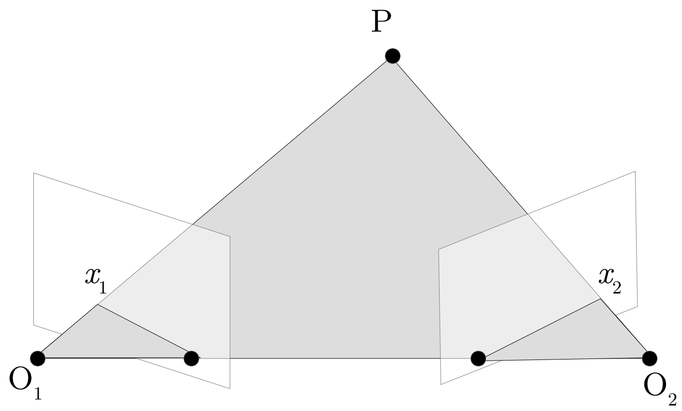
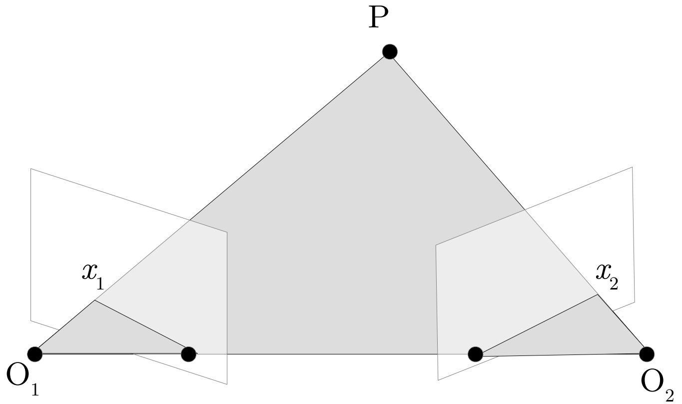
Computer Vision

Class 08



Raquel Frizera Vassallo

Fundamental Matrix



Summary

- Fundamental Matrix
- Properties
- Estimation of the Fundamental Matrix

Uncalibrated
Epipolar Geometry

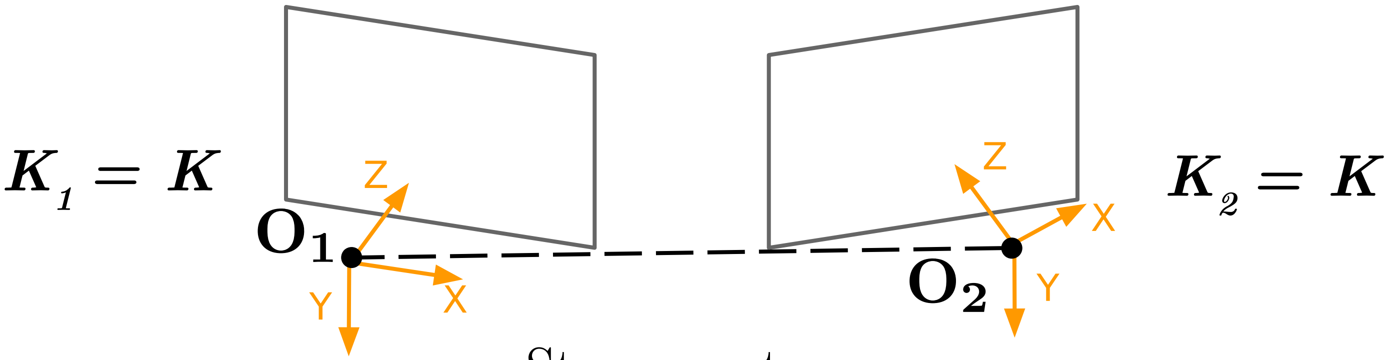
? T

? R **?**

? K ?

