Binocular Stereo

Philippos Mordohai

University of North Carolina at Chapel Hill

September 21, 2006

Outline

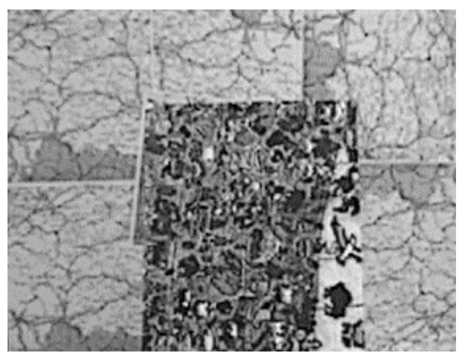
- Introduction
- Cost functions
- Challenges
- Cost aggregation
- Optimization
- Binocular stereo algorithms

Stereo Vision

- Match something
  - Feature-based algorithms
  - Area-based algorithms
- Apply constraints to help convergence
  - Smoothness/Regularization
  - Ordering
  - Uniqueness
  - Visibility
- Optimize something (typically)
  - Need energy/objective function that can be optimized

Binocular Datasets

Middlebury data (www.middlebury.edu/stereo)

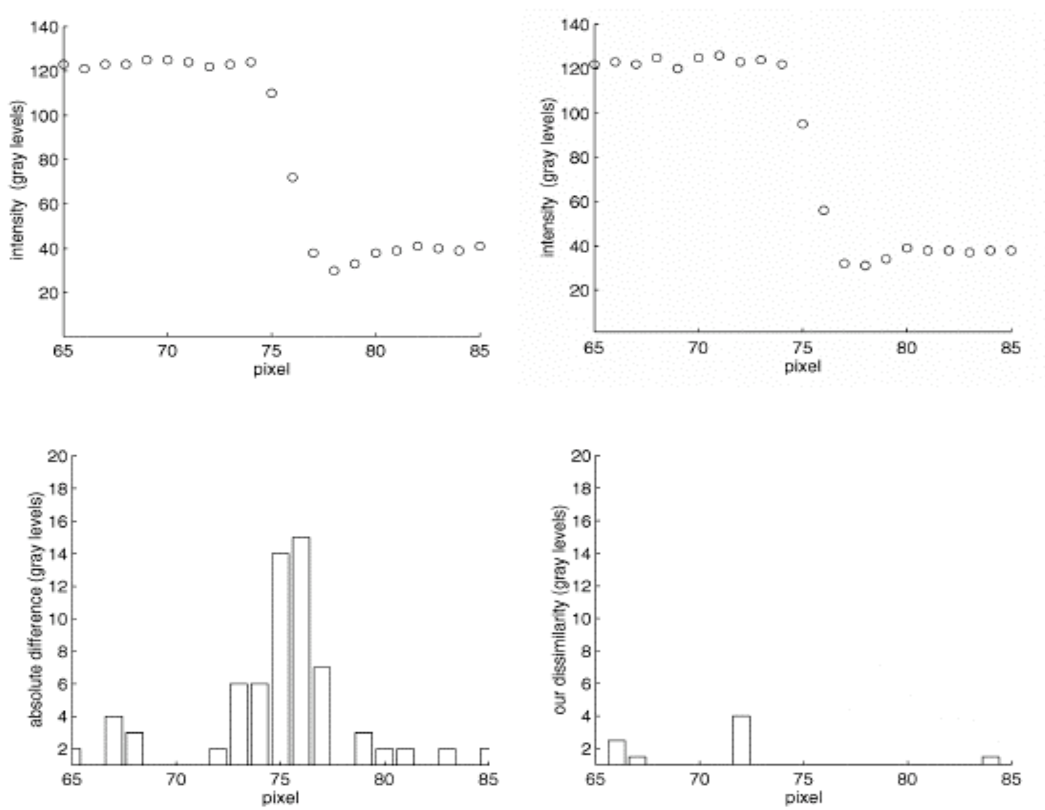


Challenges

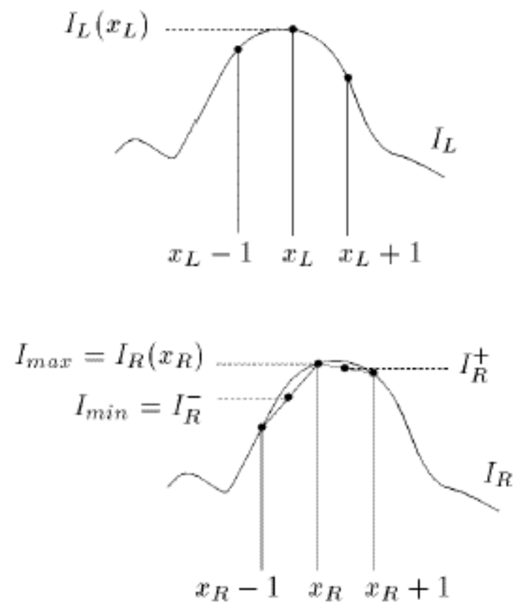
- Ill-posed inverse problem
  - Recover 3-D structure from 2-D information
- Difficulties
  - Uniform regions
  - Half-occluded pixels

Pixel Dissimilarity

- Absolute difference of intensities
- c=|I1(x,y)- I2(x-d,y)|



d



Alternative Dissimilarity Measures

- Rank and Census transforms [Zabih ECCV94]
- Rank transform:
  - Define window containing R pixels around each pixel
  - Count the number of pixels with lower intensities than center pixel in the window
  - Replace intensity with rank (0..R-1)
  - Compute SAD on rank-transformed images
- Census transform:
  - Use bit string, defined by neighbors, instead of scalar rank
- Robust against illumination changes

