GEOMETRY OF TWO OR MORE VIEWS

Václav Hlaváč

Czech Technical University, Faculty of Electrical Eng.
Center for Machine Perception, Praha

hlavac@fel.cvut.cz

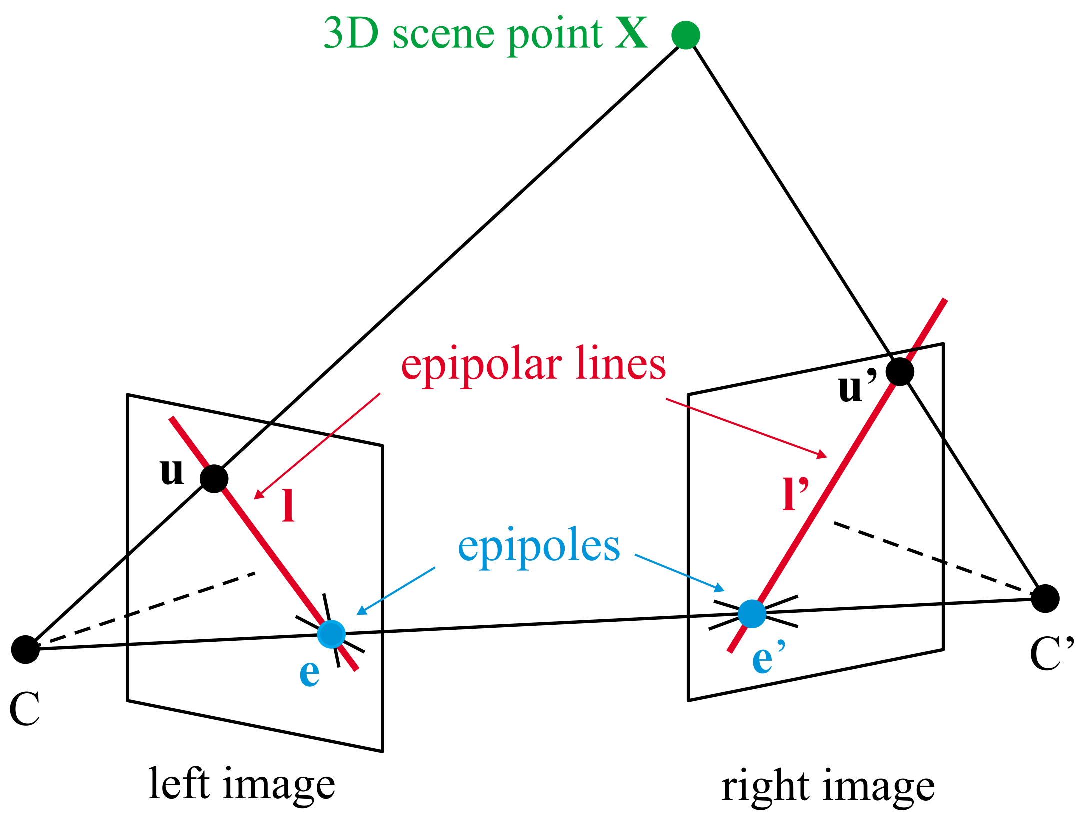
http://cmp.felk.cvut.cz/~hlavac

STEREOPSIS

- Calibration of one camera and knowledge of the co-ordinates of one image point allows us to determine a ray in space uniquely.
- If two calibrated cameras observe the same scene point $\mathbf{X}$ then its 3D co-ordinates can be computed as the intersection of two such rays. This is the basic principle of stereo vision that typically consists of three steps:

  1. Camera calibration;
  2. Establishing point correspondences between pairs of points from the left and the right images;
  3. Reconstruction of 3D co-ordinates of the points in the scene.

