Introduction to Computer Vision

Michael J. Black

Nov 2009

Stereo

Goals

- Today
  - Finish stereo
  - Start something new
- Monday
  - Finish something new

Project

- Proposals evaluated over weekend
- Due date extended to Dec 16.

Pseudocode

condense1step

% generate cumulative distribution for posterior at t-1

….

% generate a vector of uniform random numbers.

% if a the number is greater than refreshRate then

% generate a vector of uniform random numbers

% use these to search the cumulative probability

% find the indices of the corresponding particles

% for each of these particles, predict the new state

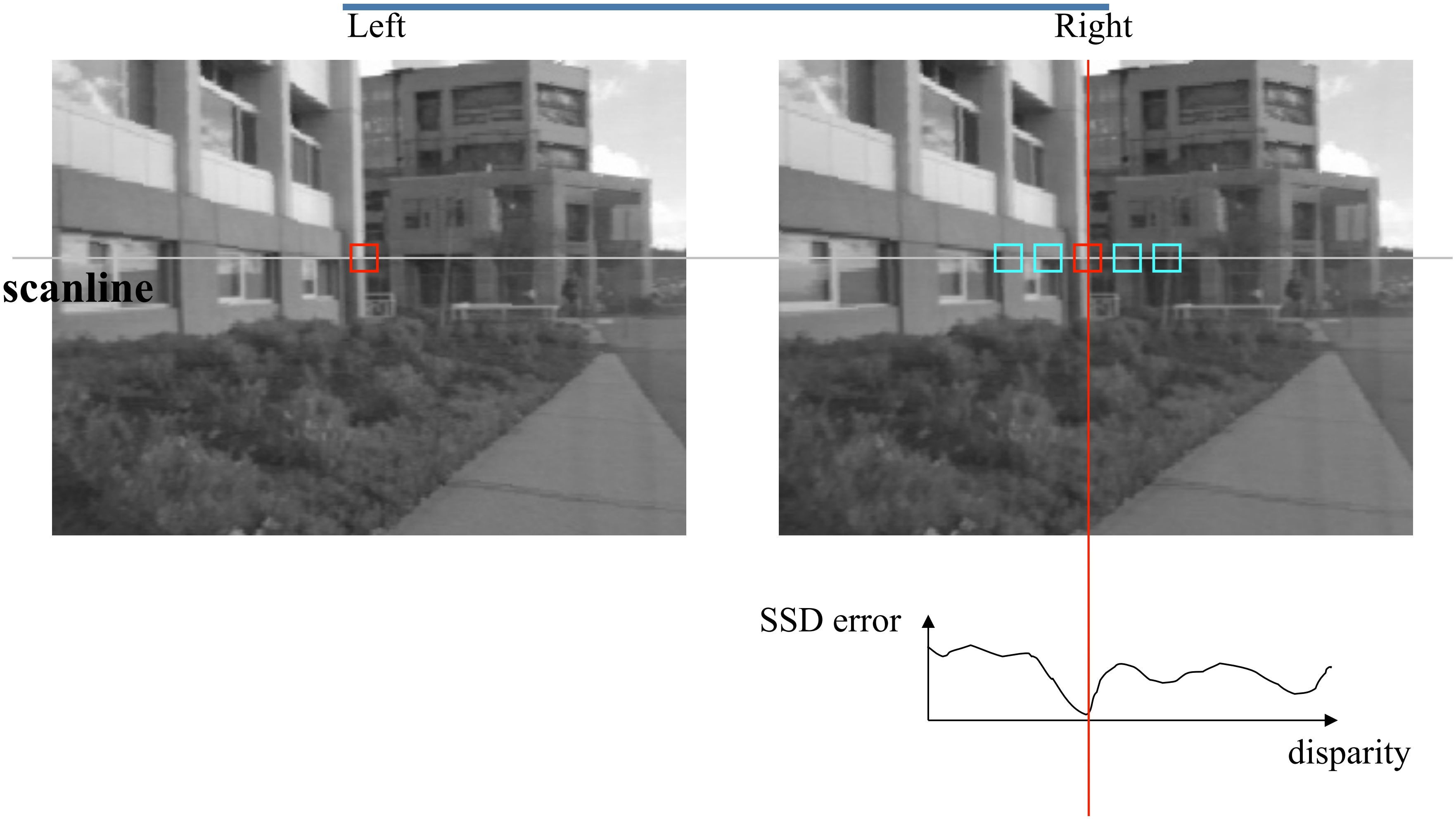
% for each of these new states compute the ***log*** likelihood

% else generate a particle at random and compute its log likelihood.

