Multiple View Geometry

Richard Hartley and Andrew Zisserman

CVPR June 1999

Part I: Single and Two View Geometry

The main points covered in this part are:

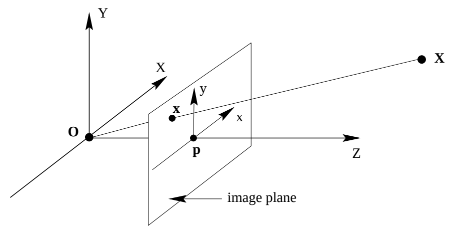
- A perspective (central) projection camera is represented by a 3×4 matrix.
- The most general perspective transformation transformation between two planes (a world plane and the image plane, or two image planes induced by a world plane) is a plane projective transformation. This can be computed from the correspondence of four (or more) points.
- The epipolar geometry between two views is represented by the fundamental matrix. This can be computed from the correspondence of seven (or more) points.

Imaging Geometry

Perspective projection

$$\lambda \begin{pmatrix} x \\ y \\ f \end{pmatrix} = \begin{pmatrix} \mathrm{X} \\ \mathrm{Y} \\ \mathrm{Z} \end{pmatrix}$$

where $\lambda = \mathrm{Z}/f$.



This can be written as a linear mapping between homogeneous coordinates (the equation is only up to a scale factor):

$$\begin{pmatrix} x \\ y \\ f \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} \mathrm{X} \\ \mathrm{Y} \\ \mathrm{Z} \\ 1 \end{pmatrix}$$

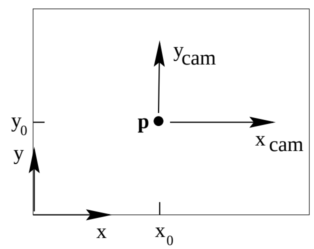
where a 3×4 projection matrix represents a map from 3D to 2D.

Image Coordinate System

Internal camera parameters

$$k_x\, x_{\mathrm{cam}} = x - x_0$$
$$k_y\, y_{\mathrm{cam}} = y - y_0$$

where the units of k_x, k_y are $[\mathrm{pixels/length}]$.



$$\mathbf{x} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \frac{1}{f} \begin{bmatrix} \alpha_x & & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix} \begin{pmatrix} x_{\mathrm{cam}} \\ y_{\mathrm{cam}} \\ f \end{pmatrix} = \mathtt{K} \begin{pmatrix} x_{\mathrm{cam}} \\ y_{\mathrm{cam}} \\ f \end{pmatrix}$$

where $\alpha_x = f k_x,\ \alpha_y = f k_y$.

Camera Calibration Matrix

$\mathtt{K}$ is a 3×3 upper triangular matrix, called the camera calibration matrix:

$$\mathtt{K} = \begin{bmatrix} \alpha_x & & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}$$

- There are four parameters:

  (i) The scaling in the image x and y directions, α_x and α_y.

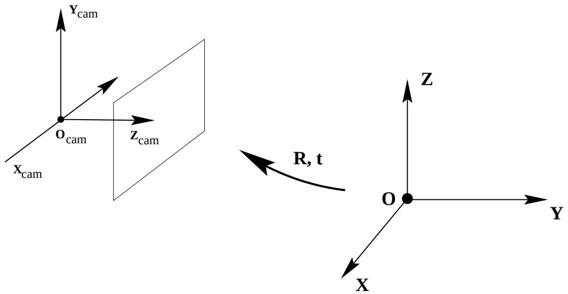
  (ii) The principal point (x_0, y_0), which is the point where the optic axis intersects the image plane.

- The aspect ratio is α_y / α_x.

World Coordinate System

External camera parameters

$$\begin{pmatrix} X_{\text{cam}} \\ Y_{\text{cam}} \\ Z_{\text{cam}} \\ 1 \end{pmatrix} = \begin{bmatrix} \mathtt{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$



Euclidean transformation between world and camera coordinates

- $\mathtt{R}$ is a 3×3 rotation matrix
- $\mathbf{t}$ is a 3×1 translation vector

Concatenating the three matrices,

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \mathtt{K} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathtt{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \mathtt{K}[\mathtt{R} | \mathbf{t}]\mathbf{X}$$

which defines the 3×4 projection matrix from Euclidean 3-space to an image as

$$\mathbf{x} = \mathtt{P}\mathbf{X} \qquad \mathtt{P} = \mathtt{K}[\mathtt{R} | \mathbf{t}] = \mathtt{K}\mathtt{R}[\mathtt{I} | \mathtt{R}^\top \mathbf{t}]$$

Note, the camera centre is at $(X, Y, Z)^\top = -\mathtt{R}^\top \mathbf{t}$.

In the following it is often only the 3×4 form of $\mathtt{P}$ that is important, rather than its decomposition.

A Projective Camera

The camera model for perspective projection is a linear map between homogeneous point coordinates

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{bmatrix} & & \\ & \mathtt{P} \ (3 \times 4) & \\ & & \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Image Point   Scene Point

$$\mathbf{x} \quad = \quad \mathtt{P} \quad \mathbf{X}$$

- The camera centre is the null-vector of $\mathtt{P}$
  e.g. if $\mathtt{P} = [\mathtt{I} | 0]$ then the centre is $\mathbf{X} = (0, 0, 0, 1)^\top$.
- $\mathtt{P}$ has 11 degrees of freedom (essential parameters).
- $\mathtt{P}$ has rank 3.

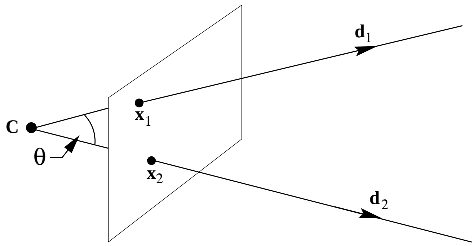
What does calibration give?

- $\mathtt{K}$ provides the transformation between an image point and a ray in Euclidean 3-space.
- Once $\mathtt{K}$ is known the camera is termed calibrated.
- A calibrated camera is a direction sensor, able to measure the direction of rays — like a 2D protractor.

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{bmatrix} \alpha_x & & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix} \begin{pmatrix} \mathrm{X}_{\mathrm{cam}} \\ \mathrm{Y}_{\mathrm{cam}} \\ \mathrm{Z}_{\mathrm{cam}} \end{pmatrix} = \mathtt{K}\mathbf{d}$$

Angle between rays

$$\cos\theta = \frac{\mathbf{d}_1.\mathbf{d}_2}{(\mathbf{d}_1.\mathbf{d}_1)^{\frac{1}{2}}(\mathbf{d}_2.\mathbf{d}_2)^{\frac{1}{2}}}$$



$$\cos\theta = \frac{\mathbf{d}_1{}^\top\mathbf{d}_2}{(\mathbf{d}_1{}^\top\mathbf{d}_1)^{1/2}(\mathbf{d}_2{}^\top\mathbf{d}_2)^{1/2}} = \frac{\mathbf{x}_1{}^\top(\mathtt{K}^{-\top}\mathtt{K}^{-1})\mathbf{x}_2}{(\mathbf{x}_1{}^\top(\mathtt{K}^{-\top}\mathtt{K}^{-1})\mathbf{x}_1)^{1/2}(\mathbf{x}_2{}^\top(\mathtt{K}^{-\top}\mathtt{K}^{-1})\mathbf{x}_2)^{1/2}}$$
$$= \frac{\mathbf{x}_1{}^\top\boldsymbol{\omega}\mathbf{x}_2}{(\mathbf{x}_1{}^\top\boldsymbol{\omega}\mathbf{x}_1)^{1/2}(\mathbf{x}_2{}^\top\boldsymbol{\omega}\mathbf{x}_2)^{1/2}}$$

where $\boldsymbol{\omega} = (\mathtt{K}\mathtt{K}^\top)^{-1}$.

Camera Calibration (Resectioning) (DLT)

Problem Statement:

Given n correspondences $\mathbf{x}_i \leftrightarrow \mathbf{X}_i$, where $\mathbf{X}_i$ is a scene point and $\mathbf{x}_i$ its image:

Compute

$\mathtt{P} = \mathtt{K}\,[\mathtt{R} | \,\mathbf{t}]$ such that $\mathbf{x}_i = \mathtt{P}\mathbf{X}_i$.

The algorithm for camera calibration has two parts:

(i) Compute the matrix $\mathtt{P}$ from a set of point correspondences.

(ii) Decompose $\mathtt{P}$ into $\mathtt{K}, \mathtt{R}$ and $\mathbf{t}$ via the $\mathtt{QR}$ decomposition.