GEOMETRY OF TWO CAMERAS

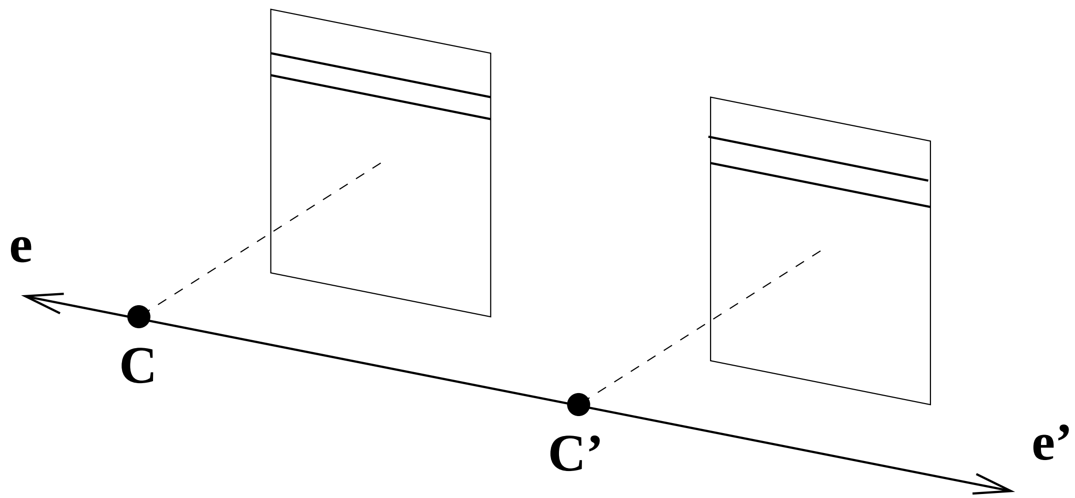


Epipoles $\mathbf{e}$, $\mathbf{e}'$, epipolar lines $\mathbf{l}$, $\mathbf{l}'$.

$\mathbf{e}$, $\mathbf{e}'$, $\mathbf{l}$, $\mathbf{l}'$, C, C', $\mathbf{X}$ lie in a single plane.

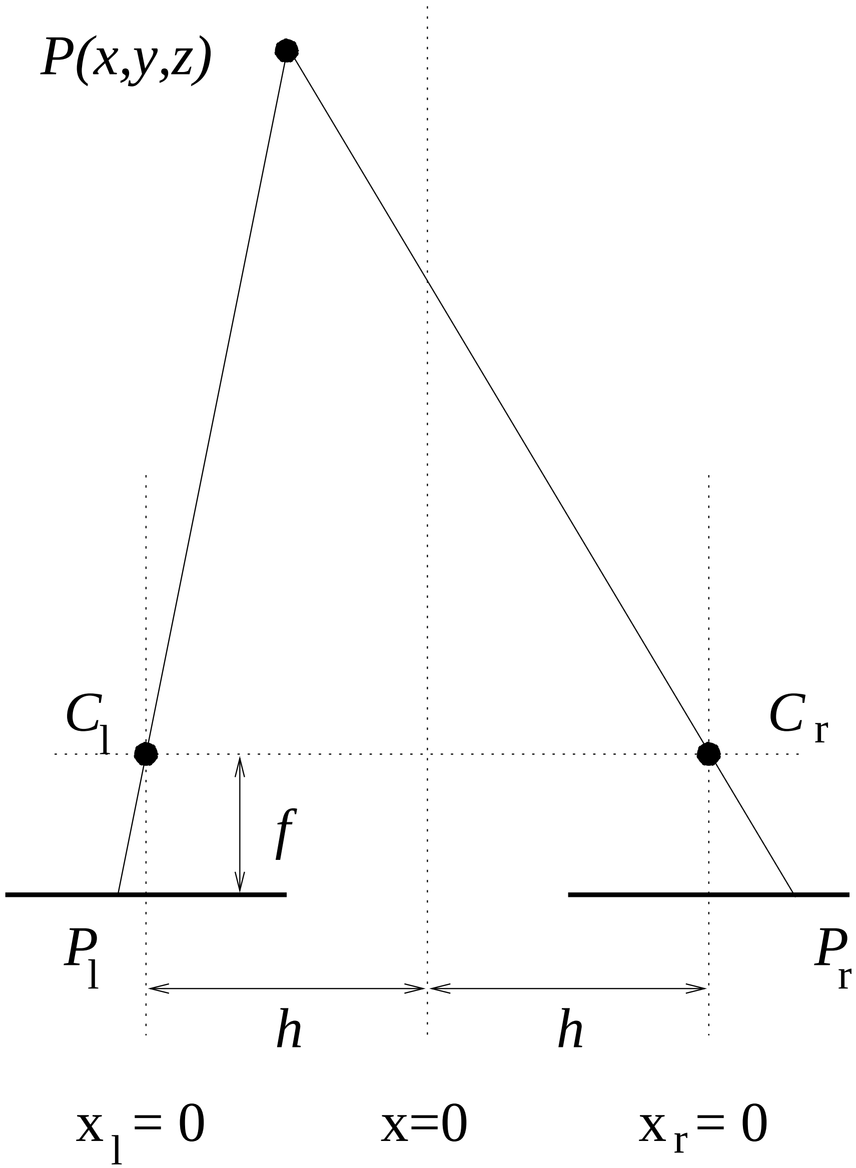
Epipolar geometry. Seeking correspondences between two 1D signals.
Bilinear relation between $\mathbf{u}$, $\mathbf{u}'$.