Rank and Census Transform Results

- Noise free, random dot stereograms
- Different gain and bias

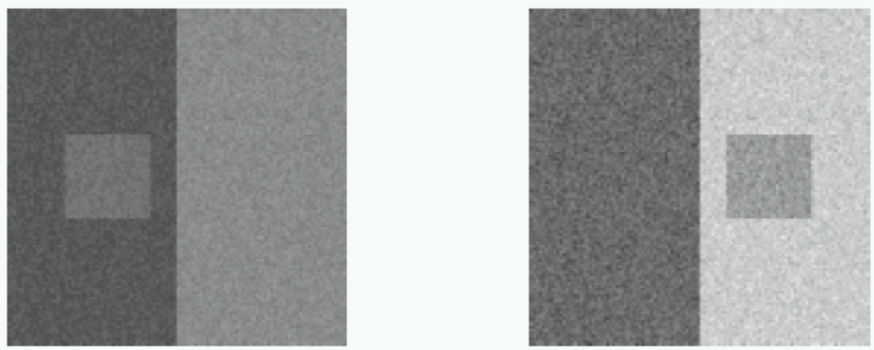


Fig. 2. Right and left random-dot stereograms

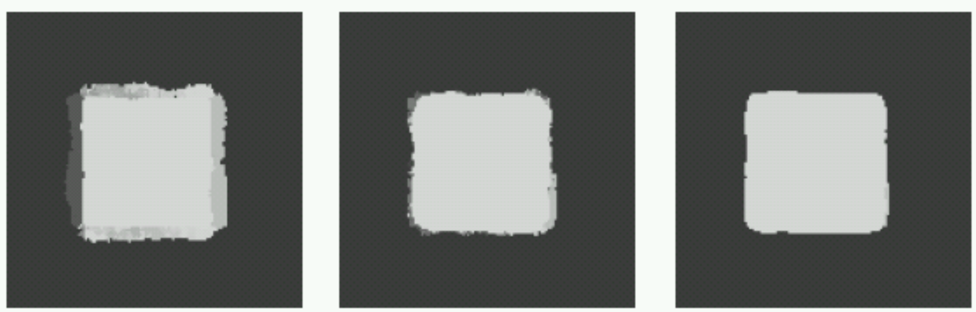


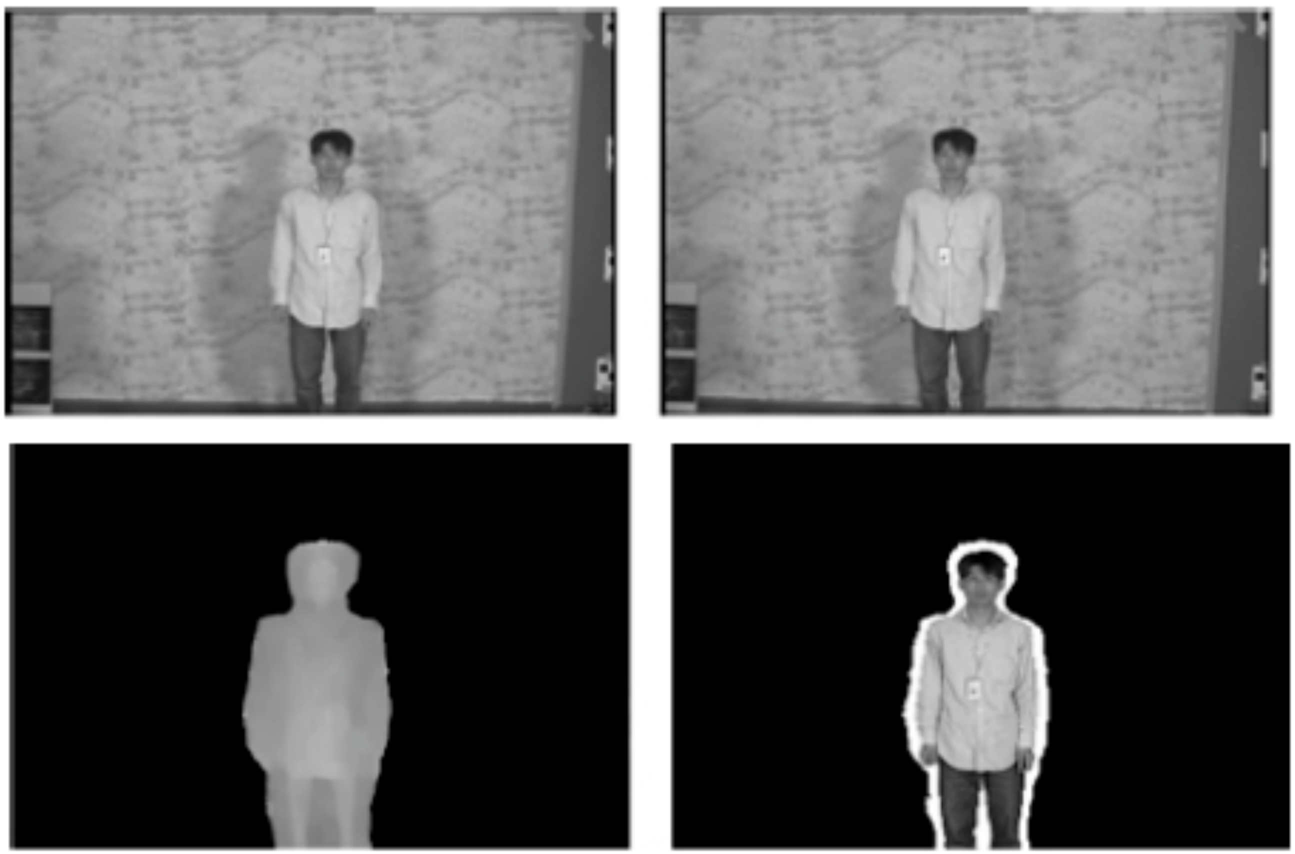
Fig. 3. Disparities from normalized correlation, rank and census transforms

Outline

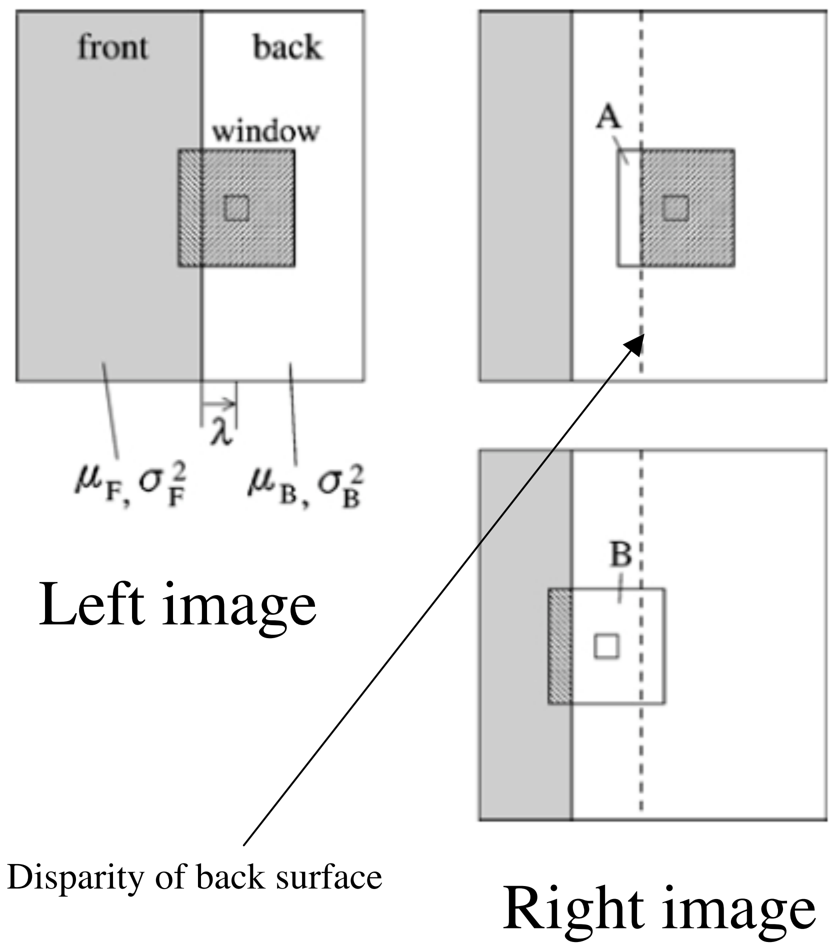
- Introduction
- Cost functions
- Challenges
- Cost aggregation
- Optimization
- Binocular stereo algorithms