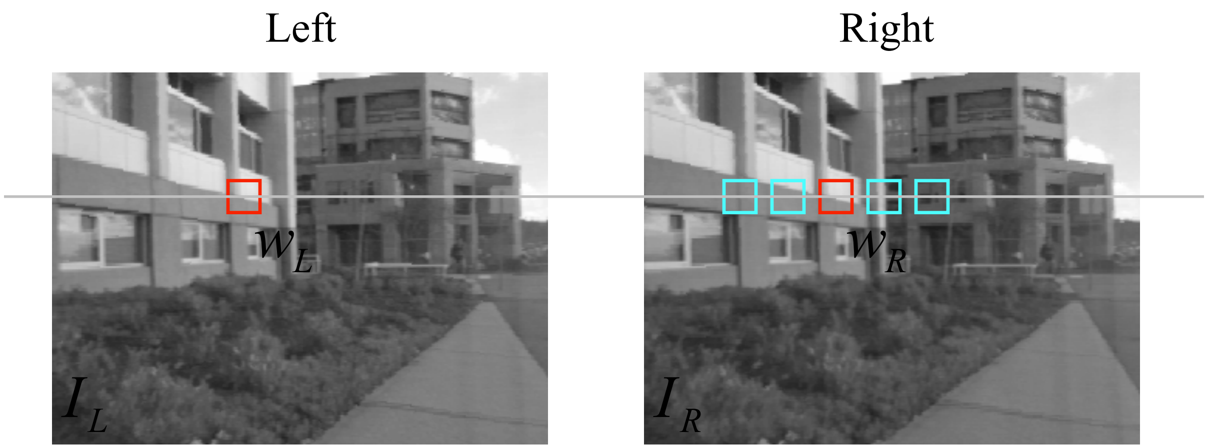
% find the maximum log likelihood and subtract it from all the other log likelihoods

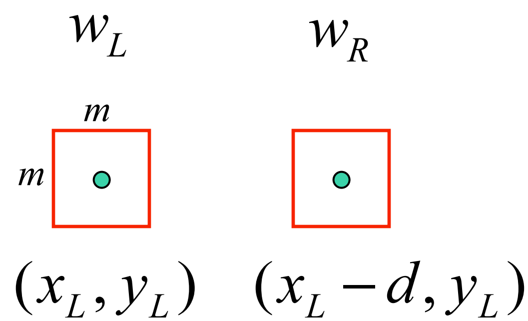
% construct the posterior at time t by exponentiating all the log likelihoods and normalizing so they sum to 1.

Correspondence Using SSD



Sum of Squared (Pixel) Differences





w_L and w_R are corresponding m by m windows of pixels.

The SSD cost measures the intensity differenceas a function of disparity :

$$SSD_r(x, y, d) = \sum_{(x', y') \in W_m(x,y)} (I_L(x', y') - I_R(x' - d, y'))^2$$

Correspondence Using SSD

Left

Disparity Map

Images courtesy of Point Grey Research

Bayesian Interpretation

$$p_M(\mathbf{d} \mid I_L, I_R)$$

How do we proceed?

Bayesian inference

Prior model $p_P(\boldsymbol{d})$

Likelihood model $p_M(I_L, I_R | \boldsymbol{d})$

Posterior model

$$p(\boldsymbol{d} \mid I_L, I_R) = k\, p_M(I_L, I_R | \boldsymbol{d})\, p_P(\boldsymbol{d})$$

Maximum a Posteriori (MAP estimate):