CANONICAL CONFIGURATION OF CAMERAS



- Epipolar lines correspond to lines in images.
- It is often used when stereo correspondence is to be determined by a human operator who will find matching points linewise to be easier.
- Any pair of images with known epipolar geometry can be converted to canonical configuration by rectification.

DISPARITY AND DEPTH



baseline $2h$

disparity $|P_l - P_r| > 0$

focal length f

Calculation of depth (similar triangles)

$$\frac{P_l}{f} = -\frac{h+x}{z}, \qquad \frac{P_r}{f} = \frac{h-x}{z}$$

$$z\,(P_r - P_l) = 2hf$$

$$z = \frac{2hf}{P_r - P_l}$$

Note: if $P_r - P_l = 0$ then $z = \infty$.

FUNDAMENTAL MATRIX (1)

Left projection $\mathbf{u}$ and right projection $\mathbf{u}'$ of the scene point $\mathbf{X}$.

$$\begin{aligned}
\mathbf{u} &\simeq [K|\mathbf{0}]\begin{bmatrix}\mathbf{X}\\1\end{bmatrix} = K\,\mathbf{X},\\
\mathbf{u}' &\simeq [K'R\,|-K'R\,\mathbf{t}]\begin{bmatrix}\mathbf{X}\\1\end{bmatrix}\\
&= K'(R\,\mathbf{X}-R\,\mathbf{t}) = K'\mathbf{X}'
\end{aligned}$$

- Coplanarity of $\mathbf{X}$, $\mathbf{X}'$ and $\mathbf{t}$.
- Distinguish co-ordinates of the left and right cameras by the subscript $_L$, $_R$.
- Vector product $\times$.

FUNDAMENTAL MATRIX (2)

Coordinates rotation
$\mathbf{X'}_R = R\,\mathbf{X'}_L$, and hence $\mathbf{X'}_L = R^{-1}\mathbf{X'}_R$.

Coplanarity constraint $\mathbf{X}_L^T(\mathbf{t} \times \mathbf{X'}_L) = 0$.

Preparing for substitution
$\mathbf{X}_L = K^{-1}\mathbf{u}$, $\mathbf{X'}_R = (K')^{-1}\mathbf{u'}$, and $\mathbf{X'}_L = R^{-1}(K')^{-1}\mathbf{u'}$.

Epipolar constraint in vector form

$$(K^{-1}\mathbf{u})^T(\mathbf{t} \times R^{-1}\,(K')^{-1}\mathbf{u'}) = 0\,.$$

Equation is homogeneous with respect to $\mathbf{t}$, so the scale is not determined.

Absolute scale cannot be recovered without ‘yardstick’.

FUNDAMENTAL MATRIX (3)

Replacement of a vector product by a matrix multiplication.

The translation vector is $\mathbf{t} = [t_x, t_y, t_z]^T$, and a skew symmetric matrix $S(\mathbf{t})$ (i.e., $S^T = -S$) can be created from it if $\mathbf{t} \neq \mathbf{0}$.

$$S(\mathbf{t}) = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}$$

Note that $\mathrm{rank}(S) = 2$ if and only if $\mathbf{t} \neq \mathbf{0}$.

FUNDAMENTAL MATRIX (4)

The vector product can be replaced by the multiplication of two matrices.

