Epipolar Geometry and the Fundamental Matrix

Review about Camera Matrix **P** (from Lecture on Calibration)

- Between the world coordinates **X**=(X_s, X_s, X_s, 1) of a scene point and the coordinates **x**=(u',v',w') of its projection, we have the following linear transformation:

$$x_{pix} = u'/w' \quad y_{pix} = v'/w' \quad \text{with}$$

$$\mathbf{x} = \mathbf{P}\,\mathbf{X}$$

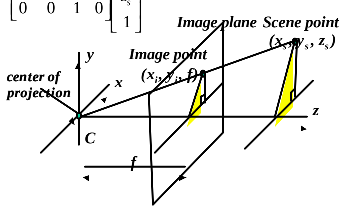
- P is a 3x4 matrix that completely represents the mapping from the scene to the image and is therefore called a "*camera*".

Central Projection

If world and image points are represented by homogeneous vectors, central projection is a linear transformation:

$$x_i = f\,\frac{x_s}{z_s} \qquad y_i = f\,\frac{y_s}{z_s}$$

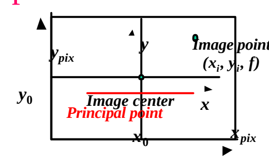
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix} \qquad x_i = u/w, \; y_i = v/w$$



Pixel Components

Transformation uses:

- principal point (x_0, y_0)
- scaling factors k_x and k_y



$$x_i = f\,\frac{x_s}{z_s} \qquad x_{pix} = k_x x_i + x_0 = f\,k_x \frac{x_s + z_s x_0}{z_s}$$

$$y_i = f\,\frac{y_s}{z_s} \qquad y_{pix} = k_y y_i + y_0 = f\,k_y \frac{y_s + z_s y_0}{z_s}$$

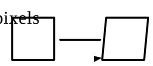
$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \begin{bmatrix} a_x & 0 & x_0 & 0 \\ 0 & a_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix} \quad \text{with} \quad \begin{matrix} a_x = f\,k_x \\ a_y = f\,k_y \end{matrix} \quad \text{then} \quad \begin{matrix} x_{pix} = u'/w' \\ y_{pix} = v'/w' \end{matrix}$$

Internal Camera Parameters

$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \begin{bmatrix} a_x & s & x_0 & 0 \\ 0 & a_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix} \quad \text{with} \quad \begin{vmatrix} a_x = f\,k_x \\ a_y = -f\,k_y \end{vmatrix} \qquad \begin{matrix} x_{pix} = u'/w' \\ y_{pix} = v'/w' \end{matrix}$$

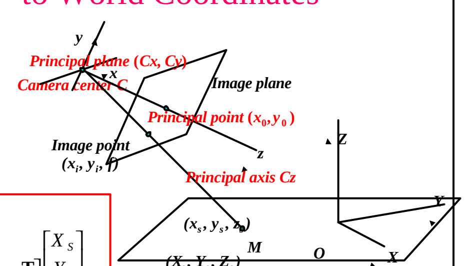
$$\begin{bmatrix} a_x & s & x_0 & 0 \\ 0 & a_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} a_x & s & x_0 \\ 0 & a_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \mathbf{K}\,[\mathbf{I}_3 \;\; | \;\; \mathbf{0}_3]$$

- a_x and a_y "focal lengths" in pixels
- x_0 and y_0 coordinates of image center in pixels



- Added parameter *S* is skew parameter
- **K** is called *calibration matrix*. **Five degrees of freedom.**
  - **K** is a 3x3 upper triangular matrix

From Camera Coordinates to World Coordinates



$$\begin{bmatrix} x_S \\ y_S \\ z_S \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ \mathbf{0}_3^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} X_S \\ Y_S \\ Z_S \\ 1 \end{bmatrix}$$

Using Camera Center Position in World Coordinates

- We can use $-\mathbf{R}\,\tilde{\mathbf{C}}$ instead of $\mathbf{T}$

$$\begin{bmatrix} x_S \\ y_S \\ z_S \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\ \mathbf{0}_3^T & 1 \end{bmatrix} \begin{bmatrix} X_S \\ Y_S \\ Z_S \\ 1 \end{bmatrix}$$

Linear Transformation from World Coordinates to Pixels

- Combine camera projection and coordinate transformation matrices into a single matrix $\mathbf{P}$

