Two-View Geometry
Multi-view geometry problems

- 3 types of information

Slide credit: Noah Snavely
Multi-view geometry problems

- **Structure**: Given projections of the same 3D point in two or more images, compute the 3D coordinates of that point.
Multi-view geometry problems

- **Stereo correspondence**: Given a point in one of the images, where could its corresponding points be in the other images?
Multi-view geometry problems

- **Motion**: Given a set of corresponding points in two or more images, compute the camera parameters.
Autonomous Aerial Navigation in Confined Indoor Environments

Shaojie Shen, Nathan Michael, Vijay Kumar
Structure: Triangulation

- Given projections of a 3D point in two or more images (with known camera matrices), find the coordinates of the point.
Structure: Triangulation

- We want to intersect the two visual rays corresponding to $x_1$ and $x_2$, but because of noise and numerical errors, they don’t meet exactly.
Triangulation: Geometric approach

- Find shortest segment connecting the two viewing rays and let \( X \) be the midpoint of that segment.
Triangulation: Linear approach

\[ \lambda_1 x_1 = P_1 X \quad x_1 \times P_1 X = 0 \quad [x_{1\times}]P_1 X = 0 \]
\[ \lambda_2 x_2 = P_2 X \quad x_2 \times P_2 X = 0 \quad [x_{2\times}]P_2 X = 0 \]

Cross product as matrix multiplication:

\[
\begin{bmatrix}
0 & -a_z & a_y \\
a_z & 0 & -a_x \\
-a_y & a_x & 0
\end{bmatrix}
\begin{bmatrix}
b_x \\
b_y \\
b_z
\end{bmatrix} = [a_{\times}]b
\]
Triangulation: Linear approach

\[
\lambda_1 x_1 = P_1 X \quad x_1 \times P_1 X = 0 \quad [x_{1x}]P_1 X = 0 \\
\lambda_2 x_2 = P_2 X \quad x_2 \times P_2 X = 0 \quad [x_{2x}]P_2 X = 0
\]

How many constraints and how many unknowns?
Triangulation: Linear approach

\[
\begin{align*}
\lambda_1 x_1 &= P_1 X \\
\lambda_2 x_2 &= P_2 X
\end{align*}
\]

\[
\begin{align*}
x_1 \times P_1 X &= 0 \\
x_2 \times P_2 X &= 0
\end{align*}
\]

\[
[x_{1\times}] P_1 X = 0 \quad [x_{2\times}] P_2 X = 0
\]

Six constraints and three unknowns
Triangulation: Linear approach

\[ \lambda_1 x_1 = P_1 X \quad x_1 \times P_1 X = 0 \quad [x_{1\times}] P_1 X = 0 \]

\[ \lambda_2 x_2 = P_2 X \quad x_2 \times P_2 X = 0 \quad [x_{2\times}] P_2 X = 0 \]

Six constraints and three unknowns

\[ AX = 0 \]
Triangulation: Linear approach

\[ \lambda_1 x_1 = P_1 X \quad x_1 \times P_1 X = 0 \quad [x_1 \times] P_1 X = 0 \]

\[ \lambda_2 x_2 = P_2 X \quad x_2 \times P_2 X = 0 \quad [x_2 \times] P_2 X = 0 \]

Six constraints and three unknowns

AX = 0

Eigen vector of \((A^\top A)\)…
Two-view geometry
Epipolar constraint

• If we observe a point $x$ in one image, where can the corresponding point $x'$ be in the other image?
Epipolar constraint

- Potential matches for $\mathbf{x}$ have to lie on the corresponding epipolar line $l'$.

- Potential matches for $\mathbf{x}'$ have to lie on the corresponding epipolar line $l$. 
Epipolar constraint

• Potential matches for \( x \) have to lie on the corresponding epipolar line \( l' \).

• Potential matches for \( x' \) have to lie on the corresponding epipolar line \( l \).
Epipolar constraint

- Potential matches for $x$ have to lie on the corresponding epipolar line $l'$.

- Potential matches for $x'$ have to lie on the corresponding epipolar line $l$. 
Epipolar geometry

- **Baseline** – line connecting the two camera centers
- **Epipolar Plane** – plane containing baseline (1D family)
- **Epipoles**
  - intersections of baseline with image planes
  - projections of the other camera center
  - vanishing points of the baseline (motion direction)
The Epipole

Photo by Frank Dellaert
Epipolar geometry

- **Baseline** – line connecting the two camera centers
- **Epipolar Plane** – plane containing baseline (1D family)
- **Epipoles**
  - intersections of baseline with image planes
  - projections of the other camera center
  - vanishing points of the baseline (motion direction)
- **Epipolar Lines** - intersections of epipolar plane with image planes (always come in corresponding pairs)
Example: Converging cameras
Example: Motion parallel to image plane

\[ e \text{ at } \infty \quad \rightarrow \quad e' \text{ at } \infty \]
Example: Motion perpendicular to image plane
Example: Motion perpendicular to image plane
Example: Motion perpendicular to image plane