For any regular matrix A, we have

$$\mathbf{t} \times A \; = S(\mathbf{t})\, A\,.$$

Thus we can rewrite the epipolar constraint in a vector form

$$(K^{-1}\mathbf{u})^T \, (S(\mathbf{t})\, R^{-1}\, (K')^{-1}\mathbf{u}') = 0\,,$$

$$\mathbf{u}^T (K^{-1})^T S(\mathbf{t}) R^{-1} (K')^{-1}\mathbf{u}' = 0\,.$$

FUNDAMENTAL MATRIX (5)

The middle part can be concentrated into a single matrix F called the fundamental matrix of two views,

$$F = (K^{-1})^T S(\mathbf{t}) R^{-1} (K')^{-1} \,.$$

With the substitution for F we finally get the bilinear relation (sometimes named after Longuet-Higgins) between any two views

$$\mathbf{u}^T \, F \, \mathbf{u}' = 0 \,.$$

It can be seen that the fundamental matrix F captures all information that can be recovered from a pair of images if the correspondence problem is solved.

RELATIVE MOTION OF THE CAMERA ESSENTIAL MATRIX E

A single camera moving in space, or two cameras with known calibration.

Known calibration matrices K, K' allows us to normalize measurement in left and right images $\breve{\mathbf{u}}$, $\breve{\mathbf{u}}'$.

$$\breve{\mathbf{u}} = K^{-1}\mathbf{u}, \quad \breve{\mathbf{u}}' = (K')^{-1}\mathbf{u}'$$

Substitute into

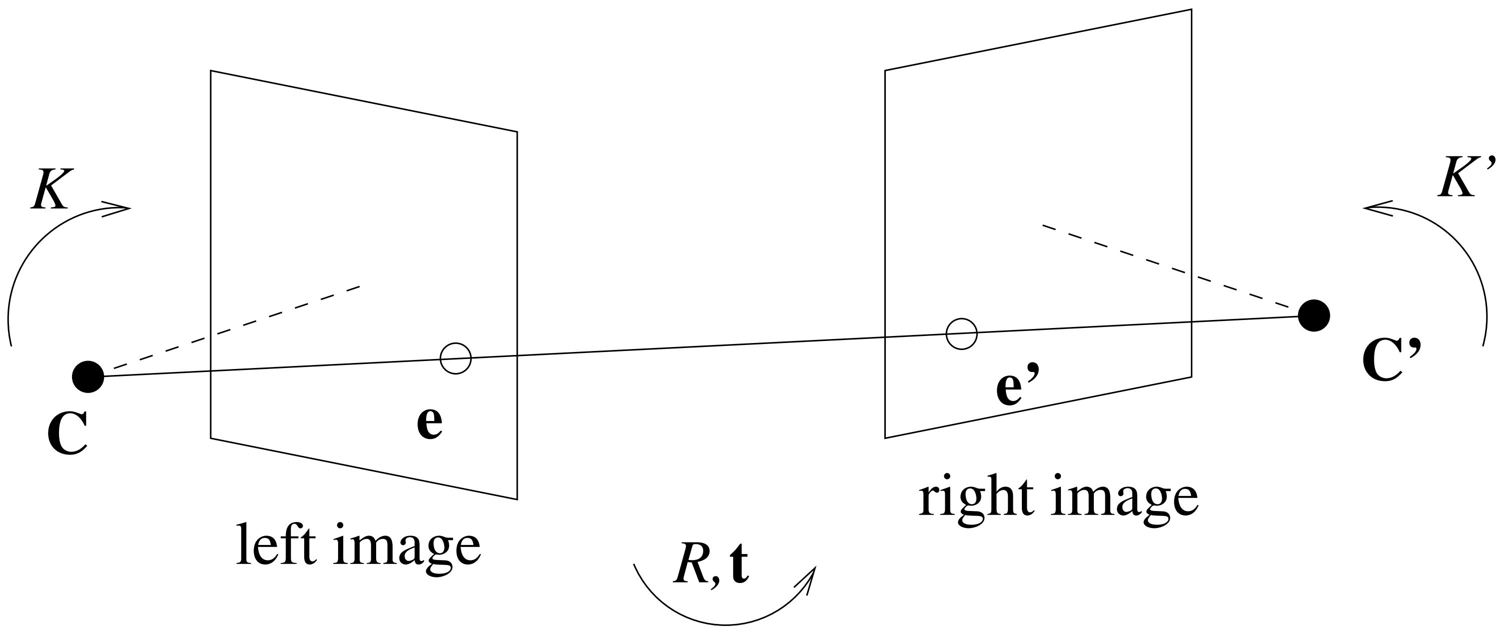
$$\mathbf{u}^T (K^{-1})^T S(\mathbf{t}) R^{-1} (K')^{-1} \mathbf{u}' = 0$$

$$\begin{aligned} \breve{\mathbf{u}}^T \, S(\mathbf{t}) R^{-1} \, \breve{\mathbf{u}}' &= 0 \\ \breve{\mathbf{u}}^T \, E \, \breve{\mathbf{u}}' &= 0 \end{aligned}$$

