Lecture 10
Threeviews and more

Olivier Faugeras

Trifocal geometry

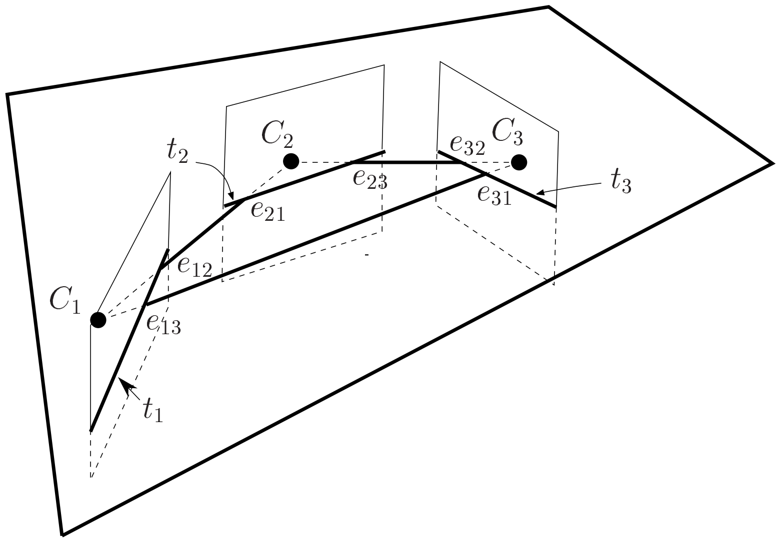


Figure 1: The trifocal geometry.

- The salient features of trifocal geometry are:
  1. The trifocal plane (when the three optical centers are not aligned).
  2. The six epipoles.
  3. The three trifocal lines.

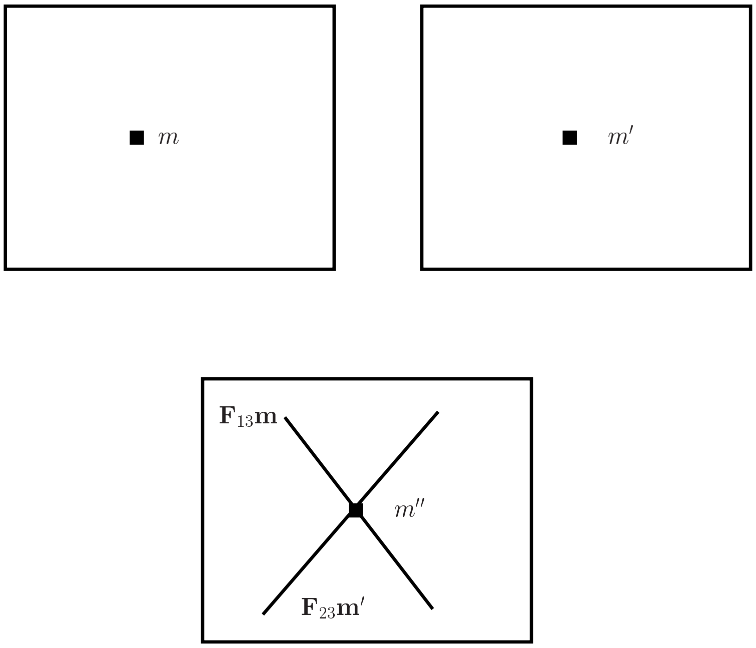


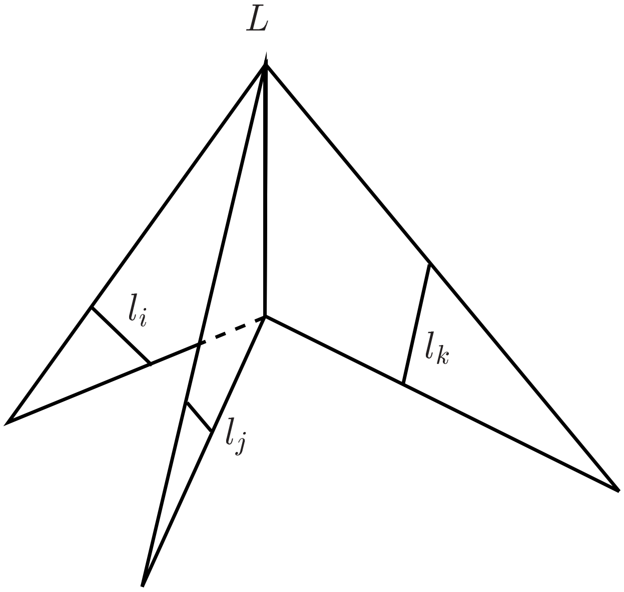
Figure 2: Transfer from views 1 and 2 to view 3.

- The knowledge of the three fundamental matrices allows in general to do prediction (also called transfer) from two of the views to the third.
- There are two cases when this fails:

1. When the points are on the trifocal lines.
2. When the optical centers are aligned.

- This is why we have to introduce the trifocal tensor.

The trifocal tensor



- Consider the projection matrices $\boldsymbol{\mathcal{P}}_i, i = 1, 2, 3$, a 3D line L with images l_i.
- Given two images l_j and l_k of L, L can be defined as the intersection (the meet) of the two planes $\boldsymbol{\mathcal{P}}_j^T \mathbf{l}_j$ and $\boldsymbol{\mathcal{P}}_k^T \mathbf{l}_k$:

$$\mathbf{L} = \boldsymbol{\mathcal{P}}_j^T \mathbf{l}_j \bigtriangleup \boldsymbol{\mathcal{P}}_k^T \mathbf{l}_k$$

The trifocal tensor (*continued*)

- Write the right-hand side of this equation explicitely in terms of the row vectors of the matrices $\boldsymbol{\mathcal{P}}_j$ and $\boldsymbol{\mathcal{P}}_k$ and the coordinates of $\mathbf{l}_j$ and $\mathbf{l}_k$:

$$\mathbf{L} = (l_j^1 \Gamma_j + l_j^2 \Lambda_j + l_j^3 \Theta_j) \vartriangle (l_k^1 \Gamma_k + l_k^2 \Lambda_k + l_k^3 \Theta_k)$$