$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \mathbf{K}[\mathbf{I}_3 \mid \mathbf{0}_3] \begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} x_S \\ y_S \\ z_S \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\ \mathbf{0}_3^T & 1 \end{bmatrix} \begin{bmatrix} X_S \\ Y_S \\ Z_S \\ 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \mathbf{K}[\mathbf{I}_3 \mid \mathbf{0}_3] \begin{bmatrix} \mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\ \mathbf{0}_3^T & 1 \end{bmatrix} \begin{bmatrix} X_S \\ Y_S \\ Z_S \\ 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \mathbf{P} \begin{bmatrix} X_S \\ Y_S \\ Z_S \\ 1 \end{bmatrix}$$

$$\boxed{\mathbf{x} = \mathbf{P}\mathbf{X}}$$

Properties of Matrix P

- Further simplification of $\mathbf{P}$:

$$\mathbf{x} = \mathbf{K}[\mathbf{I}_3 \mid \mathbf{0}_3] \begin{bmatrix} \mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\ \mathbf{0}_3^T & 1 \end{bmatrix} \mathbf{X}$$

$$[\mathbf{I}_3 \mid \mathbf{0}_3] \begin{bmatrix} \mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\ \mathbf{0}_3^T & 1 \end{bmatrix} = [\mathbf{R} \;\; -\mathbf{R}\tilde{\mathbf{C}}] = \mathbf{R}[\mathbf{I}_3 \mid -\tilde{\mathbf{C}}]$$

$$\mathbf{x} = \mathbf{K}\mathbf{R}[\mathbf{I}_3 \mid -\tilde{\mathbf{C}}]\mathbf{X}$$

$$\boxed{\mathbf{P} = \mathbf{K}\mathbf{R}[\mathbf{I}_3 \mid -\tilde{\mathbf{C}}]}$$

- $\mathbf{P}$ has 11 degrees of freedom:
  - 5 from triangular calibration matrix $\mathbf{K}$, 3 from $\mathbf{R}$ and 3 from $\tilde{\mathbf{C}}$
- $\mathbf{P}$ is a fairly general 3 x 4 matrix
  - left 3x3 submatrix $\mathbf{KR}$ is non-singular

Cross-Product in Matrix Form

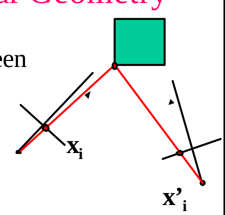
- If $\mathbf{a} = (a_1, a_2, a_3)^T$ is a 3-vector, then one can define a corresponding skew-symmetric matrix

$$[\mathbf{a}]_\times = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}$$

- The cross-product of 2 vectors $\mathbf{a}$ and $\mathbf{b}$ can be written $\mathbf{a} \times \mathbf{b} = [\mathbf{a}]_\times \mathbf{b}$
- Matrix $[\mathbf{a}]_\mathbf{x}$ is singular. Its null vector (right or left) is $\mathbf{a}$
- $\mathbf{b}\,.\,(\mathbf{a} \times \mathbf{b}) = 0 \Leftrightarrow \mathbf{b}^T [\mathbf{a}]_\times \mathbf{b} = \mathbf{0} \quad \forall \mathbf{a}, \mathbf{b}$
- $\mathbf{c}^T [\mathbf{a}]_\times \mathbf{b} = -\,\mathbf{b}^T [\mathbf{a}]_\times \mathbf{c} \Rightarrow \mathbf{F}^T [\mathbf{a}]_\times \mathbf{F} = [\mathbf{s}]_\times \forall \mathbf{F}, \mathbf{a}$

Definition of Epipolar Geometry

- Projective geometry between two views
- Independent of scene structure
- Depends only on the cameras' internal parameters and relative pose of cameras
- Fundamental matrix $\mathbf{F}$ encapsulates this geometry

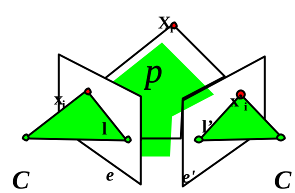


$$\boxed{\mathbf{x'}_i^{T}\, \mathbf{F}\, \mathbf{x}_i = 0}$$

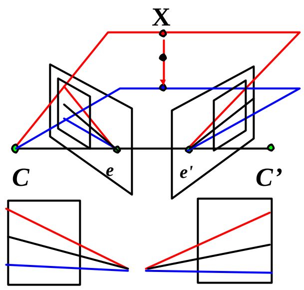
for any pair of corresponding points $\mathbf{x}_i$ and $\mathbf{x'}_i$ in the 2 images