Uncalibrated epipolar geometry

- Let's study the epipolar geometry for uncalibrated cameras
- For simplicity, we will assume that the same camera has captured both images, so that $K_1 = K_2 = K$
- The extension to different cameras will be done later



Stereo system

The Fundamental Matrix

- The epipolar constraint for uncalibrated cameras can be written by direct substitution of $x_i = K^{-1}\ x_i'$ into the epipolar constraint
- x represents image points in the metric space while x' represents the image points in the image space in pixels

$$x_i' = K x_i$$

pixel metric

$$x_i = K^{-1} x_i'$$

The Fundamental Matrix

- The epipolar constraint for uncalibrated cameras can be written by direct substitution of $x_i = K^{-1}\, x_i'$ into the epipolar constraint
- x represents image points in the metric space while x' represents the image points in the image space in pixels

$$\underset{\text{pixel}}{x_i'} = K \underset{\text{metric}}{x_i}$$

$$x_i = K^{-1} x_i'$$

$$x_2{}^T \hat{T} R x_1 = 0 \longrightarrow x_2'{}^T \underbrace{K^{-T} \hat{T} R K^{-1}}_{F} x_1' = 0$$

$$\boxed{\mathrm{F} = \mathrm{K}^{-T} \hat{T} R K^{-1}}$$

When $K = I$, the fundamental matrix F is identical to the essential matrix E

The Fundamental Matrix

- Another derivation can be obtained by elimination of the unknown depth scales λ_1, λ_2

$$\lambda_2 \boldsymbol{x}_2 = R\lambda_1 \boldsymbol{x}_1 + T$$

- Multiplying both sides by the matrix K and knowing that $x' = Kx$

The Fundamental Matrix

- Another derivation can be obtained by elimination of the unknown depth scales λ_1, λ_2

$$\lambda_2 \boldsymbol{x}_2 = R\lambda_1 \boldsymbol{x}_1 + T$$

- Multiplying both sides by the matrix K and knowing that $x' = Kx$

$$\lambda_2 K \boldsymbol{x}_2 = KR\lambda_1 \boldsymbol{x}_1 + KT \longrightarrow \lambda_2 \boldsymbol{x}'_2 = KR\lambda_1 K^{-1} \boldsymbol{x}'_1 + KT$$

The Fundamental Matrix

- Another derivation can be obtained by elimination of the unknown depth scales λ_1, λ_2

$$\lambda_2 \boldsymbol{x}_2 = R\lambda_1 \boldsymbol{x}_1 + T$$

- Multiplying both sides by the matrix K and knowing that $x' = Kx$

$$\lambda_2 K \boldsymbol{x}_2 = KR\lambda_1 \boldsymbol{x}_1 + KT \longrightarrow \lambda_2 \boldsymbol{x}_2' = KR\lambda_1 K^{-1} \boldsymbol{x}_1' + KT$$

- Now considering $T'=KT$

$$\lambda_2 \boldsymbol{x}_2' = \lambda_1 KRK^{-1} \boldsymbol{x}_1' + T'$$

- Finally, making the cross product of both sides with T' and then the inner product with x'_2

$$\boldsymbol{x}_2'^{T} \widehat{T'} KRK^{-1} \boldsymbol{x}_1' = 0$$

$$\boxed{F = \widehat{T'} KRK^{-1}}$$

Equivalence of the two expressions

- Recall that for $\mathrm{T} \in \mathbb{R}^3$ and a nonsingular matrix K

$$\mathrm{K}^{-T}\widehat{T}K^{-1} = \widehat{KT} \;\; \text{if } \det(K) = +1$$

Equivalence of the two expressions

- Recall that for $\mathrm{T} \in \mathbb{R}^3$ and a nonsingular matrix K

$$\mathrm{K}^{-T}\widehat{T}K^{-1} = \widehat{KT} \text{ if } \det(K) = +1$$

- Thus $\mathrm{F} = \mathrm{K}^{-T}\widehat{T}RK^{-1} = K^{-T}\widehat{T}K^{-1}KRK^{-1} = \widehat{T'}KRK^{-1}$

$$\boxed{\mathrm{F} = \mathrm{K}^{-T}\widehat{T}RK^{-1} = \widehat{T'}KRK^{-1}} \text{ if } \det(\mathrm{K}) = +1$$

