Announcements

• HW3 on web page

Formula for Finding Corners

Let \( I_x = \frac{\partial I}{\partial x} \) and \( I_y = \frac{\partial I}{\partial y} \)

\[
C = \left[ \begin{array}{cc}
\sum I_x^2 & \sum I_x I_y \\
\sum I_x I_y & \sum I_y^2 
\end{array} \right]
\]

Sum over a small region, the hypothetical corner

Gradient with respect to \( x \), times gradient with respect to \( y \)

Matrix is symmetric

WHY THIS?

So, to detect corners

• Filter image.
• Compute the gradient everywhere.
• We construct \( C \) in a window of some size.
• Use linear algebra to find \( \lambda_1 \) and \( \lambda_2 \).
• If \( \lambda_1 \) and \( \lambda_2 \) are both big, we have a corner.
  1. Let \( e(u,v) = \min(\lambda_1(u,v), \lambda_2(u,v)) \)
  2. \((u,v)\) is a corner local maximum of \( e(u,v) \) and \( e(u,v) > \tau \)

Need for correspondence

\[
X = \frac{d X_L}{(X_L - X_R)}
\]

\[
Z = \frac{d f}{(X_L - X_R)}
\]

\[
Z = \frac{(f X_L) X - (f X_R) (X - d)}{(X_L - X_R)}
\]

\[
X = \frac{d X_L}{(X_L - X_R)}
\]

\[
Z = \frac{d f}{(X_L - X_R)}
\]

BINOCULAR STEREO SYSTEM

Estimating Depth

<table>
<thead>
<tr>
<th>DISPARITY</th>
<th>( X - X_L )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Z = \frac{(f X_L) X - (f X_R) (X - d)}{(X_L - X_R)} )</td>
<td></td>
</tr>
<tr>
<td>( X = \frac{d X_L}{(X_L - X_R)} )</td>
<td></td>
</tr>
<tr>
<td>( Z = \frac{d f}{(X_L - X_R)} )</td>
<td></td>
</tr>
</tbody>
</table>

(Adapted from Hager)
Reconstruction: General 3-D case
Given two image measurements \( p \) and \( p' \), estimate \( P \).

- Linear Method:
  find \( P \) such that
  \[
  \begin{align*}
  \begin{bmatrix} p \times M \end{bmatrix} P &= 0 \\
  \begin{bmatrix} p' \times M \end{bmatrix} P &= 0
  \end{align*}
  \implies \begin{bmatrix} p_p \end{bmatrix} M \end{bmatrix} P = 0
  \]

- Non-Linear Method: find \( Q \) minimizing
  \[
  \delta(p, q) + \delta(p', q')
  \]
  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

Epipolar Geometry
\( P, p, p', O, O' \) are coplanar

Family of epipolar Planes
(standard approach)
Family of epipolar Planes

Family of planes $\pi$ and lines $l$ and $l'$
Intersection in $e$ and $e'$

Epipolar Constraint: Calibrated Cameras

Given that the right camera’s position and orientation relative to the left is given by a translation vector $t$ and rotation matrix $R$, what is the relation between the image of $P$ in the left image ($p$) and image of $P$ in the right image ($p'$)?

We’ll use the fact that $O$, $O'$, $p$, $p'$ are coplanar when $p$ & $p'$ are the projection of a common point.

Epipolar Constraint: Calibrated Cameras

Properties of the Essential Matrix

- $E p'$ is the epipolar line associated with $p'$.
- $E t p$ 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).

Calibration

Determine intrinsic parameters and extrinsic relation of two cameras

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

Much more on multi-view in CSE252B

Set $F_{13}$ to 1

Minimize:

$$\sum (p'_i F p_i)^2$$

under the constraint $|F|^2 = 1$. 

$$F = \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix}$$
**Example:** converging cameras
courtesy of Andrew Zisserman

**Example:** motion parallel with image plane
(simple for stereo → rectification)
courtesy of Andrew Zisserman

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

**Example:** forward motion
courtesy of Andrew Zisserman

**Rectification**
All epipolar lines are parallel in the rectified image plane.
**Image pair rectification**

Simplify stereo matching by warping the images.

Apply projective transformation so that epipolar lines correspond to horizontal scanlines.

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

Map epipole \( e \) to \((1,0,0)\) (a point at infinity).

Try to minimize image distortion.

Note that rectified images usually not rectangular.

---

**Rectification**

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

**Input Images**

---

**Features on same epipolar line**

**Rectified Images**

See Section 7.3.7 for specific method.

