{y}\mathsf{T}}]^{\mathsf{T}}$ must vanish. These subdeterminants may be expressed as polynomial expressions in the coefficients of the points $\mathbf{x}, \mathbf{x}', \mathbf{x}'', \mathbf{x}''', \mathbf{y}, \mathbf{y}', \mathbf{y}''$ and $\mathbf{y}'''$. These 24 coefficients together generate a 24-dimensional space. Thus, there is a function $f : \mathbb{R}^{24} \rightarrow \mathbb{R}^N$ for some N (equal to the number of such 31×31 subdeterminants), such that the equation matrix has rank less than 31 only on the set of zeros of the function f. Any arbitrarily chosen example (omitted) may be used to show that the function f is not identically zero. It follows that the set of point correspondences for which the set of equations has rank less than 31 is a variety in $\mathbb{R}^{24}$, and hence is nowhere dense. Thus, the set of equations generated by a general pair of point correspondences across four views has rank 31.

We now turn to the general case of n point correspondences across all four views. Note that the linear relationship that holds for two point correspondences is non-generic, but depends on the pair of correspondences. In general, therefore, given n point correspondences, there will be $\binom{n}{2}$ such relationships. This reduces the dimension of the space spanned by the set of equations to $16n - \binom{n}{2}$ as required. □

The three-view case. Similar arguments hold in the three-view case. It can be proved in the same manner that the nine equations arising from a point match contain four independent equations. This is left to an exercise (see page 432).

17.7 Choosing equations

In section 17.6 a proof was given that the singular values of the full set of equations derived from four point equations are all equal. The proof may easily be adapted to the three-view case as well (see exercise, page 432). The key point in the argument is that the two non-zero singular values of a 3×3 skew-symmetric matrix are equal. This proof may clearly be extended to apply to any of the other sets of equations derived from line or point correspondences given in sections 17.2 and 17.3.

We consider still the four-view case. The results on singular values show that it is in general advisable to include all 81 equations derived from this correspondence, rather than selecting just 16 independent equations. This will avoid difficulties with near singular situations. This conclusion is supported by experimental observation. Indeed, numerical examples show that the condition of a set of equations derived from several point correspondences is substantially better when all equations are included for each point correspondence. In this context, the condition of the equation set is given by the ratio of the first (largest) to the n-th singular value, where n is the number of linearly independent equations.

Including all 81 equations rather than just 16 means that the set of equations is larger, leading to increased complexity of solution. This can be remedied as follows. The basis for the equality of the singular values is that $\mathtt{S}_{wp} = x^i \epsilon_{ipw}$ and the other similarly defined terms are skew symmetric matrices, and hence have equal singular values. The same effect can be achieved by any other matrix $\mathtt{S}$ with equal singular values. We require only that the columns of $\mathtt{S}$ should represent lines passing through the point $\mathbf{x}$ (otherwise stated $\mathtt{S}_{wp} x^w = 0_p$). Matrix $\mathtt{S}$ will have equal singular values if its columns are orthonormal. These conditions may be achieved with $\mathtt{S}$ being a 3×2 matrix. If this is done for the point in each view the total number of equations will be reduced from $3^4 = 81$ to $2^4 = 16$, and the 16 equations will be orthogonal. A convenient way of choosing $\mathtt{S}$ is to use a Householder matrix (see section A4.1.2(p580)) as shown below.

This discussion also applies to the trifocal tensor, allowing us to reduce the number of equations from 9 to 4, while retaining equal singular values. Summarizing this discussion for the trifocal tensor case:

- *Given a point correspondence $\mathbf{x} \leftrightarrow \mathbf{x}' \leftrightarrow \mathbf{x}''$ across three views, generate equations of the form*

$$x^i \hat{l}'_{qx} \hat{l}''_{ry} \mathcal{T}_i^{qr} = 0_{xy} \textit{ for } x, y = 1, 2$$

where $\hat{l}'_{q1}$ and $\hat{l}'_{q2}$ are two lines through $\mathbf{x}'$ represented by orthonormal vectors (and similarly $\hat{l}''_{ry}$). A convenient way to find $\hat{l}'_{qx}$ and $\hat{l}''_{ry}$ is as follows.

