3D Computer Vision

Epipolar Geometry

Prof. Didier Stricker

Kaiserslautern University

http://ags.cs.uni-kl.de/

DFKI – Deutsches Forschungszentrum für Künstliche Intelligenz

http://av.dfki.de
Outline

• Previous lecture:
  ▪ Parameter estimation

• Today, a short introduction:
  ▪ Points and lines

• Two views geometry – Epipolar geometry
  ▪ Relation point/line in two views
  ▪ The geometry of two cameras
  ▪ Coplanarity constraint and epipolar lines
  ▪ Definition of the Fundamental matrix $F$ and Essential matrix $E$
Points and Lines – Recall: The Projective Plane

- **Why do we need homogeneous coordinates?**
  - Represent points at infinity
  - Homographies
  - Perspective projection
  - Multi-view relationship

- **Why is the geometric intuition?**
  - A *point* in the image is a *ray* in projective space!
Points and Lines – Recall: The Projective Plane

Homogeneous coordinates in $P^2$

$$(sx, sy, s) = (x, y, 1) \leftrightarrow (x, y) \quad s \neq 0$$

i.e. Position

Eudclidean coordinates in $\mathbb{R}^2 (s = 1)$

$$(x, y, 0) \rightarrow (\infty, \infty)$$

i.e. Direction

Point at infinity ($s = 0$)

• Why is the geometric intuition?
  - A point in the image is a ray in projective space!
What does a line in the image correspond to in projective space?

- A line is a **plane** of rays through origin.
  - all rays \((x, y, z)\) satisfy equation: \(ax + by + cz = 0\)

  In vector notation: \(0 = \begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}\)

- A line is also represented as a homogeneous 3-vector \(l\).
Points and Lines – Duality

- A line \( l \) is a homogeneous 3-vector
  - It is \( \perp \) to every point (ray) \( p \) on the line: \( l^T p = 0 \)

- What is the line \( l \) spanned by rays \( p_1 \) and \( p_2 \)?
  - \( l \) is \( \perp \) to \( p_1 \) and \( p_2 \) \( \Rightarrow \) \( l = p_1 \times p_2 \)
  - \( l \) is the plane normal

- What is the intersection of two lines \( l_1 \) and \( l_2 \)?
  - \( p \) is \( \perp \) to \( l_1 \) and \( l_2 \) \( \Rightarrow \) \( p = l_1 \times l_2 \)

- Points and lines are dual in projective space
  - Given any formula, can switch the meanings of points and lines to get another formula.
Points and Lines
– Example: Computing Vanishing Points (from Lines)

• Intersect \( p_1 q_1 \) with \( p_2 q_2 \)

\[
v = (p_1 \times q_1) \times (p_2 \times q_2)
\]

• In practice: least squares version
  - Better to use more than two lines and compute the “closest” point of intersection
  - See notes by Bob Collins for one good way of doing this: http://www-2.cs.cmu.edu/~ph/869/www/notes/vanishing.txt
Points and Lines – Cross-Product as Matrix Operation

- From a 3 element vector a skew-symmetric matrix is defined:

\[
a = \begin{pmatrix}
a_1 \\
a_2 \\
a_3
\end{pmatrix}
\]
\[
[a]_x = \begin{bmatrix}
0 & -a_3 & a_2 \\
a_3 & 0 & -a_1 \\
-a_2 & a_1 & 0
\end{bmatrix}
\]

- This allows to formulate the cross product of two vectors as matrix multiplication:

\[
a \times b = [a]_x \cdot b
\]
\[
= (a^T \cdot [b]_x)^T
\]
Points and Lines – Intersection of Parallel Lines

• Intersection of parallel lines

\[ l = (a, b, c)^T \text{ and } l' = (a, b, c')^T \]
\[ l \times l' = (b, -a, 0)^T \]

\[ [l]_x l' = \begin{bmatrix}
0 & -c & b \\
-c & 0 & -a \\
-b & a & 0
\end{bmatrix}
\begin{bmatrix}
a \\
b \\
c'
\end{bmatrix}
= \begin{bmatrix}
b(c' - c) \\
a(c - c') \\
0
\end{bmatrix}
\approx \begin{bmatrix}
b \\
-a \\
0
\end{bmatrix} \]

Skew matrix of \( l \)

Example

\[ \begin{pmatrix} b, -a \\ a, b \end{pmatrix} \]
tangent vector

\[ \begin{pmatrix} b, -a \end{pmatrix} \]
normal direction

\[ x = 1 \quad x = 2 \]
Epipolar Geometry - Camera on a Mobile Vehicle
Epipolar Geometry – Pentagon Example

left image

range map

digital surface model

right image
Epipolar Geometry - Scenarios

- A **stereo** rig consisting of two cameras
  - Two images are acquired simultaneously

- A **single** moving camera (static scene)
  - Two images are acquired sequentially

- The two scenarios are geometrically equivalent!
Epipolar Geometry - Objective

- Given two images of a scene acquired by known cameras compute the 3D position of the scene (structure recovery).