- Expand the meet operator:

$$\mathbf{L} = \mathbf{l}_j^T \begin{bmatrix} \Gamma_j \vartriangle \Gamma_k & \Gamma_j \vartriangle \Lambda_k & \Gamma_j \vartriangle \Theta_k \\ \Lambda_j \vartriangle \Gamma_k & \Lambda_j \vartriangle \Lambda_k & \Lambda_j \vartriangle \Theta_k \\ \Theta_j \vartriangle \Lambda_k & \Theta_j \vartriangle \Lambda_k & \Theta_j \vartriangle \Theta_k \end{bmatrix} \mathbf{l}_k$$

- The image l_i of L is obtained by applying the matrix $\tilde{\boldsymbol{\mathcal{P}}}_i$ to the Plcker coordinates of L:

$$\mathbf{l}_i = \tilde{\boldsymbol{\mathcal{P}}}_i (\boldsymbol{\mathcal{P}}_j^T \mathbf{l}_j \vartriangle \boldsymbol{\mathcal{P}}_k^T \mathbf{l}_k)$$

- In a more compact form:

$$\mathbf{l}_i = \boldsymbol{\mathcal{T}}_i (\mathbf{l}_j, \mathbf{l}_k)$$

The trifocal tensor (*continued*)

- To be more explicit:

$$\mathbf{l}_i = \begin{bmatrix} \mathbf{l}_j^T \mathbf{G}_i^1 \mathbf{l}_k \\ \mathbf{l}_j^T \mathbf{G}_i^2 \mathbf{l}_k \\ \mathbf{l}_j^T \mathbf{G}_i^3 \mathbf{l}_k \end{bmatrix}$$

- The three 3×3 matrices $\mathbf{G}_i^n, n = 1, 2, 3$ can be written in terms of brackets:

$$\mathbf{G}_i^1 = \begin{bmatrix} [\Lambda_i, \Theta_i, \Gamma_j, \Gamma_k] & [\Lambda_i, \Theta_i, \Gamma_j, \Lambda_k] & [\Lambda_i, \Theta_i, \Gamma_j, \Theta_k] \\ [\Lambda_i, \Theta_i, \Lambda_j, \Gamma_k] & [\Lambda_i, \Theta_i, \Lambda_j, \Lambda_k] & [\Lambda_i, \Theta_i, \Lambda_j, \Theta_k] \\ [\Lambda_i, \Theta_i, \Theta_j, \Gamma_k] & [\Lambda_i, \Theta_i, \Theta_j, \Lambda_k] & [\Lambda_i, \Theta_i, \Theta_j, \Theta_k] \end{bmatrix}$$

- Those matrices have the following (important properties):

Proposition 1 *The matrices $\mathbf{G}_i^n$ are of rank 2 and their nullspaces are the three epipolar lines in the k-th retinal plane of the three projection rays of camera i. These three lines intersect at the epipole e_{ki}.*

Proposition 2 *The nullspaces of the matrices $\mathbf{G}_i^{nT}$ are the three epipolar lines, noted $l_j^n, n = 1, 2, 3$, in the j-th retinal plane of the three projection rays of camera i. These three lines intersect at the epipole e_{ji}.*

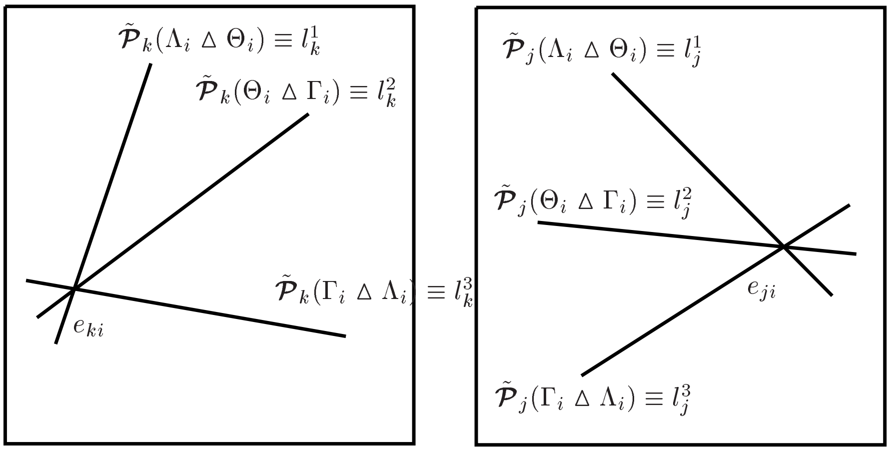


Figure 3: The lines l_j^n (resp. l_k^n), $n = 1, 2, 3$ in the nullspaces of the matrices $\mathbf{G}_i^{nT}$ (resp. $\mathbf{G}_i^n$) are the images of the three projection rays of camera i. Hence, they intersect at the epipole e_{ji} (resp. e_{ki}). The corresponding epipolar lines in camera i are obtained as $\mathcal{T}_i(l_j^n, l_k)$ (resp. $\mathcal{T}_i(l_j, l_k^n)$) for $l_k \neq l_k^n$ (resp. $l_j \neq l_j^n$).

The trifocal tensor (*continued*)

- Note that this allows us to predict the coordinates of a line l_i in image i given two images l_j and l_k of an unknown 3D line in images j and k, except in two cases:

  1. When the two planes determined by l_j and l_k are identical i.e. when l_j and l_k are corresponding epipolar lines between views j and k.
  2. When l_j and l_k are epipolar lines between views i and j and i and k, respectively. This is equivalent to saying that they are the images of the same optical ray in view i and that l_i is reduced to a point.