Equivalence of the two expressions

- Recall that for $T \in \mathbb{R}^3$ and a nonsingular matrix K

$$\mathrm{K}^{-T}\widehat{T}K^{-1} = \widehat{KT} \;\text{ if } \det(K) = +1$$

- Thus $\mathrm{F} = \mathrm{K}^{-T}\widehat{T}RK^{-1} = K^{-T}\widehat{T}K^{-1}KRK^{-1} = \widehat{T'}KRK^{-1}$

$$\boxed{\mathrm{F} = \mathrm{K}^{-T}\widehat{T}RK^{-1} = \widehat{T'}KRK^{-1}} \;\text{ if } \det(\mathrm{K}) = +1$$

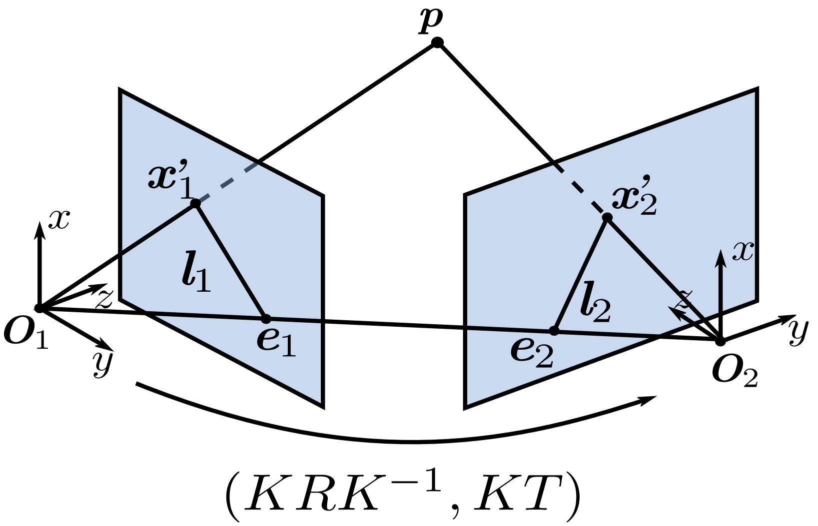
- In case $\det(\mathrm{K}) \neq 1$, one can simply scale all the matrices by a factor. In any case, we have

$$\mathrm{K}^{-T}\widehat{T}RK^{-1} \sim \widehat{T'}KRK^{-1}$$

Thus, without loss of generality, we will always assume $\det(\mathrm{K}) = 1$.

Properties of the fundamental matrix

- Epipolar lines
  - The fundamental matrix maps a point x in one image to a line in the other image



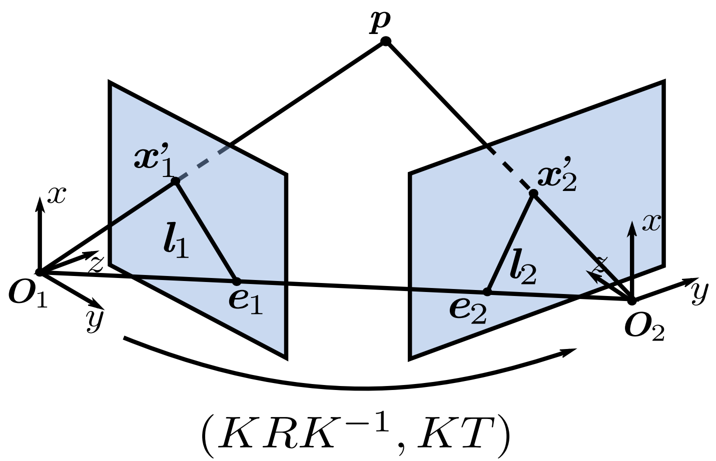
$$x_2'^{\,T} F x_1' = 0$$

$$l_1^{\,T} x_1' = 0 \qquad\qquad x_2'^{\,T} l_2 = 0$$

$$\boxed{l_1 = F^T x_2'} \qquad\qquad \boxed{l_2 = F x_1'}$$

Properties of the fundamental matrix

- Epipoles
  - The epipole is the point where the baseline (the line joining the two camera centers O_1,O_2) intersects the image plane in each view
  - The epipoles are the right and left null spaces of the fundamental matrix F
  - All the epipolar lines must pass through the epipole of each image