---

**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.

Multiple Interpretations

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

Multiple Interpretations

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:
  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:

Improvement: match windows

(Camps)

Comparing Windows: $f$ vs. $g$

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

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

(Camps)

Correspondence Search Algorithm (simple version for Cross Correlation)

For $i = 1$ to $nrows$
  for $j = 1$ to $ncols$
    best$(i,j) = -1$
    for $k = \text{mindisparity}$ to $\text{maxdisparity}$
      $c = \text{CC}(I1(i,j), I2(i,j+k), \text{winsize})$
      if ($c > \text{best}(i,j)$)
        best$(i,j) = c$
        disparities$(i,j) = k$
      end
    end
  end
end

$O(nrows \times ncols \times \text{disparities} \times \text{winx} \times \text{winy})$

Window size

- Effect of window size
  - Better results with adaptive window

(Match Metric Summary)

<table>
<thead>
<tr>
<th>MATCH METRIC</th>
<th>COMPUTATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalized Cross-Correlation (NCC)</td>
<td>$\frac{\sum_{x}I1(x)I2(x)}{\sqrt{\sum_{x}I1^2(x)\sum_{x}I2^2(x)}}$</td>
</tr>
<tr>
<td>Sum of Squared Differences (SSD)</td>
<td>$\sum_{x}I1(x) - I2(x)^2$</td>
</tr>
<tr>
<td>Normalized SSD</td>
<td>$\frac{\sum_{x}I1(x) - I2(x)^2}{\sqrt{\sum_{x}I1^2(x)\sum_{x}I2^2(x)}}$</td>
</tr>
<tr>
<td>Sum of Absolute Differences (SAD)</td>
<td>$\sum_{x}</td>
</tr>
<tr>
<td>Zero Mean SAD</td>
<td>$\sum_{x}</td>
</tr>
<tr>
<td>Rank</td>
<td>$\left</td>
</tr>
<tr>
<td>Census</td>
<td>$\sum_{x}I1(x)I2(x) - \sum_{x}I1(x) - \sum_{x}I2(x)$</td>
</tr>
</tbody>
</table>

These two are actually the same

Effect of window size (Seitz)

Better results with adaptive window


Effect of window size (Seitz)

Better results with adaptive window

Ambiguity

Lighting Conditions (Photometric Variations)

Window Shape and Forshortening

Window Shape: Fronto-parallel Configuration

Stereo results

– Data from University of Tsukuba

Results with window correlation
Results with better method

State of the art method

Ground truth


Ground truth (Seitz)

Problem of Occlusion

Ordering Constraint

Dynamic Programming (Ohta and Kanade, 1985)

Stereo matching

- Similarity measure: SSD or NCC
- Constraints:
  - epipolar
  - ordering
  - uniqueness
  - disparity limit
  - disparity gradient limit

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

(From Pollefeys)

Stereo Matching with Dynamic Programming

Dynamic programming yields the optimal path through grid. This is the best set of matches that satisfy the ordering constraint.

Every pixel on each scanline will be labeled as matching, or occluded.

Dynamic Programming

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

\[
\begin{align*}
C(i, j) = 1 \\
\end{align*}
\]

How many paths through this trellis? \(3^T\)

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

Dynamic Programming

- Used with Hidden Markov Models, Viterbi Algorithm

\[
\begin{align*}
\Pi_j &= \text{Cost of going from state } i \text{ to state } j \\
\end{align*}
\]

Suppose cost can be decomposed into stages:

Principle of Optimality for an n-stage assignment problem:

\[
C(i, j) = \min \left[ \Pi_j + C_{j+1}(i) \right]
\]

Minimum cost of path from \(t=0\) to reach state \(j\) at time \(t\).

(Slides adapted from Jim Rehg at GA Tech)
**Dynamic Programming**

\[ C_i(j) = \min \{ \Pi_j + C_{i-1}(i) \} \]
\[ b_j(i) = \arg \min \{ \Pi_j + C_{i-1}(i) \} \]

\( b(j) \) gives previous state along minimum cost path

---

**Computer Optimal Path Costs**

\[ \text{Occlusion} = \left[ \begin{array}{c} b_1 \left( \frac{c_1}{f_{ij}^y f_{ij}^x} \right) \\ \vdots \\ b_n \left( \frac{c_n}{f_{ij}^y f_{ij}^x} \right) \end{array} \right] \]

For \( i = 1 \) to \( n \):
\[ \text{Occlusion}_i = b_i \]
\[ \text{Occlusion} = \text{argmin} \{ \text{Occlusion} \} \]

---

**Back tracking to get optimal path**

```
p=1;
qu=1;
while(p<q)
do q=q+1[
    p scan(p,q); 
    if scan(p,q) then 
        \text{case 1:}
        \text{if scan(p,q) then}
            \text{if scan(p,q) then}
                \text{break;}
        \text{case 2:}
            \text{q scan(q,p);}
            \text{break;}
    \text{case 3:}
        \text{q scan(q,p);}
        \text{break;}
]
```

---

**Stereo Matching with Dynamic Programming**

**Scan across grid computing optimal cost for each node \( c(i,j) \) given its upper-left neighbors.**
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