Algorithm step 1: Compute the matrix $\mathtt{P}$ (DLT)

$$\mathbf{x}_i = \mathtt{P}\mathbf{X}_i.$$

Each correspondence generates two equations

$$x_i = \frac{p_{11}\mathrm{X}_i + p_{12}\mathrm{Y}_i + p_{13}\mathrm{Z}_i + p_{14}}{p_{31}\mathrm{X}_i + p_{32}\mathrm{Y}_i + p_{33}\mathrm{Z}_i + p_{34}} \qquad y_i = \frac{p_{21}\mathrm{X}_i + p_{22}\mathrm{Y}_i + p_{23}\mathrm{Z}_i + p_{24}}{p_{31}\mathrm{X}_i + p_{32}\mathrm{Y}_i + p_{33}\mathrm{Z}_i + p_{34}}$$

Multiplying out gives equations linear in the matrix elements of $\mathtt{P}$

$$x_i(p_{31}\mathrm{X}_i + p_{32}\mathrm{Y}_i + p_{33}\mathrm{Z}_i + p_{34}) = p_{11}\mathrm{X}_i + p_{12}\mathrm{Y}_i + p_{13}\mathrm{Z}_i + p_{14}$$
$$y_i(p_{31}\mathrm{X}_i + p_{32}\mathrm{Y}_i + p_{33}\mathrm{Z}_i + p_{34}) = p_{21}\mathrm{X}_i + p_{22}\mathrm{Y}_i + p_{23}\mathrm{Z}_i + p_{24}$$

These equations can be rearranged as

$$\begin{bmatrix} \mathrm{X} & \mathrm{Y} & \mathrm{Z} & 1 & 0 & 0 & 0 & 0 & -x\mathrm{X} & -x\mathrm{Y} & -x\mathrm{Z} & -x \\ 0 & 0 & 0 & 0 & \mathrm{X} & \mathrm{Y} & \mathrm{Z} & 1 & -y\mathrm{X} & -y\mathrm{Y} & -y\mathrm{Z} & -y \end{bmatrix} \mathbf{p} = \mathbf{0}$$

with $\mathbf{p} = (p_{11}, p_{12}, p_{13}, p_{14}, p_{21}, p_{22}, p_{23}, p_{24}, p_{31}, p_{32}, p_{33}, p_{34})^\top$ a 12-vector.

Algorithm step 1 continued

Solving for $\mathtt{P}$

(i) Concatenate the equations from ($n \geq 6$) correspondences to generate $2n$ simultaneous equations, which can be written: $\mathtt{A}\mathbf{p} = \mathbf{0}$, where $\mathtt{A}$ is a $2n \times 12$ matrix.

(ii) In general this will not have an exact solution, but a (linear) solution which minimises $|\mathtt{A}\mathbf{p}|$, subject to $|\mathbf{p}| = 1$ is obtained from the eigenvector with least eigenvalue of $\mathtt{A}^\top\mathtt{A}$. Or equivalently from the vector corresponding to the smallest singular value of the SVD of $\mathtt{A}$.

(iii) This linear solution is then used as the starting point for a non-linear minimisation of the difference between the measured and projected point:

$$\min_{\mathtt{P}} \sum_i \left((x_i, y_i) - P(\mathrm{X}_i, \mathrm{Y}_i, \mathrm{Z}_i, 1)\right)^2$$

Algorithm step 2: Decompose $\mathtt{P}$ into $\mathtt{K}, \mathtt{R}$ and $\mathbf{t}$

The first 3×3 submatrix, $\mathtt{M}$, of $\mathtt{P}$ is the product ($\mathtt{M} = \mathtt{KR}$) of an upper triangular and rotation matrix.

(i) Factor $\mathtt{M}$ into $\mathtt{KR}$ using the $\mathtt{QR}$ matrix decomposition. This determines $\mathtt{K}$ and $\mathtt{R}$.

(ii) Then

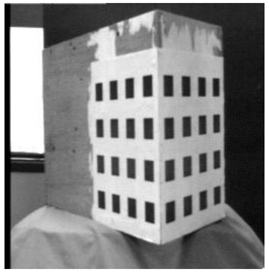
$$\mathbf{t} = \mathtt{K}^{-1}(p_{14}, p_{24}, p_{34})^\top$$

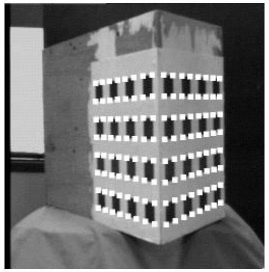
Note, this produces a matrix with an extra skew parameter s

$$\mathtt{K} = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix}$$

with $s = \tan\theta$, and θ the angle between the image axes.

Example - Calibration Object





Determine accurate corner positions by

(i) Extract and link edges using Canny edge operator.

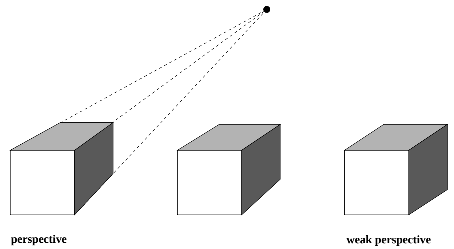
(ii) Fit lines to edges using orthogonal regression.

(iii) Intersect lines to obtain corners to sub-pixel accuracy.

The final error between measured and projected points is typically less than 0.02 pixels.

Weak Perspective

Track back, whilst zooming to keep image size fixed



The imaging rays become parallel, and the result is:

$$\mathtt{P} = \mathtt{K} \begin{bmatrix} r_{11} & r_{12} & r_{13} & * \\ r_{21} & r_{22} & r_{23} & * \\ 0 & 0 & 0 & * \end{bmatrix}$$

A generalization is the *affine camera*

The Affine Camera

$$\mathtt{P} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & t_1 \\ m_{21} & m_{22} & m_{23} & t_2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

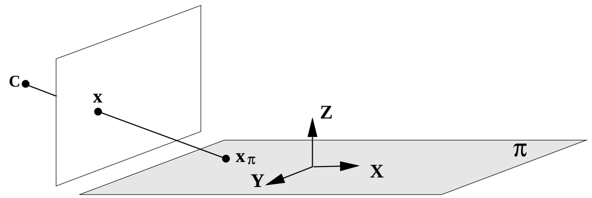
The matrix $\mathtt{M}_{2\times 3}$ has rank two.

Projection under an affine camera is a linear mapping on non-homogeneous coordinates composed with a translation:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \end{bmatrix} \begin{pmatrix} \mathtt{X} \\ \mathtt{Y} \\ \mathtt{Z} \end{pmatrix} + \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}$$

- The point $(t_1, t_2)^\top$ is the image of the world origin.
- The centre of the affine camera is at infinity.
- An affine camera has 8 degrees of freedom.
- It models weak-perspective and para-perspective.

Plane projective transformations



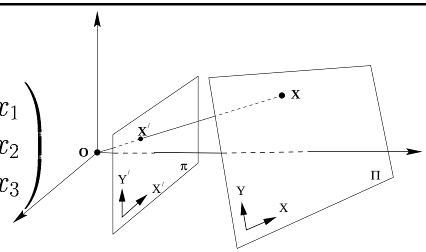
Choose the world coordinate system such that the plane of the points has zero $\mathtt{Z}$ coordinate. Then the 3×4 matrix $\mathtt{P}$ reduces to

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{pmatrix} \mathtt{X} \\ \mathtt{Y} \\ 0 \\ 1 \end{pmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{14} \\ p_{21} & p_{22} & p_{24} \\ p_{31} & p_{32} & p_{34} \end{bmatrix} \begin{pmatrix} \mathtt{X} \\ \mathtt{Y} \\ 1 \end{pmatrix}$$

which is a 3×3 matrix representing a general plane to plane projective transformation.

Projective transformations continued

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$



or $\mathbf{x}' = \mathtt{H}\mathbf{x}$, where $\mathtt{H}$ is a 3×3 non-singular homogeneous matrix.

- This is the most general transformation between the world and image plane under imaging by a perspective camera.
- It is often only the 3×3 form of the matrix that is important in establishing properties of this transformation.
- A projective transformation is also called a “homography” and a “collineation”.
- $\mathtt{H}$ has 8 degrees of freedom.

Four points define a projective transformation

Given n point correspondences $(x, y) \leftrightarrow (x', y')$

Compute $\mathtt{H}$ such that $\mathbf{x}'_i = \mathtt{H}\mathbf{x}_i$

- Each point correspondence gives two constraints

$$x' = \frac{x'_1}{x'_3} = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}, \qquad y' = \frac{x'_2}{x'_3} = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}}$$

and multiplying out generates two linear equations for the elements of $\mathtt{H}$

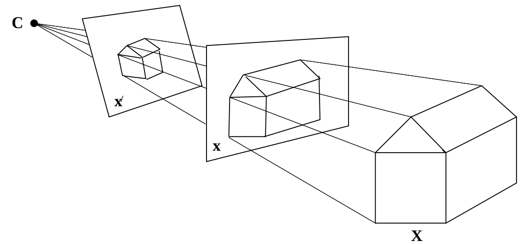
$$x'\,(h_{31}x + h_{32}y + h_{33}) = h_{11}x + h_{12}y + h_{13}$$
$$y'\,(h_{31}x + h_{32}y + h_{33}) = h_{21}x + h_{22}y + h_{23}$$

- If $n \geq 4$ (no three points collinear), then $\mathtt{H}$ is determined uniquely.
- The converse of this is that it is possible to transform any four points in general position to any other four points in general position by a projectivity.

The Cone of Rays

An image is the intersection of a plane with the cone of rays between points in 3-space and the optical centre. Any two such "images" (with the same camera centre) are related by a planar projective transformation.

$$\mathbf{x}' = \mathtt{H}\mathbf{x}$$



e.g. rotation about the camera centre

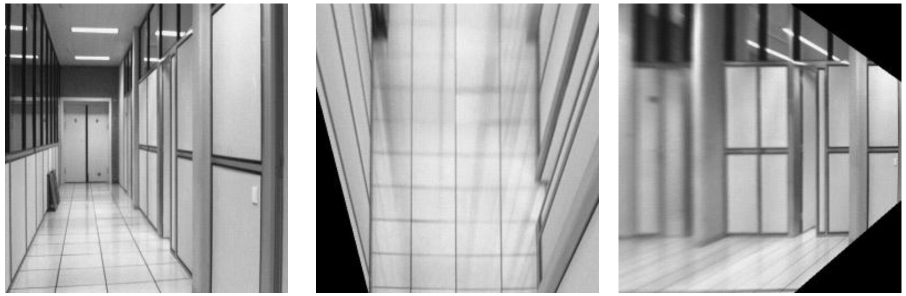
Example 1: Removing Perspective Distortion

Given: the coordinates of four points on the scene plane

Find: a projective rectification of the plane

- This rectification does not require knowledge of any of the camera's parameters or the pose of the plane.
- It is not always necessary to know coordinates for four points.