$$e_2{}^T F = 0 \qquad F e_1 = 0$$

$$\boxed{e_2 = KT = T'} \qquad \boxed{e_1 = KR^T T}$$

Properties of the fundamental matrix

- Singular Value Decomposition of F
  - F is the product of $\widehat{T}$ of rank 2 and KRK^{-1} of rank 3, so it has rank 2

$$F = U\Sigma V^T \text{ with } \Sigma = diag\{\sigma_1, \sigma_2, 0\} \text{ for some } \sigma_1, \sigma_2 \in \mathbb{R}$$

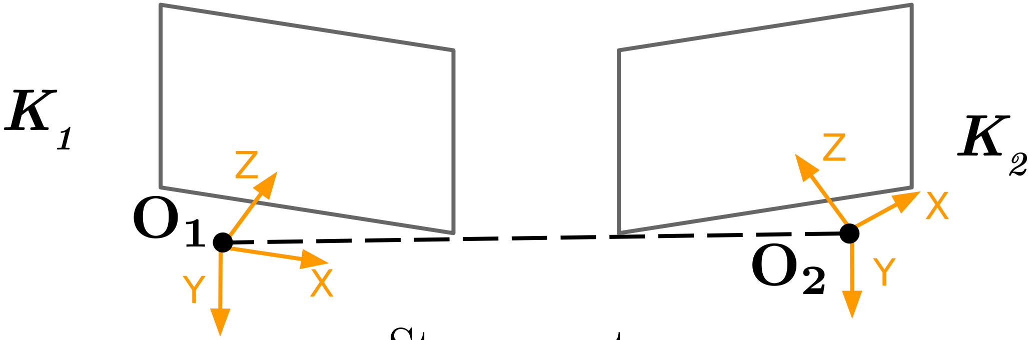
- In contrast to the essential matrix, where $\sigma_1 = \sigma_2 = \sigma$, here we have $\sigma_1 \geq \sigma_2$.
  - F can be estimated from eight or more corresponding points
  - It is not possible to recover R and T from the fundamental matrix
  - F has 8 DOF(degress of freedom) but it is composed by the matrices K (5 DOF), R (3 DOF) and T (2 DOF: three elements defined up to a scalar factor).
  - From 8 DOF it is not possible to recover 10 DOF: K (5 DOF) + R (3 DOF) + T (2 DOF)

For Different Calibration Matrices

$$F = {K_2}^{-T} \widehat{T} R {K_1}^{-1}$$

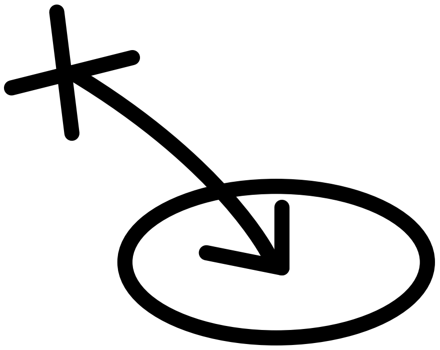
or

$$F = \widehat{T'} K_2 R {K_1}^{-1} \quad \text{with} \quad T' = K_2 T$$



Stereo system

Estimating the Fundamental Matrix



Eight-point linear algorithm

- Find such F that the epipolar error is minimized

$$min_F \sum_{j=1}^{n} {x'^{j}_2}^T \mathrm{F}\, x'^{j}_1$$

- Fundamental matrix can be estimated up to scale

$${x'_2}^T F x'_1 = 0 \longrightarrow [x'_1x'_2, x'_1y'_2, x'_1z'_2, y'_1x'_2, y'_1y'_2, y'_1z'_2, z'_1x'_2, z'_1y'_2, z'_1z'_2]^T F^S = 0$$

$$\mathrm{a} = x'_1 \otimes x'_2 = [x'_1x'_2, x'_1y'_2, x'_1z'_2, y'_1x'_2, y'_1y'_2, y'_1z'_2, z'_1x'_2, z'_1y'_2, z'_1z'_2]^T$$

$$F^s \doteq [f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8, f_9]^T$$

$$\boxed{\mathrm{a}^T F^s = 0}$$

Eight-point linear algorithm

- For $n \geq 8$, collect and stack the constraints from all points

$$\chi = [a^1, a^2, ..., a^n]^T \longrightarrow \chi F^s = 0$$

$$min_F \sum_{j=1}^{n} x_2'^{j\,T} F\, x_1'^j \longrightarrow min_{F^s} \|\chi F^s\|^2$$

- Solution is the eigenvector associated with smallest eigenvalue of $\chi^T \chi$.