maximize $p(\boldsymbol{d} \mid I_L, I_R)$

Measurement model

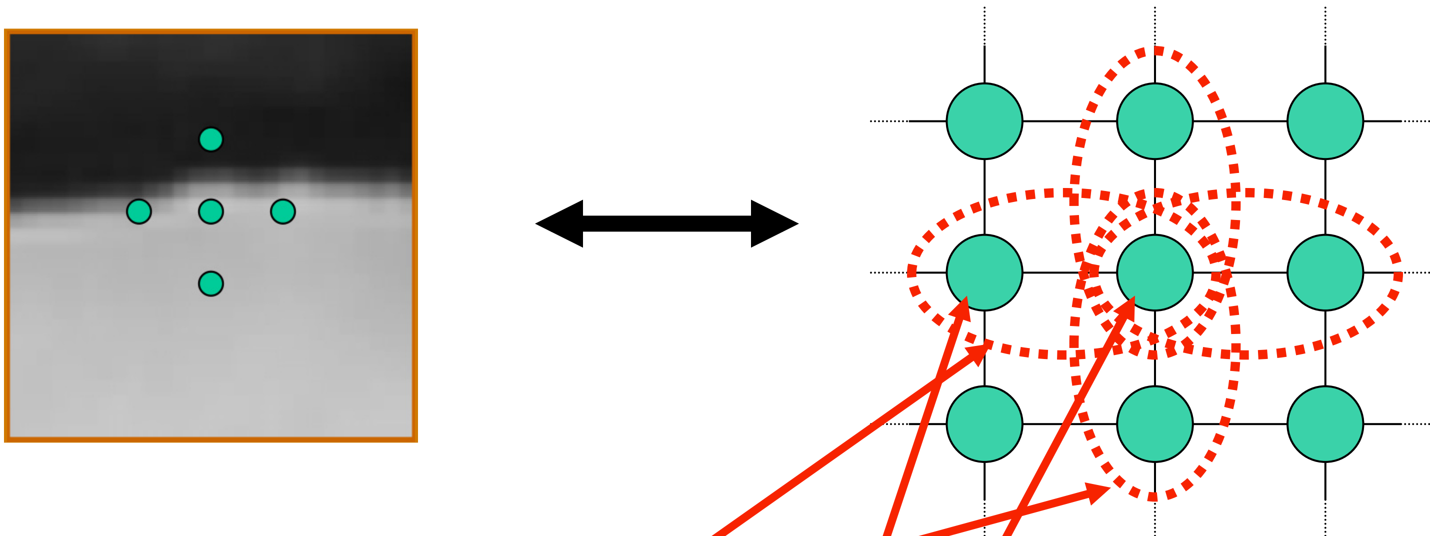
Likelihood of intensity correspondence

$$p_M(I_L, I_R|\mathbf{d}) = \frac{1}{Z_M} e^{-E_0(x,y;d)}$$

$$E_0(x, y; d) = \rho(I_L(x' + d, y') - I_R(x', y'))$$

Corresponds to Gaussian noise for quadratic *r*

Pairwise Markov Random Fields



Image

$$p(\mathbf{x}) = \frac{1}{Z} \prod_{\substack{\text{neighbors} \\ (x_i, x_j)}} \Psi(x_i, x_j)$$

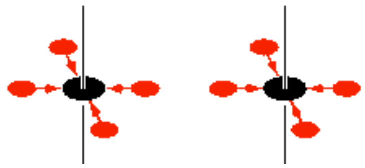
Clique “potential”

Markov Random Field

Probability distribution on disparity field *d*(*x*,*y*)

$$p(\mathbf{x}) = \frac{1}{Z} \prod_{\substack{\text{neighbors} \\ (x_i, x_j)}} \Psi(x_i, x_j)$$