Example 2: Synthetic Rotations



The synthetic images are produced by projectively warping the original image so that four corners of an imaged rectangle map to the corners of a rectangle. Both warpings correspond to a synthetic rotation of the camera about the (fixed) camera centre.

Two View Geometry

- Cameras $\mathtt{P}$ and $\mathtt{P}'$ such that

$$\mathbf{x} = \mathtt{P}\mathbf{X} \qquad \mathbf{x}' = \mathtt{P}'\mathbf{X}$$

- Baseline between the cameras is non-zero.

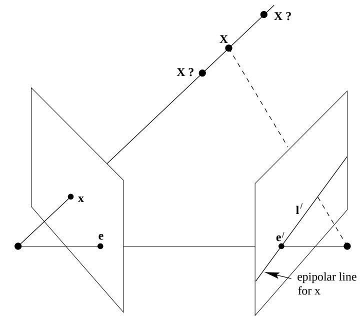
Given an image point in the first view, where is the corresponding point in the second view?

What is the relative position of the cameras?

What is the 3D geometry of the scene?

Correspondence Geometry

Given the image of a point in one view, what can we say about its position in another?



- A point in one image “generates” a line in the other image.
- This line is known as an epipolar line, and the geometry which gives rise to it is known as epipolar geometry.

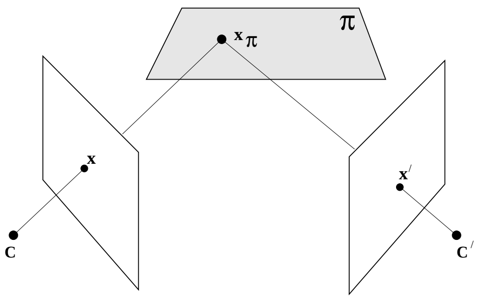
Images of Planes

Projective transformation between images induced by a plane

$$\mathbf{x} = \mathtt{H}_{1\pi}\mathbf{x}_\pi \qquad \mathbf{x}' = \mathtt{H}_{2\pi}\mathbf{x}_\pi$$

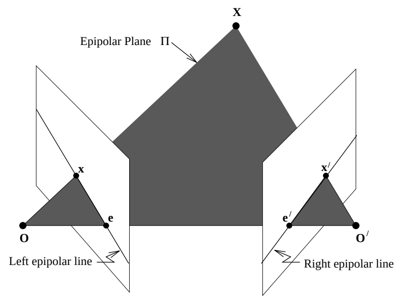
$$\mathbf{x}' = \mathtt{H}_{2\pi}\mathbf{x}_\pi$$

$$= \mathtt{H}_{2\pi}\mathtt{H}_{1\pi}^{-1}\mathbf{x} = \mathtt{H}\mathbf{x}$$



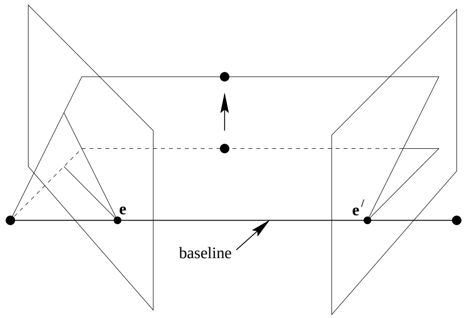
- $\mathtt{H}$ can be computed from the correspondence of four points on the plane.

Epipolar Geometry



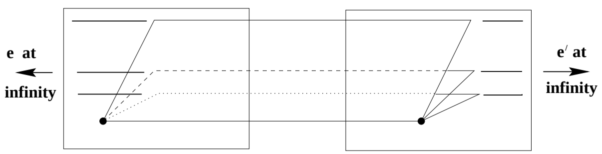
- The epipolar line $\mathbf{l}'$ is the image of the ray through $\mathbf{x}$.
- The epipole $\mathbf{e}'$ is the point of intersection of the line joining the camera centres—the baseline—with the image plane.
- The epipole is also the image in one camera of the centre of the other camera.
- All epipolar lines intersect in the epipole.

Epipolar pencil



As the position of the 3D point $\mathbf{X}$ varies, the epipolar planes "rotate" about the baseline. This family of planes is known as an epipolar pencil. All epipolar lines intersect at the epipole.

Epipolar geometry example



Epipolar geometry depends only on the relative pose (position and orientation) and internal parameters of the two cameras, i.e. the position of the camera centres and image planes. It does not depend on structure (3D points external to the camera).

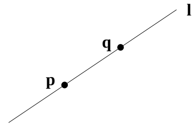
Homogeneous Notation Interlude

- A line $\mathbf{l}$ is represented by the homogeneous 3-vector

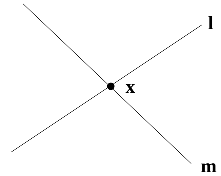
$$\mathbf{l} = \begin{pmatrix} l_1 \\ l_2 \\ l_3 \end{pmatrix}$$

for the line $l_1 x + l_2 y + l_3 = 0$. Only the ratio of the homogeneous line coordinates is significant.

- point on line: $\mathbf{l}.\mathbf{x} = 0$ or $\mathbf{l}^\top \mathbf{x} = 0$ or $\mathbf{x}^\top \mathbf{l} = 0$



- two points define a line: $\mathbf{l} = \mathbf{p} \times \mathbf{q}$



- two lines define a point: $\mathbf{x} = \mathbf{l} \times \mathbf{m}$

Matrix notation for vector product

The vector product $\mathbf{v} \times \mathbf{x}$ can be represented as a matrix multiplication

$$\mathbf{v} \times \mathbf{x} = [\mathbf{v}]_\times \mathbf{x}$$

where

$$[\mathbf{v}]_\times = \begin{bmatrix} 0 & -v_z & v_y \\ v_z & 0 & -v_x \\ -v_y & v_x & 0 \end{bmatrix}$$

- $[\mathbf{v}]_\times$ is a 3×3 skew-symmetric matrix of rank 2.
- $\mathbf{v}$ is the null-vector of $[\mathbf{v}]_\times$, since $\mathbf{v} \times \mathbf{v} = [\mathbf{v}]_\times \mathbf{v} = \mathbf{0}$.

Algebraic representation - the Fundamental Matrix

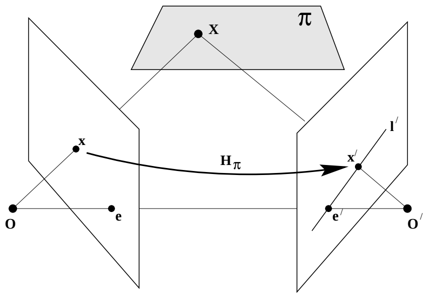
$$\mathbf{x}'^{\top}\, \mathtt{F}\, \mathbf{x} = 0 \qquad \mathbf{l}' = \mathtt{F}\,\mathbf{x}$$

- $\mathtt{F}$ is a 3×3 rank 2 homogeneous matrix
- $\mathtt{F}^{\top}\mathbf{e}' = \mathbf{0}$
- It has 7 degrees of freedom
- Counting: $2 \times 11 - 15 = 7$.
- Compute from 7 image point correspondences

Properties of F

- $\mathtt{F}$ is a rank 2 homogeneous matrix with 7 degrees of freedom.
- Point correspondence: If $\mathbf{x}$ and $\mathbf{x}'$ are corresponding image points, then $\mathbf{x}'^{\top}\mathtt{F}\mathbf{x} = 0$.
- Epipolar lines:
  - ◇ $\mathbf{l}' = \mathtt{F}\mathbf{x}$ is the epipolar line corresponding to $\mathbf{x}$.
  - ◇ $\mathbf{l} = \mathtt{F}^{\top}\mathbf{x}'$ is the epipolar line corresponding to $\mathbf{x}'$.
- Epipoles:
  - ◇ $\mathtt{F}\mathbf{e} = \mathbf{0} \qquad \mathtt{F}^{\top}\mathbf{e}' = \mathbf{0}$
- Computation from camera matrices $\mathtt{P}, \mathtt{P}'$:
  - ◇ $\mathtt{F} = [\mathtt{P}'\mathbf{C}]_{\times}\mathtt{P}'\mathtt{P}^{+}$, where $\mathtt{P}^{+}$ is the pseudo-inverse of $\mathtt{P}$, and $\mathbf{C}$ is the centre of the first camera. Note, $\mathbf{e}' = \mathtt{P}'\mathbf{C}$.
  - ◇ Canonical cameras, $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}], \mathtt{P}' = [\mathtt{M} \mid \mathbf{m}]$, $\mathtt{F} = [\mathbf{e}']_{\times}\mathtt{M} = \mathtt{M}^{-\top}[\mathbf{e}]_{\times}$, where $\mathbf{e}' = \mathbf{m}$ and $\mathbf{e} = \mathtt{M}^{-1}\mathbf{m}$.

Fundamental matrix - sketch derivation



Step 1: Point transfer via a plane $\mathbf{x}' = \mathtt{H}_{\pi}\mathbf{x}$

Step 2 : Construct the epipolar line $\mathbf{l}' = \mathbf{e}' \times \mathbf{x}' = [\mathbf{e}']_{\times}\mathbf{x}'$

$\mathbf{l}' = [\mathbf{e}']_{\times}\mathtt{H}_{\pi}\mathbf{x} = \mathtt{F}\mathbf{x}$

$\mathtt{F} = [\mathbf{e}']_{\times}\mathtt{H}_{\pi}$

This shows that $\mathtt{F}$ is a 3×3 rank 2 matrix.

