Stereo

Lecture 16

Copyright Mubarak Shah 2003
Shape From X

• Recovery of 3D (shape) from one or two (2D images).
Shape From X

- Stereo
- Motion
- Shading
- Photometric Stereo
- Texture
- Contours
- Silhouettes
- Defocus

Copyright Mubarak Shah 2003
Shape from Stereo
Stereo

Figure 6.1: Stereo geometry.

\[
\frac{Z + f}{Z} = \frac{x_1 + x_2 + B}{B}, \quad Z = \frac{fB}{x_1 + x_2},
\]

B=Baseline  
f=focal length  
C1 and C2=Camera Centers  
\(x_1 + x_2 = \text{disparity} = d\)  
X1, X2=Image location in left and right cameras
Stereo Pairs and Depth Maps (from Szeliski’s book)
Rectification

(1)

(2)
Correspondence using Search

Criterion function:

\[ Z = \frac{fB}{d} \]
Correlation Based Stereo Methods

- Disparity map can be constructed based on a correlation measure

\[
\text{SSD} = \sum \sum (I_{left} - I_{right})^2 \quad \text{Sum of squares difference}
\]

\[
\text{AD} = \sum \sum |I_{left} - I_{right}| \quad \text{Absolute difference}
\]

\[
\text{CC} = \sum \sum I_{left} I_{right} \quad \text{Cross correlation}
\]

\[
\text{NC} = \frac{\sum \sum (I_{left} I_{right})}{\sqrt{\sum \sum I_{left} I_{right}}} \quad \text{Normalized Correlation}
\]

\[
\text{MC} = \frac{1}{64 \sigma_{left} \sigma_{right}} \sum \sum (I_{left} - \mu_{left})(I_{right} - \mu_{right}) \quad \text{Mutual Correlation}
\]

Alper Yilmaz, Mubarak shah, Fall 2011 UCF
Correlation

• Similarity/Dissimilarity Measures
  – Sum of Squares Difference (SSD)
  – Normalized Correlation
  – Mutual Correlation
  – Mutual information  \[ I(x, y) = \sum \sum p(x, y) \log \frac{p(x, y)}{p_1(x)p_2(y)} \]

• Use
  – Gray levels
  – Laplacian of Gaussian
  – Gradient magnitude
Block Matching

- Can be used for
  - Computing MPEG motion vectors
  - Optical flow
  - Stereo (displacement limited to only x-axis)
  - Image matching
Block Matching

• For each 8X8 block, centered around pixel (x,y) in right image, $B_k$
  – Obtain 16X16 block in left, centered around (x,y), $B_{k-1}$
  – Compute Sum of Squares Differences (SSD) between 8X8 block, $B_k$, and all possible 8X8 blocks in $B_{k-1}$
  – The 8X8 block in $B_{k-1}$ centered around $(x’,y’)$, which gives the least SSD is the match
  – The displacement vector (disparity, optical flow) is given by $u=x-x’; v=y-y’$
Sum of Squares Differences (SSD)

\[(u(x, y), v(x, y)) = \arg \min_{u,v=-4...4} \sum_{i=0}^{-7} \sum_{j=0}^{-7} (f_k(x+i, y+j) - f_{k-1}(x+i+u, y+j+v))^2\]
Minimum Absolute Difference (MAD)

\((u(x, y), v(x, y)) = \arg \min_{u,v=-4\ldots4} \sum_{i=0}^{7-7} \sum_{j=0}^{7-7} |(f_k(x+i, y+j) - f_{k-1}(x+i+u, y+j+v))|\)
Maximum Matching Pixel Count (MPC)

$$T(x, y; u, v) = \begin{cases} 
1 & \text{if } |f_k(x, y) - f_{k-1}(x+u, y+v)| \leq t \\
0 & \text{otherwise}
\end{cases}$$

$$(u(x, y), v(x, y)) = \arg \max_{u, v = -4 \ldots 4} \sum_{i=0}^{-7} \sum_{j=0}^{-7} T(x+i, y+j; u, v)$$
Cross Correlation

\[(u(x, y), v(x, y)) = \arg \max_{u,v=-4...4} \sum_{i=0}^{-7} \sum_{j=0}^{-7} (f_k(x+i, y+j).f_{k-1}(x+i+u, y+j+v))\]
Normalized Correlation

\[
(u,v) = \arg \max_{u,v=-4...4} \left\{ \frac{\sum_{t=0}^{-7} \sum_{j=0}^{-7} ((f_{k}(x+i,y+j)-\mu_1)(f_{k-1}(x+i+u,y+j+v)-\mu_2))}{\sqrt{\left(\sum_{i=0}^{-7} \sum_{j=0}^{-7} (f_{k}(x+i,y+j)-\mu_1)^2\right)\left(\sum_{i=0}^{-7} \sum_{j=0}^{-7} (f_{k-1}(x+i+u,y+j+v)-\mu_2)^2\right)}} \right\}
\]

and \(\mu_2\) are the means of patch-1 and patch-2 respectively.
Mutual Correlation

\[
(u(x, y), v(x, y)) = \arg \max_{u, v = -4 \ldots 4} \frac{1}{64\sigma_1 \sigma_2} \sum_{i=0}^{7} \sum_{j=0}^{7} \left( f_k(x + i, y + j) - \mu_1 \right) \cdot f_{k-1}(x + i + u, y + j + v) - \mu_2 \]

Sigma and mu are standard deviation and mean of patch-1 and patch-2 respectively
Barnard’s Stereo Method

• Similar intensity
  – Similar to brightness constraint
• Smoothness of disparity

\[
E = \sum_{i=-1}^{1} \sum_{j=-1}^{1} \| I_{left}(x+i, y+j) - I_{right}(x+i + D_x(x, y), y+j) \| + \lambda \| \nabla D(x, y) \|
\]

\[
\nabla D(x, y) = \sum_{i=-1}^{1} \sum_{j=-1}^{1} |D(x+i, y+j) - D(x, y) |
\]
Barnard’s Stereo Method

• Energy can be minimized using brute force search
  – Let max allowed disparity is 10 pixels
  – For 128x128 image for 10 possible levels of disparity
    • There $10^{16384}$ possible disparity values
  – We can select any minimization technique
  – Barnard choose simulated annealing
Simulated Annealing

- Select a random state \( S \) (disparities)
- Select a high temperature
  - Select random \( S' \)
  - Compute \( \Delta E = E(S') - E(S) \)
  - If \( \Delta E < 0 \) \( S \leftarrow S' \)
  - Else
    - \( P \leftarrow \exp(-\Delta E/T) \)
    - \( X \leftarrow \text{random}(0,1) \)
      - If \( X < P \) then \( S \leftarrow S' \)
    - If no decrease in several iterations lower \( T \)
Examples
Stereo results
– Data from University of Tsukuba

Scene

Ground truth
Results with window correlation

Window-based matching
(best window size)

Ground truth
Results with better method

State of the art method

Ground truth
Applications of Stereo (from Szeliski’s book)

(a) input image, (b) computed depth map, and (c) new view generation (d) view morphing between two images (e–f) 3D face modeling (g) Virtual reality (h–j) building 3D surface models
Reading Material

• Fundamental of Computer Vision
  – 6.2.1, 6.2.4 and 6.2.5

• Computer Vision: Algorithms and Applications, Richard Szeliski
  – Chapter 11