Stereo
(Part 2)

Introduction to Computer Vision
CSE 152
Lecture 9
Announcements

• Homework 3 is due May 9, 11:59 PM
• Reading:
  – Chapter 7: Stereopsis
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
Rectification

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

Under perspective projection, the mapping from a plane to a plane is given by a 2D projective transformation (homography)

\[
\begin{bmatrix}
x_L \\
y_L \\
w_L
\end{bmatrix} = H_L
\begin{bmatrix}
u_L \\
v_L \\
1
\end{bmatrix}
\]
Rectification

Under perspective projection, the mapping from a plane to a plane is given by a 2D projective transformation (homography)

Two images – Two projective transformations
Epipolar Rectification

- Create pair of virtual cameras
  - Virtual cameras have the same camera centers as real cameras
  - Both virtual cameras have the same:
    - Camera rotation matrix $R$
    - Camera calibration matrix $K$
- Rectification transformation matrices
  \[ H = K_{\text{virtual}} R_{\text{virtual}} R_{\text{real}}^T K_{\text{real}}^{-1} \]
Image pair rectification

Simplify stereo matching by warping the images

Apply projective transformation so that epipolar lines correspond to horizontal scanlines

\[ H \]

\[ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = He \]

H should map epipole e to (1,0,0), a point at infinity on the x-axis

H should minimize image distortion

Note that rectified images are usually not rectangular

See textbook for complete method
Rectification
Given a pair of images, transform both images so that epipolar lines are scan lines.

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

Rectified Images
See Section 7.2.1 for specific method

epipolar lines run parallel with the \( x \)-axis and are aligned between two views (no \( y \) disparity)
Rectification

Original

Rectified
Rectification

• Epipolar lines
Rectification
Polar Rectification

Homography-based Rectification

Polar Rectification

Alternative epipolar rectification method that minimizes pixel distortion
Polar Rectification

Epipoles are in images (white dot on ball)

Homography-based rectification is not possible
Features on same epipolar line
Mobi: Stereo-based navigation
Epipolar correspondence
Symbolic Map
Using epipolar & constant Brightness constraints for stereo matching

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

(Seitz)
Finding Correspondences

$W(p_1)$

$W(p_r)$
Comparing Windows:

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

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

Most popular

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</td>
<td>$\sum_{u,v} \frac{(I_1(u,v) - I_1)(I_2(u + d, v) - I_2)}{\sqrt{\sum_{u,v} (I_1(u,v) - I_1)^2 \cdot \sum_{u,v} (I_2(u + d, v) - I_2)^2}}$</td>
</tr>
<tr>
<td>Sum of Squared Differences</td>
<td>$\sum_{u,v} (I_1(u,v) - I_2(u + d, v))^2$</td>
</tr>
<tr>
<td>Normalized SSD</td>
<td>$\sum_{u,v} \left( \frac{(I_1(u,v) - I_1)}{\sqrt{\sum_{u,v} (I_1(u,v) - I_1)^2}} - \frac{(I_2(u + d, v) - I_2)}{\sqrt{\sum_{u,v} (I_2(u + d, v) - I_2)^2}} \right)^2$</td>
</tr>
<tr>
<td>Sum of Absolute Differences</td>
<td>$\sum_{u,v}</td>
</tr>
<tr>
<td>Zero Mean SAD</td>
<td>$\sum_{u,v}</td>
</tr>
<tr>
<td>Rank</td>
<td>$I'<em>k(u,v) = \sum</em>{m,n} I_k(m,n) &lt; I_k(u,v) \quad \sum_{u,v} (I'_1(u,v) - I'_2(u + d, v))$</td>
</tr>
<tr>
<td>Census</td>
<td>$I'<em>k(u,v) = BITSTRING</em>{m,n}(I_k(m,n) &lt; I_k(u,v)) \quad \sum_{u,v} HAMMING(I'_1(u,v), I'_2(u + d, v))$</td>
</tr>
</tbody>
</table>

These two are actually the same
Stereo results

– Data from University of Tsukuba

(Scene) (Seitz)
Results with window correlation

Window-based matching
(best window size)

Ground truth
(Seitz)
Results with better method

Using global optimization

Boykov et al., Fast Approximate Energy Minimization via Graph Cuts,
International Conference on Computer Vision, September 1999.

(Seitz)

Ground truth
State of the Art Results

Using neural networks

[Sparse stereo disparity map densification using hierarchical image segmentation](https://example.com), 13th International Symposium on Mathematical Morphology.

Ground truth
Some Issues

- Epipolar ordering
- Ambiguity
- Window size
- Window shape
- Lighting
- Half occluded regions
A challenge: Multiple Interpretations

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

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

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

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

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

• Epipolar ordering
• Ambiguity
• Window size
• Window shape
• Lighting
• Half occluded regions
Window size

- Effect of window size

Better results with *adaptive window*


(Seitz)
Some Issues

- Epipolar ordering
- Ambiguity
- Window size
- Window shape
- Lighting
- Half occluded regions
Window Shape and Forshortening
Window Shape: Fronto-parallel Configuration
Some Issues

- Epipolar ordering
- Window size
- Ambiguity
- Window shape
- Lighting
- Half occluded regions
Lighting Conditions (Photometric Variations)

\[ W(P_1) \]

\[ W(P_r) \]
Some Issues

• Epipolar ordering
• Ambiguity
• Window size
• Window shape
• Lighting
• Half occluded regions
Half occluded regions
## Summary of Stereo Constraints

<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 rectified based on epipolar geometry such that stereo matches lie along one-dimensional scanlines. 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>Disparities vary smoothly (i.e. disparity gradient is small) over most of the image. This assumption is violated 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 violated by sloping and creased surfaces.</td>
</tr>
<tr>
<td>Feature Similarity</td>
<td>Corresponding features must be similar (e.g. edges must have roughly the same length and orientation).</td>
</tr>
<tr>
<td>Structural Grouping</td>
<td>Corresponding feature groupings and their connectivity must be consistent.</td>
</tr>
</tbody>
</table>
Next Lecture

• Early vision: multiple images
  – Structure from motion

• Reading:
  – Chapter 8: Structure from Motion