Stereo

Lecture 16

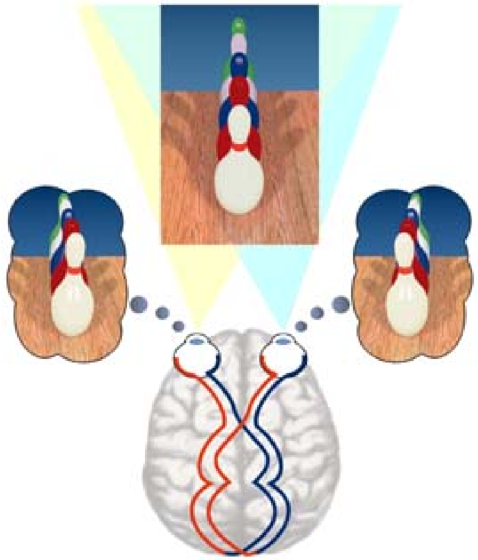
Shape From X

- Recovery of 3D (shape) from one or two (2D images).

Shape From X

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

Shape from Stereo



(a) (b) (c)

Stereo

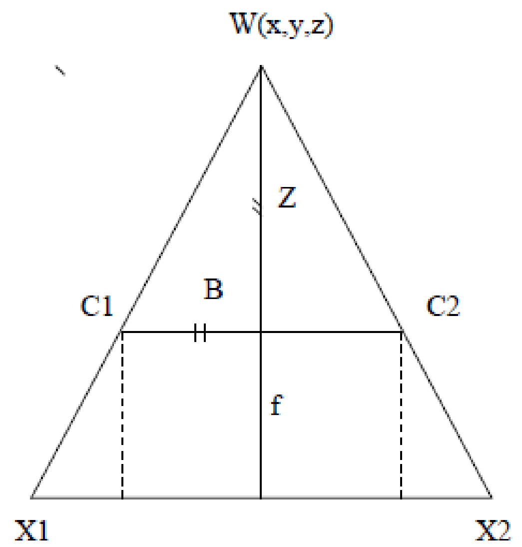


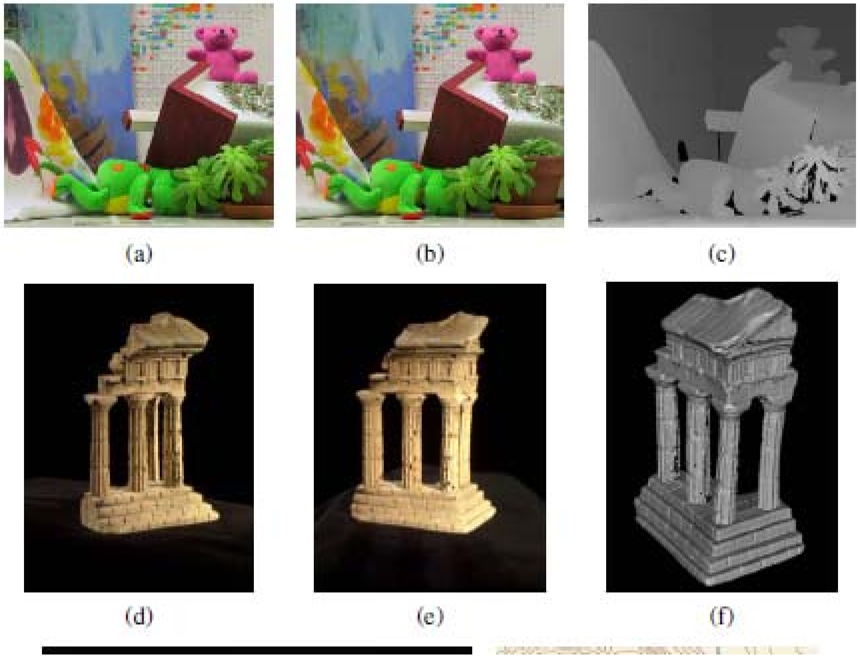
Figure 6.1: Stereo geometry.

