Announcements

• HW3 on web page

Stereo

Computer Vision I
CSE252A
Lecture 15

Epipolar Geometry

P, p, p', O, O' are coplanar

• Epipolar Plane
• Epipolar Lines
• Epipoles
• Baseline

Family of epipolar Planes

Family of planes π and lines l and l'
Intersection in e and e'

Epipolar Constraint: Calibrated Cameras

O, O', p, p' are coplanar

\[ \overrightarrow{p} \cdot (\overrightarrow{O} \times \overrightarrow{O'}) = 0 \]

\[ \overrightarrow{p} \cdot \left[ \overrightarrow{t} \times (\overrightarrow{R} \overrightarrow{p}) \right] = 0 \]

\[ p = (u, v, 1)^T \]

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

\[ M = \left( \begin{array}{ccc} 1 & 0 & 0 \\ 0 & -R^T & 0 \\ 0 & 0 & 1 \end{array} \right) \]

\[ M' = \left( \begin{array}{ccc} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{array} \right) \]

Properties of the Essential Matrix

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

• \( E \) p' is the epipolar line associated with p'.
• \( E'p \) is the epipolar line associated with p.
• \( E \) is singular
• \( e'=0 \) and \( E'e=0 \).
• \( E \) 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++

\[
\begin{bmatrix}
(w^l, u^l, 1) \\
(w^r, u^r, 1)
\end{bmatrix} \begin{bmatrix}
F_{11} & F_{12} & F_{13} \\
F_{21} & F_{22} & F_{23} \\
F_{31} & F_{32} & F_{33}
\end{bmatrix} \begin{bmatrix}
1 \\
w^l' \\
w^r'
\end{bmatrix} = 0
\]

Set \( F_{33} \) to 1

Minimize:

\[
\sum_{i=1}^{n} (p_i - F_i p_i')^2
\]

under the constraint

\( |F| = 1 \).

Example: converging cameras

Rectification

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

Input Images

Rectified Images

See Trucco & Verri 7.3.7 or Forsyth & Ponce 11.1.1 for more details

Multiple Interpretations

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

Correspondence: Photometric constraint

- Same world point has same intensity in both images (Constant Brightness Constraint)
  - Lambertian fronto-parallel
  - Issues:
    - Noise
    - Specularity
    - Foreshortening
Using epipolar & constant Brightness constraints for stereo matching

For each epipolar line
- compare with every pixel on same epipolar line in right image
- pick pixel with minimum match cost
This will never work, so:

Improvement: match windows
(Seitz)

Finding Correspondences

Slide Window to different disparities to find best match

Best Match amounts to minimizing (or maximizing some) Match Metric

Comparing Windows:

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

(Camps)

Correspondence Search Algorithm

For i = 1:rows
for j = 1:ncols
best(i,j) = -1
for k = mindisparity:maxdisparity

c = Match_Metric(I1(i,j),I2(i,j+k),winsize)
if (c > best(i,j))
best(i,j) = c
disparities(i,j) = k
end
end
end

O(rows * ncols * disparities * winx * 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>[ \sum \frac{(I(x,y) - \mu_I)(g(x,y) - \mu_G)}{\sigma_I \sigma_G} ]</td>
</tr>
<tr>
<td>Type of Squared Differences (SSD)</td>
<td>[ \sum (f(x,y) - g(x,y))^2 ]</td>
</tr>
<tr>
<td>Normalized SSD</td>
<td>[ \sum \frac{(f(x,y) - g(x,y))^2}{\sigma_f^2 + \sigma_g^2} ]</td>
</tr>
<tr>
<td>Type of Absolute Differences (SAD)</td>
<td>[ \sum</td>
</tr>
<tr>
<td>Zero Mean SAD</td>
<td>[ \sum</td>
</tr>
<tr>
<td>Rank</td>
<td>[ \sum \frac{1}{</td>
</tr>
<tr>
<td>Census</td>
<td>[ \sum w(x,y) ]</td>
</tr>
</tbody>
</table>

These two are actually the same
Stereo results
– Data from University of Tsukuba

Results with window correlation
Window-based matching
(best window size)

Results with better method
State of the art method

Some Issues
• Window size
• Window shape
• Lighting
• Ambiguity
• Half occluded regions

Window size
• Effect of window size

Window Shape and Forshortening
Window Shape: Fronto-parallel Configuration

Lighting Conditions (Photometric Variations)