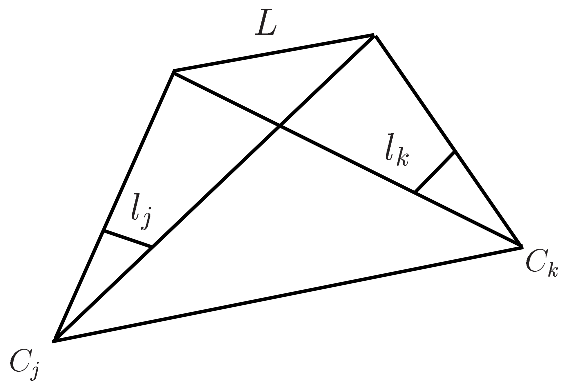


Figure 4: When l_j and l_k are corresponding epipolar lines, the two planes $\mathcal{P}_j^T \mathbf{l}_j$ and $\mathcal{P}_k^T \mathbf{l}_k$ are identical and therefore $\mathcal{T}_i(\mathbf{l}_j, \mathbf{l}_k) = \mathbf{0}$.

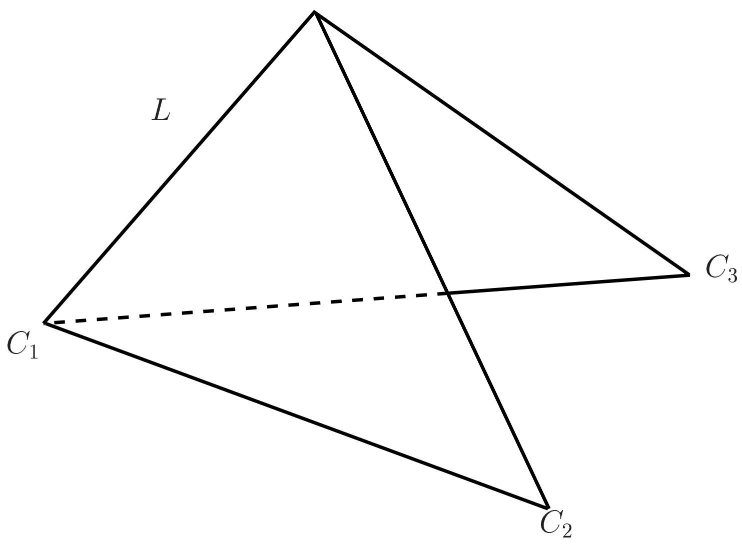


Figure 5: When l_j and l_k are epipolar lines with respect to view i, the line l_i is reduced to a point, hence $\mathcal{T}_i(\mathbf{l}_j, \mathbf{l}_k) = \mathbf{0}$.

A complementary approach: the trilinearities

- Let us write the reprojection equations as:

$$\begin{bmatrix} z_1\mathbf{\Gamma}_1 - x_1\mathbf{\Theta}_1 \\ y_1\mathbf{\Theta}_1 - z_1\mathbf{\Lambda}_1 \\ z_2\mathbf{\Gamma}_2 - x_2\mathbf{\Theta}_2 \\ y_2\mathbf{\Theta}_2 - z_2\mathbf{\Lambda}_2 \\ z_3\mathbf{\Gamma}_3 - x_3\mathbf{\Theta}_3 \\ y_3\mathbf{\Theta}_3 - z_3\mathbf{\Lambda}_3 \end{bmatrix} \mathbf{M} = \mathbf{A}\mathbf{M} = \mathbf{0}$$

- This system of linear equations must have at least one nonzero solution: the 6×4 matrix $\mathbf{A}$ must be of rank less than or equal to 3.
- All 4×4 minors have to be equal to 0.

A complementary approach: the trilinearities (*continued*)

- They fall into two categories:
  1. Those which contain two rows arising from one camera and two rows from another.
  2. Those which contain two rows arising from one camera, one row from a second camera and one row arising from the third camera.
- The three determinants of the first type yield the epipolar constraints.
- There are twelve determinants of the second type which fall into three groups noted I, II, and III, depending upon which camera provides two rows:

Group I rows (1,2,3,5), (1,2,3,6), (1,2,4,5), (1,2,4,6).

Group II rows (1,3,4,5), (1,3,4,6), (2,3,4,5), (2,3,4,6).

Group III rows (1,3,5,6), (2,3,5,6), (1,4,5,6), (2,4,5,6).

A complementary approach: the trilinearities (*continued*)

- A complete description:

Group I

$T_1^I = (x_1\mathbf{\Lambda}_1 \triangle \mathbf{\Theta}_1 + y_1\mathbf{\Theta}_1 \triangle \mathbf{\Gamma}_1 + z_1\mathbf{\Gamma}_1 \triangle \mathbf{\Lambda}_1) \triangle (z_2\mathbf{\Gamma}_2 - x_2\mathbf{\Theta}_2) \triangle (z_3\mathbf{\Gamma}_3 - x_3\mathbf{\Theta}_3)$
$T_2^I = (x_1\mathbf{\Lambda}_1 \triangle \mathbf{\Theta}_1 + y_1\mathbf{\Theta}_1 \triangle \mathbf{\Gamma}_1 + z_1\mathbf{\Gamma}_1 \triangle \mathbf{\Lambda}_1) \triangle (y_2\mathbf{\Theta}_2 - z_2\mathbf{\Lambda}_2) \triangle (z_3\mathbf{\Gamma}_3 - x_3\mathbf{\Theta}_3)$
$T_3^I = (x_1\mathbf{\Lambda}_1 \triangle \mathbf{\Theta}_1 + y_1\mathbf{\Theta}_1 \triangle \mathbf{\Gamma}_1 + z_1\mathbf{\Gamma}_1 \triangle \mathbf{\Lambda}_1) \triangle (z_2\mathbf{\Gamma}_2 - x_2\mathbf{\Theta}_2) \triangle (y_3\mathbf{\Theta}_3 - z_3\mathbf{\Lambda}_3)$
$T_4^I = (x_1\mathbf{\Lambda}_1 \triangle \mathbf{\Theta}_1 + y_1\mathbf{\Theta}_1 \triangle \mathbf{\Gamma}_1 + z_1\mathbf{\Gamma}_1 \triangle \mathbf{\Lambda}_1) \triangle (y_2\mathbf{\Theta}_2 - z_2\mathbf{\Lambda}_2) \triangle (y_3\mathbf{\Theta}_3 - z_3\mathbf{\Lambda}_3)$