- Compute the SVD of χ and choose F^S to be the right singular vector associated with the smallest singular value, i.e. the last column of V.

Eight-point linear algorithm

- Compute SVD of F recovered from data

$$F = U\Sigma V^T \longrightarrow \Sigma = diag\,(\sigma_1, \sigma_2, \sigma_3)$$

- Project onto the essential manifold:

$$\Sigma' = diag\,(\sigma_1, \sigma_2, 0) \longrightarrow F = U\Sigma' V^T$$

Normalization

- Since image coordinates x'_1 and x'_2 are measured in pixels, the individual entries of the matrix $\mathcal{X}$ can vary by two orders of magnitude.
- Thus normalization should be done!
- This can be done, in each image, by transforming the points by an affine matrix H_i, so that the resulting points have zero mean and unit variance.
- The "normalized" coordinates, $\tilde{\boldsymbol{x}}_i$ can be obtained by:

$$\tilde{\boldsymbol{x}}_i \doteq H_i \boldsymbol{x}'_i = \begin{bmatrix} 1/\sigma_{x_i} & 0 & -\mu_{x_i}/\sigma_{x_i} \\ 0 & 1/\sigma_{y_i} & -\mu_{y_i}/\sigma_{y_i} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix}$$

$$\underbrace{\mu_{x_i}}_{\text{mean}} \doteq \frac{1}{n}\sum_{j=1}^{n}(x'_i)^j, \quad \underbrace{\sigma_{x_i}}_{\text{standard deviation}} \doteq \sqrt{\frac{1}{n}\sum_{j=1}^{n}[(x'_i)^j - \mu_{x_i}]^2};$$

with $i = 1, 2$ and μ_{y_i} and σ_{y_i} are defined similarly.

Normalization

- The normalized coordinates and the epipolar constrain become

$$\tilde{\boldsymbol{x}}_1 = H_1 \boldsymbol{x}'_1 \qquad \tilde{\boldsymbol{x}}_2 = H_2 \boldsymbol{x}'_2. \qquad \boldsymbol{x}'^{T}_2 F \boldsymbol{x}'_1 = \tilde{\boldsymbol{x}}^T_2 \underbrace{H_2^{-T} F H_1^{-1}}_{\tilde{F}} \tilde{\boldsymbol{x}}_1 = 0$$

- Use the eight-point linear algorithm with the new image pairs $(\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_2)$ to estimate:

$$\tilde{F} \doteq H_2^{-T} F H_1^{-1}$$

- Then recover

$$F = H_2^T \tilde{F} H_1$$

Improvement by nonlinear optimization

- The optimal estimate of the fundamental matrix F can be obtained by minimizing the reprojection error

$$\phi(\boldsymbol{x}, F) \doteq \sum_{j=1}^{n} \|\tilde{\boldsymbol{x}}'^{j}_{1} - \boldsymbol{x}'^{j}_{1}\|_2^2 + \|\tilde{\boldsymbol{x}}'^{j}_{2} - \boldsymbol{x}'^{j}_{2}\|_2^2$$

$$\tilde{\boldsymbol{x}}'^{j}_{i} = \boldsymbol{x}'^{j}_{i} + w^{j}_{i}$$

noisy measurements

ideal measurements

$$\boldsymbol{x}'^{jT}_{2} F \boldsymbol{x}'^{j}_{2} = 0$$

with w_i^j having a Normal distribution $\mathcal{N}(0, \sigma^2)$

- Substituting this model into the epipolar constraint, we obtain

$$\tilde{\boldsymbol{x}}'^{jT}_{2} F \tilde{\boldsymbol{x}}'^{j}_{1} = w^{jT}_{2} F \tilde{\boldsymbol{x}}'^{j}_{1} + \tilde{\boldsymbol{x}}'^{jT}_{2} F w_1 + \cancel{w^{jT}_{2} F w^{j}_{1}}$$

