Stereo II

Computer Vision I
CSE252A
Lecture 15

Announcements

• HW3 has been posted
• Makeup lecture: Thursday, some time after 5:00. Waiting to find a room. Check piazza, hopefully later today.

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

Mars Exploratory Rovers: Spirit and Opportunity

BINOCULAR STEREO SYSTEM
Estimating Depth

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

• Linear Method: find $P$ such that
  \[
  \begin{align*}
  p & = MP \\
  p' & = M'P
  \end{align*}
  \]
  \[
  \Leftrightarrow \begin{bmatrix} p \mid M \end{bmatrix} P = 0
  \]

• Non-Linear Method: find $Q$ minimizing
  \[
  \min Q \left( \frac{d(p, q) + d(p', q')}{2} \right)
  \]
  where $q = MQ$ and $q' = M'Q$
Two Approaches

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

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

Random Dot Stereograms

Need for correspondence

• 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 $l$.
Skew Symmetric Matrix & Cross Product

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

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

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

Epipolar Constraint: Calibrated Case

$$
\bar{r} \cdot (\bar{a} \times \bar{b}) = 0
$$

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

$$
\begin{bmatrix}
F_{11} & F_{12} & F_{13} \\
F_{21} & F_{22} & F_{23} \\
F_{31} & F_{32} & F_{33}
\end{bmatrix}
$$

Here, $F$ is the Essential Matrix.

- Consider 8 points $(u_i, v_i), (u_i', v_i')$
- Set $F_{33}$ to 1

Calibration

Determine intrinsic parameters and extrinsic relation of two cameras.

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

- Alternatively, view this as a system of homogeneous equations in $F_{13}$ to $F_{33}$
- Solve as Eigenvector corresponding to the smallest eigenvalue of matrix created from the image data.

Equivalent to solving

$$
\sum_{i=1}^{n} (p'_i \mathcal{F} p_i)^2
$$

Minimize under the constraint $\mathcal{F}^T \mathcal{F} = 1$.
Epipolar geometry example

Properties of the Essential Matrix

- $Ep'$ is the epipolar line associated with $p'$.
- $E^Tp$ is the epipolar line associated with $p$.
- $Ee'=0$ and $E^Te=0$.
- $E$ is singular.
- $E$ has two equal non-zero singular values (Huang and Faugeras, 1989).

The Fundamental Matrix

The epipolar constraint is given by: $p^TEp'=0$ with $E = \mathcal{F}$ 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:

$$(Aq)^TE(A'q') = q^T(A^TEA')q = q^TFq' = 0$$

where $F$ is called the Fundamental Matrix.

Can estimate $F$ using 8 point algorithm WITHOUT CALIBRATION.

Example: converging cameras

Example: motion parallel with image plane

Example: forward motion

Example: forward motion
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

$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = H \cdot e$
map epipole $e$ to $(1,0,0)$
try to minimize image distortion
Note that rectified images usually not rectangular

Rectification

Mobi: Stereo-based navigation

Epipolar correspondence

Symbolic Map
Multiple Interpretations

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

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

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

Comparing Windows:

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

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

(Camps)

Correspondence Search Algorithm

```
For i = 1:nrows
  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
end
```

O(nrows * 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>$\frac{\sum_{(x,y) \in R} (f(x,y) - \mu_f)(g(x,y) - \mu_g)}{\sigma_f \sigma_g}$</td>
</tr>
<tr>
<td>Sum of Squared Differences (SSD)</td>
<td>$\sum_{(i,j) \in R} (f(i,j) - g(i,j))^2$</td>
</tr>
<tr>
<td>Normalized SSD</td>
<td>$\frac{\sum_{(i,j) \in R} (f(i,j) - \mu_f)(g(i,j) - \mu_g)}{\sigma_f \sigma_g}$</td>
</tr>
<tr>
<td>Sum of Absolute Differences (SAD)</td>
<td>$\sum_{(i,j) \in R}</td>
</tr>
<tr>
<td>Zero Mean SAD</td>
<td>$\sum_{(i,j) \in R}</td>
</tr>
<tr>
<td>Census</td>
<td>$\sum_{(i,j) \in R} \text{match}(f(i,j), g(i,j))$</td>
</tr>
</tbody>
</table>

These two are actually the same
Stereo results

Data from University of Tsukuba

Scene

Ground truth

Results with window correlation

Window-based matching (best window size)

Ground truth

(Seitz)