Group II

$T_1^{II} = (z_1\mathbf{\Gamma}_1 - x_1\mathbf{\Theta}_1) \triangle (x_2\mathbf{\Gamma}_2 \triangle \mathbf{\Lambda}_2 + y_2\mathbf{\Lambda}_2 \triangle \mathbf{\Theta}_2 + z_2\mathbf{\Theta}_2 \triangle \mathbf{\Gamma}_2) \triangle (z_3\mathbf{\Gamma}_3 - x_3\mathbf{\Theta}_3)$
$T_2^{II} = (y_1\mathbf{\Theta}_1 - z_1\mathbf{\Lambda}_1) \triangle (x_2\mathbf{\Gamma}_2 \triangle \mathbf{\Lambda}_2 + y_2\mathbf{\Lambda}_2 \triangle \mathbf{\Theta}_2 + z_2\mathbf{\Theta}_2 \triangle \mathbf{\Gamma}_2) \triangle (z_3\mathbf{\Gamma}_3 - x_3\mathbf{\Theta}_3)$
$T_3^{II} = (z_1\mathbf{\Gamma}_1 - x_1\mathbf{\Theta}_1) \triangle (x_2\mathbf{\Gamma}_2 \triangle \mathbf{\Lambda}_2 + y_2\mathbf{\Lambda}_2 \triangle \mathbf{\Theta}_2 + z_2\mathbf{\Theta}_2 \triangle \mathbf{\Gamma}_2) \triangle (y_3\mathbf{\Theta}_3 - z_3\mathbf{\Lambda}_3)$
$T_4^{II} = (y_1\mathbf{\Theta}_1 - z_1\mathbf{\Lambda}_1) \triangle (x_2\mathbf{\Gamma}_2 \triangle \mathbf{\Lambda}_2 + y_2\mathbf{\Lambda}_2 \triangle \mathbf{\Theta}_2 + z_2\mathbf{\Theta}_2 \triangle \mathbf{\Gamma}_2) \triangle (y_3\mathbf{\Theta}_3 - z_3\mathbf{\Lambda}_3)$

Group III

$T_1^{III} = (z_1\mathbf{\Gamma}_1 - x_1\mathbf{\Theta}_1) \triangle (z_2\mathbf{\Gamma}_2 - x_2\mathbf{\Theta}_2) \triangle (x_3\mathbf{\Gamma}_3 \triangle \mathbf{\Lambda}_3 + y_3\mathbf{\Lambda}_3 \triangle \mathbf{\Theta}_3 + z_3\mathbf{\Theta}_3 \triangle \mathbf{\Gamma}_3)$
$T_2^{III} = (z_1\mathbf{\Gamma}_1 - x_1\mathbf{\Theta}_1) \triangle (y_2\mathbf{\Theta}_2 - z_2\mathbf{\Lambda}_2) \triangle (x_3\mathbf{\Gamma}_3 \triangle \mathbf{\Lambda}_3 + y_3\mathbf{\Lambda}_3 \triangle \mathbf{\Theta}_3 + z_3\mathbf{\Theta}_3 \triangle \mathbf{\Gamma}_3)$
$T_3^{III} = (y_1\mathbf{\Theta}_1 - z_1\mathbf{\Lambda}_1) \triangle (z_2\mathbf{\Gamma}_2 - x_2\mathbf{\Theta}_2) \triangle (x_3\mathbf{\Gamma}_3 \triangle \mathbf{\Lambda}_3 + y_3\mathbf{\Lambda}_3 \triangle \mathbf{\Theta}_3 + z_3\mathbf{\Theta}_3 \triangle \mathbf{\Gamma}_3)$
$T_4^{III} = (y_1\mathbf{\Theta}_1 - z_1\mathbf{\Lambda}_1) \triangle (y_2\mathbf{\Theta}_2 - z_2\mathbf{\Lambda}_2) \triangle (x_3\mathbf{\Gamma}_3 \triangle \mathbf{\Lambda}_3 + y_3\mathbf{\Lambda}_3 \triangle \mathbf{\Theta}_3 + z_3\mathbf{\Theta}_3 \triangle \mathbf{\Gamma}_3)$

Table 1: Definitions of the trilinearities.

Relation with the trilinear tensors

- The coefficients that appear in the trilinearities of the three groups I, II and III are exactly those which appear in $\boldsymbol{\mathcal{T}}_1$, $\boldsymbol{\mathcal{T}}_2$, and $\boldsymbol{\mathcal{T}}_3$, respectively.
- Consider the mappings which, to a pixel m associates the line $\langle m, e_n \rangle, n = 1, 2, 3$ where e_n is the point represented by the vector $\mathbf{e}_n$ of the canonical basis.
- In image plane i and if $m_i \neq e_n$, the line $\langle m_i, e_n \rangle$ is the trace of the plane $y_i \Theta_i - z_i \Lambda_i \, (n = 1)$, $z_i \Gamma_i - x_i \Theta_i \, (n = 2)$ and $x_i \Lambda_i - y_i \Gamma_i \, (n = 3)$.
- Given the pixels m_j and m_k of the two image planes j and k we can associate in general the line $\boldsymbol{\mathcal{T}}_i(\mathbf{m}_j \times \mathbf{e}_n, \mathbf{m}_k \times \mathbf{e}_m)$ of the image plane i. To each pixel m_i of this image plane, we can then associate the value $\mathbf{m}_i^T \boldsymbol{\mathcal{T}}_i(\mathbf{m}_j \times \mathbf{e}_n, \mathbf{m}_k \times \mathbf{e}_m)$ which is zero iff m_i belongs to the line.