$$\frac{Z+f}{Z} = \frac{x_1 + x_2 + B}{B}, \qquad Z = \frac{fB}{x_1 + x_2},$$

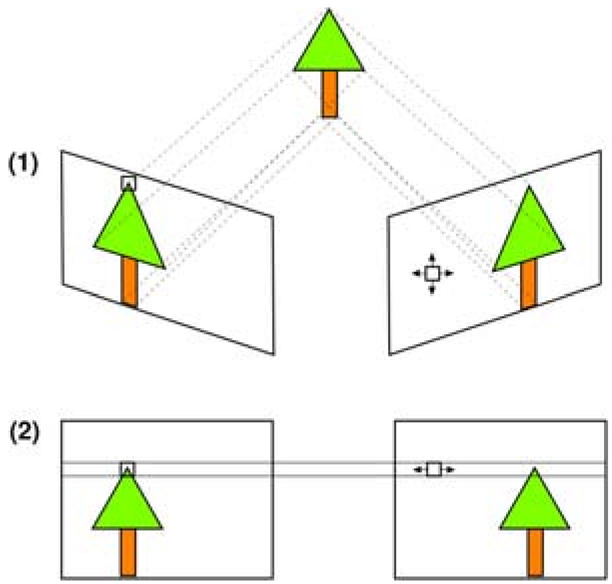
B=Baseline
f=focal length
C1 and C2=Camera Centers
X1, X2=Image location in left and right cameras

$$x_1 + x_2 = \text{disparity} = d$$

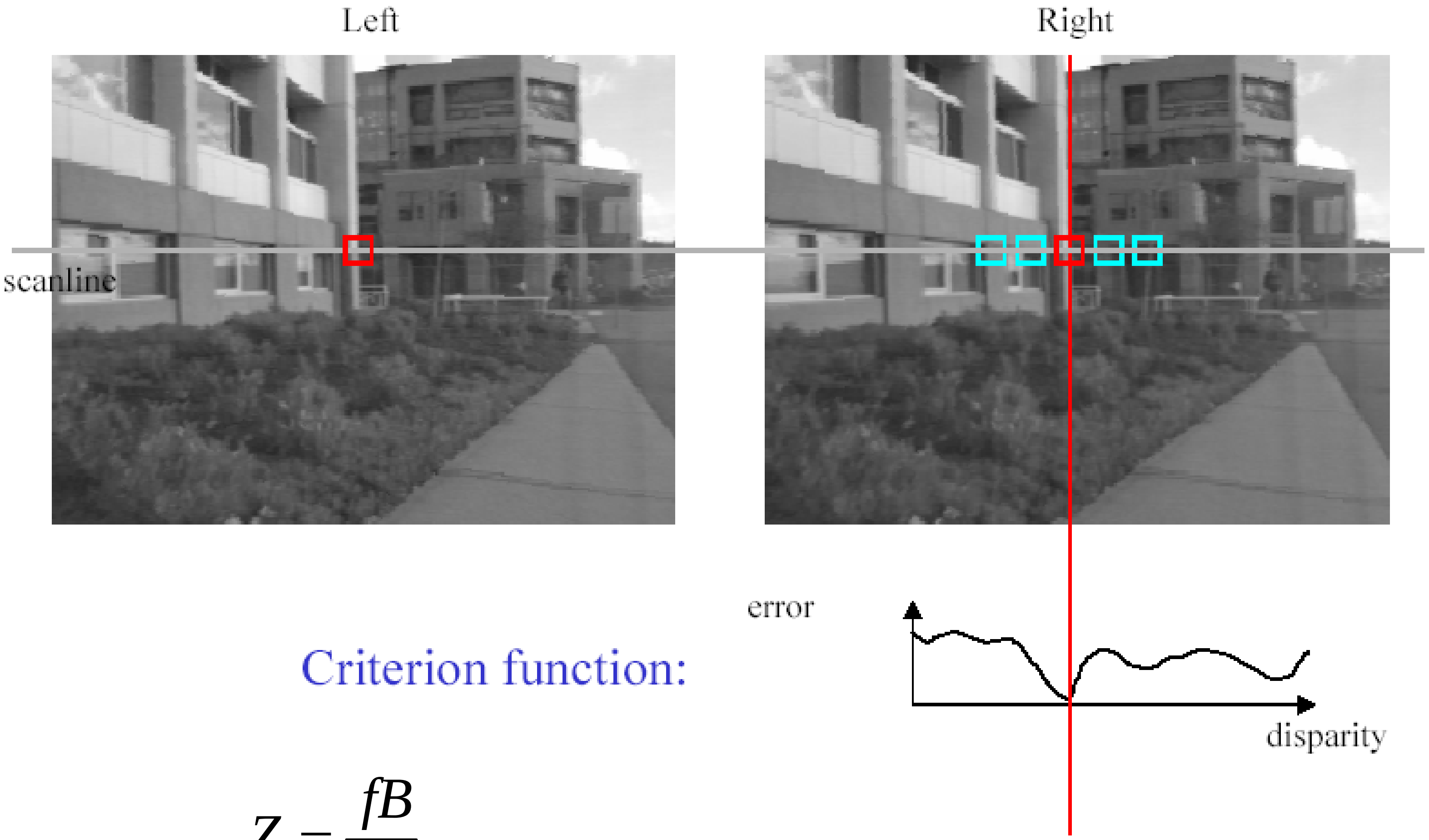
Stereo Pairs and Depth Maps (from Szeliski's book)