- Points move along lines radiating from the epipole: “focus of expansion”
- Epipole is the principal point
Epipolar constraint example
Epipolar constraint: Calibrated case

- Assume that the intrinsic and extrinsic parameters of the cameras are known
- We can multiply the projection matrix of each camera (and the image points) by the inverse of the calibration matrix to get *normalized* image coordinates
The vectors $Rx$, $t$, and $x'$ are coplanar.
Epipolar constraint: Calibrated case

\[ x' \cdot [t \times (Rx)] = 0 \quad \rightarrow \quad x'^T E x = 0 \quad \text{with} \quad E = [t_x]R \]

**Essential Matrix**
(Longuet-Higgins, 1981)

The vectors \( Rx, t, \) and \( x' \) are coplanar
Triangulation: Linear approach

\[
\begin{align*}
\lambda_1 x_1 &= P_1 X \\
\lambda_2 x_2 &= P_2 X \\
x_1 \times P_1 X &= 0 \\
x_2 \times P_2 X &= 0 \\
[x_{1\times}]P_1 X &= 0 \\
[x_{2\times}]P_2 X &= 0
\end{align*}
\]

Cross product as matrix multiplication:

\[
a \times b = \begin{bmatrix}
0 & -a_z & a_y \\
 a_z & 0 & -a_x \\
- a_y & a_x & 0 \\
\end{bmatrix}
\begin{bmatrix}
b_x \\
b_y \\
b_z \\
\end{bmatrix} = [a_{\times}]b
\]
Epipolar constraint: Calibrated case

- $E x$ is the epipolar line associated with $x$ ($l' = E x$)
- $E^T x'$ is the epipolar line associated with $x'$ ($l = E^T x'$)
- $E e = 0$ and $E^T e' = 0$
- $E$ is singular (rank two)
- $E$ has five degrees of freedom
Epipolar constraint: Uncalibrated case

- The calibration matrices $K$ and $K'$ of the two cameras are unknown.
- We can write the epipolar constraint in terms of unknown normalized coordinates:

$$\hat{x}'^T E \hat{x} = 0 \quad \hat{x} = K^{-1} x, \quad \hat{x}' = K'^{-1} \hat{x}'$$
Epipolar constraint: Uncalibrated case

\[ \hat{x}'^T E \hat{x} = 0 \quad \Rightarrow \quad x'^T F x = 0 \quad \text{with} \quad F = K'^{-T} E K^{-1} \]

\[ \hat{x} = K^{-1} x \]

\[ \hat{x}' = K'^{-1} x' \]

Fundamental Matrix
(Faugeras and Luong, 1992)
Epipolar constraint: Uncalibrated case

\[
\hat{x}'^T E \hat{x} = 0 \quad \Rightarrow \quad x'^T F x = 0 \quad \text{with} \quad F = K'^{-T} E K^{-1}
\]

- \( F x \) is the epipolar line associated with \( x \) \((l' = F x)\)
- \( F^T x' \) is the epipolar line associated with \( x' \) \((l' = F^T x')\)
- \( F e = 0 \) and \( F^T e' = 0 \)
- \( F \) is singular (rank two)
- \( F \) has seven degrees of freedom
The eight-point algorithm

\[ x = (u, v, 1)^T, \quad x' = (u', v', 1) \]

\[
\begin{bmatrix}
  u' & v' & 1
\end{bmatrix}
\begin{bmatrix}
  f_{11} & f_{12} & f_{13} \\
  f_{21} & f_{22} & f_{23} \\
  f_{31} & f_{32} & f_{33}
\end{bmatrix}
\begin{bmatrix}
  u \\
  v \\
  1
\end{bmatrix} = 0
\]

\[
\begin{bmatrix}
  u'u' & u'v & u' & v'u & v'v & v' & u & v & 1
\end{bmatrix}
\begin{bmatrix}
  f_{11} \\
  f_{12} \\
  f_{13} \\
  f_{21} \\
  f_{22} \\
  f_{23} \\
  f_{31} \\
  f_{32} \\
  f_{33}
\end{bmatrix} = 0
\]

Minimize:

\[
\sum_{i=1}^{N} (x_i'^T F x_i)^2
\]

under the constraint

\[ ||F||^2 = 1 \]
The eight-point algorithm

- Meaning of error $\sum_{i=1}^{N} (x_i'^T F x_i)^2$:
  
  sum of squared \textit{algebraic} distances between points $x_i'$ and epipolar lines $Fx_i$ (or points $x_i$ and epipolar lines $F^Tx_i$)