Relation between Image Points x_i and x_i'

- Camera centers C and C', scene point $\mathbf{X}_i$, image points $\mathbf{x}_i$ and $\mathbf{x'}_i$ belong to a common *epipolar plane* p
- *Epipoles e and e'*
  - *On baseline CC'*
- *Epipolar lines* $\mathbf{l}$ *and* $\mathbf{l'}$

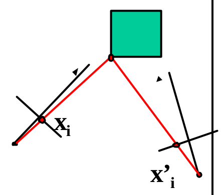


Pencils of Epipolar Lines



Computation of **F**

- **F** can be computed from correspondences between image points alone
- No knowledge of camera internal parameters required
- No knowledge of relative pose required

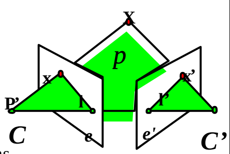


$$\mathbf{x'}_i^T \mathbf{F} \mathbf{x}_i = 0$$

for any pair of corresponding points $\mathbf{x}_i$ and $\mathbf{x'}_i$ in the 2 images

Finding the Fundamental Matrix from Known Cameras **P** and **P'** (Outline)

- Pick up an image point **x** in camera **P**
- Find one scene point **X** on ray of **x** in camera **P**
- Find the image **x'** of **X** in camera **P'**
- Find epipole **e'** as image of **C** in camera **P'**, is epipole = **P'C**
- Find epipolar line **l'** from **e' to x'** in **P'** as function of **x**
- The fundamental matrix **F** is defined by **l'=F x**
- **x'** belongs to **l'**, so $\mathbf{x'}^T \mathbf{l'} = 0$, so $\mathbf{x'}^T \mathbf{F} \mathbf{x} = 0$
- The fundamental matrix **F** is alternately defined by $\mathbf{x'}^T \mathbf{F} \mathbf{x} = 0$

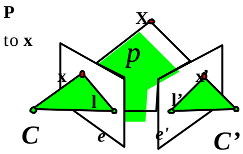


Finding the Fundamental Matrix from Known Cameras **P** and **P'** (Details)

- Pick up an image point **x** in camera **P**
- Find one scene point on ray from **C** to **x**
  - Point $\mathbf{X} = \mathbf{P}^+\mathbf{x}$ satisfies $\mathbf{x} = \mathbf{PX}$
    - $\mathbf{P}^+ = \mathbf{P}^T (\mathbf{P}\mathbf{P}^T)^{-1}$, so $\mathbf{PX} = \mathbf{P}\mathbf{P}^T (\mathbf{P}\mathbf{P}^T)^{-1} \mathbf{x} = \mathbf{x}$
- Image of this point in camera **P'** is $\mathbf{x'} = \mathbf{P'X} = \mathbf{P'}\mathbf{P}^+\mathbf{x}$
- Image of **C** in camera **P'** is epipole **e'** = **P'C**
- Epipolar line of **x** in **P'** is

$$\mathbf{l'} = (\mathbf{e'}) \times (\mathbf{P'}\mathbf{P}^+\mathbf{x}) = [\mathbf{e'}]_\times \mathbf{P'}\mathbf{P}^+\mathbf{x}$$

- **l'=F x** defines **F** fundamental matrix $\Rightarrow \mathbf{F} = [\mathbf{P'C}]_\times \mathbf{P'}\mathbf{P}^+$
- **x'** belongs to **l'**, so $\mathbf{x'}^T \mathbf{l'} = 0$, so $\mathbf{x'}^T \mathbf{F} \mathbf{x} = 0$



Properties of Fundamental Matrix **F**

- Matrix 3X3 (since $\mathbf{x'}^T \mathbf{F} \mathbf{x} = \mathbf{0}$)
- If **F** is fundamental matrix of camera pair (**P, P'**) then the fundamental matrix **F'** of camera pair (**P', P**) is equal to $\mathbf{F}^T$
  - $\mathbf{x}^T \mathbf{F'} \mathbf{x'} = 0$ implies $\mathbf{x'}^T \mathbf{F'}^T \mathbf{x} = 0$, so $\mathbf{F'} = \mathbf{F}^T$
- Epipolar line of **x** is $\mathbf{l'} = \mathbf{F}\mathbf{x}$
- Epipolar line of **x'** is $\mathbf{l} = \mathbf{F}^T \mathbf{x'}$