$$p_P(\mathbf{d}) = \frac{1}{Z_P} e^{-E_P(\mathbf{d})}$$

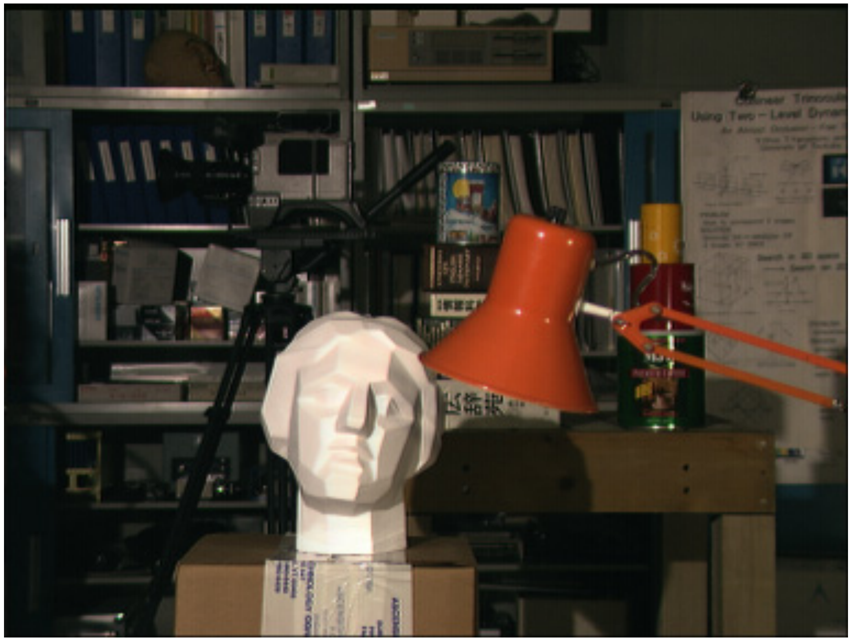


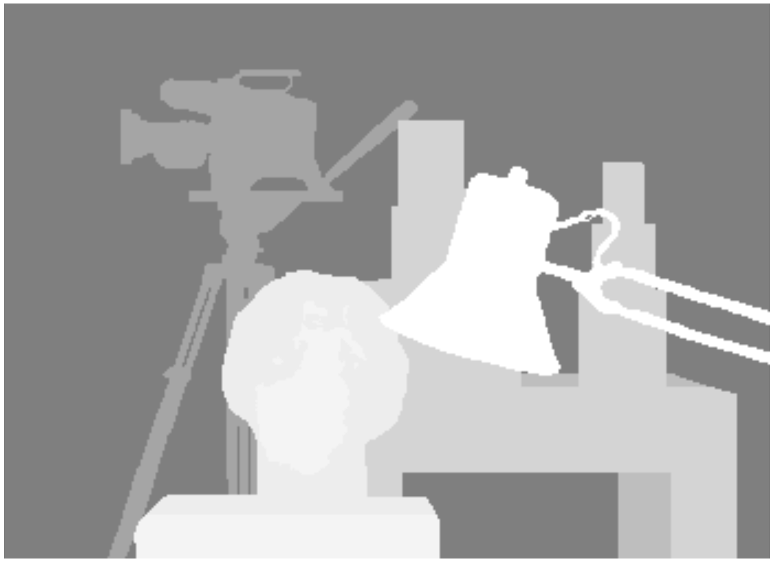
$$E_P(\mathbf{d}) = \sum_{x,y} \rho_P(d_{x+1,y} - d_{x,y}) + \rho_P(d_{x,y+1} - d_{x,y})$$

Enforces *smoothness* or *coherence* on field

Stereo results

– Data from University of Tsukuba



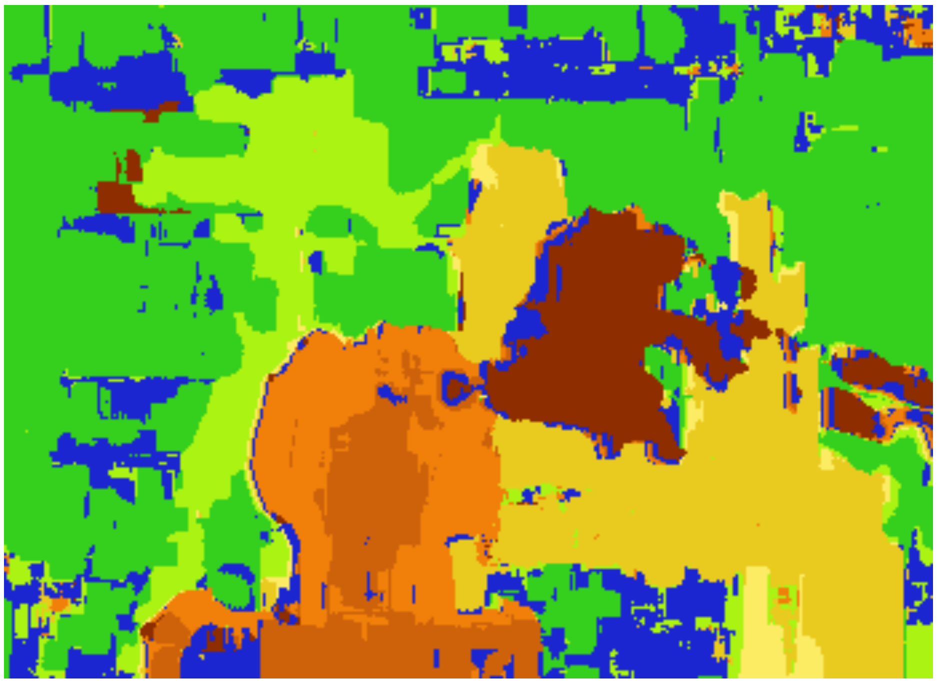


Scene

Ground truth

(Seitz)

Results with window correlation



Window-based matching (best window size)

Ground truth

(Seitz)

Results with better method

Reasonably good method

Ground truth

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

(Seitz)

Results with better method

Best method on Middlebury stereo site
http://vision.middlebury.edu/stereo/eval/.

Ground truth

(Seitz)

Choice?

- Markov random field models of images
  - Detail and novel research
- Object recognition
  - High level overview of the current field