Vision 3D artificielle
Session 2: Essential and fundamental matrices, their computation, RANSAC algorithm

Pascal Monasse monasse@imagine.enpc.fr

IMAGINE, École des Ponts ParisTech

Contents

Contents

Compact matrix multiplication formulas

- Block matrix multiplication

$$A\begin{pmatrix} B_1 & B_2 \end{pmatrix} = \begin{pmatrix} AB_1 & AB_2 \end{pmatrix} \quad A\begin{pmatrix} B_1 & \cdots & B_n \end{pmatrix} = \begin{pmatrix} AB_1 & \cdots & AB_n \end{pmatrix}$$

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} B = \begin{pmatrix} A_1 B \\ A_2 B \end{pmatrix} \qquad \begin{pmatrix} A_1^T \\ \vdots \\ A_m^T \end{pmatrix} B = \begin{pmatrix} A_1^T B \\ \vdots \\ A_m^T B \end{pmatrix}$$

- Both matrices split into blocks

$$\begin{pmatrix} A_1 & A_2 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = A_1 B_1 + A_2 B_2$$

$$\begin{pmatrix} A_1 & \cdots & A_k \end{pmatrix} \begin{pmatrix} B_1^T \\ \vdots \\ B_k^T \end{pmatrix} = A_1 B_1^T + \cdots + A_k B_k^T$$

Vector product

- Definition

$$\mathbf{a} \times \mathbf{b} = [\mathbf{a}]_\times \mathbf{b} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \times \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} yz' - zy' \\ zx' - xz' \\ xy' - yx' \end{pmatrix}$$

$$[\mathbf{a}]_\times = \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix}$$

- Properties: bilinear, antisymmetric.
- Link with determinant

$$\mathbf{a}^T(\mathbf{b} \times \mathbf{c}) = |a \quad b \quad c|$$

- Composition

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{a}^T\mathbf{c})\mathbf{b} - (\mathbf{b}^T\mathbf{c})\mathbf{a}$$

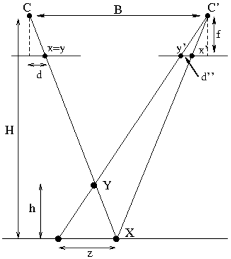
- Composition with isomorphism M

$$(M\mathbf{a}) \times (M\mathbf{b}) = |M|\, M^{-T}(\mathbf{a} \times \mathbf{b}) \quad [M\mathbf{a}]_\times = |M|\, M^{-T}[\mathbf{a}]_\times M^{-1}$$

Contents

Triangulation

Fundamental principle of stereo vision



$$h = \frac{z}{B/(H-h)} \simeq \frac{z}{B/H}, \; z = d'' \frac{H}{f}.$$

f focal length.

H distance optical center-ground.

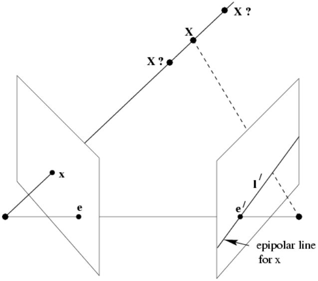
B distance between optical centers (baseline).

Goal

Given two rectified images, point correspondences and computation of their apparent shift (disparity) gives information about relative depth of the scene.

Epipolar constraints

Rays from matching points must intersect in space



- The vectors $\overrightarrow{C\mathbf{x}}$, $\overrightarrow{C'\mathbf{x}'}$ and T are coplanar. We write it in camera 1 coordinate frame: $\mathbf{x}$, $R\mathbf{x}'$ and T coplanar,

  $$\begin{vmatrix} \mathbf{x} & T & R\mathbf{x}' \end{vmatrix} = 0,$$

  which we can write:

  $$\mathbf{x}^T (T \times R\mathbf{x}') = 0.$$

- We note $[T]_\times \mathbf{x} = T \times \mathbf{x}$ and we get the equation

  $$\mathbf{x}^T E\mathbf{x}' = 0 \text{ with } E = [T]_\times R$$

  (Longuet-Higgins 1981)

Epipolar constraints

- E is the essential matrix but deals with points expressed in camera coordinate frame.
- Converting to pixel coordinates requires multiplying by the inverse of camera calibration matrix K: $\mathbf{x}_{cam} = K^{-1}\mathbf{x}_{image}$
- We can rewrite the epipolar constraint as:

  $$\mathbf{x}^T F \mathbf{x}' = 0 \text{ with } F = K^{-T} E K'^{-1} = K^{-T} [T]_\times R K'^{-1}$$

  (Faugeras 1992)
- F is the fundamental matrix. The progress is important: we can constrain the match without calibrating the cameras!
- It can be easily derived formally, by expressing everything in camera 2 coordinate frame:

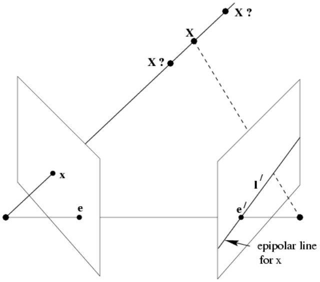
  $$\lambda \mathbf{x} = K(R\mathbf{X} + T) \quad \lambda' \mathbf{x}' = K' \mathbf{X}$$

  We remove the 5 unknowns $\mathbf{X}$, λ and λ' from the system

  $$\lambda K^{-1}\mathbf{x} = \lambda' R K'^{-1} \mathbf{x}' + T \Rightarrow \lambda T \times (K^{-1}\mathbf{x}) = \lambda' [T]_\times R K'^{-1} \mathbf{x}'$$

  followed by scalar product with $K^{-1}\mathbf{x}$

Anatomy of the fundamental matrix



Glossary:

- $e = KT$ satisfies $e^T F = 0$, that is the left epipole
- $e' = K'R^{-1}T$ satisfies $Fe' = 0$, that is the right epipole
- $F\mathbf{x}'$ is the epipolar line (in left image) associated to $\mathbf{x}'$
- $F^T\mathbf{x}$ is the epipolar line (in right image) associated to $\mathbf{x}$

- Observe that if $T = 0$ we get $F = 0$, that is, no constraints: without displacement of optical center, no 3D information.
- The constraint is important: it is enough to look for the match of point $\mathbf{x}$ along its associated epipolar line (1D search).

Theorem

A 3×3 matrix is a fundamental matrix iff it has rank 2

Example

Image 1

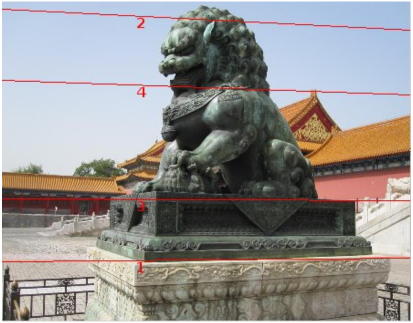


Image 2

Contents

Singular Value Decomposition

Theorem (SVD)

Let A be an $m \times n$ matrix. We can decompose A as:

$$A = U\Sigma V^T = \sum_{i=1}^{\min(m,n)} \sigma_i U_i V_i^T$$

with Σ diagonal $m \times n$ matrix and $\sigma_i = \Sigma_{ii} \geq 0$, U $(m \times m)$ and V $(n \times n)$ composed of orthonormal columns.

- The rank of A is the number of non-zero σ_i
- An orthonormal basis of the kernel of A is composed of

$$\{V_i : \sigma_i = 0\} \cup \{V_i : i = \min(m,n) + 1 \ldots \max(m,n)\}$$

Theorem (Thin SVD)

If $m \geq n$, U $m \times n$ and

$$A = U \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix} V^T$$

If $m \leq n$, V $n \times m$ and

$$A = U \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \end{pmatrix} V^T$$

Singular Value Decomposition

- Proof:
  1. Orthonormal diagonalization of $A^T A = V \Sigma^2 V^T$
  2. Write $U_i = AV_i / \sigma_i$ if $\sigma_i \neq 0$.
  3. Check that $U_i^T U_j = \delta_{ij}$.
  4. Complement the U_i by orthonormal vectors.
  5. Check $A = U \Sigma V^T$ by comparison on the basis formed by V_i.
- Implementation: efficient algorithm but:
  As much as we dislike the use of black-box routines, we need to ask you to accept this one, since it would take us too far afield to cover its necessary background material here.
  Numerical Recipes

Contents

Computation of F

- The 8 point method (actually 8+) is the simplest as it is linear.
- We write the epipolar constraint for the 8 correspondences

$$\mathbf{x_i}^T F \mathbf{x'_i} = 0 \Leftrightarrow A_i^T f = 0 \text{ with } f = \begin{pmatrix} f_{11} & f_{12} & f_{13} & f_{21} & \dots & f_{33} \end{pmatrix}^T$$

- Each one is a linear equation in the unknown f.
- f has 8 independent parameters, since scale is indifferent.
- We impose the constraint $\|f\| = 1$:

$$\min_A \|Af\|^2 \text{ subject to } \|f\|^2 = 1 \text{ with } A = \begin{pmatrix} A_1^T \\ \vdots \\ A_8^T \end{pmatrix}$$

- Solution: f is an eigenvector of $A^T A$ associated to its smallest eigenvalue (can be recovered from SVD of A).
- Constraint: to enforce rank 2 of F, we can decompose it as SVD, put $\sigma_3 = 0$ and recompose.

Computation of F

- Enforcing constraint $\det F = 0$ after minimization is not optimal.
- The 7 point method imposes that from the start.
- We get linear system $Af = 0$ with A of size 7×9.
- Let f_1, f_2 be 2 free vectors of the kernel of A (from SVD).
- Look for a solution $f_1 + xf_2$ with $\det F = 0$.
- $\det(F_1 + xF_2) = P(x)$ with P polynomial of degree 3, we get 1 or 3 solutions.
- The main interest is not computing F with fewer points (we have many more in general, which is anyway better for precision), but we have fewer chances of selecting false correspondences.
- By the way, how to ensure we did not incorporate bad correspondences in the equations?