More Properties of **F**

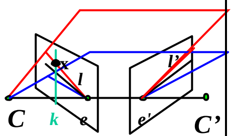
- Epipole **e'** is left null space of **F**, and **e** is right null space.
  - All epipolar lines **l'** contains epipole **e'**, so $\mathbf{e'}^T \mathbf{l'} = 0$, i.e. $\mathbf{e'}^T \mathbf{F} \mathbf{x} = 0$ for all **x**. Therefore $\mathbf{e'}^T \mathbf{F} = \mathbf{0}$. Similarly $\mathbf{e}^T \mathbf{F}^T \mathbf{x'} = 0$ implies $\mathbf{e}^T \mathbf{F}^T = \mathbf{0}$, therefore $\mathbf{F}\mathbf{e} = \mathbf{0}$
- **F** is of rank 2 because $\mathbf{F} = [\mathbf{e'}]_x \mathbf{P'}\mathbf{P}^+$ and $[\mathbf{e'}]_x$ is of rank 2
- **F** has 7 degrees of freedom
  - There are 9 elements, but scaling is not significant
  - Det **F** = 0 removes one degree of freedom

Mapping between Epipolar Lines (a Homography)

- Define $\mathbf{x}$ as intersection between line $\mathbf{l}$ and a line $\mathbf{k}$ ($\mathbf{k}$ does not pass through epipole $\mathbf{e}$):

$$\mathbf{x} = \mathbf{k} \times \mathbf{l} = [\mathbf{k}]_{\times}\mathbf{l}$$

$$\mathbf{l}' = \mathbf{F}\,\mathbf{x} = \mathbf{F}[\mathbf{k}]_{\times}\mathbf{l}$$



- Line $\mathbf{e}$ does not pass through point $\mathbf{e}$

$$\mathbf{l}' = \mathbf{F}\,\mathbf{x} = \mathbf{F}[\mathbf{e}]_{\times}\mathbf{l}$$

- Similarly

$$\mathbf{l} = \mathbf{F}^{T}\,\mathbf{x}' = \mathbf{F}^{T}[\mathbf{e}']_{\times}\mathbf{l}'$$

Retrieving Camera Matrices **P** and **P'** from Fundamental Matrix **F**

- General form of **P** is $\mathbf{P} = \mathbf{K}\,\mathbf{R}\left[\mathbf{I}_3 \;\; | \;\; -\tilde{\mathbf{C}}\right]$
- Select world coordinates as camera coordinates of first camera, select focal length = 1, and count pixels from the principal point. Then $\mathbf{P} = [\,\mathbf{I}_3 \,|\, \mathbf{0}]$
- Then $\mathbf{P}' = [\mathbf{S}\,\mathbf{F} \,|\, \mathbf{e}']$ with **S** any skew-symmetric matrix is a solution. Proof:
  - $\mathbf{x}'^{T}\,\mathbf{F}\,\mathbf{x} = \mathbf{X}^{T}\,\mathbf{P}'^{T}\,\mathbf{F}\,\mathbf{P}\,\mathbf{X}$
  - $\mathbf{P}'^{T}\,\mathbf{F}\,\mathbf{P} = [\mathbf{S}\,\mathbf{F} \,|\, \mathbf{e}']^{T}\,\mathbf{F}\,[\,\mathbf{I}_3 \,|\, \mathbf{0}\,]$ is skew-symmetric
  - For any skew-symmetric matrix **S'** and any **X**, $\mathbf{X}^{T}\,\mathbf{S}'\,\mathbf{X} = 0$
- $\mathbf{S} = [\mathbf{e}']_{\times}$ is a good choice. Therefore $\mathbf{P}' = [[\mathbf{e}']_{\times}\,\mathbf{F} \,|\, \mathbf{e}']$

$[\mathbf{S}\,\mathbf{F} \,|\, \mathbf{e}']^{T}\,\mathbf{F}\,[\,\mathbf{I}_3 \,|\, \mathbf{0}\,]$ is skew-symmetric

$$[\mathbf{S}\mathbf{F}|\mathbf{e}']^{T}\,\mathbf{F}\;[\mathbf{I}_3\,|\mathbf{0}_3] = \begin{bmatrix} \mathbf{F}^{T}\mathbf{S}^{T}\mathbf{F} & \mathbf{0}_3 \\ \mathbf{e}'^{T}\,\mathbf{F} & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{F}^{T}\mathbf{S}^{T}\mathbf{F} & \mathbf{0}_3 \\ \mathbf{0}_3^{\;T} & 0 \end{bmatrix}$$