Systematic Errors of Area-based Stereo

- Ambiguous matches in textureless regions
- Surface over-extension [Okutomi IJCV02]

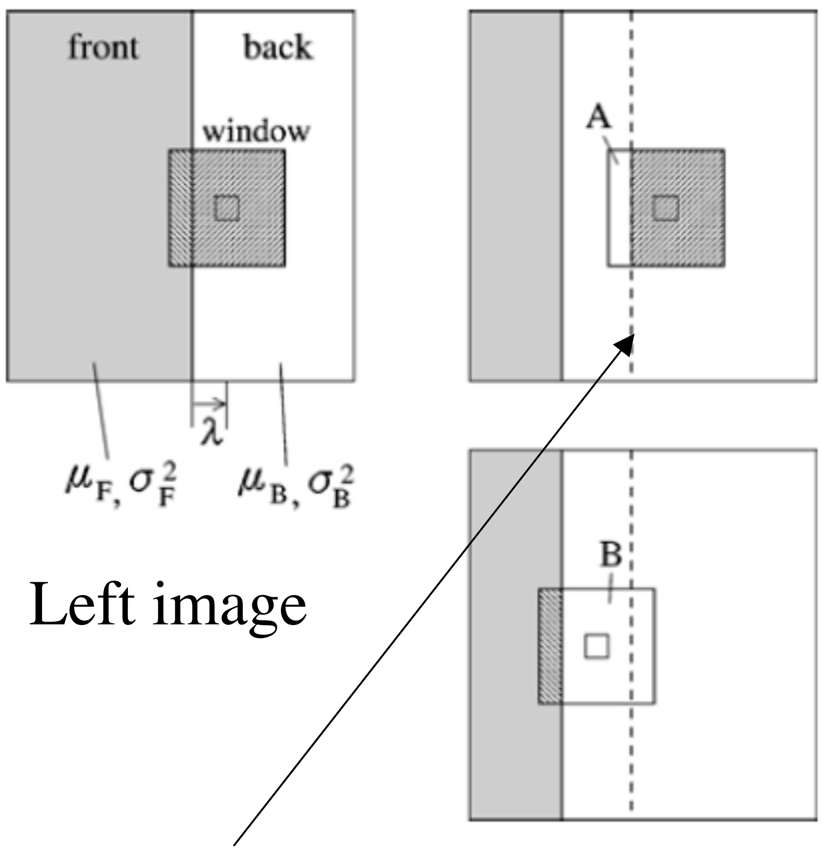


Surface Over-extension



- Expected value of $E[(x-y)^2]$ for x in left and y in right image is:
- Case A: $\sigma_F^2 + \sigma_B^2 + (\mu_F - \mu_B)^2$ for $w/2-\lambda$ pixels in each row
- Case B: $2\,\sigma_B^2$ for $w/2+\lambda$ pixels in each row

Surface Over-extension



Left image

Disparity of back surface

Right image

- Discontinuity perpendicular to epipolar lines

$$\left(\sigma_F^2 + \sigma_B^2 + (\mu_F - \mu_B)^2\right)\left(\frac{w}{2} - \lambda\right) w$$
$$= 2\sigma_B^2 \left(\frac{w}{2} + \lambda\right) w.$$

$$\lambda = \frac{\sigma_F^2 - \sigma_B^2 + (\mu_F - \mu_B)^2}{\sigma_F^2 + 3\sigma_B^2 + (\mu_F - \mu_B)^2} \frac{w}{2}.$$

- Discontinuity parallel to epipolar lines

$$\lambda = \frac{\sigma_F^2 - \sigma_B^2}{\sigma_F^2 + \sigma_B^2} \frac{w}{2}.$$