Plane induced homographies given F

Given the fundamental matrix $\mathtt{F}$ between two views, the homography induced by a world plane is

$$\mathtt{H} = [\mathbf{e}']_{\times}\mathtt{F} + \mathbf{e}'\mathbf{v}^{\top}$$

where $\mathbf{v}$ is the inhomogeneous 3-vector which parametrizes the 3-parameter family of planes.

e.g. compute plane from 3 point correspondences.

Given a homography $\hat{\mathtt{H}}$ induced by a particular world plane, then a homography induced by any plane may be computed as

$$\mathtt{H} = \hat{\mathtt{H}} + \mathbf{e}'\mathbf{v}^{*\top}$$

Projective Reconstruction from 2 views

Statement of the problem

<u>Given</u>

Corresponding points $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ in two images.

<u>Find</u>

Cameras $\mathtt{P}$ and $\mathtt{P}'$ and 3D points $\mathbf{X}_i$ such that

$$\mathbf{x}_i = \mathtt{P}\mathbf{X}_i \quad ; \quad \mathbf{x}'_i = \mathtt{P}\mathbf{X}_i$$

Projective ambiguity of reconstruction

- Solution is not unique without camera calibration
- Solution is unique up to a projective mapping :

$$\begin{aligned} \mathtt{P} &\mapsto \mathtt{P}\mathtt{H}^{-1} \\ \mathtt{P}' &\mapsto \mathtt{P}'\mathtt{H}^{-1} \\ \mathbf{X}_i &\mapsto \mathtt{H}\mathbf{X}_i \end{aligned}$$

Then verify

$$\begin{aligned} \mathbf{x}_i &= (\mathtt{P}\mathtt{H}^{-1})(\mathtt{H}\mathbf{X}_i) = \mathtt{P}\mathbf{X}_i \\ \mathbf{x}'_i &= (\mathtt{P}'\mathtt{H}^{-1})(\mathtt{H}\mathbf{X}_i) = \mathtt{P}'\mathbf{X}_i \end{aligned}$$

- Same problem holds however many views we have

Projective Distortion demo

< Projective distortion demo >

Basic Theorem

Given sufficiently many points to compute unique fundamental matrix :

- 8 points in general position
- 7 points not on a ruled quadric with camera centres

Then 3D points may be constructed from two views **Up to a 3D projective transformation**

- Except for points on the line between the camera centres.

Details of Projective Reconstruction - Computation of $\mathtt{F}$.

Methods of computation of $\mathtt{F}$ left until later

Several methods are available :

(i) Normalized 8-point algorithm
(ii) Algebraic minimization
(iii) Minimization of epipolar distance
(iv) Minimization of symmetric epipolar distance
(v) Maximum Likelihood (Gold-standard) method.
(vi) Others , . . .

Steps of projective reconstruction

Reconstruction takes place in the following steps :

- Compute the fundamental matrix $\mathtt{F}$ from point correspondences
- Factor the fundamental matrix as

$$\mathtt{F} = [\mathbf{t}]_\times \mathtt{M}$$

- The two camera matrices are

$$\mathtt{P} = [\mathtt{I} \mid \mathbf{0}] \text{ and } \mathtt{P}' = [\mathtt{M} \mid \mathbf{t}] \ .$$

- Compute the points $\mathbf{X}_i$ by triangulation

Factorization of the fundamental matrix

SVD method

(i) Define

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

(ii) Compute the SVD

$$\mathtt{F} = \mathtt{UDV}^\top \text{ where } \mathtt{D} = \mathrm{diag} r, s, 0$$

(iii) Factorization is

$$\mathtt{F} = (\mathtt{UZU}^\top)(\mathtt{UZDV}^\top)$$

- Simultaneously corrects $\mathtt{F}$ to a singular matrix.

Factorization of the fundamental matrix

Direct formula

Let $\mathbf{e}'$ be the epipole.

$$\text{Solve } \mathbf{e}'^{\top}\mathtt{F} = 0$$

Specific formula

$$\mathtt{P} = [\mathtt{I} \mid \mathbf{0}] \;\; ; \;\; \mathtt{P}' = [[\mathbf{e}']_{\times}\mathtt{F} \mid \mathbf{e}'] = [\mathtt{M} \mid \mathbf{e}']$$

This solution is identical to the SVD solution.

Non-uniqueness of factorization

- Factorization of the fundamental matrix is not unique.
- General formula : for varying $\mathbf{v}$ and λ

$$\mathtt{P} = [\mathtt{I} \mid \mathbf{0}] \;\; ; \;\; \mathtt{P}' = [\mathtt{M} + \mathbf{e}'\mathbf{v}^{\top} \mid \lambda\mathbf{e}']$$

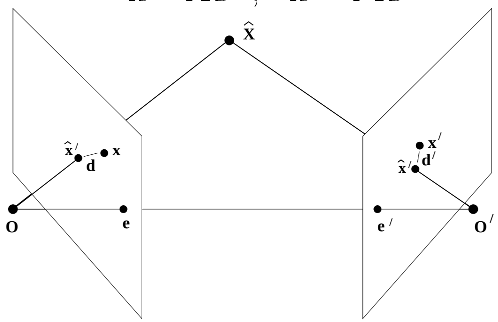
- Difference factorizations give configurations varying by a projective transformation.
- 4-parameter family of solutions with $\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$.

Triangulation

Triangulation :

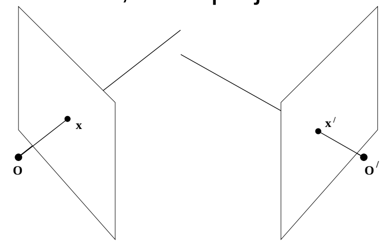
- Knowing $\mathtt{P}$ and $\mathtt{P}'$
- Knowing $\mathbf{x}$ and $\mathbf{x}'$
- Compute $\mathbf{X}$ such that

$$\mathbf{x} = \mathtt{P}\mathbf{X} \;\; ; \;\; \mathbf{x}' = \mathtt{P}'\mathbf{X}$$

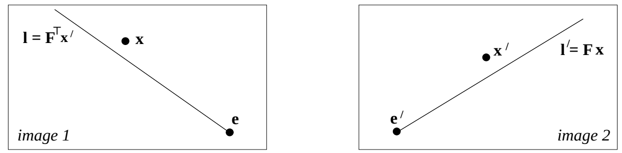


Triangulation in presence of noise

- In the presence of noise, back-projected lines do not intersect.



Rays do not intersect in space



Measured points do not lie on corresponding epipolar lines

Which 3D point to select ?

- Mid-point of common perpendicular to the rays ?
  - **Not** a good choice in projective environment.
  - Concepts of mid-point and perpendicular are meaningless under projective distortion.
- Weighted point on common perpendicular, weighted by distance from camera centres ?
  - Distance is also undefined concept.
- Some algebraic distance ?
  - Write down projection equations and solve ?
  - Linear least squares solution.
  - Minimizes nothing meaningful.

Linear triangulation methods

- Direct analogue of the linear method of camera resectioning.
- Given equations

$$\mathbf{x} = \mathtt{P}\mathbf{X}$$
$$\mathbf{x}' = \mathtt{P}'\mathbf{X}$$

- $\mathbf{p}^{i\top}$ are the rows of $\mathtt{P}$.
- Write as linear equations in $\mathbf{X}$

$$\begin{bmatrix} x\mathbf{p}^{3\top} - \mathbf{p}^{1\top} \\ y\mathbf{p}^{3\top} - \mathbf{p}^{2\top} \\ x'\mathbf{p}'^{3\top} - \mathbf{p}'^{1\top} \\ y\mathbf{p}'^{3\top} - \mathbf{p}'^{2\top} \end{bmatrix} \mathbf{X} = \mathbf{0}$$

- Solve for $\mathbf{X}$.
- Generalizes to point match in several images.
- Minimizes no meaningful quantity – not optimal.

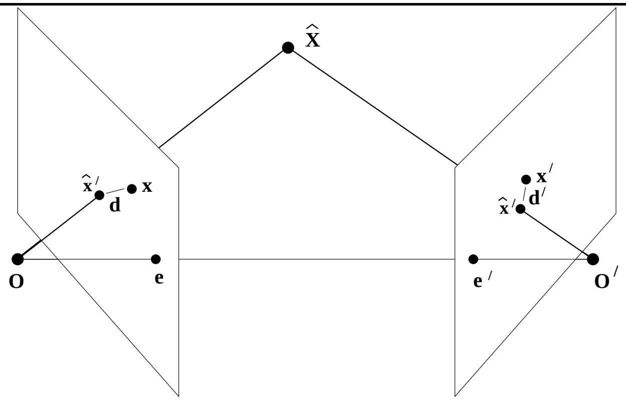
Problem statement

- Assume camera matrices are given without error, up to projective distortion.
- Hence $\mathtt{F}$ is known.
- A pair of matched points in an image are given.
- Possible errors in the position of matched points.
- Find 3D point that minimizes suitable error metric.
- Method must be invariant under 3D projective transformation.

Minimizing geometric error

- Point $\mathbf{X}$ in space maps to projected points $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}'$ in the two images.
- Measured points are $\mathbf{x}$ and $\mathbf{x}'$.
- Find $\mathbf{X}$ that minimizes difference between projected and measured points.

Geometric error . . .