Relation with the trilinear tensors (*continued*)

- For various values of i, n, m we obtain all the trilinearities in the three groups. More precisely, we have the following table:

Group I

$T_1^I = \mathbf{m}_1^T \boldsymbol{\mathcal{T}}_1(\mathbf{m}_2 \times \mathbf{e}_2, \mathbf{m}_3 \times \mathbf{e}_2)$
$T_2^I = \mathbf{m}_1^T \boldsymbol{\mathcal{T}}_1(\mathbf{m}_2 \times \mathbf{e}_1, \mathbf{m}_3 \times \mathbf{e}_2)$
$T_3^I = \mathbf{m}_1^T \boldsymbol{\mathcal{T}}_1(\mathbf{m}_2 \times \mathbf{e}_2, \mathbf{m}_3 \times \mathbf{e}_1)$
$T_4^I = \mathbf{m}_1^T \boldsymbol{\mathcal{T}}_1(\mathbf{m}_2 \times \mathbf{e}_1, \mathbf{m}_3 \times \mathbf{e}_1)$

Group II

$T_1^{II} = \mathbf{m}_2^T \boldsymbol{\mathcal{T}}_2(\mathbf{m}_1 \times \mathbf{e}_2, \mathbf{m}_3 \times \mathbf{e}_2)$
$T_2^{II} = \mathbf{m}_2^T \boldsymbol{\mathcal{T}}_2(\mathbf{m}_1 \times \mathbf{e}_1, \mathbf{m}_3 \times \mathbf{e}_2)$
$T_3^{II} = \mathbf{m}_2^T \boldsymbol{\mathcal{T}}_2(\mathbf{m}_1 \times \mathbf{e}_2, \mathbf{m}_3 \times \mathbf{e}_1)$
$T_4^{II} = \mathbf{m}_2^T \boldsymbol{\mathcal{T}}_1(\mathbf{m}_1 \times \mathbf{e}_1, \mathbf{m}_3 \times \mathbf{e}_1)$

Group III

$T_1^{III} = \mathbf{m}_3^T \boldsymbol{\mathcal{T}}_3(\mathbf{m}_1 \times \mathbf{e}_2, \mathbf{m}_2 \times \mathbf{e}_2)$
$T_2^{III} = \mathbf{m}_3^T \boldsymbol{\mathcal{T}}_3(\mathbf{m}_1 \times \mathbf{e}_1, \mathbf{m}_2 \times \mathbf{e}_2)$
$T_3^{III} = \mathbf{m}_3^T \boldsymbol{\mathcal{T}}_3(\mathbf{m}_1 \times \mathbf{e}_2, \mathbf{m}_2 \times \mathbf{e}_1)$
$T_4^{III} = \mathbf{m}_3^T \boldsymbol{\mathcal{T}}_3(\mathbf{m}_1 \times \mathbf{e}_1, \mathbf{m}_2 \times \mathbf{e}_1)$

Table 2: Relations between the trilinearities and the trifocal tensors.

Where does this help?

- **Points in the trifocal plane** We know that the prediction from the fundamental matrices fail. It does not with the trilinearities.
- let us assume that $\mathbf{m}_1 = a_1\,\mathbf{e}_{1,2} + b_1\,\mathbf{e}_{1,3}$, $\mathbf{m}_2 = a_2\,\mathbf{e}_{2,3} + b_2\,\mathbf{e}_{2,1}$, $\mathbf{m}_3 = a_3\,\mathbf{e}_{3,1} + b_3\,\mathbf{e}_{3,2}$. We have:

Proposition 3 *The restriction of any of the 12 trilinearities to the trifocal plane is either identically 0 or proportional to* $a_1a_2a_3 + b_1b_2b_3$.

Where does this help? (*continued*)

- **Optical centers aligned** We know that the prediction from the fundamental matrices fail. It does not with the trilinearities.
-

Proposition 4 *Let m_2 and m_3 be two points in images 2 and 3 such that m_2 and m_3 satisfy the epipolar constraint. Let l_2, l_3 be two lines going through the points m_2 and m_3, respectively and such that l_k, $k = 2, 3$ does not go through the epipole e_{k1}. Then, the point of intersection m_1 of the line $\boldsymbol{\mathcal{T}}(\mathbf{l}_2, \mathbf{l}_3)$ with the epipolar line in image 1 of m_2 (which is identical to the epipolar line of m_3) is well defined and is the point in image 1 corresponding to the pair m_2, m_3).*

Estimation of the trifocal tensor

- The situation is much less advanced than in the twoviews case.
- This is because the (projective) geometry of a triplet of cameras has $3 * 11 - 15 = 18$ degrees of freedom whereas one trifocal tensor depends upon 26 parameters.
- Hence the components of the tensor satisfy 8 algebraic constraints.
- Let $\mathbf{e}_n, n = 1, 2, 3$ be the canonical basis of $\mathbb{R}^3$ and let us consider the four lines $\boldsymbol{\mathcal{T}}(\mathbf{e}_{k_2}, \mathbf{e}_{k_3})$, $\boldsymbol{\mathcal{T}}(\mathbf{e}_{l_2}, \mathbf{e}_{k_3})$, $\boldsymbol{\mathcal{T}}(\mathbf{e}_{k_2}, \mathbf{e}_{l_3})$ and $\boldsymbol{\mathcal{T}}(\mathbf{e}_{l_2}, \mathbf{e}_{l_3})$ where the indexes k_2 and l_2 (resp. k_3 and l_3) are different.

Estimation of the trifocal tensor (*continued*)

- These four lines can be chosen in nine different ways satisfy an algebraic constraint which is detailed in the following theorem.