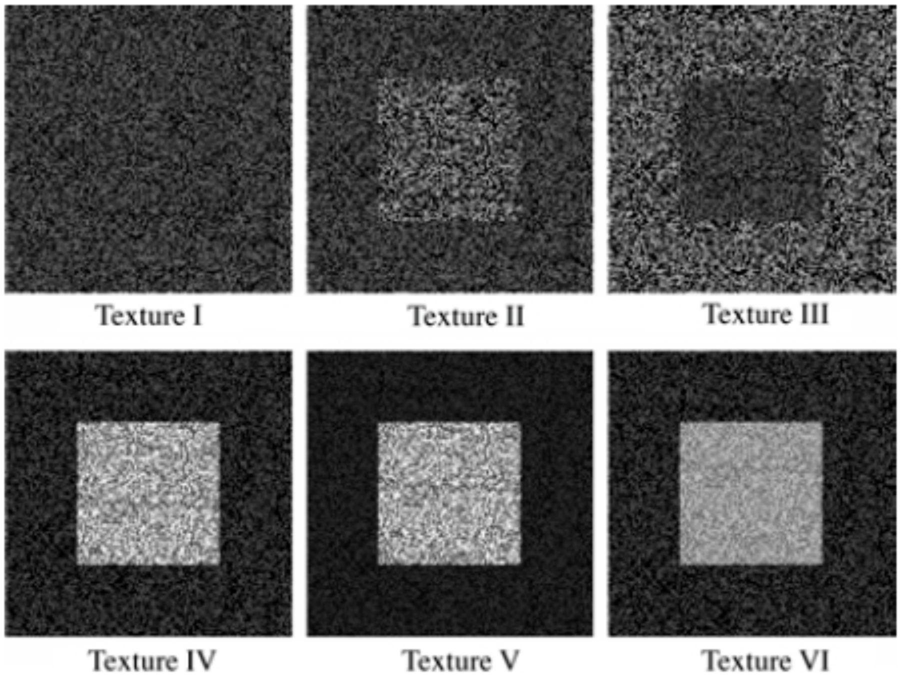
Over-extension and shrinkage

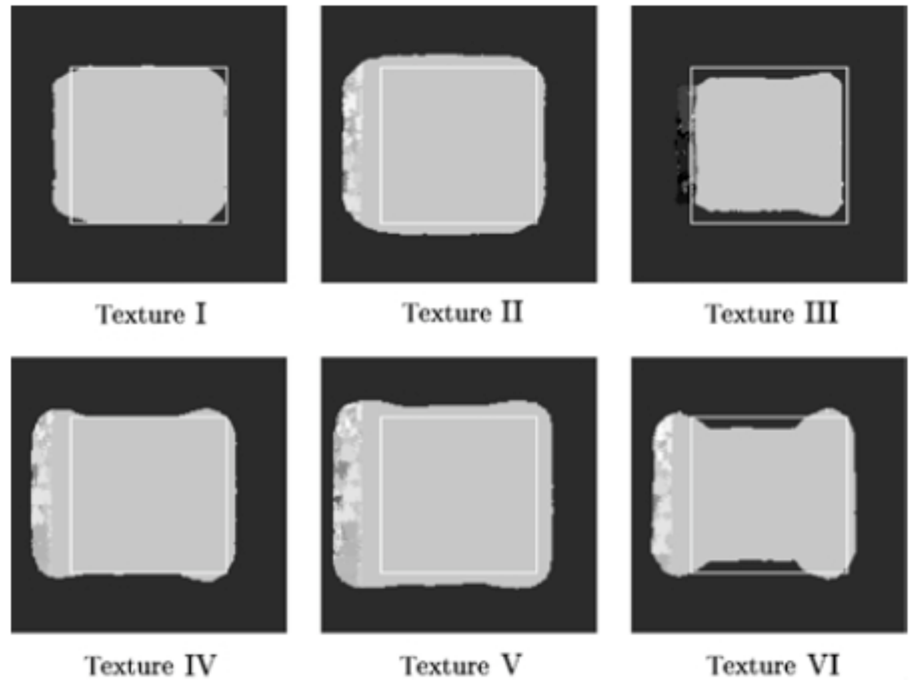
- Turns out that: $$-\frac{w}{6} \leq \lambda \leq \frac{w}{2}$$

  for discontinuities perpendicular to epipolar lines
- And: $$-\frac{w}{2} \leq \lambda \leq \frac{w}{2}$$

  for discontinuities parallel to epipolar lines

Random Dot Stereogram Experiments



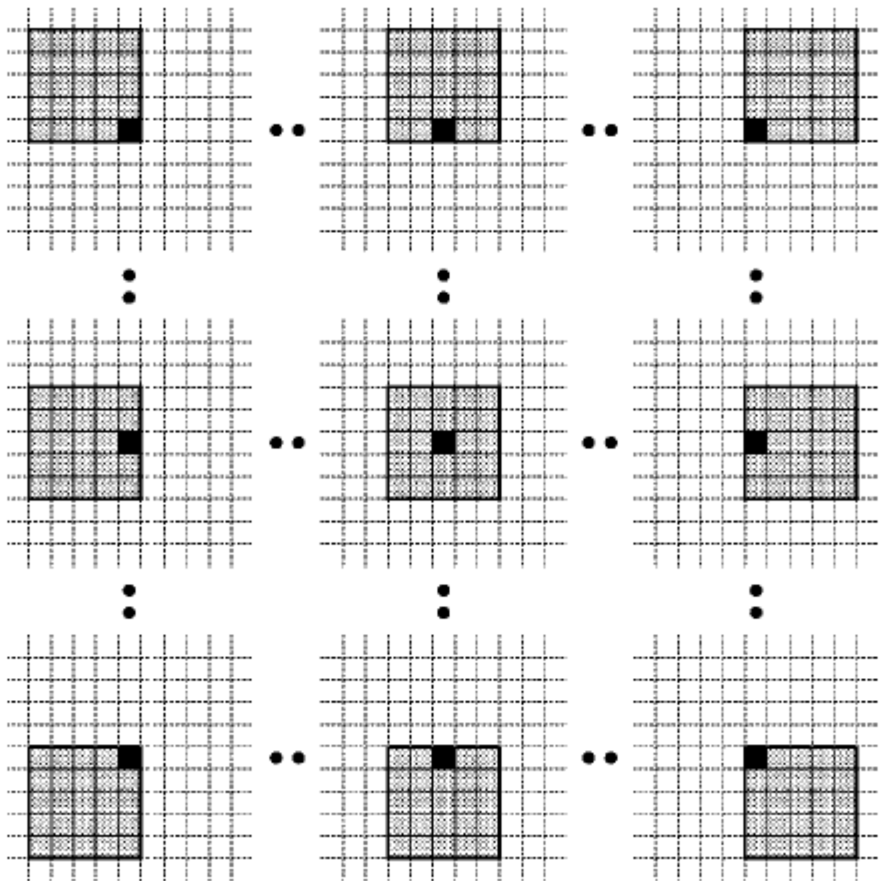


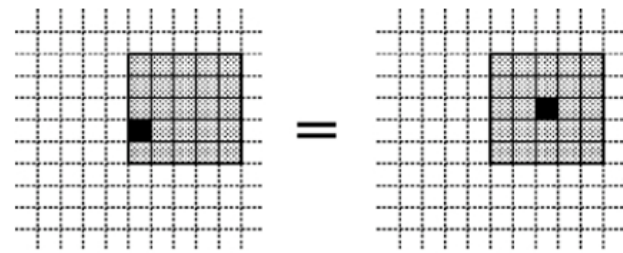
	Front surface		Back surface	
	μ_F	σ_F	μ_B	σ_B
I	100	50	100	50
II	100	100	100	50
III	100	50	100	100
IV	200	50	100	50
V	200	50	100	25
VI	200	25	100	50

Random Dot Stereogram Experiments

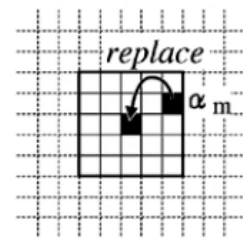
	Perpendicular		Parallel		Perpendicular		Parallel		Perpendicular		Parallel	
w	Theoretical	Actual	Theoretical	Actual	Theoretical	Actual	Theoretical	Actual	Theoretical	Actual	Theoretical	Actual
	Texture I				Texture II				Texture III			
7	0.00	0.03	0.00	0.05	1.50	1.53	2.10	2.08	−0.81	−0.69	−2.10	−1.87
11	0.00	0.08	0.00	0.13	2.36	2.35	3.30	3.32	−1.27	−1.04	−3.30	−3.09
17	0.00	0.09	0.00	0.13	3.64	3.75	5.10	5.29	−1.96	−2.00	−5.10	−5.00
25	0.00	0.12	0.00	0.42	5.36	5.20	7.50	7.73	−2.88	−3.00	−7.50	−7.53
35	0.00	0.40	0.00	−0.33	7.50	7.50	10.50	10.25	−4.04	−4.50	−10.50	−10.75
	Texture IV				Texture V				Texture VI			
7	1.75	1.81	0.00	0.23	2.89	2.87	2.10	2.27	1.57	1.49	−2.10	−1.87
11	2.75	2.74	0.00	0.56	4.54	4.60	3.30	3.61	2.47	2.52	−3.30	−3.02
17	4.25	4.32	0.00	0.68	7.02	7.11	5.10	5.47	3.81	3.88	−5.10	−4.87
25	6.25	6.15	0.00	0.65	10.33	10.20	7.50	7.75	5.50	5.80	−7.50	−7.43
35	8.75	9.00	0.00	0.90	14.46	14.45	10.50	10.80	7.84	8.00	−10.50	−10.35

Offset Windows





(a)



- Equivalent to using min nearby cost
- Result: loss of depth accuracy

Discontinuity Detection

- Use offset windows only where appropriate
  - Bi-modal distribution of SSD
  - Pixel of interest different than mode within window

Outline

- Introduction
- Cost functions
- Challenges
- Cost aggregation
- Optimization
- Binocular stereo algorithms