Normalization

- The 8 point algorithm "as is" yields very imprecise results
- Hartley (1997): *In Defense of the Eight-Point Algorithm*
- *Explanation*: the scales of coefficients of F are very different. F_{11}, F_{12}, F_{21} and F_{22} are multiplied by $x_i x_i'$, $x_i y_i'$, $y_i x_i'$ and $y_i y_i'$, that can reach 10^6. On the contrary, F_{13}, F_{23}, F_{31} and F_{32} are multiplied by x_i, y_i, x_i' and y_i' that are of order 10^3. F_{33} is multiplied by 1.
- The scales being so different, A is badly conditioned.
- Solution: normalize points so that coordinates are of order 1.

$$N = \begin{pmatrix} 10^{-3} & & \\ & 10^{-3} & \\ & & 1 \end{pmatrix}, \tilde{x}_i = N x_i, \tilde{x'}_i = N x_i'$$

- We find $\tilde{F}$ for points $(\tilde{x}_i, \tilde{x'}_i)$ then $F = N^T \tilde{F} N$

Computation of E

- E depends on 5 parameters (3 for R+3 for T-1 for scale)
- A 3×3 matrix E is essential iff its singular values are 0 and two equal positive values. It can be written:

$$2EE^TE - \mathrm{tr}(EE^T)E = 0, \quad \det E = 0$$

- 5 point algorithm (Nister, 2004)
- We have $Ae = 0$, A of size 5×9, we get a solution of the form

$$E = xX + yY + zZ + W$$

  with X, Y, Z, W a basis of the kernel of A (SVD)
- The contraints yield 10 polynomial equations of degree 3 in x, y, z
- 1) Gauss pivot to eliminate terms of degree 2+ in x, y, then $B(z)\begin{pmatrix} x & y & 1 \end{pmatrix}^T = 0$, that is $\det B(z) = 0$, degree 10.
  2) Gröbner bases. 3) $C(z)\begin{pmatrix} 1 & x & y & x^2 & xy & \dots & y^3 \end{pmatrix}^T = 0$ and $\det C(z) = 0$.

Contents

RANSAC algorithm

- How to solve a problem of parameter estimation in presence of outliers? This is the framework of robust estimation.
- Example: regression line of plane points (x_i, y_i) with for certain i bad data (not simply imprecise).
- Correct data are called inliers and incorrect outliers. Hypothesis: inliers are coherent while outliers are random.
- RANdom SAmple Consensus (Fishler&Bolles, 1981):
  1. Select k samples out of n, k being the minimal number to estimate uniquely a model.
  2. Compute model and count samples among n explained by model at precision σ.
  3. If this number is larger than the most coherent one until now, keep it.
  4. Back to 1 if we have iterations left.
- Example: $k = 2$ for a plane regression line.

RANSAC for fundamental matrix

- Choose $k = 7$ or $k = 8$
- Classify (x_i, x_i') inlier/outlier as a function of the distance of x_i' to epipolar line associated to x_i $(F^T x_i)$.
- $k = 7$ is better, because we have fewer chances to select an outlier. In that case, we can have 3 models by sample. We test the 3 models.

RANSAC: number of iterations

- Suppose there are m inliers.
- The probability of having an uncontaminated sample of k inliers is $(m/n)^k$
- We require the probability that $N_{\texttt{iter}}$ samples are bad to be lower than $\beta = 1\%$:

$$\left(1 - (m/n)^k\right)^{N_{\texttt{iter}}} \leq \beta$$

- Therefore we need

$$N_{\texttt{iter}} \geq \left\lceil \frac{\log \beta}{\log(1 - (m/n)^k)} \right\rceil.$$

- m is unknown, but a lower bound is the best number of inliers found so far.
- $\Rightarrow$ recompute $N_{\texttt{iter}}$ each time a better model is found.

Conclusion

- Epipolar constraint:
  1. Essential matrix E (calibrated case)
  2. Fundamental matrix F (non calibrated case)
- F can be computed with the 7- or 8-point algorithm.
- Computation of E is much more complicated (5-point algorithm)
- Removing outliers through RANSAC algorithm.

Practical session: RANSAC algorithm for F computation

Objective: Fundamental matrix computation with RANSAC algorithm.

- Get initial program from the website.
- Write the core of function `ComputeF`. Use RANSAC algorithm (update $N_{\texttt{iter}}$ dynamically), based on 8-point algorithm. Solve the linear system estimating F from 8 matches. Do not forget normalization! Hint: it is easier to use SVD with a square matrix. For that, add the 9th equation $0^T f = 0$.
- After RANSAC, refine resulting F with least square minimization based on all inliers.
- Write the core of `displayEpipolar`: when user clicks, find in which image (left or right). Display this point and show associated epipolar line in other image.