- Nonlinear approach: minimize sum of squared \textit{geometric} distances

$$\sum_{i=1}^{N} \left[ d^2(x_i', F x_i) + d^2(x_i, F^T x_i') \right]$$
Problem with eight-point algorithm

\[
\begin{bmatrix}
u'v & u'v & u' & v'u & v'v & v' & u & v
\end{bmatrix}
\begin{bmatrix}
f_{11} \\
f_{12} \\
f_{13} \\
f_{21} \\
f_{22} \\
f_{23} \\
f_{31} \\
f_{32}
\end{bmatrix} = -1
\]
Problem with eight-point algorithm

<table>
<thead>
<tr>
<th></th>
<th>250906.36</th>
<th>183269.57</th>
<th>921.81</th>
<th>200931.10</th>
<th>146766.13</th>
<th>738.21</th>
<th>272.19</th>
<th>198.81</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2692.28</td>
<td>131633.03</td>
<td>176.27</td>
<td>6196.73</td>
<td>302975.59</td>
<td>405.71</td>
<td>15.27</td>
<td>746.79</td>
</tr>
<tr>
<td></td>
<td>416374.23</td>
<td>871684.30</td>
<td>935.47</td>
<td>408110.89</td>
<td>854384.92</td>
<td>916.90</td>
<td>445.10</td>
<td>931.81</td>
</tr>
<tr>
<td></td>
<td>191183.60</td>
<td>171759.40</td>
<td>410.27</td>
<td>416435.62</td>
<td>374125.90</td>
<td>893.65</td>
<td>465.99</td>
<td>418.65</td>
</tr>
<tr>
<td></td>
<td>48988.86</td>
<td>30401.76</td>
<td>57.89</td>
<td>298604.57</td>
<td>185309.58</td>
<td>352.87</td>
<td>846.22</td>
<td>525.15</td>
</tr>
<tr>
<td></td>
<td>164786.04</td>
<td>546559.67</td>
<td>813.17</td>
<td>1998.37</td>
<td>6628.15</td>
<td>9.86</td>
<td>202.65</td>
<td>672.14</td>
</tr>
<tr>
<td></td>
<td>116407.01</td>
<td>2727.75</td>
<td>138.89</td>
<td>169941.27</td>
<td>3982.21</td>
<td>202.77</td>
<td>838.12</td>
<td>19.64</td>
</tr>
<tr>
<td></td>
<td>135384.58</td>
<td>75411.13</td>
<td>198.72</td>
<td>411350.03</td>
<td>229127.78</td>
<td>603.79</td>
<td>681.28</td>
<td>379.48</td>
</tr>
</tbody>
</table>

\[
\begin{bmatrix}
  f_{11} \\
  f_{12} \\
  f_{13} \\
  f_{21} \\
  f_{22} \\
  f_{23} \\
  f_{31} \\
  f_{32}
\end{bmatrix} = -1
\]

Poor numerical conditioning
Can be fixed by rescaling the data
The normalized eight-point algorithm

(Hartley, 1995)

• Center the image data at the origin, and scale it so the mean squared distance between the origin and the data points is 2 pixels
• Use the eight-point algorithm to compute $F$ from the normalized points
• Enforce the rank-2 constraint (for example, take SVD of $F$ and throw out the smallest singular value)
• Transform fundamental matrix back to original units: if $T$ and $T'$ are the normalizing transformations in the two images, than the fundamental matrix in original coordinates is $T'^TFT$
Comparison of estimation algorithms

<table>
<thead>
<tr>
<th></th>
<th>8-point</th>
<th>Normalized 8-point</th>
<th>Nonlinear least squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Av. Dist. 1</td>
<td>2.33 pixels</td>
<td>0.92 pixel</td>
<td>0.86 pixel</td>
</tr>
<tr>
<td>Av. Dist. 2</td>
<td>2.18 pixels</td>
<td>0.85 pixel</td>
<td>0.80 pixel</td>
</tr>
</tbody>
</table>
fundamental matrix song
From epipolar geometry to camera calibration

• Estimating the fundamental matrix is known as “weak calibration”

• If we know the calibration matrices of the two cameras, we can estimate the essential matrix: \( E = K'\mathbf{T} FK \)

• The essential matrix gives us the relative rotation and translation between the cameras, or their extrinsic parameters
Calibrated case

\[ x' \cdot [t \times (Rx)] = 0 \quad \Rightarrow \quad x'^TEx = 0 \quad \text{with} \quad E = [t_x]R \]
5-point algorithm

• Feature tracking
• RANSAC on 5 feature matches to get E
• Extract R and t
• continue...