Ambiguity

Problem of Occlusion

Stereo Constraints

Stereo matching

<table>
<thead>
<tr>
<th>CONSTRAINT</th>
<th>BRIEF DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-D Epipolar Search</td>
<td>Arbitrary images of the same scene may be matched based on epipolar geometry even if they mismatch in a 2D or 3D vision scene. This reduces the computational complexity and also reduces the likelihood of false matches.</td>
</tr>
<tr>
<td>Monotonic Ordering</td>
<td>Points along an epipolar scanline appear in the same order in both stereo images, assuming that all objects in the scene are approximately the same distance from the cameras.</td>
</tr>
<tr>
<td>Image Brightness Constancy</td>
<td>Assuming Lambertian surfaces, the brightness of corresponding points in stereo images are the same.</td>
</tr>
<tr>
<td>Match Uniqueness</td>
<td>For every point in one stereo image, there is at most one corresponding point in the other image.</td>
</tr>
<tr>
<td>Disparity Continuity</td>
<td>Disparity varies smoothly (i.e., depth varies smoothly) in most regions of the image. This assumption is broken at object boundaries.</td>
</tr>
<tr>
<td>Disparity Limit</td>
<td>The search space may be reduced significantly by limiting the disparity range, reducing both computational complexity and the likelihood of false matches.</td>
</tr>
<tr>
<td>Fronto-Parallel Surfaces</td>
<td>The implicit assumption made by area-based matching is that objects have fronto-parallel surfaces (i.e., depth is constant within the region of local support). This assumption is broken at object boundaries.</td>
</tr>
<tr>
<td>Feature Similarity</td>
<td>Corresponding features must be similar in terms of intensity, texture, and orientation.</td>
</tr>
<tr>
<td>Structural Grouping</td>
<td>Corresponding feature groupings and their connectivity must be consistent.</td>
</tr>
</tbody>
</table>

Constraints
- epipolar
- ordering
- uniqueness
- disparity limit
- disparity gradient limit

Trade-off
- Matching cost (data)
- Discontinuities (prior)

Similarity measure (SSD or NCC)

Optimal path (dynamic programming)

(From Pollefeys)

(From G. Hager)
Dynamic Programming

- Efficient algorithm for solving sequential decision (optimal path) problems. Cost associated with each arc.

\[
\begin{array}{c|c|c|c}
& i=1 & i=2 & i=3 \\
\hline
j=1 & 1 & 1 & 1 \\
\hline
j=2 & 2 & 2 & 2 \\
\hline
j=3 & 3 & 3 & 3 \\
\hline
\end{array}
\]

Using Dynamic Programming, can find optimal path in \(O(MT)\) time (here \(M=3\)).

For Stereo, \(t\) can denote pixel coordinates across an epipolar line in one image. \(i\) can denote the disparity to the other epipolar line.

\[
\begin{align*}
\Pi_i &= \text{Cost of going from state } i \text{ to state } j \\
\text{Suppose cost can be decomposed into stages:} \\
\Pi_j &= \text{Cost of going from state } i \text{ to state } j \\
\end{align*}
\]
Dynamic Programming

$$C_i(j) = \min \{C_{i-1} + C_{i-1}(i)\}$$

$$b_i(j) = \arg \min \{C_{i-1} + C_{i-1}(i)\}$$

$$b_t(j) \text{ gives previous state along minimum cost path}$$

So, $$b_t(3) = 3$$

$$b_t(2) = 2$$

$$b_t(1) = 2$$

Min cost path

1. Iteratively, compute minimum cost to reach all nodes
2. Recursively, starting with the node at time t-max, select lowest cost terminal node, and backtrack along path

Compute Optimal Path Costs

```
A Maximum Likelihood Stereo Algorithm', Cox, Hingorani, Rao, Maggs, Computer Vision & Image Understanding, 63, 3, pp. 542-567.
```

Back tracking to get optimal path

Code example:

```c
// Compute optimal path costs

int ComputeOptimalPathCosts(int i, int j, int C[i][j], int M[i][j])

// Dynamic Programming algorithm
```
Stereo Matching with Dynamic Programming

C(i,j) is minimum of
1. C(i-1,j-1) + match-cost of pixel L(i) & R(i)
2. C(i-1,j) + occlusion-penalty
3. C(i,j-1) + occlusion-penalty

Scan across grid computing optimal cost for each node given its upper-left neighbors.

Once C(i,j) is completely calculated:
Backtrack from the terminal to get the optimal path.
Some Challenges & Problems

- Photometric issues:
  - specularities
  - strongly non-Lambertian BRDF’s

- Surface structure
  - lack of texture
  - repeating texture within horopter bracket

- Geometric ambiguities
  - as surfaces turn away, difficult to get accurate reconstruction
    (affine approximate can help)
  - at the occluding contour, likelihood of good match but incorrect
    reconstruction

Many variations

- Subpixel interpolation
- Probabilistic framework
- Creases
- Occlusion penalties

Variations on Binocular Stereo

1. Trinocular Stereopsis
2. Helmholtz Reciprocity Stereopsis

Trinocular Epipolar Constraints

\[
\begin{align*}
 p_1^T e_{313} n_2 &= 0 \\
p_1^T e_{313} n_3 &= 0 \\
p_1^T e_{313} n_4 &= 0 \\
e_1^T e_{323} e_{12} &= e_1^T e_{323} e_{13} = e_1^T e_{323} e_{21} = 0
\end{align*}
\]

These constraints are not independent!