Theorem 1 *The bilinear mapping $\boldsymbol{\mathcal{T}}$ satisfies the 9 algebraic constraints of degree 6:*

$$\begin{aligned} &| \boldsymbol{\mathcal{T}}(\mathbf{e}_{k_2}, \mathbf{e}_{k_3})\boldsymbol{\mathcal{T}}(\mathbf{e}_{k_2}, \mathbf{e}_{l_3})\boldsymbol{\mathcal{T}}(\mathbf{e}_{l_2}, \mathbf{e}_{l_3}) || \boldsymbol{\mathcal{T}}(\mathbf{e}_{k_2}, \mathbf{e}_{k_3})\boldsymbol{\mathcal{T}}(\mathbf{e}_{l_2}, \mathbf{e}_{k_3})\boldsymbol{\mathcal{T}}(\mathbf{e}_{l_2}, \mathbf{e}_{l_3}) | - \\ &\qquad | \boldsymbol{\mathcal{T}}(\mathbf{e}_{l_2}, \mathbf{e}_{k_3})\boldsymbol{\mathcal{T}}(\mathbf{e}_{k_2}, \mathbf{e}_{l_3})\boldsymbol{\mathcal{T}}(\mathbf{e}_{l_2}, \mathbf{e}_{l_3}) || \boldsymbol{\mathcal{T}}(\mathbf{e}_{k_2}, \mathbf{e}_{k_3})\boldsymbol{\mathcal{T}}(\mathbf{e}_{l_2}, \mathbf{e}_{k_3})\boldsymbol{\mathcal{T}}(\mathbf{e}_{k_2}, \mathbf{e}_{l_3}) | = 0 \end{aligned}$$

- The only practical method today is to minimize the quadratic criterion:

$$C(\boldsymbol{\mathcal{T}}) = \sum_{i=1}^{n} \|\mathbf{l_i} \times \boldsymbol{\mathcal{T}}(\mathbf{l}'_i, \mathbf{l}_i'')\|^2$$

where the triplets (l_i, l'_i, l_i'') are correspondences.

- From the tensor, we can recover the perspective projection matrices.

Bundle adjustment

- Once we have an initial estimate of the perspective matrices $\boldsymbol{\mathcal{P}}_i$, we want to refine this estimate by minimizing the reprojection error:

$$\begin{aligned} D &= \sum_i \sum_j (d(\mathbf{m}_{ij}, \boldsymbol{\mathcal{P}}_i \mathbf{M}_j))^2 \\ &= \sum_i \sum_j \varepsilon_{ij}^2 \end{aligned}$$

- If we have M points and N views, this is a sum of NM terms functions of $11N - 15 + 3M$ variables.

- Let $\hat{\mathbf{m}}_{ij}$ be the 2×1 vector representing the reprojected jth point in the ith view:

$$\hat{\mathbf{m}}_{ij} \simeq \boldsymbol{\mathcal{P}}_i \mathbf{M}_j$$

Bundle adjustment (*continued*)

- Consider the function

$$f \ : \ \mathbb{R}^{11N+3M} \longrightarrow \mathbb{R}^{NM}$$

such that

$$f(\boldsymbol{\mathcal{P}}_1, \dots, \boldsymbol{\mathcal{P}}_N, \mathbf{M}_1, \dots, \mathbf{M}_M) = [\varepsilon_{11}, \dots, \varepsilon_{MN}]^T$$

where:

$$\varepsilon_{ij} = \frac{1}{2}\|\hat{\mathbf{m}}_{ij} - \mathbf{m}_{ij}\|^2$$

- Given an initial vector $\mathbf{x}_0 = [\boldsymbol{\mathcal{P}}_1^0, \dots, \boldsymbol{\mathcal{P}}_N^0, \mathbf{M}_1^0, \dots, \mathbf{M}_M^0]$, we look for a vector $\boldsymbol{\Delta}$ such that $\|f(\mathbf{x}_0 + \boldsymbol{\Delta})\|$ is as small as possible.

- Use the usual first order approximation:

$$f(\mathbf{x}_0 + \boldsymbol{\Delta}) \simeq f(\mathbf{x}_0) + Df(\mathbf{x}_0) \cdot \boldsymbol{\Delta}$$

Bundle adjustment (*continued*)

- We want to find $\mathbf{\Delta}$ that minimizes $\|f(\mathbf{x}_0) + Df(\mathbf{x}_0) \cdot \mathbf{\Delta}\|$:

$$\mathbf{\Delta} = -(\mathbf{J}^T\mathbf{J})^{-1}\mathbf{J}^T f(\mathbf{x}_0)$$

  where the jacobian matrix $\mathbf{J} = Df(\mathbf{x}_0)$. These equations are called the *normal* equations.

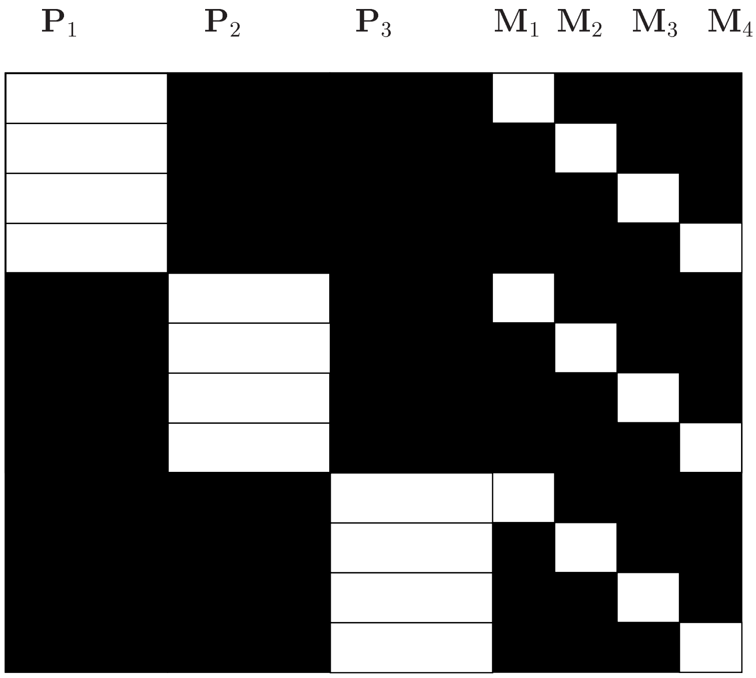


Figure 6: The jacobian matrix is sparse. For example, the first row corresponds to ε_{11} and only depends upon the variables $\boldsymbol{\mathcal{P}}_1$ and $\mathbf{M}_1$, $\mathbf{m}_{11}$ being constant. The partial derivatives of ε_{11} with respect to the other variables are 0.