- Basic principle:
  - 3D Points are triangulated from corresponding image points
  - Determine 3D point at intersection of two back-projected rays
Triangulation - Correspondences

- Corresponding points are images of the same scene point.
Triangulation

• For each point in the first image determine the corresponding point in the second image
  (this is a search problem)

• For each pair of matched points determine the 3D point by triangulation
  (this is an estimation problem)
3D Reconstruction – General Outline

- **Epipolar geometry**
  - The geometry of two cameras
  - Reduce the correspondence problem to a line search

- **Structure-from-Motion (SfM)**
  - Incremental reconstruction of scene geometry (structure) and the camera pose (motion)
  - Triangulation

- **Dense reconstruction**
Epipolar Geometry – Notation

- The two cameras with projection matrices $P$ and $P'$, and a 3D point $X$ with the corresponding 2D points in the images ($x$ and $x'$):

\[
x = PX \quad x' = PX
\]

- $P$: 3x4 camera matrix
- $X$: 4D vector
- $x$: 3D vector

**Warning:**
- For equations involving homogeneous quantities ‘=’ means ‘equal up to scale’!
Epipolar Geometry – Epipolar Line

• Given an image point in one view, where is the corresponding point in the other view?

- A point in one view “generates” an **epipolar line** in the other view.
- The corresponding point lies on this line.
Epipolar Geometry – Epipolar Line

- Epipolar constraint
  - Reduces correspondence problem to 1D search along an epipolar line!
Epipolar geometry is a consequence of the coplanarity of the camera centers and scene point.

- The camera centres, corresponding points and scene point lie in a single plane, known as the **epipolar plane**.
Epipolar Geometry – Nomenclature

- The epipolar line $l'$ is the image of the ray through $x$.
- The epipole $e$ is the point of intersection of the line joining the camera centres with the image plane.
  - This line is the baseline for a stereo rig.
  - The translation vector for a moving camera.
- The epipole is the image of the centre of the other camera:
  - $e = PC'$, $e' = P'C$
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.

(a pencil is a one parameter family)
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 the scene structure (3D points external to the camera).
Epipolar Geometry
– Example: Converging Cameras

Note, epipolar lines are in general not parallel
Epipolar Geometry
– Example: Parallel Cameras
Epipolar Geometry – Algebraic Representation

- We know that the epipolar geometry defines a mapping:

\[ x \rightarrow l' \]

- The map only depends on the cameras \( P, P' \) (not on structure of the scene).

- It will be shown that the map is \textit{linear} and can be written as \( l' = Fx \), where \( F \) is the \textit{Fundamental matrix} \((3 \times 3)\).
Epipolar Geometry – Algebraic Representation

- Derivation of the algebraic expression \( I' = Fx \):

- Outline
  - Step 1: for a point \( x \) in the first image back project a ray with camera \( P \)
  - Step 2: choose two points on the ray and project into the second image with camera \( P' \)
  - Step 3: compute the line through the two image points using the relation \( I' = p \times q \)
**Recall**: camera matrices:

\[ P = K[R|t] \]

- Calibration matrix (intrinsic parameters)
- Rotation from world to camera coordinate frame
- Translation from world to camera coordinate frame

- First camera: \( P = K[I|0] \)
  - World coordinate frame aligned with first camera
  - Now looking for the relative orientation between the first camera and second camera

- Second camera: \( P' = K'[R|t] \)
Epipolar Geometry – Algebraic Representation

**Step 1**: For a point \( x \) in the first image back project a ray with camera

\[
P = K[I|0]
\]

A point \( x \) back projects to a ray

\[
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix} = zK^{-1}
\begin{pmatrix}
x \\
y \\
1
\end{pmatrix} = zK^{-1}x
\]

Where \( z \) is the point’s depth, since

\[
X(z) = \begin{pmatrix}
zK^{-1}x \\
1
\end{pmatrix}
\]

satisfies

\[
PX(z) = K[I|0]X(z) = x
\]
**Epipolar Geometry – Algebraic Representation**

**Step 2:** Choose two points on the ray and project onto the second image with camera $P'$

Consider two points on the ray:

- $z = 0$ is the camera centre $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$
- $z = \infty$ is the point at infinity $\begin{pmatrix} K^{-1}x \\ 0 \end{pmatrix}$

Project these two points onto the second view

$$P' \begin{pmatrix} 0 \\ 1 \end{pmatrix} = K'[R|t] \begin{pmatrix} 0 \\ 1 \end{pmatrix} = K't \quad P' \begin{pmatrix} K^{-1}x \\ 0 \end{pmatrix} = K' [R|t] \begin{pmatrix} K^{-1}x \\ 0 \end{pmatrix} = K' RK^{-1}x$$
Epipolar Geometry – Algebraic Representation

Step 3: Compute the line through the two image points using the relation \( l' = p \times q \)

- Compute the line through the points:
  \[
  l' = (K't) \times (K'R K^{-1}x)
  \]

- Using the identity:

  \[
  (Ma) \times (Mb) = M^{-T}(a \times b)
  \]

  where \( M^{-T} = (M^{-1})^T = (M^T)^{-1} \)

  \[
  l' = K'^{-T}(t \times (RK^{-1}x)) = K'^{-T}[t] \times RK^{-1}x
  \]

  \[
  l' = Fx
  \]

- Points \( x \) and \( x' \) correspond \((x \leftrightarrow x')\), then \( x'^T l' = 0 \)

  \[
  x'^T Fx = 0
  \]
Epipolar Geometry – The Fundamental Matrix $F$

- $F$ is the unique $3 \times 3$ rank 2 matrix that satisfies $x'^T F x = 0$ for all $x \leftrightarrow x'$

  - **Epipolar lines:** $l' = Fx$ and $l = F^T x'$
  
  - **Epipoles:** on all epipolar lines, thus $e'^T F x = 0$, $\forall x \Rightarrow e'^T F = 0$, similarly $Fe = 0$

- $F$ has 7 DOF, i.e. $3 \times 3 - 1$ (homogeneous, up to scale) - 1 ($\text{rank}(F) = 2$)

- $F$ is a correlation, projective mapping from a point $x$ to a line $l' = Fx$ (not a proper correlation, i.e. not invertible)
Epipolar Geometry – Epipolar Lines from $F$

- Fundamental matrix $F$ relates (homogeneous) pixel coordinates of corresponding points:
  \[ x'^T F x = 0 \]

- The transformation $Fx$ provides all potential positions of corresponding point $x'$ ($F$ has rank 2):
  \[ l' = Fx \]

- Homogenous coordinates $x'$ of points on the line $l'$ are computed in normal form by:
  \[ l'^T \cdot x = (l'_x \quad l'_y \quad l'_z) \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = l'_x \cdot x' + l'_y \cdot y' + l'_z = 0 \]

- Which provides the linear form:
  \[ y'(x') = \frac{-l'_x \cdot x' - l'_z}{l'_y} \]
Epipolar Geometry – Essential Matrix

- **Fundamental matrix:**
  - For points in **pixel** coordinates

  \[
  \begin{pmatrix} x'_{\text{pix}} & y'_{\text{pix}} & 1 \end{pmatrix} F \begin{pmatrix} x_{\text{pix}} \\ y_{\text{pix}} \\ 1 \end{pmatrix} = 0
  \]

- **Essential matrix:**
  - For points in **normalized image** coordinates

  \[
  \begin{pmatrix} x'_{\text{n}} & y'_{\text{n}} & 1 \end{pmatrix} E \begin{pmatrix} x_{\text{n}} \\ y_{\text{n}} \\ 1 \end{pmatrix} = 0
  \]

  - Transformation between pixel coordinates and camera coordinates using calibration matrix \( K \) (lecture 2):
    \[
    x_{\text{pix}} = K x_{\text{n}} \implies x_{\text{n}} = K^{-1} x_{\text{pix}}
    \]
    \[
    x'_{\text{pix}}^T F x_{\text{pix}} = (K' x'_{\text{n}})^T F (K x_{\text{n}}) = x'_{\text{n}}^T K'^T F K x_{\text{n}} = x'_{\text{n}}^T E x_{\text{n}} \implies E = K'^T F K
    \]
  
  - Moreover, we know that: 
    \[
    F = K'^{-T} [t] \times R K^{-1}
    \]
    \[
    \implies E = [t] \times R
    \]
Epipolar Geometry – Coplanarity Constraint

- Epipolar lines result from coplanarity constraint:

\[(x' \times t) \cdot x = (x' \times t)^T x = \begin{vmatrix} x' & t_x & x \\ y' & t_y & y \\ z' & t_z & z \end{vmatrix} = \det \begin{pmatrix} x' & t_x & x \\ y' & t_y & y \\ z' & t_z & z \end{pmatrix} = 0\]

- $x$, $x'$ and $t$ form a plane, therefore the coplanarity equation holds true!

