Stereo Vision II

Introduction to Computer Vision
CSE 152
Lecture 14

Binocular Stereopsis: Mars
Given two images of a scene where relative locations of cameras are known, estimate depth of all common scene points.

Two images of Mars

Stereo Vision Outline

- Offline: Calibrate cameras & determine “epipolar geometry”
- Online
  1. Acquire stereo images
  2. Rectify images to convenient epipolar geometry
  3. Establish correspondence
  4. Estimate depth

Binocular Stereo System
Estimating Depth
2D world with 1-D image plane

Constants:
Baseline: d
Focal length: f

Disparity: (X_L - X_R)

Reconstruction: General 3-D case
Given two image measurements p and p', estimate P.

- Linear Method: find P such that
  \[ \begin{align*}
  p \cdot MP &= 0 \\
  p' \cdot M'P &= 0
  \end{align*} \]
  Where M is camera matrix

- Non-Linear Method: find Q minimizing
  \[ d(p, q) + d(p', q') \]
  where q=MQ and q'=M'Q

The search for correspondence:
Where do you look?
Two Approaches

1. Feature-Based
   - From each image, process “monocular” image to obtain image features or cues (e.g., corners, lines).
   - Establish correspondence between the detected features.

2. Area-Based
   - Directly compare image regions between the two images.

Human Stereopsis: Binocular Fusion

How are the correspondences established?

Julesz (1971): Is the mechanism for binocular fusion a monocular process or a binocular one??

• There is anecdotal evidence for the latter (camouflage).

Random dot stereograms provide an objective answer

Random Dot Stereograms

• Potential matches for \( p \) have to lie on the corresponding epipolar line \( l' \).
• Potential matches for \( p' \) have to lie on the corresponding epipolar line \( l \).
1. From image of known calibration fixture, determine intrinsic parameters and extrinsic relation of two cameras.
2. Compute the relative position and orientation of the two cameras from $R_x, R_y, t_x, t_y$
3. Compute the Essential Matrix $E = [t]R$

**Skew Symmetric Matrix & Cross Product**

- The cross product $a \times b$ of two vectors $a$ and $b$ can be expressed as a matrix vector product $[a]b$ where $[a]$ is the skew symmetric matrix:

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

- A matrix $S$ is skew symmetric iff $S = -S^T$

**Two ways to estimate the Essential Matrix**

1. Calibration-based
2. Eight-Point Algorithm

**The Eight-Point Algorithm (Longuet-Higgins, 1981)**

$$[u, v \mid E_{11} E_{12} E_{13} \begin{bmatrix} u' \\ v' \end{bmatrix}] = 0$$

- Set $E_{13}$ to 1
- Consider 8 points $(u_i, v_i), (u'_i, v'_i), i = 1, 8$

Solve $E_{12}$ to $E_{11}$

**The Essential Matrix**
Properties of the Essential Matrix

\[ \mathbf{p}' \mathbf{E} \mathbf{p} = 0 \text{ with } \mathbf{E} = [t, \mathbf{R}] \]

• \( \mathbf{E} \mathbf{p}' \) is the epipolar line associated with \( \mathbf{p}' \).
• \( \mathbf{E} \mathbf{p} \) is the epipolar line associated with \( \mathbf{p} \).
• \( \mathbf{E} \mathbf{e}' = 0 \) and \( \mathbf{E} \mathbf{e} = 0 \).
• \( \mathbf{E} \) is singular.
• \( \mathbf{E} \) has two equal non-zero singular values (Huang and Faugeras, 1989).

The Eight-Point Algorithm (Longuet-Higgins, 1981)

Much more on multi-view in CSE252B!!

\[
\begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = 0
\]

\[
\begin{bmatrix} (u', v', x, y, u, v, x', y') \end{bmatrix} = \begin{bmatrix} 1 \end{bmatrix}
\]

• Following the text, view this as a system of homogeneous equations in \( F_{ij} \) to \( F_{ij} \).
• Solve as eigenvector corresponding to the smallest Eigenvalue of matrix created from the image data.

Equivalent to solving

\[
\sum_{i} (\mathbf{p}_i \mathbf{E} \mathbf{p}_i)^{-1}
\]

under the constraint

\[ \mathbf{E} \mathbf{E}^T = 1. \]

The Fundamental Matrix

The epipolar constraint is given by: \( \mathbf{p}' \mathbf{E} \mathbf{p} = 0 \) with \( \mathbf{E} = [t, \mathbf{R}] \)

where \( \mathbf{p} \) and \( \mathbf{p}' \) are 3-D coordinates of the image coordinates of points in the two images.

Without calibration, we can still identify corresponding points in two images, but we can’t convert to 3-D coordinates. However, the relationship between the calibrated coordinates \( (\mathbf{p}, \mathbf{p}') \) and uncalibrated image coordinates \( (\mathbf{q}, \mathbf{q}') \) can be expressed as \( \mathbf{p} = \mathbf{A} \mathbf{q} \) and \( \mathbf{p}' = \mathbf{A}' \mathbf{q}' \).

Therefore, we can express the epipolar constraint as:

\[ (\mathbf{A} \mathbf{q})^T \mathbf{E}(\mathbf{A}' \mathbf{q}') = \mathbf{q}^T (\mathbf{A}^T \mathbf{E} \mathbf{A}') \mathbf{q}' = 0 \]

where \( \mathbf{F} \) is called the Fundamental Matrix.

Can estimate \( \mathbf{F} \) using 8 point algorithm WITHOUT CALIBRATION

Epipolar geometry example

Example: converging cameras

Example: motion parallel with image plane

(simple for stereo → rectification)
Example: forward motion

courtesy of Andrew Zisserman