Bundle adjustment (*continued*)

- $\mathbf{J}^T\mathbf{J}$ is $(11N+3M)\times(11N+3M)$ which is in general big, but it has a nice block structure.

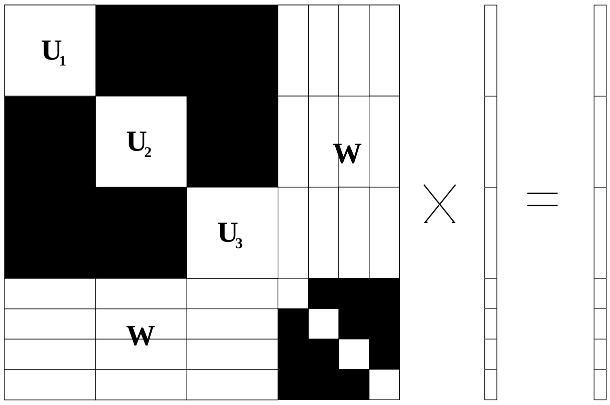


Figure 7: Normal equations

Bundle adjustment (*continued*)

- Let us define:

$$\begin{aligned}
\mathbf{U}_i &= \sum_j \left(\frac{\partial \hat{\mathbf{m}}_{ij}}{\partial \boldsymbol{\mathcal{P}}_i}\right)^T \frac{\partial \hat{\mathbf{m}}_{ij}}{\partial \boldsymbol{\mathcal{P}}_i} \\
\mathbf{V}_j &= \sum_i \left(\frac{\partial \hat{\mathbf{m}}_{ij}}{\partial \mathbf{M}_j}\right)^T \frac{\partial \hat{\mathbf{m}}_{ij}}{\partial \mathbf{M}_j} \\
\mathbf{W}_{ij} &= \left(\frac{\partial \hat{\mathbf{m}}_{ij}}{\partial \boldsymbol{\mathcal{P}}_i}\right)^T \frac{\partial \hat{\mathbf{m}}_{ij}}{\partial \mathbf{M}_j} \\
\varepsilon(\boldsymbol{\mathcal{P}}_i) &= \sum_j \left(\frac{\partial \hat{\mathbf{m}}_{ij}}{\partial \boldsymbol{\mathcal{P}}_i}\right)^T (\hat{\mathbf{m}}_{ij} - \mathbf{m}_{ij})) \\
\varepsilon(\mathbf{M}_j) &= \sum_i \left(\frac{\partial \hat{\mathbf{m}}_{ij}}{\partial \mathbf{M}_j}\right)^T (\hat{\mathbf{m}}_{ij} - \mathbf{m}_{ij}))
\end{aligned}$$

- The normal equations can be rewritten compactly:

$$\begin{pmatrix} \mathbf{U} & \mathbf{W} \\ \mathbf{W}^T & \mathbf{V} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Delta}(\boldsymbol{\mathcal{P}}) \\ \boldsymbol{\Delta}(\mathbf{M}) \end{pmatrix} = \begin{pmatrix} \varepsilon(\boldsymbol{\mathcal{P}}) \\ \varepsilon(\mathbf{M}) \end{pmatrix}$$

the matrices $\mathbf{U}$, $\mathbf{V}$, the vectors $\boldsymbol{\Delta}(\boldsymbol{\mathcal{P}})$, $\boldsymbol{\Delta}(\mathbf{M})$, $\varepsilon(\boldsymbol{\mathcal{P}})$ and $\varepsilon(\mathbf{M})$ have an interesting block structure.

Bundle adjustment (*continued*)

- Assume that $\mathbf{V}$ is invertible and multiply both sides with:

$$\begin{pmatrix} \mathbf{I} & -\mathbf{W}\mathbf{V}^{-1} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$$

- We obtain:

$$\begin{pmatrix} \mathbf{U} - \mathbf{W}\mathbf{V}^{-1}\mathbf{W}^T & \mathbf{0} \\ \mathbf{W}^T & \mathbf{V} \end{pmatrix} \begin{pmatrix} \mathbf{\Delta}(\boldsymbol{\mathcal{P}}) \\ \mathbf{\Delta}(\mathbf{M}) \end{pmatrix} = \begin{pmatrix} \varepsilon(\boldsymbol{\mathcal{P}}) - \mathbf{W}\mathbf{V}^{-1}\varepsilon(\mathbf{M}) \\ \varepsilon(\mathbf{M}) \end{pmatrix}$$

- This splits into two groups of equations:

$$(\mathbf{U} - \mathbf{W}\mathbf{V}^{-1}\mathbf{W}^T)\mathbf{\Delta}(\boldsymbol{\mathcal{P}}) = \varepsilon(\boldsymbol{\mathcal{P}}) - \mathbf{W}\mathbf{V}^{-1}\varepsilon(\mathbf{M})$$

which can be solved for $\mathbf{\Delta}(\boldsymbol{\mathcal{P}})$.