- Note that vectors $x$, $x'$ and $t$ should be in the same coordinate system (we choose the one of camera $C'$).
Epipolar Geometry
– Essential Matrix from Coplanarity Constraint

- Rotation between different normalized coordinate systems of cameras (C to C').

\[ x'_n = Rx_n = R \begin{pmatrix} x_n \\ y_n \\ 1 \end{pmatrix} \quad x'_n = \begin{pmatrix} x'_n \\ y'_n \\ 1 \end{pmatrix} \]

- Due to coplanarity equation:

\[ (x'_n \times t)^T x'_C = 0 \]
\[ \Leftrightarrow (x'_n \times t)^T Rx_n = 0 \]
\[ \Leftrightarrow x'_n^T [t] \times Rx_n = 0 \]
\[ \Rightarrow E = [t] \times R \]

- \( x_n \) expressed in coordinate system of C'

- \( t \) is in coordinate system of C'

- Baseline/Translation \( t \)
Epipolar Geometry – Example: Parallel Cameras

- Compute the fundamental matrix for a parallel camera stereo rig:

  \[ P = K[I|0] \quad P' = K'[R|t] \]

  \[ K = K' = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad R = I \quad t = \begin{pmatrix} t_x \\ 0 \\ 0 \end{pmatrix} \]

  \[ F = K'^{-T}[t] \times RK^{-1} = \begin{bmatrix} 1/f & 0 & 0 \\ 0 & 1/f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -t_x \\ 0 & t_x & 0 \end{bmatrix} \begin{bmatrix} 1/f & 0 & 0 \\ 0 & 1/f & 0 \\ 0 & 0 & 1 \end{bmatrix} = \lambda \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \]

  \[ \lambda = \frac{t_x}{f^2} \]

  But we are in homogeneous space
Epipolar Geometry – Example: Parallel Cameras

\[ x'^T F x = \begin{pmatrix} x' & y' & 1 \end{pmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0 \]

- Reduce to \( y = y' \), i.e. raster correspondence (horizontal scan-lines)

- \( F \) is a rank 2 matrix:
  - The epipole \( e \) is the null-space vector (kernel) of \( F \):
    \[ F e = 0 \]
  - In our case:
    \[
    \begin{bmatrix}
    0 & 0 & 0 \\
    0 & 0 & -1 \\
    0 & 1 & 0 
    \end{bmatrix}
    \begin{pmatrix}
    1 \\
    0 \\
    0 
    \end{pmatrix} = 0 \\
    e = \begin{pmatrix}
    1 \\
    0 \\
    0 
    \end{pmatrix}
    \]

What is the geometric interpretation?
Epipolar Geometry
– Example: Forward Translating Camera

• Compute the $F$ for forward moving camera:

- $P = K[I|0]$; $P' = K'[R|t]$
- $K = K' = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}$; $R = I$; $t = \begin{pmatrix} 0 \\ 0 \\ t_z \end{pmatrix}$

$$F = K'^{-T}[t]_x RK^{-1} = \begin{bmatrix} 1/f & 0 & 0 \\ 0 & 1/f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -t_z & 0 \\ t_z & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1/f & 0 & 0 \\ 0 & 1/f & 0 \\ 0 & 0 & 1 \end{bmatrix} = \lambda \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\lambda = \frac{t_z}{f^2}$$

But we are in homogeneous space
From $I' = Fx$, the epipolar line for the point $x = (x, y, 1)^T$ is:

$$l' = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}$$

The points $(x, y, 1)^T$ and $(0,0,1)^T$ lie on this line.
Epipolar Geometry – Fundamental and Essential Matrix

- **Fundamental matrix** $F$: relative orientation between two views + camera parameters
  - 7 degree of freedom ($F$ has rank 2, $\det(F) = 0$)
  - General case: intrinsic parameters **not** available

- **Essential Matrix** $E$: relative orientation between two views
  - 5 degree of freedom
  - Calibrated case: intrinsic parameters available
THANK YOU!