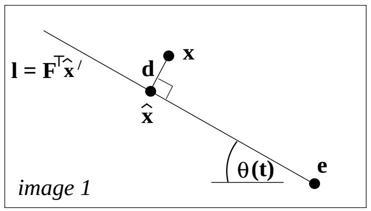
Cost function

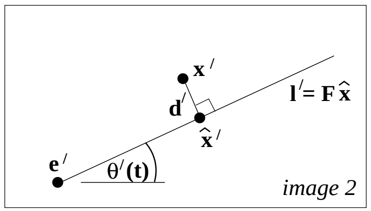
$$\mathcal{C}(\mathbf{X}) = d(\mathbf{x}, \hat{\mathbf{x}})^2 + d(\mathbf{x}', \hat{\mathbf{x}}')^2$$

Different formulation of the problem

Minimization problem may be formulated differently:

- Minimize
$$d(\mathbf{x}, \mathbf{l})^2 + d(\mathbf{x}', \mathbf{l}')^2$$
- $\mathbf{l}$ and $\mathbf{l}'$ range over all choices of corresponding epipolar lines.
- $\hat{\mathbf{x}}$ is the closest point on the line $\mathbf{l}$ to $\mathbf{x}$.
- Same for $\hat{\mathbf{x}}'$.





Minimization method

Our strategy for minimizing cost function is as follows

(i) Parametrize the pencil of epipolar lines in the first image by a parameter t. Epipolar line is $\mathbf{l}(t)$.

(ii) Using the fundamental matrix $\mathtt{F}$, compute the corresponding epipolar line $\mathbf{l}'(t)$ in the second image.

(iii) Express the distance function $d(\mathbf{x}, \mathbf{l}(t))^2 + d(\mathbf{x}', \mathbf{l}'(t))^2$ explicitly as a function of t.

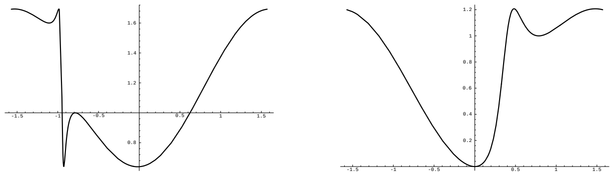
(iv) Find the value of t that minimizes this function.

Minimization method ...

- Find the minimum of a function of a single variable, t.
- Problem in elementary calculus.
- Derivative of cost reduces to a 6-th degree polynomial in t.
- Find roots of derivative explicitly and compute cost function.
- Provides global minimum cost (guaranteed best solution).
- Details : See Hartley-Sturm “Triangulation”.

Multiple local minima

- Cost function may have local minima.
- Shows that gradient-descent minimization may fail.

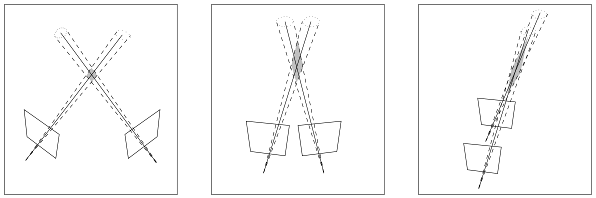


Left : Example of a cost function with three minima.

Right : Cost function for a perfect point match with two minima.

This page left empty

Uncertainty of reconstruction



Uncertainty of reconstruction. The shape of the uncertainty region depends on the angle between the rays.

This page left empty

Computation of the Fundamental Matrix

Basic equations

Given a correspondence

$$\mathbf{x} \leftrightarrow \mathbf{x}'$$

The basic incidence relation is

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

May be written

$$x'x f_{11} + x'y f_{12} + x' f_{13} + y'x f_{21} + y'y f_{22} + y' f_{23} + x f_{31} + y f_{32} + f_{33} = 0 \ .$$

Single point equation - Fundamental matrix

Gives an equation :

$$(x'x, x'y, x', y'x, y'y, y', x, y, 1) \begin{pmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{pmatrix} = 0$$

where

$$\mathbf{f} = (f_{11}, f_{12}, f_{13}, f_{21}, f_{22}, f_{23}, f_{31}, f_{32}, f_{33})^{\top}$$

holds the entries of the Fundamental matrix

Total set of equations

$$\mathtt{A}\mathbf{f} = \begin{bmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_n x_n & x'_n y_n & x'_n & y'_n x_n & y'_n y_n & y'_n & x_n & y_n & 1 \end{bmatrix} \begin{pmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{pmatrix} = \mathbf{0}$$

Solving the Equations

- Solution is determined up to scale only.
- Need 8 equations $\Rightarrow$ 8 points
- 8 points $\Rightarrow$ unique solution
- > 8 points $\Rightarrow$ least-squares solution.

Least-squares solution

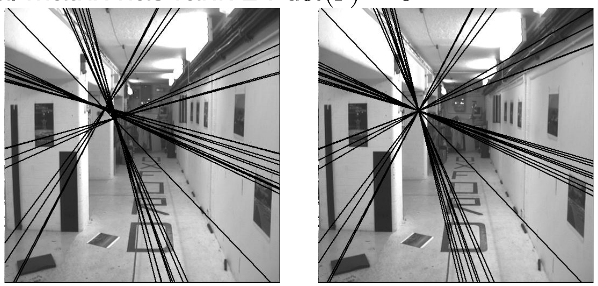
(i) Form equations $\mathtt{A}\mathbf{f} = \mathbf{0}$.
(ii) Take SVD : $\mathtt{A} = \mathtt{UDV}^\top$.
(iii) Solution is last column of $\mathtt{V}$ (corresp : smallest singular value)
(iv) Minimizes $||\mathtt{A}\mathbf{f}||$ subject to $||\mathbf{f}|| = 1$.

Computing $\mathtt{F}$ from 7 points

- $\mathtt{F}$ has 9 entries but is defined only up to scale.
- Singularity condition $\det \mathtt{F} = 0$ gives a further constraint.
- $\mathtt{F}$ has 3 rows $\Longrightarrow \det \mathtt{F} = 0$ is a cubic constraint.
- $\mathtt{F}$ has only 7 degrees of freedom.
- It is possible to solve for $\mathtt{F}$ from just 7 point correspondences.

The singularity constraint

Fundamental matrix has rank 2 : $\det(\mathtt{F}) = 0$.



Left : Uncorrected $\mathtt{F}$ – epipolar lines are not coincident.

Right : Epipolar lines from corrected $\mathtt{F}$.

7-point algorithm

Computation of $\mathtt{F}$ from 7 point correspondences

(i) Form the 7×9 set of equations $\mathtt{A}\mathbf{f} = 0$.
(ii) System has a 2-dimensional solution set.
(iii) General solution (use SVD) has form

$$\mathbf{f} = \lambda \mathbf{f}_0 + \mu \mathbf{f}_1$$

(iv) In matrix terms

$$\mathtt{F} = \lambda \mathtt{F}_0 + \mu \mathtt{F}_1$$

(v) Condition $\det \mathtt{F} = 0$ gives cubic equation in λ and μ.
(vi) Either one or three real solutions for ratio $\lambda : \mu$.

Correcting $\mathtt{F}$ using the Singular Value Decomposition

If $\mathtt{F}$ is computed linearly from 8 or more correspondences, singularity condition does not hold.

SVD Method

(i) SVD : $\mathtt{F} = \mathtt{UDV}^\top$

(ii) $\mathtt{U}$ and $\mathtt{V}$ are orthogonal, $\mathtt{D} = \mathrm{diag}(r, s, t)$.

(iii) $r \geq s \geq t$.

(iv) Set $\mathtt{F}' = \mathtt{U}\,\mathrm{diag}(r, s, 0)\,\mathtt{V}^\top$.

(v) Resulting $\mathtt{F}'$ is singular.

(vi) Minimizes the Frobenius norm of $\mathtt{F} - \mathtt{F}'$

(vii) $\mathtt{F}'$ is the "closest" singular matrix to $\mathtt{F}$.

The normalized 8-point algorithm

Raw 8-point algorithm performs badly in presence of noise.

Normalization of data

- 8-point algorithm is sensitive to origin of coordinates and scale.
- Data must be translated and scaled to "canonical" coordinate frame.
- Normalizing transformation is applied to both images.
- Translate so centroid is at origin
- Scale so that RMS distance of points from origin is $\sqrt{2}$.
- "Average point" is $(1, 1, 1)^\top$.

Complete 8-point algorithm

8 point algorithm has two steps :

(i) **Linear solution.** Solve $\mathtt{A}\mathbf{f} = 0$ to find $\mathtt{F}$.

(ii) **Constraint enforcement.** Replace $\mathtt{F}$ by $\mathtt{F}'$.

Warning This algorithm is unstable and should never be used with unnormalized data (see next slide).

Normalized 8-point algorithm

(i) **Normalization:** Transform the image coordinates :

$$\hat{\mathbf{x}}_i = \mathtt{T}\mathbf{x}_i$$
$$\hat{\mathbf{x}}'_i = \mathtt{T}'\mathbf{x}'_i$$

(ii) **Solution:** Compute $\mathtt{F}$ from the matches $\hat{\mathbf{x}}_i \leftrightarrow \hat{\mathbf{x}}'_i$

$$\hat{\mathbf{x}}'^\top_i \widehat{\mathtt{F}} \hat{\mathbf{x}}_i = 0$$

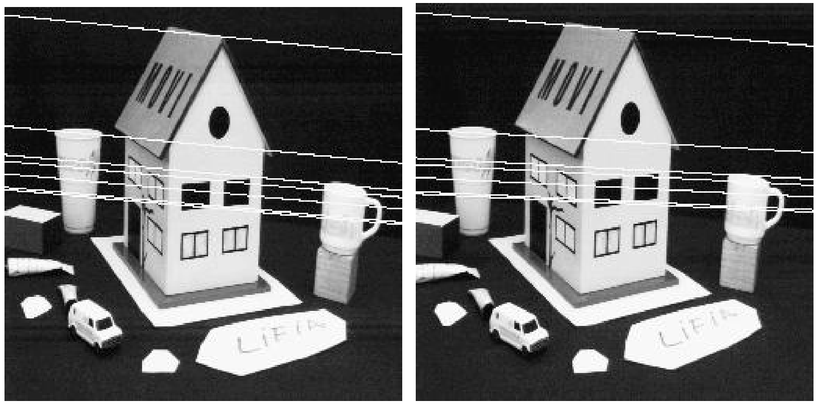
(iii) **Singularity constraint :** Find closest singular $\widehat{\mathtt{F}}'$ to $\widehat{\mathtt{F}}$.

