Lecture 6
Stereo Systems
Multi-view geometry

Professor Silvio Savarese
Computational Vision and Geometry Lab
Lecture 6
Stereo Systems
Multi-view geometry

- Stereo systems
  - Rectification
  - Correspondence problem
- Multi-view geometry
  - The SFM problem
  - Affine SFM

Reading:

- [AZ] Chapter: 18 “N view computational methods”
- [FP] Chapters: 7 “Stereopsis”
- [FP] Chapters: 8 “Structure from Motion”
Epipolar geometry

- Epipolar Plane
- Baseline
- Epipolar Lines

\[ e, e' \] = intersections of baseline with image planes
\[ p, p' \] = projections of the other camera center
Epipolar Constraint

\[ p^T E p' = 0 \]

**E = Essential Matrix**  
(Longuet-Higgins, 1981)
Essential matrix

$$E = [T_x] \cdot R$$

$$E = \begin{bmatrix}
0 & -T_z & T_y \\
T_z & 0 & -T_x \\
-T_y & T_x & 0
\end{bmatrix}$$
Epipolar Constraint

\[ F = \text{Fundamental Matrix} \]

(Faugeras and Luong, 1992)
• Epipolar lines are horizontal
• Epipoles go to infinity
• v-coordinates are equal

\[
p = \begin{bmatrix} p_u \\ p_v \\ 1 \end{bmatrix} \quad p' = \begin{bmatrix} p'_u \\ p_v \\ 1 \end{bmatrix}
\]
Parallel image planes

Hint:
\[ R = I \quad T = (T, 0, 0) \]

E = ?

\[ K_1 = K_2 = \text{known} \]
\[ x \text{ parallel to } O_1O_2 \]
Essential matrix for parallel images

\[
E = \begin{bmatrix} T_x \end{bmatrix} \cdot R
\]

\[
E = \begin{bmatrix}
0 & -T_z & T_y \\
T_z & 0 & -T_x \\
-T_y & T_x & 0 \\
\end{bmatrix}
\]

\[
R = \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & -T \\
0 & T & 0 \\
\end{bmatrix}
\]

\[
T = \begin{bmatrix} T & 0 & 0 \end{bmatrix}
\]

\[
R = I
\]
What are the directions of epipolar lines?

\[ l = E \hat{p}' = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -T \\ 0 & T & 0 \end{bmatrix} \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -T \\ T v' \end{bmatrix} \text{ horizontal!} \]
Parallel image planes

How are $p$ and $p'$ related?

$$p^T \cdot E \cdot p' = 0$$
How are $p$ and $p'$ related?

$$A = A' = 0 \Rightarrow v = v'$$

$$A = A'$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Parallel image planes
Parallel image planes

Rectification: making two images “parallel”

Why it is useful? • Epipolar constraint \( v = v' \)
• New views can be synthesized by linear interpolation
Rectification: making two images “parallel”
Application: view morphing

Rectification
From its reflection!
Deep view morphing

D. Ji, J. Kwon, M. McFarland, S. Savarese, CVPR 2017
Deep view morphing

D. Ji, J. Kwon, M. McFarland, S. Savarese, CVPR 2017
Deep view morphing

D. Ji, J. Kwon, M. McFarland, S. Savarese, CVPR 2017
Why are parallel images useful?

- Makes triangulation easy
- Makes the correspondence problem easier
Point triangulation

Disparity is inversely proportional to depth $z$!

$$p = \begin{bmatrix} p_u \\ p_v \\ 1 \end{bmatrix} \quad p' = \begin{bmatrix} p'_u \\ p'_v \\ 1 \end{bmatrix}$$

Disparity = $p_u - p'_u \propto \frac{B \cdot f}{z}$

$[\text{Eq. 1}]$
Disparity is inversely proportional to depth $z$!
Disparity maps

http://vision.middlebury.edu/stereo/

\[ p_u - p'_u \propto \frac{B \cdot f}{z} \]

Eq. 1

Stereo pair

Disparity map / depth map
Why are parallel images useful?

- Makes triangulation easy
- Makes the correspondence problem easier
Correspondence problem

Given a point in 3D, discover corresponding observations in left and right images [also called binocular fusion problem]
When images are rectified, this problem is much easier!
Correspondence problem

• A Cooperative Model (Marr and Poggio, 1976)

• Correlation Methods (1970–)

• Multi-Scale Edge Matching (Marr, Poggio and Grimson, 1979-81)

[FP] Chapters: 7
Correlation Methods (1970–)

Where is $\vec{p}$?

Where is $\vec{p}'$?

$\vec{p} = \begin{bmatrix} \bar{u} \\ \bar{v} \\ 1 \end{bmatrix}$

$\vec{p}' = \begin{bmatrix} \bar{u}' \\ \bar{v} \\ 1 \end{bmatrix}$
Correlation Methods (1970–)

Where is $\bar{p}'$?