E captures all the information about the relative motion from the first to the second position of the calibrated camera.

PROPERTIES OF ESSENTIAL MATRIX E

- The essential matrix E has rank 2.
- Let $\mathbf{t}$ be the translational vector, and $\mathbf{t}' = R\,\mathbf{t}$. There holds $E\,\mathbf{t}' = 0$ and $\mathbf{t}^T E = 0$.



PROPERTIES OF ESSENTIAL MATRIX E

SVD decomposes E as $E = UDV^T$ for a diagonal D;

$$D = \begin{bmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

ROTATION R AND TRANSLATION $\mathbf{t}$ from E

[Hartley 1992] We have seen $E = S(\mathbf{t})R^{-1}$.

$$\breve{\mathbf{u}}^T\, S(\mathbf{t})R^{-1}\, \breve{\mathbf{u}}' = 0\,, \quad \breve{\mathbf{u}}'^T\, RS(\mathbf{t})\, \breve{\mathbf{u}} = 0\,, \quad E = RS(\mathbf{t})$$

$$G = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad Z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The rotation matrix R can be calculated using SVD: $E = UDV^T$.

$$R = UG\,V^T \text{ or } R = UG^T\,V^T$$

Components of the translation vector can be derived from the matrix $S(\mathbf{t})$ expressed as 3×3 matrix.

$$S(\mathbf{t}) = VZ\,V^T$$

PROPERTIES OF THE FUNDAMENTAL MATRIX F

- rank$(E) = 2$. As $F = (K^{-1})^T E K'^{-1}$ and the calibration matrices are regular $\Rightarrow$ F rank$(F) = 2$.
- Consider two epipoles $\mathbf{e}, \mathbf{e}'$.

$$\mathbf{e}^T F = 0 \text{ and } F\,\mathbf{e}' = 0$$

- SVD of the fundamental matrix gives $F = UDV^T$, where

$$D = \begin{bmatrix} k_1 & 0 & 0 \\ 0 & k_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; k_1 \neq k_2 \neq 0$$

ESTIMATING F, 8 POINT ALGORITHM

- Epipolar geometry has 7 degrees of freedom. Epipoles $\mathbf{e}$, $\mathbf{e}'$ have 2 co-ordinates each (giving 4 dof), while another three come from the mapping of any three epipolar lines in the first image to the second image.
- Thus the correspondence of 7 points in left and right images enables the establishment of the fundamental matrix F using a nonlinear algorithm, numerically unstable.
- If there are eight non-coplanar corresponding points available then the linear eight point algorithm. Usually its overconstrained and robust version is used.

8-POINT ALGORITHM (2)

$$\mathbf{u}_i{}^T \, F \, \mathbf{u}'_i = 0 \, , \quad \mathbf{u}^T = [u_i, v_i, 1]$$

The 3×3 fundamental matrix F has only eight unknowns as it is only known up to scale $\Rightarrow$ 8 correspondences.

$$[u_i, \; v_i, \; 1] \; F \begin{bmatrix} u'_i \\ v'_i \\ 1 \end{bmatrix} = 0$$

8-POINT ALGORITHM (3)

Rewriting the elements of the fundamental matrix as a column vector with nine elements $\mathbf{f}^T = [f_{11}, f_{12}, \ldots, f_{33}]$, can be rewritten as a system of linear equations

$$\begin{bmatrix} u_i u'_i & u_i v'_i & u_i & v_i u'_i & v_i v'_i & v_i & u'_i & v'_i & 1 \\ & & & & \vdots & & & & \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{12} \\ \vdots \\ f_{33} \end{bmatrix} = A\ \mathbf{f} = 0$$