Can be disconsidered because w_1^j and w_2^j are much smaller than $\tilde{\boldsymbol{x}}'^{j}_{1}$ and $\tilde{\boldsymbol{x}}'^{j}_{2}$

Improvement by nonlinear optimization

- An approximate cost function that only takes into account the first-order effects of the noise is given by:

$$\phi(F) = \sum_{j=1}^{n} \frac{\left(\tilde{\boldsymbol{x}}'^{jT}_{2} F \tilde{\boldsymbol{x}}'^{j}_{1}\right)^2}{\left\|\widehat{e}_3 F \tilde{\boldsymbol{x}}'^{j}_{1}\right\|^2 + \left\|\tilde{\boldsymbol{x}}'^{jT}_{2} F \widehat{e}_3\right\|^2}$$

with $e_3 = [0,0,1]$

- The solution to the minimization of $\phi(F)$ calls for nonlinear optimization techniques.

The normalized 8-point algorithm for F

Objective:

For a given set of image correspondences (x'^j_1, x'^j_2), $j = 1, 2, \ldots, n$ $(n \geq 8)$, this algorithm finds the fundamental matrix F that minimizes the epipolar constraint:

$$(\boldsymbol{x}'^j_2)^T F \boldsymbol{x}'^j_1 = 0, \quad j = 1, 2, \ldots, n$$

Algorithm:

1. Normalize the points: $\tilde{\boldsymbol{x}}_1 = H_1 \boldsymbol{x}'_1$ $\quad \tilde{\boldsymbol{x}}_2 = H_2 \boldsymbol{x}'_2$ $\quad \tilde{F} \doteq H_2^{-T} F H_1^{-1}$

2. Compute a first approximation of the fundamental matrix such that $\|\chi \tilde{F}^s\|^2$ is minimized by

$$\mathrm{a} = \tilde{\boldsymbol{x}}_1 \otimes \tilde{\boldsymbol{x}}_2 \qquad \chi = [\mathrm{a}^1, \mathrm{a}^2, \ldots, \mathrm{a}^n]^T \qquad \chi \tilde{F}^s = 0$$

computing the SVD of χ and define $\tilde{F}^s$ to be the ninth column of V. Unstack the nine elements of $\tilde{F}^s$ into a square 3 x 3 matrix $\tilde{F}$.

3. Impose the rank-2 constraint and set the fundamental matrix to be $\tilde{F} = U \mathrm{diag}\{\sigma_1, \sigma_2, 0\} V^T$

4. Denormalize: $F = H_2^T \tilde{F} H_1$

5. Optimize using as cost function:

$$\phi(F) = \sum_{j=1}^{n} \frac{\left(\boldsymbol{x}'^{jT}_2 F \boldsymbol{x}'^j_1\right)^2}{\left\|\widehat{e}_3 F \boldsymbol{x}'^j_1\right\|^2 + \left\|\boldsymbol{x}'^{jT}_2 F \widehat{e}_3\right\|^2} \quad \text{with } e_3 = [0,0,1]$$

Automatically estimate the fundamental matrix F using RANSAC

Objective: Compute the fundamental matrix F between two images.

Algorithm:

1. **Interest points:** Compute interest points in each image.
2. **Putative correspondences:** Compute a set of interest point matches based on proximity and similarity of their intensity neighborhood
3. **RANSAC robust estimation:** Repeat for N samples, where N is determined adaptively.
   (a) Select a random sample of 8 correspondences and compute the fundamental matrix F.
   (b) Calculate the distance $d_\perp$ for each putative correspondence.
   (c) Compute the number of inliers consistent with F by the number of correspondences for which $d_\perp < t$ pixels.

Choose the F with the largest number of inliers. Re-estimate F from all correspondences classified as inliers.