Rectification



Correspondence using Search



Criterion function:

$$Z = \frac{fB}{d}$$

Correlation Based Stereo Methods

- Disparity map can be constructed based on a correlation measure

$$SSD = \sum \sum \left(I_{left} - I_{right}\right)^2$$ Sum of squares difference

$$NC = \frac{\sum \sum \left(I_{left}.I_{right}\right)}{\sqrt{\sum \sum I_{left}.I_{right}}}$$

Normalized Correlation

$$AD = \sum \sum \left| \left(I_{left} - I_{right}\right) \right|$$ Absolute difference

$$CC = \sum \sum I_{left} I_{right}$$ Cross correlation

$$MC = \frac{1}{64\sigma_{left}\sigma_{right}} \sum \sum \left(I_{left} - \mu_{left1}\right)\left(I_{right} - \mu_{right}\right)$$

Mutual Correlation

Mean

Standard Deviation

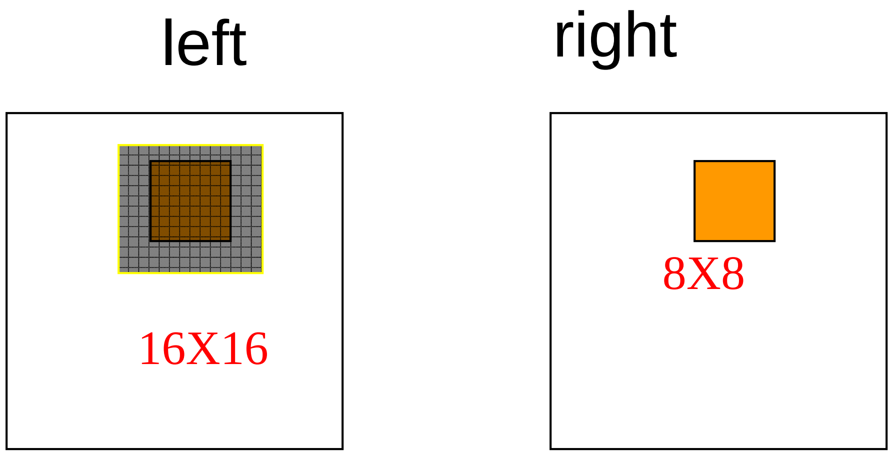
Correlation

- 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)

$$(u(x,y),v(x,y)) = \operatorname{argmin}_{u,v=-4\ldots4} \sum_{i=0}^{-7} \sum_{j=0}^{-7} \left(f_k(x+i, y+j) - f_{k-1}(x+i+u, y+j+v) \right)^2$$

Minimum Absolute Difference (MAD)

$$(u(x,y),v(x,y)) = \arg\min_{u,v=-4\ldots4} \sum_{i=0}^{-7}\sum_{j=0}^{-7}\left|\left(f_k(x+i,y+j) - f_{k-1}(x+i+u,y+j+v)\right)\right|$$

Maximum Matching Pixel Count (MPC)

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

$$(u(x, y), v(x, y)) = \arg\max_{u,v=-4\ldots4} \sum_{i=0}^{-7} \sum_{j=0}^{-7} T(x+i, y+j; u, v)$$

Cross Correlation

$$(u(x,y),v(x,y)) = \arg\max_{u,v=-4\ldots4} \sum_{i=0}^{-7}\sum_{j=0}^{-7}\left(f_k(x+i,y+j).f_{k-1}(x+i+u,y+j+v)\right)$$

Normalized Correlation

$$(u,v) = \arg\max_{u,v=-4\ldots4} \frac{\sum_{i=0}^{-7}\sum_{j=0}^{-7}\left((f_k(x+i,y+j)-\mu_1).(f_{k-1}(x+i+u,y+j+v)-\mu_2)\right)}{\sqrt{\left(\sum_{i=0}^{-7}\sum_{j=0}^{-7}(f_k(x+i,y+j)-\mu_1)^2\right)\left(\sum_{i=0}^{-7}\sum_{j=0}^{-7}(f_{k-1}(x+i+u,y+j+v)-\mu_2)^2\right)}}$$

and μ_2 are the means of patch-1 and patch-2 respectively.

