Stereo

Single image stereogram, by Niklas Een
Stereo

• Given two images from different viewpoints
  – How can we compute the depth of each point in the image?
  – Based on *how much each pixel moves* between the two images
Epipolar geometry

Two images captured by a purely horizontal translating camera (rectified stereo pair)

\[ x_2 - x_1 = \text{the disparity of pixel } (x_1, y_1) \]
Stereo matching algorithms

Match Pixels in Conjugate Epipolar Lines

• Assume brightness constancy
• This is a tough problem
• Numerous approaches
  – A good survey and evaluation: http://www.middlebury.edu/stereo/
Your basic stereo algorithm

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

Improvement: match **windows**
Window size

Effect of window size

- Smaller window
  +
  -
- Larger window
  +
  -

Better results with adaptive window

Stereo results

- Data from University of Tsukuba
- Similar results on other images without ground truth
Results with window search

Window-based matching (best window size)  Ground truth
Better methods exist...

State of the art method

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

For the latest and greatest: http://www.middlebury.edu/stereo/
Stereo as energy minimization

What defines a good stereo correspondence?
1. Match quality
   - Want each pixel to find a good match in the other image
2. Smoothness
   - If two pixels are adjacent, they should (usually) move about the same amount
Stereo as energy minimization

• Find disparity map $d$ that minimizes an energy function $E(d)$

• Simple pixel / window matching

$$E(d) = \sum_{(x,y) \in I} C(x, y, d(x, y))$$

$$C(x, y, d(x, y)) = \text{SSD distance between windows } I(x, y) \text{ and } J(x + d(x,y), y)$$
Stereo as energy minimization

\[ I(x, y) \]

\[ J(x, y) \]

\[ y = 141 \]

\[ C(x, y, d); \text{the disparity space image (DSI)} \]
Stereo as energy minimization

Simple pixel / window matching: choose the minimum of each column in the DSI independently:

\[ d(x, y) = \arg \min_{d'} C(x, y, d') \]
Stereo as energy minimization

Better objective function

\[ E(d) = E_d(d) + \lambda E_s(d) \]

- Match cost
- Smoothness cost

Want each pixel to find a good match in the other image

Adjacent pixels should (usually) move about the same amount
Stereo as energy minimization

\[ E(d) = E_d(d) + \lambda E_s(d) \]

match cost: \[ E_d(d) = \sum_{(x,y) \in I} C(x, y, d(x, y)) \]

smoothness cost: \[ E_s(d) = \sum_{(p,q) \in \mathcal{E}} V(d_p, d_q) \]

\[ \mathcal{E} : \text{set of neighboring pixels} \]

4-connected neighborhood \hspace{1cm} 8-connected neighborhood
Smoothness cost

\[ E_s(d) = \sum_{(p,q) \in \mathcal{E}} V(d_p, d_q) \]

How do we choose \( V \)?

\[ V(d_p, d_q) = |d_p - d_q| \]

\( L_1 \) distance

\[ V(d_p, d_q) = \begin{cases} 
0 & \text{if } d_p = d_q \\
1 & \text{if } d_p \neq d_q 
\end{cases} \]

“Potts model”
Dynamic programming

\[ E(d) = E_d(d) + \lambda E_s(d) \]

Can minimize this independently per scanline using dynamic programming (DP)

\[ D(x, y, d) : \text{minimum cost of solution such that } d(x,y) = d \]

\[ D(x, y, d) = C(x, y, d) + \min_{d'} \{ D(x - 1, y, d') + \lambda |d - d'| \} \]
Dynamic programming

Finds “smooth” path through DPI from left to right
Dynamic Programming
Dynamic programming

Can we apply this trick in 2D as well?

No: \( d_{x-1,y-1} \) and \( d_{x-1,y} \) and \( d_{x,y-1} \) may depend on different values of \( d_{x-1,y-1} \).
Stereo as a minimization problem

\[ E(d) = E_d(d) + \lambda E_s(d) \]

The 2D problem has many local minima
- Gradient descent doesn’t work well

And a large search space
- \( n \times m \) image w/ \( k \) disparities has \( k^{nm} \) possible solutions
- Finding the global minimum is NP-hard in general

Good approximations exist… we’ll see this soon
Questions?
Depth from disparity

\[ \text{disparity} = x - x' = \frac{\text{baseline} \times f}{z} \]
Real-time stereo

Nomad robot searches for meteorites in Antartica
http://www.frc.ri.cmu.edu/projects/meteorobot/index.html

Used for robot navigation (and other tasks)
• Several software-based real-time stereo techniques have been developed (most based on simple discrete search)
Stereo reconstruction pipeline

Steps

• Calibrate cameras
• Rectify images
• Compute disparity
• Estimate depth

What will cause errors?

• Camera calibration errors
• Poor image resolution
• Occlusions
• Violations of brightness constancy (specular reflections)
• Large motions
• Low-contrast image regions
Active stereo with structured light

- Project "structured" light patterns onto the object
  - simplifies the correspondence problem
Laser scanning

- Project a single stripe of laser light
- Scan it across the surface of the object
- This is a very precise version of structured light scanning

Digital Michelangelo Project
http://graphics.stanford.edu/projects/mich/
Laser scanned models

The Digital Michelangelo Project, Levoy et al.
Laser scanned models

The Digital Michelangelo Project, Levoy et al.
Laser scanned models

The Digital Michelangelo Project, Levoy et al.
Laser scanned models

The Digital Michelangelo Project, Levoy et al.