- We then substitute in the second equation which yields:

$$\mathbf{\Delta}(\mathbf{M}) = \mathbf{V}^{-1}(\varepsilon(\mathbf{M}) - \mathbf{W}^T\varepsilon(\boldsymbol{\mathcal{P}}))$$

Bundle adjustment (*continued*)

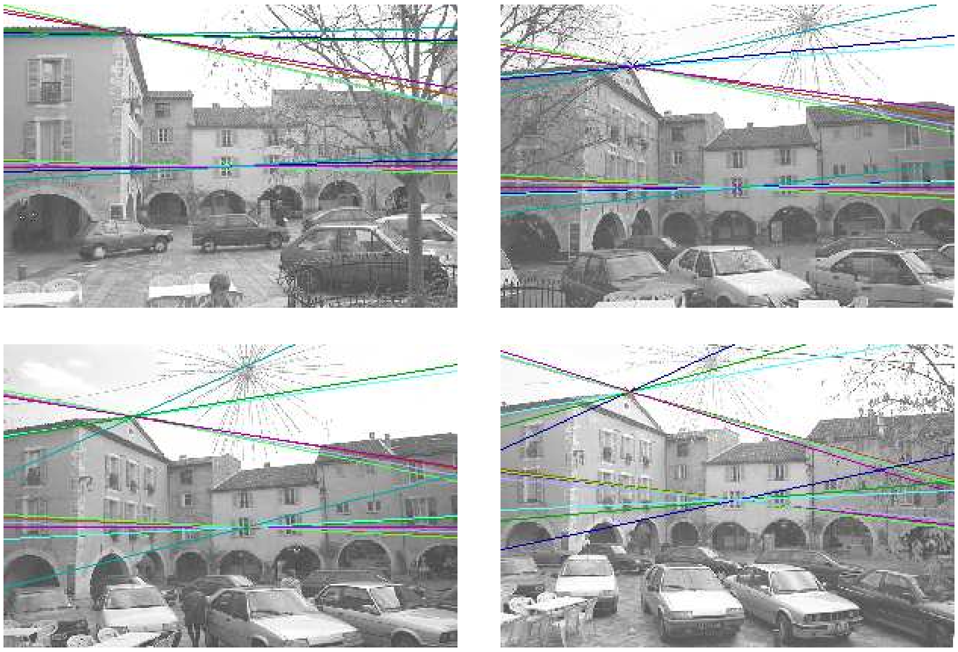
- The matrix $\mathbf{A} = \mathbf{U} - \mathbf{W}\mathbf{V}^{-1}\mathbf{W}^T$ and the vector $\mathbf{b} = \varepsilon(\boldsymbol{\mathcal{P}}) - \mathbf{W}\mathbf{V}^{-1}\varepsilon(\mathbf{M})$ inherit simple block structures:

$$\begin{aligned}
\mathbf{A}_{ij} &= \delta_{ij}\mathbf{U}_i - \sum_k \mathbf{W}_{ik}\mathbf{V}_k^{-1}\mathbf{W}_{jk}^T \\
\mathbf{b}_i &= \varepsilon(\boldsymbol{\mathcal{P}}_i) - \sum_j \mathbf{W}_{ij}\mathbf{V}_j^{-1}\varepsilon(\mathbf{M}_j)
\end{aligned}$$

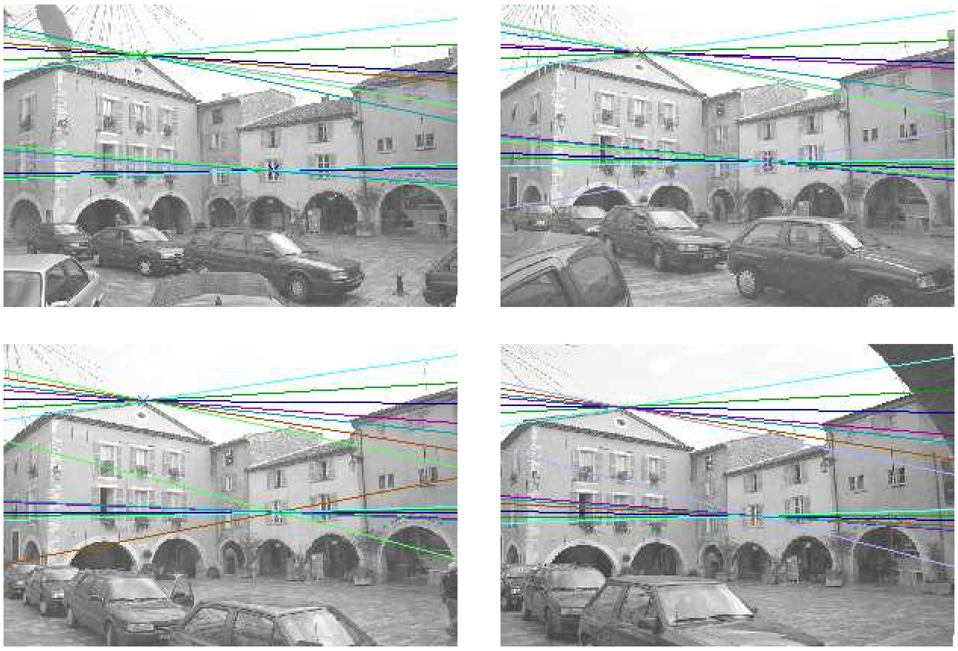
- We have replaced the problem of inverting a $(11N + 3M) \times (11N + 3M)$ by the problem of inverting N 11×11 matrices and M 3×3 matrices.

Projective calibration

- All perspective projection matrices are obtained using one of the previous methods: this provides a projective calibration which can then be used for stereo.



Projective calibration (*continued*)



Affine and Euclidean calibration

- Affine and Euclidean information can be recovered from a priori information and/or from self calibration (constant intrinsic parameters).
- An interesting variation of the bundle adjustment method consists of

  1. Estimate the intrinsic parameters by self-calibration.
  2. Estimate the relative intrinsic parameters from the computed E-matrices.
  3. Do a bundle adjustment on the intrinsic and relative extrinsic parameters.