Mutual Correlation

$$(u(x,y),v(x,y)) = \arg\max_{u,v=-4\ldots4} \frac{1}{64\sigma_1\sigma_2} \sum_{i=0}^{-7} \sum_{j=0}^{-7} \left(f_k(x+i,y+j) - \mu_1 \right) . f_{k-1}(x+i+u, y+j+v) - \mu_2 \right)$$

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

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

$$\nabla D(x,y) = \sum_{i=-1}^{1}\sum_{j=-1}^{1}\left| D(x+i, y+j) - D(x,y)\right|$$

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

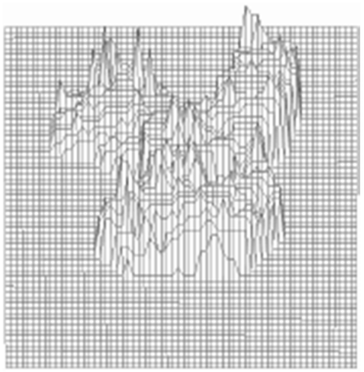
Simulated Annealing

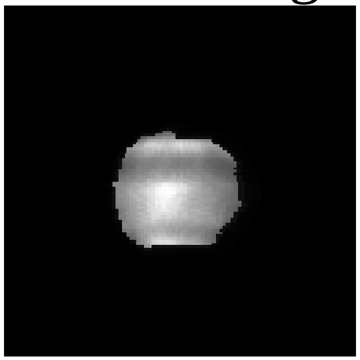
- Select a random state S (disparities)
- Select a high temperature
  - Select random S'
  - Compute $\Delta E = E(S') - E(S)$
  - If ($\Delta E < 0$) $S \leftarrow S'$
  - Else
    - $P \leftarrow \exp(-\Delta E/T)$
    - $X \leftarrow$ random(0,1)
      - If $X < P$ then $S \leftarrow S'$
    - If no decrease in several iterations lower T