(iv) **Denormalization:** $\mathtt{F} = \mathtt{T}'^\top \widehat{\mathtt{F}}' \mathtt{T}$.

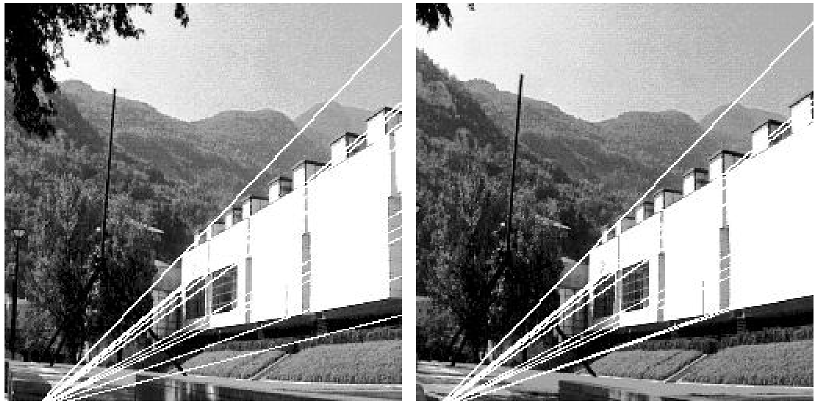
Comparison of Normalized and Unnormalized Algorithms

Statue images

Lifia House images

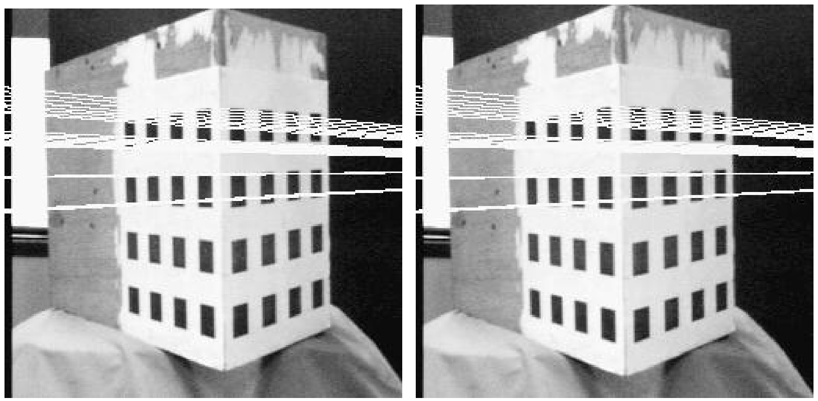


Grenoble Museum



Oxford Basement

Calibration object



Testing methodology

(i) Point matches found in image pairs and outliers discarded.

(ii) Fundamental matrix was found from varying number (n) of points.

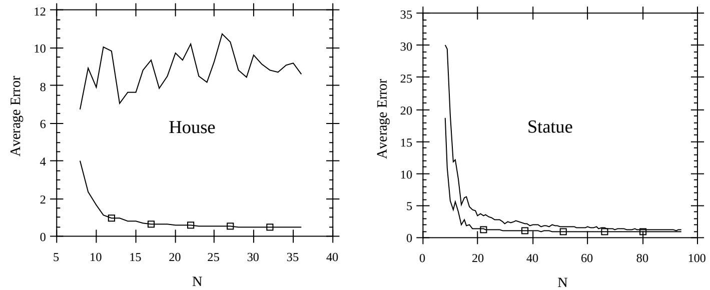
(iii) `F` was tested against other matched points not used to compute it.

(iv) Distance of a point from predicted epipolar line was the metric.

(v) 100 trials for each value of n.

(vi) Average error is plotted against n.

Comparison of normalized and unnormalized 8-point algorithms.

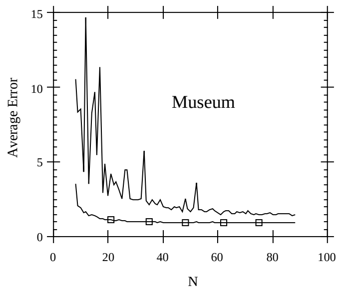


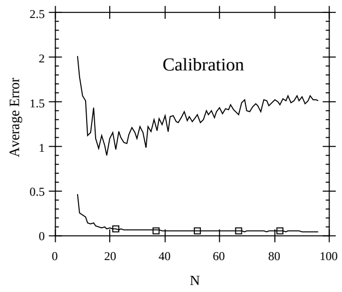
Distance of points from epipolar lines :

Top : Unnormalized 8-point algorithm.

Bottom : Normalized 8-point algorithm.

Comparison of normalized and unnormalized 8-point algorithms.



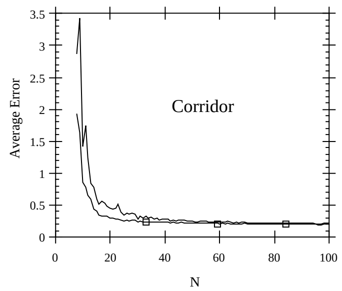


Distance of points from epipolar lines :

Top : Unnormalized 8-point algorithm.

Bottom : Normalized 8-point algorithm.

Comparison of normalized and unnormalized 8-point algorithms.



Distance of points from epipolar lines :

Top : Unnormalized 8-point algorithm.

Bottom : Normalized 8-point algorithm.

Illustration of Effect of Normalization

Similar problem : Computation of a 2D projective transformation given point matches in two images.

(i) Homography is computed from 5 noisy point matches.

(ii) Homography is applied to a further (6th) point

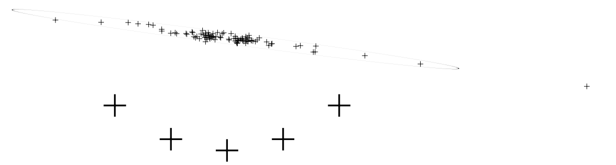
(iii) Noise level approximately equal to with of lines in the crosses (next page)

(iv) Repeated 100 times.

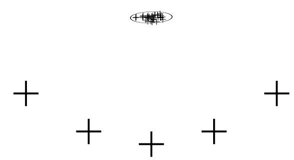
(v) Spread of the transformed 6th point is shown in relation to the 5 data points.

(vi) 95% ellipses are plotted.

Normalization and 2D homography computation

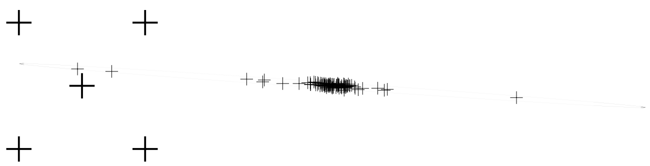


Unnormalized data

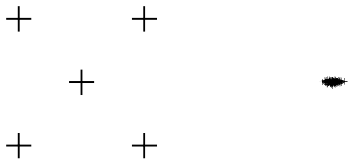


Normalized data

Normalization and 2D homography computation



Unnormalized data

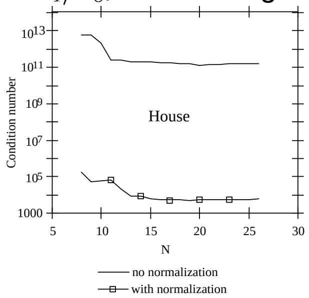


Normalized data

Algebraic Minimization Algorithm

Condition number

- Bad condition number the reason for poor performance.
- Condition number $= \kappa_1/\kappa_8$, ratio of singular values of $\mathtt{A}$.



- Bad conditioning acts as a noise amplifier.
- Normalization improves the condition number by a factor of 10^8.
- **Reference** : Hartley "In defence of the 8-point algorithm".

The algebraic minimization algorithm

Enforcing the singularity constraint

- SVD method minimizes $||\mathtt{F}' - \mathtt{F}||$.
- simple and rapid.
- Not optimal
- Treats all entries of $\mathtt{F}$ equally.
- However, some entries of $\mathtt{F}$ are more tightly constrained by the data.

The Algebraic Method

- Minimize $||\mathtt{A}\mathbf{f}'||$ subject to $||\mathbf{f}'|| = 1$ **AND** $\det \mathtt{F}' = 0$.
- $\det \mathtt{F}' = 0$ is a cubic constraint.
- Requires an iterative solution.
- However, simple iterative method works.

