Stereo III

Computer Vision I
CSE252A
Lecture 16

Announcements
• HW1 being returned
• HW3 assigned and due date extended until 11/27/12
• No office hours today
• No class on Thursday 12/6
• Extra class on Tuesday 12/4 at 6:30PM in WLH Room 2112

Random Dot Stereograms

The Eight-Point Algorithm (Longuet-Higgins, 1981)
Much more on multi-view in CSE252B

\( \begin{align*}
 F_{11} & = F_{22} \\
 F_{12} & = F_{21} \\
 F_{31} & = F_{32} \\
 F_{33} & = 1 \\
 \end{align*} \)

\( \begin{bmatrix}
 u_i & v_i & 1
\end{bmatrix} = 0 \\
 \begin{bmatrix}
 u_i' & v_i' & 1
\end{bmatrix} = 0 \\
\end{align*} \)

Here, \( F \) is the Essential Matrix

\( \begin{align*}
 F_{11} & = F_{22} \\
 F_{12} & = F_{21} \\
 F_{31} & = F_{32} \\
 F_{33} & = 1 \\
 \end{align*} \)

Consider 8 points \( (u_i,v_i, 1) \), \( (u'_i,v'_i, 1) \)

Set \( F_{13} \) to 1

Solve for \( F_{12} \) and \( F_{22} \)

For more than 8 points, solve using linear least squares

Properties of the Essential Matrix

\[ p^T \mathcal{E} p' = 0 \quad \text{with} \quad \mathcal{E} = [t, \mathcal{R}] \]

• \( \mathcal{E} p' \) is the epipolar line associated with \( p' \).
• \( \mathcal{E}^T p \) is the epipolar line associated with \( p \).
• \( \mathcal{E} e = 0 \) and \( \mathcal{E}^T e = 0 \).
• \( \mathcal{E} \) is singular.
• \( \mathcal{E} \) has two equal non-zero singular values (Huang and Faugeras, 1989).
The Fundamental Matrix

The epipolar constraint is given by: \( p' \xi p = 0 \) with \( \xi = [x, y]^T \)
where \( p \) and \( p' \) 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 \((p, p')\) and uncalibrated image coordinates \((q, q')\) can be expressed as \( p = Aq \) and \( p' = A'q' \).

Therefore, we can express the epipolar constraint as:
\[
(Aq)^T F (A'q') = q^T (A^T E A')q' = q^T F q' = 0
\]
where \( F \) is called the Fundamental Matrix.

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

Example: forward motion

Need for correspondence

RANSAC

Slides shamelessly taken from Frank Dellaert and Marc Pollefeys and modified

Motivation

- Estimating motion models
- Typically: points in two images
- Candidates:
  - Translation
  - Affine
  - Homography
Mosaicing: Homography Estimate with RANSAC

Fundamental Matrix Estimate with RANSAC

www.cs.cmu.edu/~dellaert/mosaicking

Simpler Example

- Fitting a straight line

- Inliers
- Outliers

Discard Outliers

- No point with \( d > t \)
- RANSAC:
  - RANdom SAmple Consensus
  - Fischler & Bolles 1981
  - Copes with a large proportion of outliers

Main Idea

- Select 2 points at random
- Fit a line
- “Support” = number of inliers
- Line with most inliers wins

Why will this work?

- Plot of lines with inliers and outliers
RANSAC

Objective
Robust fit of model to data set $S$ which contains outliers

Algorithm
(i) Randomly select a sample of $s$ data points from $S$ and instantiate the model from this subset.
(ii) Determine the set of data points $S_i$ which are within a distance threshold $t$ of the model. The set $S_i$ is the consensus set of samples and defines the inliers of $S$.
(iii) If the size of $S_i$ is greater than some threshold $T$, re-estimate the model using all the points in $S_i$ and terminate
(iv) If the size of $S_i$ is less than $T$, select a new subset and repeat the above.
(v) After $N$ trials the largest consensus set $S_i$ is selected, and the model is re-estimated using all the points in the subset $S_i$.

How many samples?
Choose $N$ (number of samples) so that, with probability $p$, at least one random sample is free from outliers. e.g. $p=0.99$

$$\left(1 - (1 - e)^s\right)^N = 1 - p$$

$$N = \log(1 - p)/\log\left(1 - (1 - e)^s\right)$$

$e$: proportion of outliers
$s$: Number of points needed for the model

<table>
<thead>
<tr>
<th>proportion of outliers $e$</th>
<th>2%</th>
<th>5%</th>
<th>10%</th>
<th>20%</th>
<th>25%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>11</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>4</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td>19</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>5</td>
<td>9</td>
<td>13</td>
<td>17</td>
<td>34</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>6</td>
<td>12</td>
<td>17</td>
<td>26</td>
<td>57</td>
<td>146</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>7</td>
<td>16</td>
<td>24</td>
<td>37</td>
<td>97</td>
<td>293</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>8</td>
<td>20</td>
<td>33</td>
<td>54</td>
<td>163</td>
<td>588</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>9</td>
<td>26</td>
<td>44</td>
<td>78</td>
<td>272</td>
<td>1177</td>
<td></td>
</tr>
</tbody>
</table>

Acceptable consensus set?
• Typically, terminate when inlier ratio reaches expected ratio of inliers

$$T = (1 - e)N$$

What’s a good
• How do we "Determine the set of data points $S_i$ which are within a distance threshold $t$ of the model" when estimating the Fundamental Matrix?