4. **Non-linear estimation:** Re-estimate F minimizing a cost function:

$$\phi(F) = \sum_{j=1}^{n} \frac{\left(\tilde{\boldsymbol{x}'}_2^{jT} F \tilde{\boldsymbol{x}'}_1^{j}\right)^2}{\left\|\hat{e}_3 F \tilde{\boldsymbol{x}'}_1^{j}\right\|^2 + \left\|\tilde{\boldsymbol{x}'}_2^{jT} F \hat{e}_3\right\|^2} \quad \text{with } e_3 = [0,0,1]$$

5. **If wanted, do a guided matching:** Further interest point correspondences are now determined using the estimated F to define a search strip about the epipolar line.

Automatically estimate the fundamental matrix F using RANSAC and the minimum number of 7 correspondences

Objective Compute the fundamental matrix between two images.

Algorithm

(i) **Interest points:** Compute interest points in each image.
(ii) **Putative correspondences:** Compute a set of interest point matches based on proximity and similarity of their intensity neighbourhood.
(iii) **RANSAC robust estimation:** Repeat for N samples, where N is determined adaptively as in algorithm 4.5(p121):
   (a) Select a random sample of 7 correspondences and compute the fundamental matrix $\mathtt{F}$ as described in section 11.1.2. There will be one or three real solutions.
   (b) Calculate the distance $d_\perp$ for each putative correspondence.
   (c) Compute the number of inliers consistent with $\mathtt{F}$ by the number of correspondences for which $d_\perp < t$ pixels.
   (d) If there are three real solutions for $\mathtt{F}$ the number of inliers is computed for each solution, and the solution with most inliers retained.

   Choose the $\mathtt{F}$ with the largest number of inliers. In the case of ties choose the solution that has the lowest standard deviation of inliers.
(iv) **Non-linear estimation:** re-estimate $\mathtt{F}$ from all correspondences classified as inliers by minimizing a cost function, e.g. (11.6), using the Levenberg–Marquardt algorithm of section A6.2(p600).
(v) **Guided matching:** Further interest point correspondences are now determined using the estimated $\mathtt{F}$ to define a search strip about the epipolar line.

The last two steps can be iterated until the number of correspondences is stable.

Algorithm 11.4. *Algorithm to automatically estimate the fundamental matrix between two images using RANSAC.*

11.1.2 The minimum case – seven point correspondences

The equation $\mathbf{x}_i'^{\mathsf{T}}\mathtt{F}\mathbf{x}_i = 0$ gives rise to a set of equations of the form $\mathtt{A}\mathbf{f} = \mathbf{0}$. If $\mathtt{A}$ has rank 8, then it is possible to solve for $\mathbf{f}$ up to scale. In the case where the matrix $\mathtt{A}$ has rank seven, it is still possible to solve for the fundamental matrix by making use of the singularity constraint. The most important case is when only 7 point correspondences are known (other cases are discussed in section 11.9). This leads to a 7×9 matrix $\mathtt{A}$, which generally will have rank 7.

The solution to the equations $\mathtt{A}\mathbf{f} = \mathbf{0}$ in this case is a 2-dimensional space of the form $\alpha\mathtt{F}_1 + (1-\alpha)\mathtt{F}_2$, where α is a scalar variable. The matrices $\mathtt{F}_1$ and $\mathtt{F}_2$ are obtained as the matrices corresponding to the generators $\mathbf{f}_1$ and $\mathbf{f}_2$ of the right null-space of $\mathtt{A}$. Now, we use the constraint that $\det \mathtt{F} = 0$. This may be written as $\det(\alpha\mathtt{F}_1 + (1-\alpha)\mathtt{F}_2) = 0$. Since $\mathtt{F}_1$ and $\mathtt{F}_2$ are known, this leads to a cubic polynomial equation in α. This polynomial equation may be solved to find the value of α. There will be either one or three real solutions (the complex solutions are discarded [Hartley-94c]). Substituting back in the equation $\mathtt{F} = \alpha\mathtt{F}_1 + (1-\alpha)\mathtt{F}_2$ gives one or three possible solutions for the fundamental matrix.