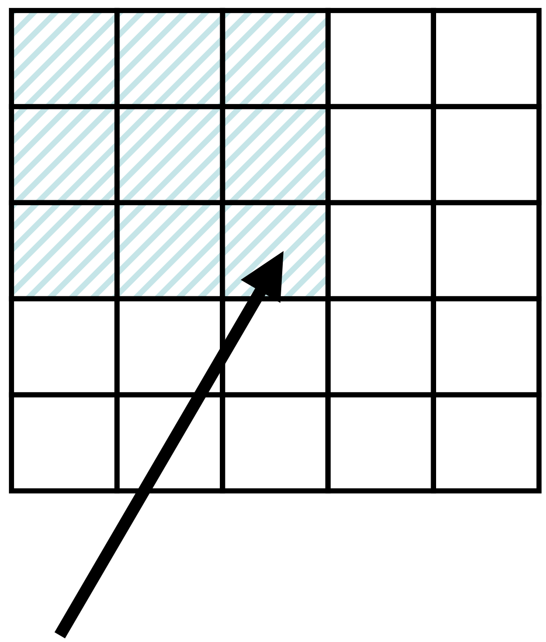
Compact Windows

- [Veksler CVPR03]: Adapt windows size based on:
  - Average matching error per pixel
  - Variance of matching error
  - Window size (to bias towards larger windows)

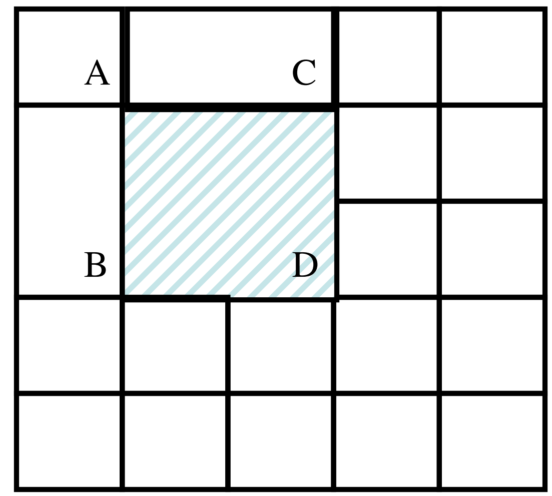
$$C_d(W) = \overline{e} + \alpha \cdot var(e) + \frac{\beta}{\sqrt{W} + \gamma}.$$

- Pick window that minimizes cost

Integral Image



Sum of shaded part



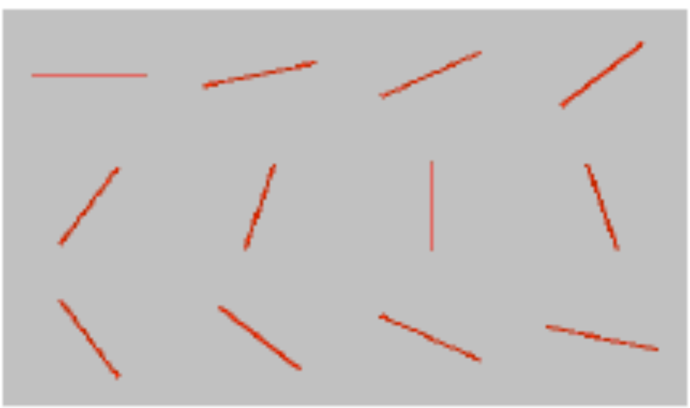
Shaded area = A+D-B-C
Independent of size

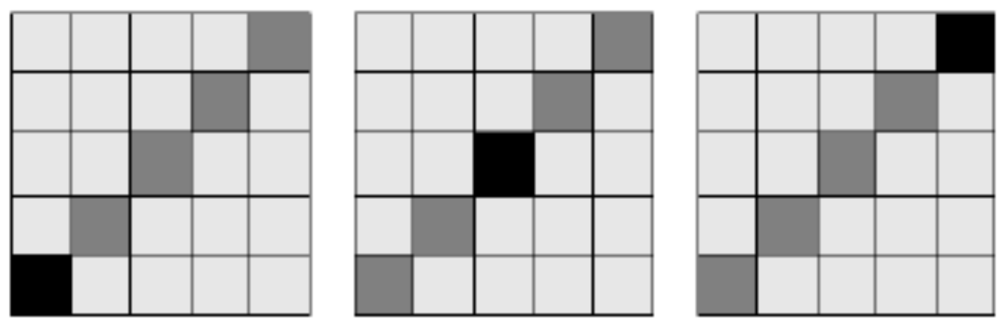
Compute an integral image for pixel dissimilarity at each possible disparity

Results using Compact Windows

Rod-shaped filters

- Instead of square windows aggregate cost in rod-shaped shiftable windows [Kim CVPR05]
- Search for one that minimizes the cost (assume that it is an iso-disparity curve)
- Typically use 36 orientations





Locally Adaptive Support

Apply weights to contributions of neighboring pixels according to similarity and proximity [Yoon CVPR05]

(a) left support window (b) right support window (c) color difference between (a) and (b)

Locally Adaptive Support

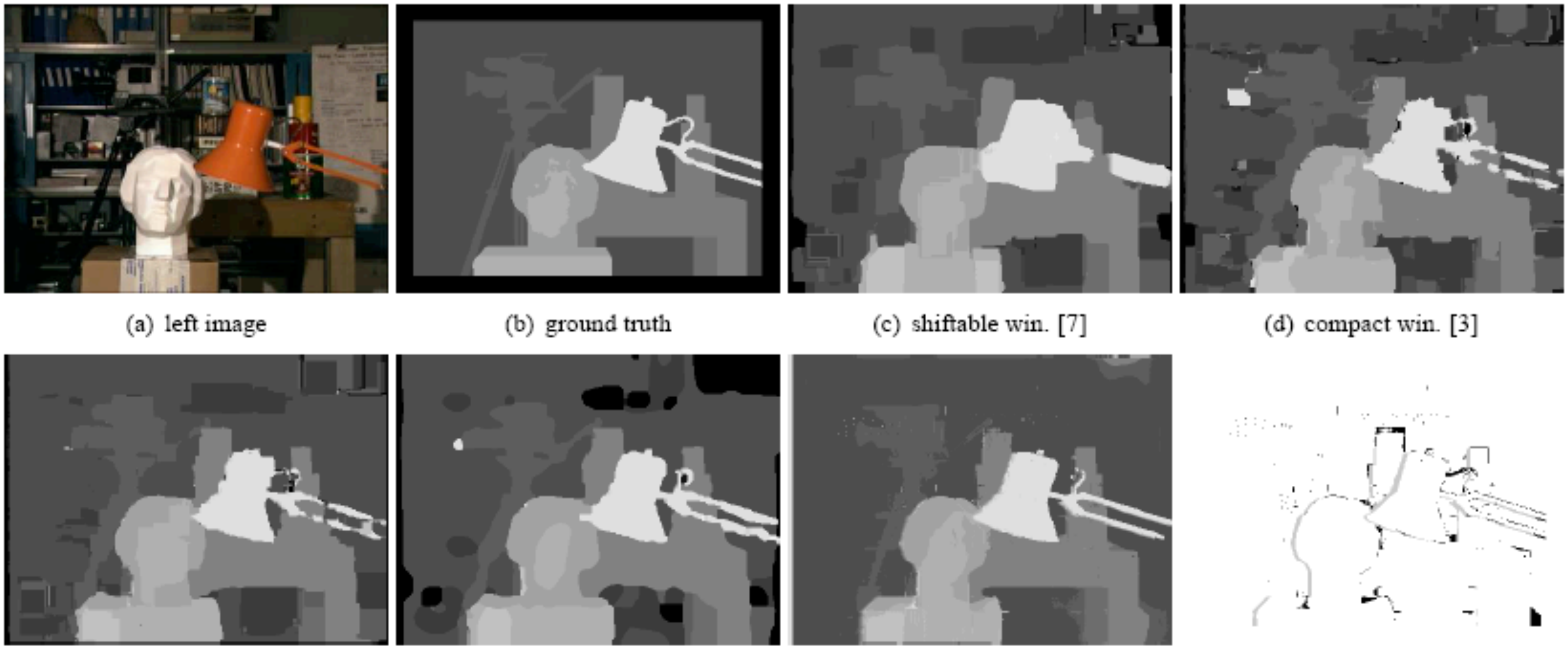
- Similarity in CIE Lab color space:

$$\Delta c_{pq} = \sqrt{(L_p - L_q)^2 + (a_p - a_q)^2 + (b_p - b_q)^2}$$

- Proximity: Euclidean distance

- Weights: $$w(p,q) = k \cdot \exp\left(-(\frac{\Delta c_{pq}}{\gamma_c} + \frac{\Delta g_{pq}}{\gamma_p})\right)$$

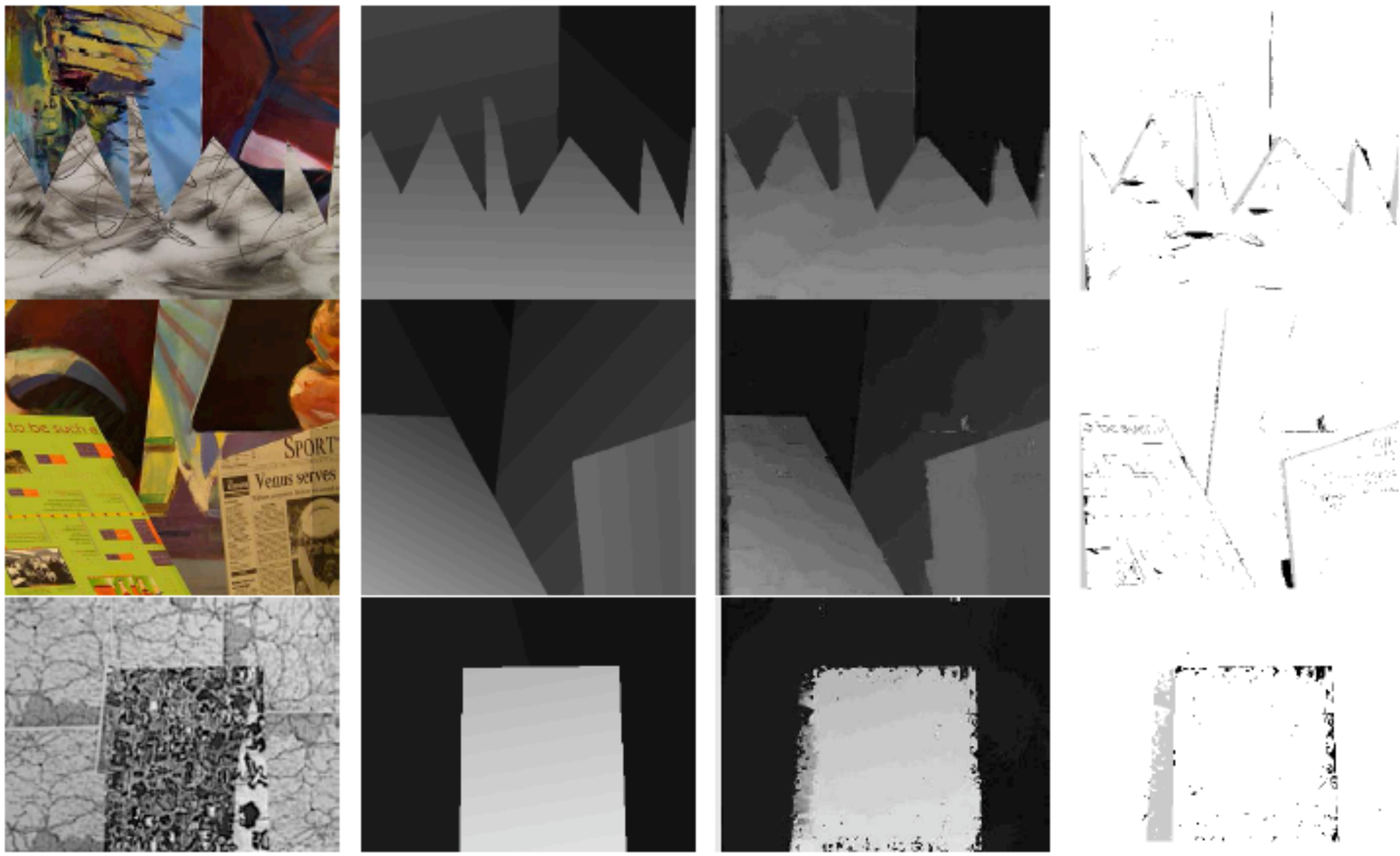
Locally Adaptive Support: Results



(a) left image (b) ground truth (c) shiftable win. [7] (d) compact win. [3]

(e) variable win. [4] (f) Bay. diff. [19] (g) our result (h) bad pixels (error > 1)

Locally Adaptive Support: Results



(a) left image (b) ground truth (c) our result (d) bad pixels (error > 1)

Outline

- Introduction
- Cost functions
- Challenges
- Cost aggregation
- Optimization
- Binocular stereo algorithms

Constraints

- Results of un-sophisticated local operators still noisy
- Optimization required
- Need constraints
  - Smoothness
  - Ordering
  - Uniqueness
  - Visibility
- Energy function

Ordering Constraint

- If A is on the left of B in reference image => the match for A has to be on the left of the match of B in target image
- Violated by thin objects
- But, useful for dynamic programming