Solution assuming known epipole

- We may write $\mathtt{F} = \mathtt{M}[\mathbf{e}]_\times$, where $\mathbf{e}$ is epipole.
- $\mathtt{F}$ of this form is singular.
- Assume $\mathbf{e}$ is known, find $\mathtt{M}$.
- Write $\mathtt{F} = \mathtt{M}[\mathbf{e}]_\times$ as $\mathbf{f} = \mathtt{E}\mathbf{m}$

$$\begin{pmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \\ f_{31} \\ f_{32} \\ f_{33} \end{pmatrix} = \begin{bmatrix} 0 & -e_3 & e_2 & & & & & & \\ e_3 & 0 & -e_1 & & & & & & \\ -e_2 & e_1 & 0 & & & & & & \\ & & & 0 & -e_3 & e_2 & & & \\ & & & e_3 & 0 & -e_1 & & & \\ & & & -e_2 & e_1 & 0 & & & \\ & & & & & & 0 & -e_3 & e_2 \\ & & & & & & e_3 & 0 & -e_1 \\ & & & & & & -e_2 & e_1 & 0 \end{bmatrix} \begin{pmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{21} \\ m_{22} \\ m_{23} \\ m_{31} \\ m_{32} \\ m_{33} \end{pmatrix}$$

Solution assuming known epipole - continued

- Write $\mathtt{A}\mathbf{f} = \mathtt{A}\mathtt{E}\mathbf{m}$.
- Minimize $||\mathtt{A}\mathtt{E}\mathbf{m}||$ subject to $||\mathtt{E}\mathbf{m}|| = 1$.
- This is a linear least-squares estimation problem.
- Non-iterative algorithm involving SVD is possible
- Reference : Hartley, “Minimizing Algebraic Error”, Royal Society Proceedings, 1998.

Iterative Algebraic Estimation

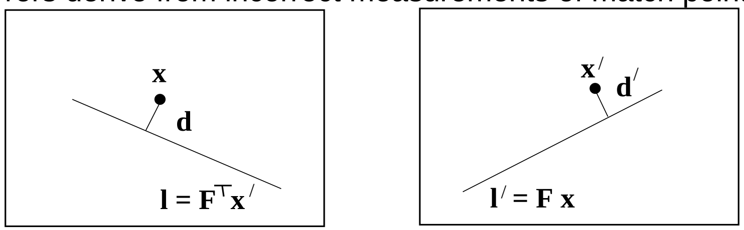
Find the fundamental matrix $\mathtt{F}$ that minimizes the algebraic error $||\mathtt{A}\mathbf{f}||$ subject to $||\mathbf{f}|| = 1$ and $\det \mathtt{F} = 0$.

- **Concept :** Vary epipole $\mathbf{e}$ to minimize the algebraic error $||\mathtt{A}\mathbf{f}'|| = ||\mathtt{A}\mathtt{E}\mathbf{m}||$.
- **Remark :** Each choice of epipole $\mathbf{e}$ defines a minimum error vector $\mathtt{A}\mathtt{E}\mathbf{m}$ as above.
- Use Levenberg-Marquardt method to minimize this error.
- Simple 3×9 minimization problem.
  - 3 inputs – the coordinates of the epipole
  - 9 outputs – the algebraic error vector $\mathtt{A}\mathbf{f}' = \mathtt{A}\mathtt{E}\mathbf{m}$.
- Each step requires estimation of $\mathbf{m}$ using SVD method.
- Tricks can be used to avoid SVD (see Hartley-Royal-Society).

Minimization of Geometric Error

Minimization of Geometric Error

- Algebraic error vector $\mathtt{A}\mathbf{f}$ has no clear geometric meaning.
- Should be minimizing geometric quantities.
- Errors derive from incorrect measurements of match points.



- We should be measuring distances from epipolar lines.

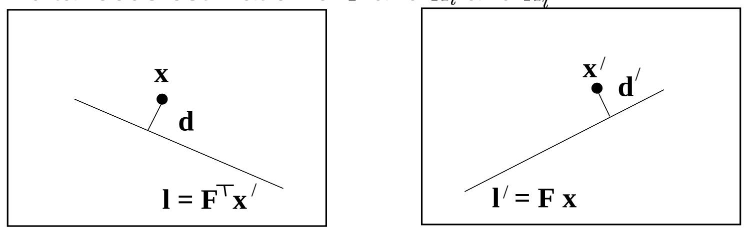
The Gold Standard (ML) Method

- Assumes a Gaussian distributed noise.
- Measured correspondences $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$.
- Estimated correspondences $\hat{\mathbf{x}}_i$ and $\hat{\mathbf{x}}'_i$

$$\sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2$$

subject to $\hat{\mathbf{x}}_i'^{\top} \mathtt{F} \hat{\mathbf{x}}_i = 0$ exactly for some $\mathtt{F}$.

- Simultaneous estimation of $\mathtt{F}$ and $\hat{\mathbf{x}}_i$ and $\hat{\mathbf{x}}'_i$.



The Gold Standard (ML) Method

Minimizing the Gold-Standard error function.

- Initial 3D reconstruction :

$$\mathtt{P} = [\mathtt{I} \mid \mathbf{0}]$$
$$\mathtt{P}' = [\mathtt{M} \mid \mathbf{t}]$$
$$\mathbf{X}_i = (\mathrm{X}_i, \mathrm{Y}_i, 1, \mathrm{T}_i)^{\top}$$

- Compute $\hat{\mathbf{x}}_i = \mathtt{P}\mathbf{X}_i = (\mathrm{X}_i, \mathrm{Y}_i, 1)^{\top}$ and $\hat{\mathbf{x}}'_i = \mathtt{P}'\mathbf{X}_i$.
- Iterate over $\mathtt{P}'$ and $\mathbf{X}_i = (\mathrm{X}_i, \mathrm{Y}_i, 1, \mathrm{T}_i)^{\top}$ to minimize cost function :

$$\sum_i d(\mathbf{x}_i, \hat{\mathbf{x}}_i)^2 + d(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)^2$$

- Total of $3n + 12$ parameters.
  - 12 parameters for the camera matrix $\mathtt{P}'$
  - 3 parameters for each point $\mathbf{X}_i$.
  - Once $\mathtt{P}' = [\mathtt{M} \mid \mathbf{t}]$ is found, compute $\mathtt{F} = [\mathbf{t}]_{\times}\mathtt{M}$.

Sparse Levenberg-Marquardt

Reference : Hartley- Azores

(i) Coordinates of $\mathbf{X}_i$ do not affect $\hat{\mathbf{x}}_j$ or $\hat{\mathbf{x}}'_j$ (for $i \neq j$).
(ii) Sparse LM takes advantage of sparseness
(iii) Linear time in n (number of points).
(iv) Reference : Hartley- Azores

Parametrization of rank-2 matrices

Both epipoles as parameters. The resulting form of $\mathtt{F}$ is

$$\mathtt{F} = \begin{bmatrix} a & b & \alpha a + \beta b \\ c & d & \alpha c + \beta d \\ \alpha' a + \beta' c & \alpha' b + \beta' d & \alpha' \alpha a + \alpha' \beta b + \beta' \alpha c + \beta' \beta d \end{bmatrix}$$

Parametrization of rank-2 matrices

- Estimation of $\mathtt{F}$ may be done by parameter minimization.
- Parametrize $\mathtt{F}$ such that $\det \mathtt{F} = 0$.
- Various parametrizations have been used.

Overparametrization.

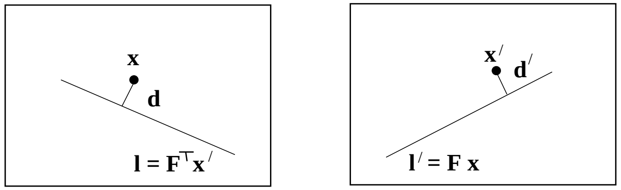
- Write $\mathtt{F} = [\mathbf{t}]_\times \mathtt{M}$.
- 3 parameters for $\mathbf{t}$ and 9 for $\mathtt{M}$.

Epipolar parametrization.

$$\mathtt{F} = \begin{bmatrix} a & b & \alpha a + \beta b \\ c & d & \alpha c + \beta d \\ e & f & \alpha e + \beta f \end{bmatrix}$$

To achieve a minimum set of parameters, set one of the elements, for instance f to 1.

Epipolar distance



- Point correspondence $\mathbf{x}' \leftrightarrow \mathbf{x}$:
- Point $\mathbf{x} \mapsto$ epipolar line $\mathtt{F}\mathbf{x}$.
- Epipolar distance is distance of point $\mathbf{x}'$ to epipolar line $\mathtt{F}\mathbf{x}$.
- Write $\mathtt{F}\mathbf{x} = (\lambda, \mu, \nu)^\top$ and $\mathbf{x}' = (x', y', 1)^\top$.
- Distance is

$$d(\mathbf{x}', \mathtt{F}\mathbf{x}) = \mathbf{x}'^\top \mathtt{F}\mathbf{x}(\lambda^2 + \mu^2)^{-1/2}$$

Epipolar distance - continued

- Epipolar distance may be written as

$$d(\mathbf{x}', \mathrm{F}\mathbf{x}) = \frac{\mathbf{x}'^{\top}\mathrm{F}\mathbf{x}}{((\mathrm{F}\mathbf{x})_1^2 + (\mathrm{F}\mathbf{x})_2^2)^{1/2}}$$

- Total cost function :

$$\sum_i d(\mathbf{x}'_i, \mathrm{F}\mathbf{x}_i)^2 = \sum_i \frac{(\mathbf{x}'^{\top}_i\mathrm{F}\mathbf{x}_i)^2}{(\mathrm{F}\mathbf{x}_i)_1^2 + (\mathrm{F}\mathbf{x}_i)_2^2}$$

- Total cost : sum over all $\mathbf{x}'_i \leftrightarrow \mathbf{x}_i$.
- Minimize this cost function over parametrization of F.

Luong / Zhang's other error function

$$\sum_i \frac{(\mathbf{x}'^{\top}_i\mathrm{F}\mathbf{x}_i)^2}{(\mathrm{F}\mathbf{x}_i)_1^2 + (\mathrm{F}\mathbf{x}_i)_2^2 + (\mathrm{F}^{\top}\mathbf{x}'_i)_1^2 + (\mathrm{F}^{\top}F\mathbf{x}'_i)_2^2}$$

Represents a first-order approximation to geometric error.