$\bar{p} = \begin{bmatrix} \bar{u} \\ \bar{v} \\ 1 \end{bmatrix}$

$\bar{p}' = \begin{bmatrix} \bar{u}' \\ \bar{v} \\ 1 \end{bmatrix}$
Correlation Methods (1970–)

What’s the problem with this?
Window-based correlation

- Pick up a window $W$ around $\overline{p} = (\overline{u}, \overline{v})$
- Build vector $w$
Window-based correlation

Example: \( \mathbf{W} \) is a 3x3 window in red
\( \mathbf{w} \) is a 9x1 vector
\( \mathbf{w} = [100, 100, 100, 90, 100, 20, 150, 150, 145]^T \)

- Pick up a window \( \mathbf{W} \) around \( \overline{p} = (\overline{u}, \overline{v}) \)
- Build vector \( \mathbf{w} \)
- Slide the window \( \mathbf{W} \) along \( \mathbf{v} = \overline{V} \) in image 2 and compute \( \mathbf{w}'(u) \) for each \( u \)
- Compute the dot product \( \mathbf{w}^T \mathbf{w}'(u) \) for each \( u \) and retain the max value
Window-based correlation

Example: $W$ is a 3x3 window in red

$w$ is a 9x1 vector

$w = [100, 100, 100, 90, 100, 20, 150, 150, 145]^T$

What's the problem with this?
Changes of brightness/exposure

Changes in the mean and the variance of intensity values in corresponding windows!
Find $u$ that maximizes:

$$\frac{(w - \bar{w})^T (w'(u) - \bar{w}')}{{\|w - \bar{w}\|}{\|w'(u) - \bar{w}'\|}}$$

[Eq. 2]

$\bar{w}$ = mean value within $W$ located at $u^{bar}$ in image 1

$\bar{w}'(u)$ = mean value within $W$ located at $u$ in image 2
Example

Image 1

Image 2

\[ \vec{v} \]

scanline

NCC

Credit slide S. Lazebnik
Effect of the window’s size

- Smaller window
  - More detail
  - More noise

- Larger window
  - Smoother disparity maps
  - Less prone to noise

Window size = 3
Window size = 20

Credit slide S. Lazebnik
Issues

• Fore shortening effect

• Occlusions
Issues

• To reduce the effect of foreshortening and occlusions, it is desirable to have small $B/z$ ratio!

• However, when $B/z$ is small, small errors in measurements imply large error in estimating depth
Issues

- Homogeneous regions
Issues

• Repetitive patterns
Correspondence problem is difficult!

- Occlusions
- Fore shortening
- Baseline trade-off
- Homogeneous regions
- Repetitive patterns

Apply non-local constraints to help enforce the correspondences
Non-local constraints

• Uniqueness
  – For any point in one image, there should be at most one matching point in the other image

• Ordering
  – Corresponding points should be in the same order in both views

• Smoothness
  – Disparity is typically a smooth function of $x$ (except in occluding boundaries)
Lecture 6
Stereo Systems
Multi-view geometry

- Stereo systems
  - Rectification
  - Correspondence problem
- Multi-view geometry
  - The SFM problem
  - Affine SFM
Structure from motion problem

Courtesy of Oxford Visual Geometry Group
Given $m$ images of $n$ fixed 3D points

\[ x_{ij} = M_i X_j, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n \]
From the $m \times n$ observations $x_{ij}$, estimate:

- $m$ projection matrices $M_i$
- $n$ 3D points $X_j$
Affine structure from motion
(simpler problem)

From the $m \times n$ observations $x_{ij}$, estimate:

- $m$ projection matrices $M_i$ (affine cameras)
- $n$ 3D points $X_j$
**Perspective**

\[
x = M X = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} X = \begin{bmatrix} m_1 X \\ m_2 X \\ m_3 X \end{bmatrix}
\]

\[
x^E = \left( \frac{m_1 X}{m_3 X}, \frac{m_2 X}{m_3 X} \right)^T
\]

**Affine**

\[
x = M X = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} X = \begin{bmatrix} m_1 X \\ m_2 X \\ 1 \end{bmatrix}
\]

\[
M = \begin{bmatrix} A & b \\ v & 1 \end{bmatrix} = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}
\]

\[
x^E = (m_1 X, m_2 X)^T = \begin{bmatrix} A & b \end{bmatrix} X = \begin{bmatrix} A & b \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = AX^E + b
\]

\[
X^E = \begin{bmatrix} x \\ y \\ z \end{bmatrix}
\]

magnification

[Eq. 3]
Affine cameras

For the affine case (in Euclidean space)

\[ x_{ij} = A_i X_j + b_i \]  

[Eq. 4]
The Affine Structure-from-Motion Problem

Given $m$ images of $n$ fixed points $X_i$ we can write

$$x_{ij} = A_i X_j + b_i \quad \text{for } i = 1, \ldots, m \text{ and } j = 1, \ldots, n$$

Problem: estimate $m$ matrices $A_i$, $m$ matrices $b_i$, and the $n$ positions $X_i$ from the $m \times n$ observations $x_{ij}$.

How many equations and how many unknown?

$2m \times n$ equations in $8m + 3n - 9$ unknowns
The Affine Structure-from-Motion Problem

Two approaches:

- Algebraic approach (affine epipolar geometry; estimate $F$; cameras; points)

- Factorization method
Next lecture

Multiple view geometry:
Affine and Perspective structure from Motion