- $\mathbf{e}'^{T}\,\mathbf{F} = \mathbf{0}$ because **e'** is left null space of **F**
- $\mathbf{F}^{T}\,\mathbf{S}^{T}\,\mathbf{F}$ is skew-symmetric for any **F** and any skew-symmetric **S**

Essential Matrix **E**

- Specialization of fundamental matrix for calibrated cameras and normalized coordinates
  - $\mathbf{x} = \mathbf{P}\,\mathbf{X} = \mathbf{K}\,[\mathbf{R} \,|\, \mathbf{T}]\,\mathbf{X}$
  - Normalize coordinates: $\mathbf{x}_0 = \mathbf{K}^{-1}\,\mathbf{x} = [\mathbf{R} \,|\, \mathbf{T}]\,\mathbf{X}$
- Consider pair of normalized cameras
  - $\mathbf{P} = [\mathbf{I} \,|\, \mathbf{0}]$, $\mathbf{P}' = [\mathbf{R} \,|\, \mathbf{T}]$
- Then we compute $\mathbf{F} = [\mathbf{P}'\mathbf{C}]_{\times}\mathbf{P}'\,\mathbf{P}^{+}$

$$[\mathbf{P}'\mathbf{C}]_{\times} = [[\mathbf{R}\,|\,\mathbf{T}][0\,0\,0\,1]^{T}]_{\times} = [\mathbf{T}]_{\times}$$

$$\mathbf{P}^{+} = \begin{bmatrix} \mathbf{I}_3 \\ \mathbf{0}_3^{T} \end{bmatrix},\; \mathbf{P}'\mathbf{P}^{+} = \mathbf{R} \Rightarrow \boxed{\mathbf{F} = [\mathbf{T}]_{\times}\mathbf{R} \equiv \mathbf{E}}$$

Essential Matrix and Fundamental Matrix

- The defining equation for essential matrix is $\mathbf{x}_0'^{T}\,\mathbf{E}\,\mathbf{x}_0 = \mathbf{0}$, with
  - $\mathbf{x}_0 = \mathbf{K}^{-1}\,\mathbf{x}$
  - $\mathbf{x}_0' = \mathbf{K}'^{-1}\,\mathbf{x}'$
- Therefore $\mathbf{x}'^{T}\,\mathbf{K}'^{-T}\,\mathbf{E}\,\mathbf{K}^{-1}\,\mathbf{x} = \mathbf{0}$
- Comparing with $\mathbf{x}'^{T}\,\mathbf{F}\,\mathbf{x} = \mathbf{0}$, we get

$$\mathbf{E} = \mathbf{K}'^{T}\,\mathbf{F}\mathbf{K}$$

Computing Fundamental Matrix from Point Correspondences

- The fundamental matrix is defined by the equation $\mathbf{x}'^{T}_{\mathbf{i}}\,\mathbf{F}\,\mathbf{x}_{\mathbf{i}} = 0$ for any pair of corresponding points $\mathbf{x}_i$ and $\mathbf{x}'_i$ in the 2 images
- The equation for a pair of points $(x, y, 1)$ and $(x', y', 1)$ is: $x'x\,f_{11} + x'y\,f_{12} + x'f_{13} + y'x\,f_{21} + y'y\,f_{22} + y'f_{23} + x\,f_{31} + y\,f_{32} + f_{33} = 0$
- For n point matches:

$$\mathbf{A}\,\mathbf{f} = \begin{bmatrix} x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_nx_n & x'_ny_n & x'_n & y'_nx_n & y'_ny_n & y'_n & x_n & y_n & 1 \end{bmatrix}\mathbf{f} = 0$$

Computing Fundamental Matrix from Point Correspondences

- We have a homogeneous set of equations **A f** = 0
- **f** can be determined only up to a scale, so there are 8 unknowns, and at least 8 point matchings are needed
  - hence the name "8 point algorithm"
- The least square solution is the singular vector corresponding the smallest singular value of **A**, i.e. the last column of **V** in the SVD $\mathbf{A} = \mathbf{U}\,\mathbf{D}\,\mathbf{V}^T$

Next Class

- 3D Reconstruction from Multiple Views

References

- Multiple View Geometry in Computer Vision, R. Hartley and A. Zisserman, Cambridge University Press, 2000.