  (i) *Find Householder matrices h'_{qx} and h''_{ry} such that $x'^q h'_{qx} = \delta_{3x}$ and $x''^r h'_{ry} = \delta_{3y}$.*

  (ii) *For $x, y = 1, 2$ set $\hat{l}'_{qx} = h'_{qx}$ and $\hat{l}''_{ry} = h''_{ry}$.*

It is evident that essentially this method will work for all the types of equations summarized in table 17.1(p417) and table 17.2(p421).

This chapter suggests that the most basic relations are the point–line–line correspondence equation $x^i l'_q l''_r \mathcal{T}_i^{qr} = 0$ in the three-view case, and the line-correspondence equation $l_p l'_q l''_r l'''_s Q^{pqrs} = 0$ for four views. Indeed numerical robustness may be enhanced by reducing other correspondences to this type of correspondence, for carefully selected lines.

17.8 Closure

17.8.1 The literature

Although using a slightly different approach, this chapter summarizes previous results of [Triggs-95] and Faugeras and Mourrain [Faugeras-95a] on the derivation of multilinear relationships between corresponding image coordinates. The formulae for relations between mixed point and line correspondences are extensions of the results of [Hartley-95b, Hartley-97a].

The enumeration of the complete set of multilinear relations given in table 17.1(*p*417) and table 17.2(*p*421), formulae for the multifocal tensors, and the analysis of the number of independent equations derived from point correspondences, are adapted from [Hartley-95a]. A similar analysis of the multifocal tensors has also appeared in [Heyden-98].

The quadrifocal tensor was probably first discovered by [Triggs-95]. The quadrilinear constraints and their associated tensor have been described in several papers [Triggs-95, Faugeras-95a, Shashua-95b, Heyden-95b, Heyden-97c].

The double (or Grassmann–Cayley) algebra was introduced into the computer vision literature in [Carlsson-93], see also [Faugeras-95a, Faugeras-97] for further applications.

An algorithm for computing the quadrifocal tensor and an algorithm for reconstructing it based on the reduced tensor was given in ([Heyden-95b, Heyden-97c]). A later paper [Hartley-98c] refined this algorithm.

17.8.2 Notes and exercises

(i) Determine the properties of the *affine quadrifocal tensor*, i.e. the quadrifocal tensor computed from affine camera matrices. In particular using the determinant definition of the tensor (17.21–*p*419), verify the number of non-zero elements given in table 17.4.

(ii) Show that the 9 linear equations (17.11–*p*415) derived from a single point correspondence $\mathbf{x} \leftrightarrow \mathbf{x}' \leftrightarrow \mathbf{x}''$ across three views contains 4 linearly independent equations. Furthermore, let the equations be written as $\mathtt{At} = \mathbf{0}$ where $\mathtt{A}$ is a 9×27 matrix. Then the 4 non-zero singular values of $\mathtt{A}$ are equal. Unlike the four-view case, the equations resulting from different point matches are linearly independent, so n point matches produce $4n$ independent equations.

(iii) If a canonical affine basis is chosen for the image coordinates such that three corresponding points have coordinates $(0, 0), (1, 0), (0, 1)$ in each view, then the resulting tensors have a simpler form. These "reduced tensors" have a greater number of zero elements than the general form of the tensors. For example, in the case of the reduced fundamental matrix the diagonal elements are zero, and the reduced trifocal tensor has only 15 nonzero entries. Also the tensors are specified by fewer parameters, e.g. four in the case of the reduced fundamental matrix, as effectively the basis points specify the other parameters. Further details are given in [Heyden-95b, Heyden-95a].

(iv) Show that if the four camera centres are coplanar then the quadrifocal tensor has 28 geometric degrees of freedom.