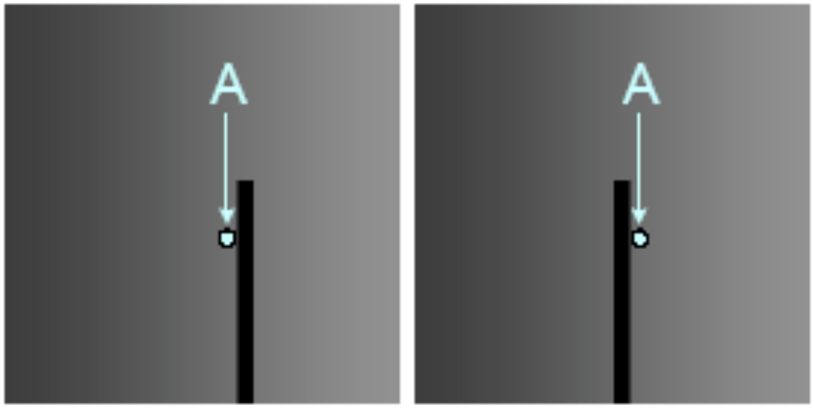
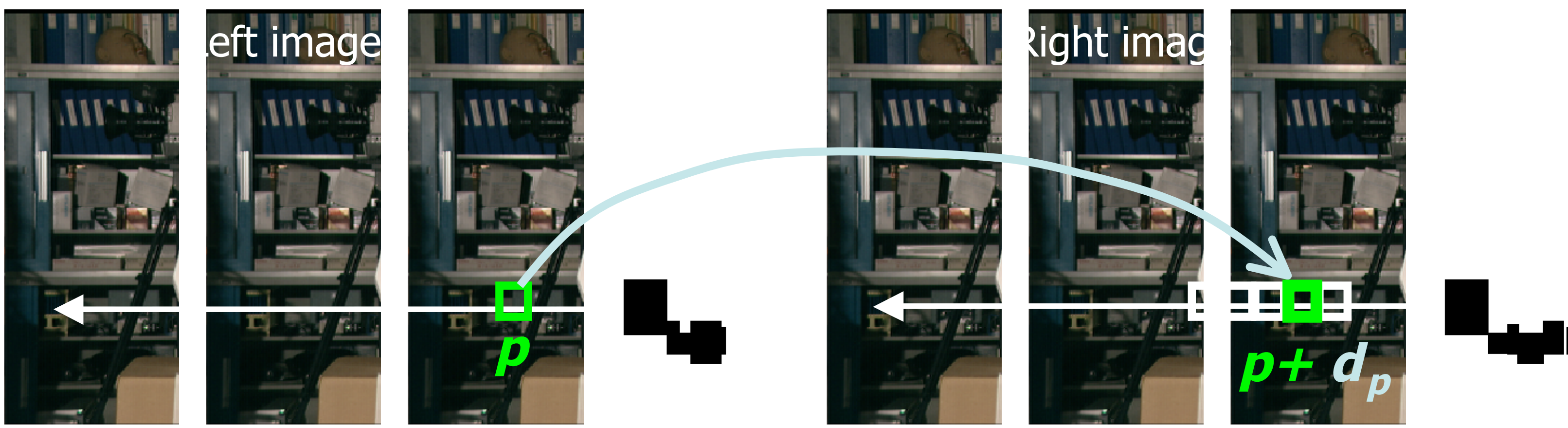
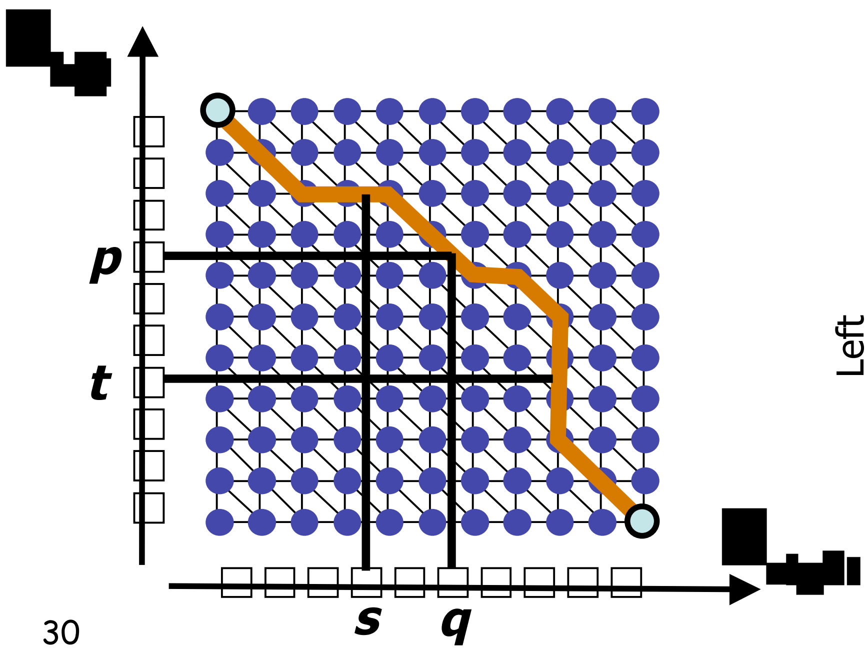
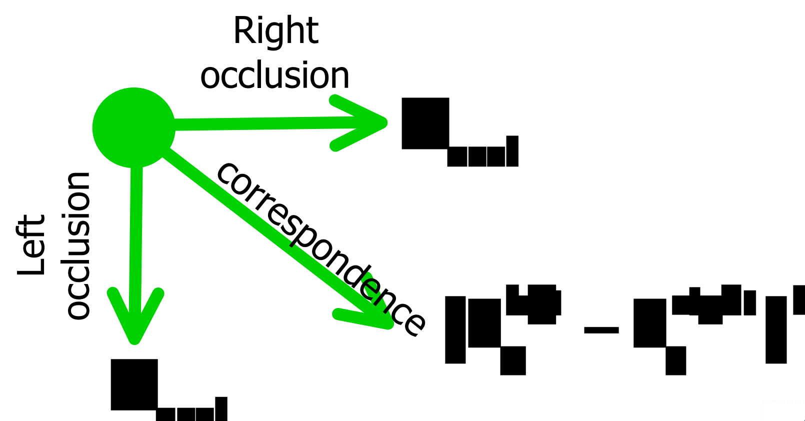