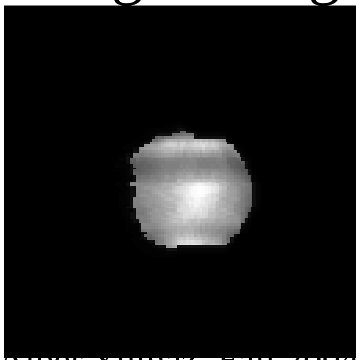
Examples

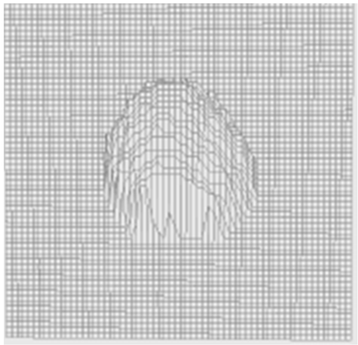

Left Image


Right Image


Depth Map



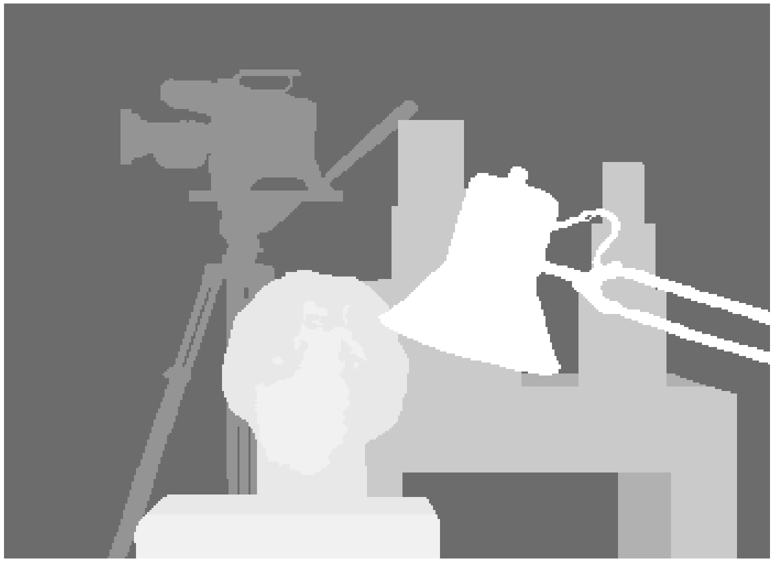




Stereo results

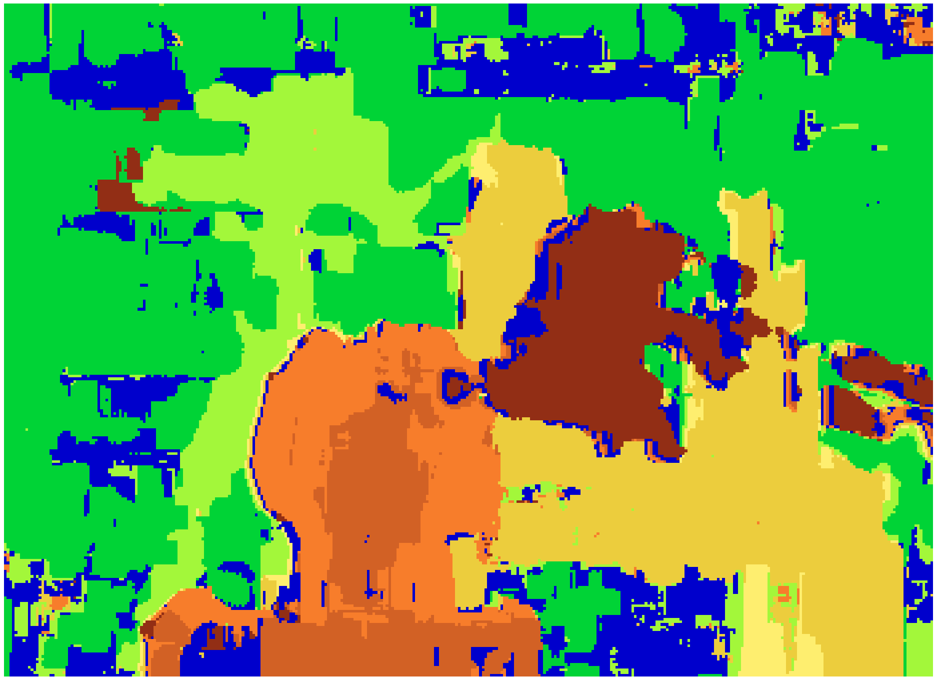
– Data from University of Tsukuba

Scene



Ground truth

Results with window correlation



Window-based matching
(best window size)

Ground truth

Results with better method

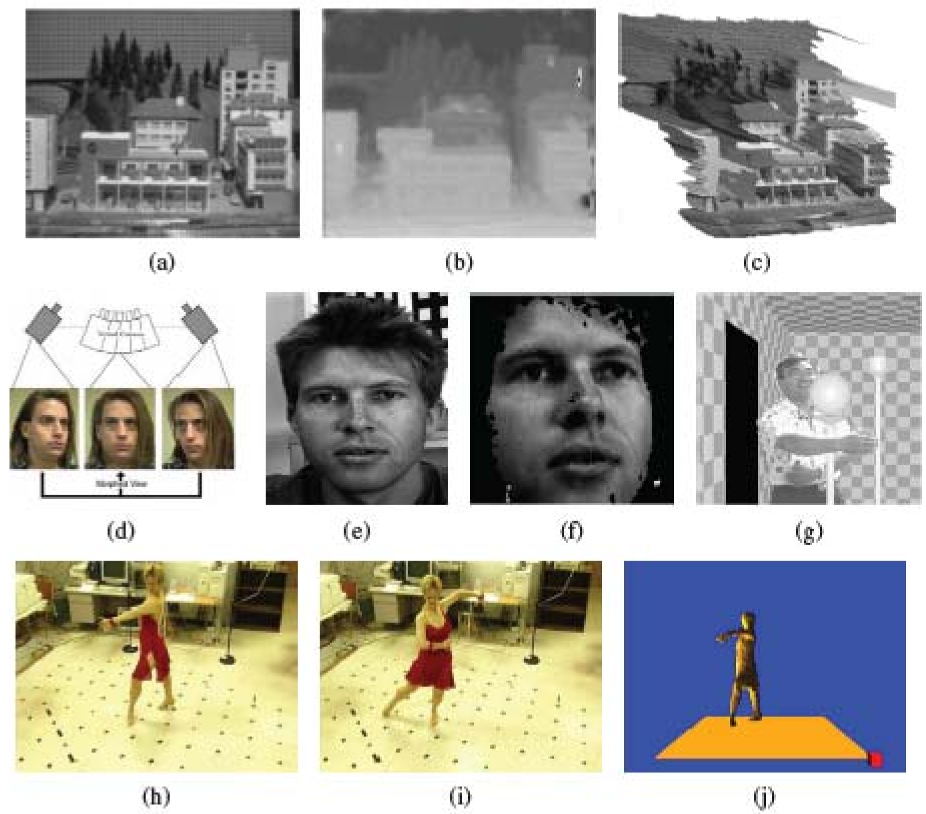
State of the art method

Ground truth

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

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