Symmetric epipolar distance

- Epipolar distance function is not symmetric.
- Prefer sum of distances in both images.
- Symmetric cost function is

$$d(\mathbf{x}', \mathrm{F}\mathbf{x})^2 + d(\mathbf{x}, \mathrm{F}^{\top}\mathbf{x}')^2$$

- Sum over all points :

$$\text{Cost } = \sum_i (\mathbf{x}'^{\top}_i\mathrm{F}\mathbf{x}_i)^2 \left(\frac{1}{(\mathrm{F}\mathbf{x}_i)_1^2 + (\mathrm{F}\mathbf{x}_i)_2^2} + \frac{1}{(\mathrm{F}^{\top}\mathbf{x}'_i)_1^2 + (\mathrm{F}^{\top}\mathbf{x}'_i)_2^2} \right)$$

Problem

- Points near the epipole have a disproportionate influence.
- Small deviation in point makes big difference to epipolar line.

This page left empty

More Algorithm Comparison

Experimental Evaluation of the Algorithms

Three of the algorithms compared

(i) The normalized 8-point algorithm

(ii) Minimization of algebraic error whilst imposing the singularity constraint

(iii) The Gold Standard geometric algorithm

Experimental procedure

(i) Find matched points in image pair.

(ii) Select n matched points at random

(iii) Compute the fundamental matrix

(iv) Compute epipolar distance for all other points.

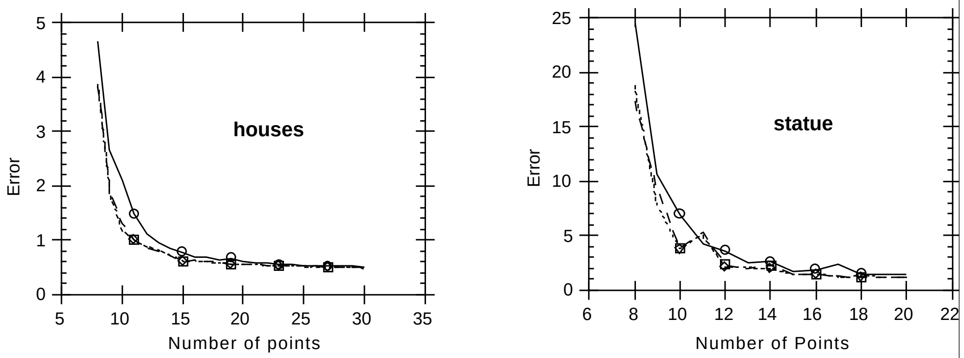
(v) Repeat 100 times for each n and collect statistics.

The error is defined as

$$\frac{1}{N}\sum_{i}^{N}(d(\mathbf{x}'_i, \mathtt{F}\mathbf{x}_i) + d(\mathbf{x}_i, \mathtt{F}^\top\mathbf{x}'_i))$$

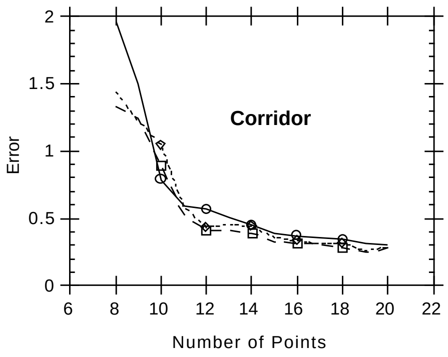
i.e Average symmetric epipolar distance.

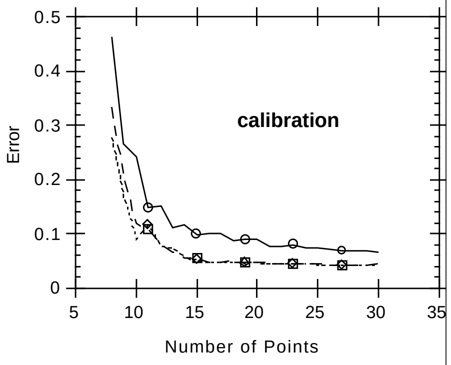
Results



Normalized 8-point, geometric and algebraic error.

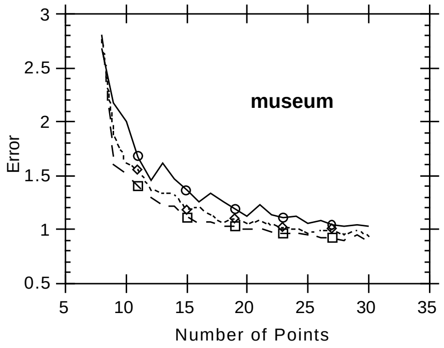
Results - continued





Normalized 8-point, geometric and algebraic error.

Results - continued



Normalized 8-point, geometric and algebraic error.

Covariance computation

Covariance of $\mathtt{P}'$

To compute the covariance matrix of the entries of $\mathbf{P}'$:

(i) Define $\mathbf{Q} = (\mathbf{x}_i, \mathbf{x}'_i)^\top$ – vector of meaurments in both images.

(ii) Compute the derivative matrices

$$\mathtt{A}_i = [\partial \hat{\mathbf{Q}}_i / \partial \mathbf{P}'] \text{ and } \mathtt{B}_i = [\partial \hat{\mathbf{Q}}_i / \partial \mathbf{X}_i]$$

(iii) Compute in steps

$$\mathtt{U} = \sum_i \mathtt{A}_i{}^\top \Sigma_{\mathbf{Q}_i}^{-1} \mathtt{A}_i$$
$$\mathtt{V}_i = \mathtt{B}_i{}^\top \Sigma_{\mathbf{Q}_i}^{-1} \mathtt{B}_i$$
$$\mathtt{W}_i = \mathtt{A}_i{}^\top \Sigma_{\mathbf{Q}_i}^{-1} \mathtt{B}_i$$
$$\Sigma_{\mathbf{P}'} = (\mathtt{U} - \sum_i \mathtt{W}_i \mathtt{V}_i^{-1} \mathtt{W}_i{}^\top)^+ \text{ (pseudo-inverse)}$$

Covariance of F

To compute the covariance of $\mathtt{F}$:

(i) Given $\mathtt{P}' = [\mathtt{M} \mid \mathbf{m}]$, then $\mathtt{F} = [\mathbf{m}]_\times \mathtt{M}$.
(ii) Compute $\mathtt{J} = \partial \mathtt{F} / \partial \mathtt{P}'$.
(iii)

$$\Sigma_{\mathtt{F}} = \mathtt{J} \Sigma_{\mathtt{P}'} \mathtt{J}^\top$$

Covariance of epipolar line corresponding to $\mathbf{x}$

(i) Epipolar line is $\mathbf{l} = \mathtt{F}\mathbf{x}$.
(ii) Given $\mathbf{x}$ and $\Sigma_{\mathtt{F}}$ compute $\mathtt{J} = \partial \mathbf{l} / \partial \mathtt{F}$
(iii)

$$\Sigma_{\mathbf{l}} = \mathtt{J} \Sigma_{\mathtt{F}} \mathtt{J}^\top$$

The envelope of epipolar lines

- May compute envelope of epipolar lines.

$$\mathtt{C} = \bar{\mathbf{l}}\, \bar{\mathbf{l}}^\top - k^2 \Sigma_{\mathbf{l}}$$

- $\mathtt{C}$ is a hyperbola that contains the epipolar line with a given probability α.
- k^2 chosen such that $F_2^{-1}(k^2) = \alpha$,
- $F_2(k^2)$ represents the cumulative χ^2_2 distribution,
- with probability α the lines lie within this region.

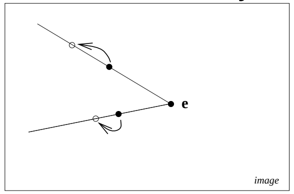
$<$ epipolar line demonstration here $>$

Special cases of F-computation

Using special motions can simplify the computation of the fundamental matrix.

Pure translation

- Can assume $\mathtt{P} = [\mathtt{I} | \mathbf{0}]$ and $\mathtt{P}' = [\mathtt{I} | \mathbf{t}]$.
- $\mathtt{F} = [\mathbf{t}]_\times$.
- $\mathtt{F}$ is skew-symmetric – has 2 dof.
- Being skew-symmetric, automatically has rank 2.



For a pure translation the epipole can be estimated from the image motion of two points.

Cameras with the same principal plane

- Principal plane of the camera is the third row of P.
- Cameras have the same third row.
- Affine cameras - last row is $(0,0,0,1)^\top$.

Simple correspondences exist :

$$(x', y', 0)\mathtt{F}(x, y, 0)^\top = 0$$

for any $(x', y', 0)^\top$ and $(x, y, 0)^\top$.

F has the following form :

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

Points on a ruled quadric

(i) If all the points and the two camera centres lie on a ruled quadric, then there are three possible fundamental matrices.

(ii) points lie in a plane. The correspondences $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ lead to a 3-parameter family of possible fundamental matrices F (note, one of the parameters accounts for scaling the matrix so there is only a *two*-parameter family of homogeneous matrices).

(iii) Two cameras at the same point :

- The fundamental matrix does not exist.
- There is no such thing as an epipolar plane, and epipolar lines are not defined.
- Correspondences $\mathbf{x}'_i \leftrightarrow \mathbf{x}_i$ give at least a 2-parameter family of F.

Degeneracies

Correspondences are degenerate if they satisfy more than one F.

$$\mathbf{x}_i{}^\top \mathtt{F}_1 \mathbf{x}_i{}^\top = 0 \quad \text{and} \quad \mathbf{x}'_i \mathtt{F}_2 \mathbf{x}_i = 0 \qquad (1 \le i \le n) \ .$$

This page left empty