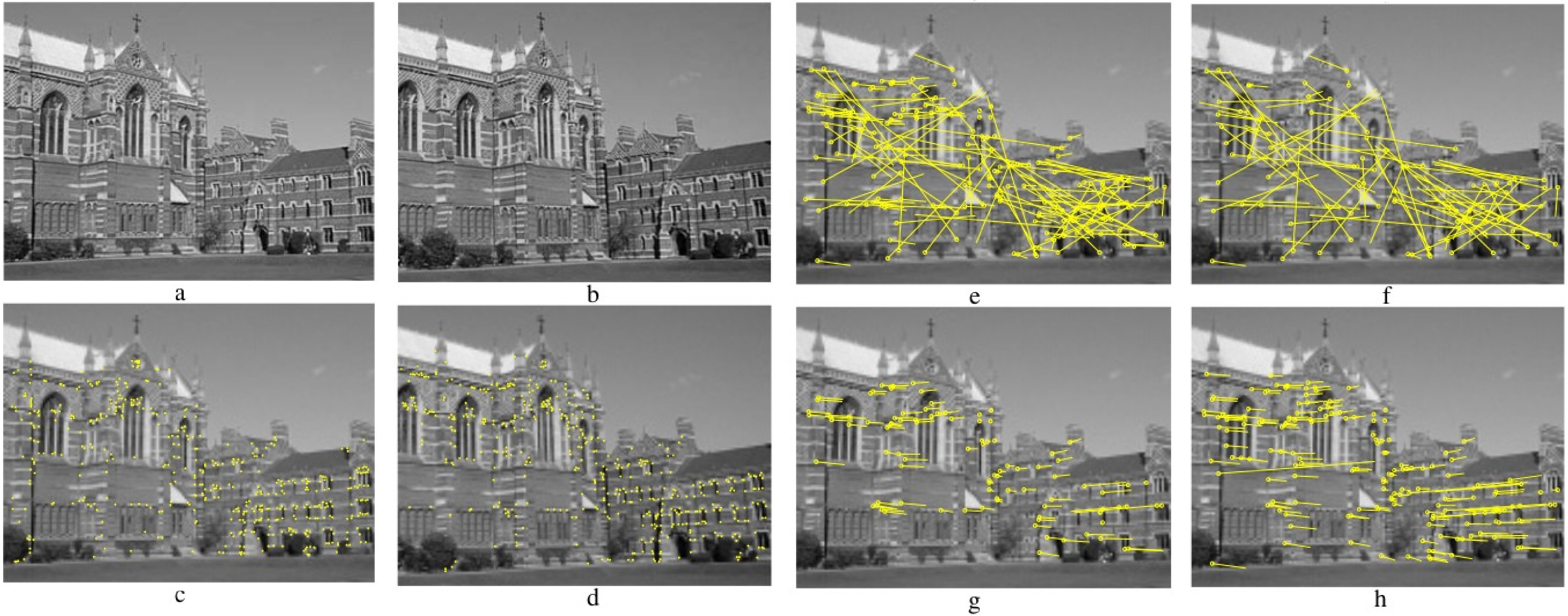


Fig. 11.4. **Automatic computation of the fundamental matrix between two images using RANSAC.** *(a) (b) left and right images of Keble College, Oxford. The motion between views is a translation and rotation. The images are* 640×480 *pixels. (c) (d) detected corners superimposed on the images. There are approximately 500 corners on each image. The following results are superimposed on the left image: (e) 188 putative matches shown by the line linking corners, note the clear mismatches; (f) outliers – 89 of the putative matches. (g) inliers – 99 correspondences consistent with the estimated* $\mathtt{F}$*; (h) final set of 157 correspondences after guided matching and MLE. There are still a few mismatches evident, e.g. the long line on the left.*

Credits

- Yi Ma, Stefano Soatto, Jana Kosecka e S. Shankar Sastry. **An Invitation to 3D Vision: From Images to Geometric Models.** Springer, ISBN 0387008934

- Richard Hartley and Andrew Zisserman. **Multiple View Geometry in Computer Vision.** Cambridge, ISBN 0521623049