Using RANSAC to estimate the Fundamental Matrix
• What is the model?
• How many “points” are needed, and where do they come from?
• What distance do we use to compute the consensus set?
• How often do outliers occur
**Rectification**

Given a pair of images, transform both images so that epipolar lines are scan lines.

**Image pair rectification**

- simplify stereo matching by warping the images
- Apply projective transformation $H$ so that epipolar lines correspond to horizontal scanlines
- Map epipole $e$ to $(1,0,0)$
- Try to minimize image distortion
- Note that rectified images usually not rectangular

**The ordering constraint**

The order of corresponding points along a pair of epipolar lines should be the same.

Note: As seen in previous slides, doesn’t always hold, but usually does and can vastly speed up computation

**Finding Correspondences**

Comparing Windows:

\[
SSD = \sum_{(i,j) \in R} (f(i,j) - g(i,j))^2
\]

\[
C_{fp} = \sum_{(i,j) \in R} f(i,j)g(i,j)
\]

For each window, match to closest window on epipolar line in other image.

(Camps)
Greedy Correspondence Search Algorithm

For \( i = 1: \text{nrows} \)
for \( j = 1: \text{ncols} \)
    best\((i,j)\) = -1
    for \( k = \text{mindisparity}: \text{maxdisparity} \)
        \( c = \text{Match\_Metric}(I_1(i,j), I_2(i,j+k), \text{winsize}) \)
        if \( (c > \text{best}(i,j)) \)
            \( \text{best}(i,j) = c \)
            \( \text{disparities}(i,j) = k \)
    end
end
end
Complexity \( O(\text{nrows} * \text{ncols} * \text{disparities} * \text{winx} * \text{winy}) \)

Match Metric Summary

<table>
<thead>
<tr>
<th>MATCH METRIC</th>
<th>DEFINITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalized Cross-Correlation (NCC)</td>
<td>( \frac{\sum_{x=0}^{\text{winsize}} \sum_{y=0}^{\text{winsize}} (I_1(x,y) - \mu_1)(I_2(x+k,y) - \mu_2)}{\sqrt{\sum_{x=0}^{\text{winsize}} \sum_{y=0}^{\text{winsize}} (I_1(x,y) - \mu_1)^2} \sqrt{\sum_{x=0}^{\text{winsize}} \sum_{y=0}^{\text{winsize}} (I_2(x+k,y) - \mu_2)^2}} )</td>
</tr>
<tr>
<td>Sum of Squared Differences (SSD)</td>
<td>( \sum_{x=0}^{\text{winsize}} \sum_{y=0}^{\text{winsize}} (I_1(x,y) - I_2(x+k,y))^2 )</td>
</tr>
<tr>
<td>Normalized SSD</td>
<td>( \frac{\sum_{x=0}^{\text{winsize}} \sum_{y=0}^{\text{winsize}} (I_1(x,y) - I_2(x+k,y))^2}{\sum_{x=0}^{\text{winsize}} \sum_{y=0}^{\text{winsize}} (I_1(x,y) - \mu_1)^2} )</td>
</tr>
<tr>
<td>Sum of Absolute Differences (SAD)</td>
<td>( \sum_{x=0}^{\text{winsize}} \sum_{y=0}^{\text{winsize}}</td>
</tr>
<tr>
<td>Zero Mean SAD</td>
<td>( \sum_{x=0}^{\text{winsize}} \sum_{y=0}^{\text{winsize}}</td>
</tr>
<tr>
<td>Census</td>
<td>( \sum_{x=0}^{\text{winsize}} \sum_{y=0}^{\text{winsize}} \text{census}(I_1(x,y), I_2(x+k,y)) )</td>
</tr>
</tbody>
</table>

Some Issues

- Ambiguity
- Uniqueness
- Ordering
- Window size
- Window shape
- Lighting
- Half occluded regions
Ambiguity

The window in left image might end up matching either of the two windows in the right image. They look very similar.

Uniqueness

Each feature on left epipolar line match one and only one feature on right epipolar line.

The two window in left image might are very likely to match the same window in the right image. So, the right patch image patch doesn’t have a unique match.

The ordering constraint

The order of corresponding points along a pair of epipolar lines should be the same.

Note: As seen in previous slides, doesn’t always hold, but usually does and can vastly speed up computation.

Window size

- Effect of window size

Better results with adaptive window


Window Shape and Forshortening
Lighting Conditions (Photometric Variations)

$W(P_L)$

$W(P_R)$

Problem of Occlusion: Half Occluded Regions

Occluded in Right Image

Occluded in Left Image

Left Image

Left Center of Projection

Right Image

Right Center of Projection