Image from Sun et al. CVPR05

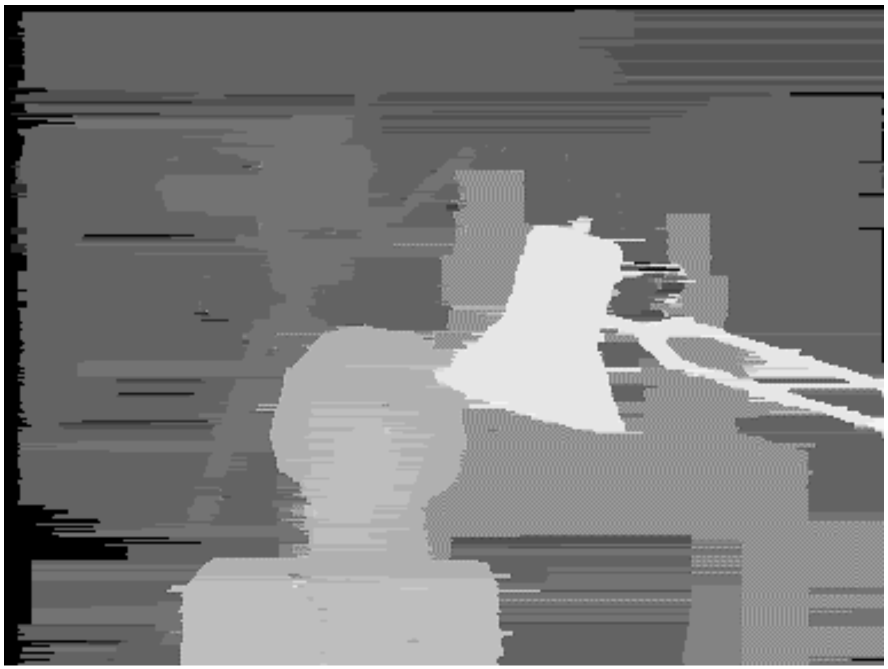
Dynamic Programming

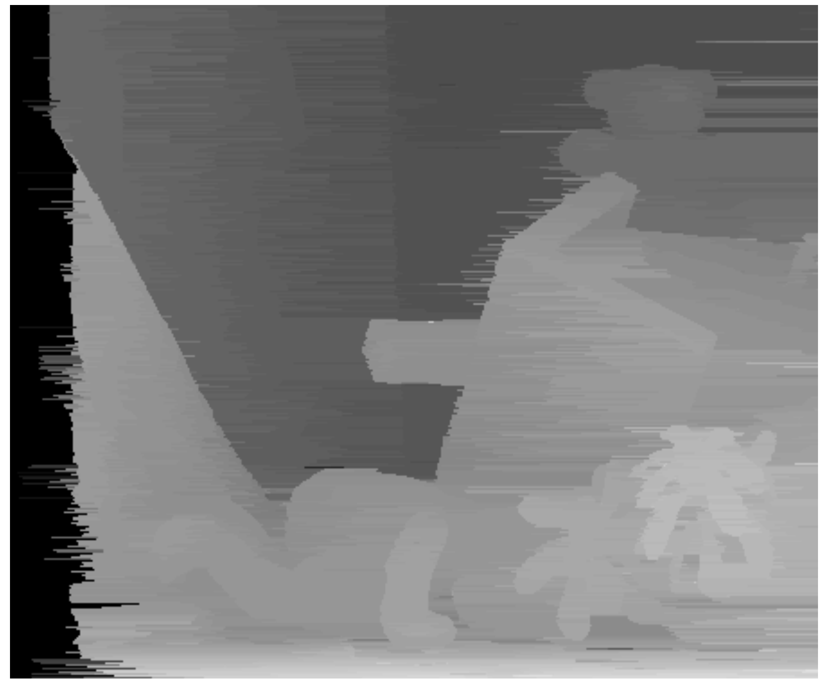






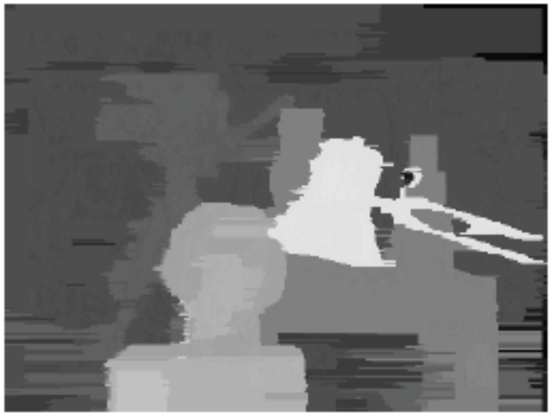
Results using Dynamic Programming





Dynamic Programming without the Ordering Constraint

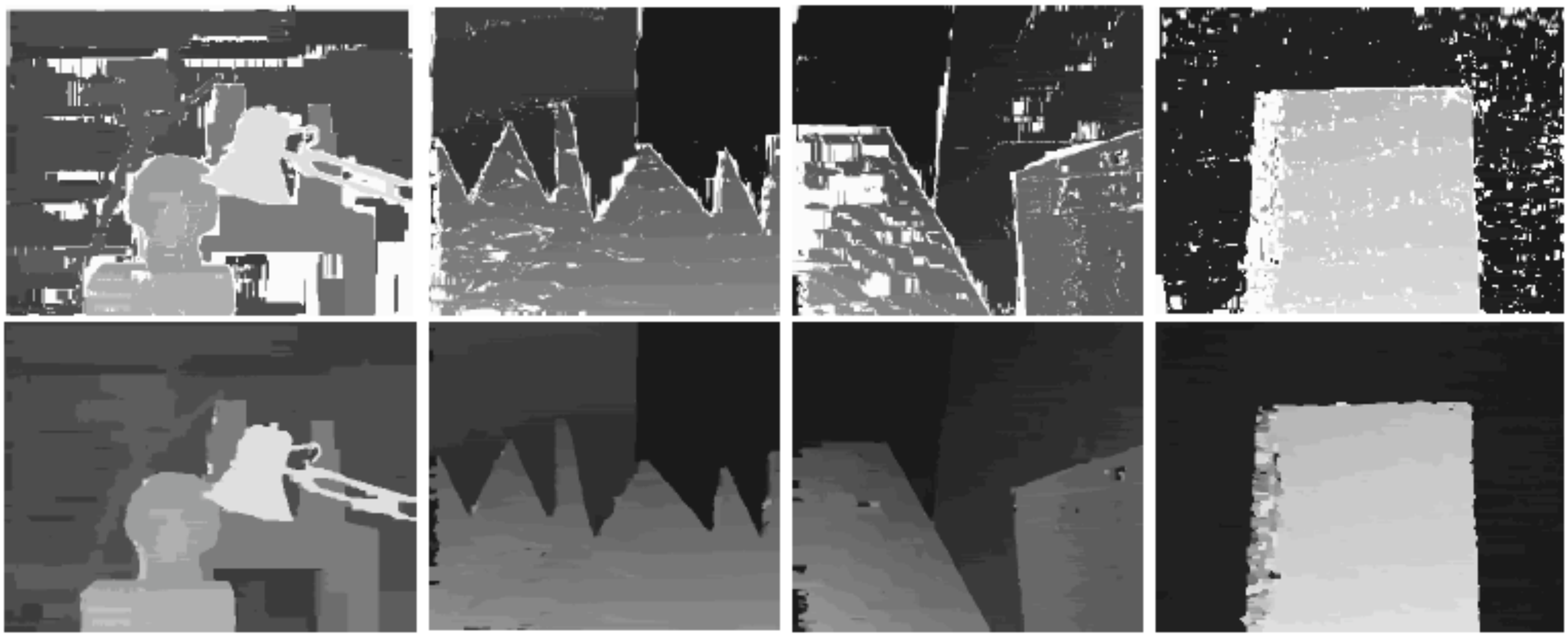
- Two Pass Dynamic Programming [Kim CVPR05]
  - Use reliable matches found with rod-shaped filters as "ground control points"
  - No ordering
  - Second pass along columns to enforce inter-scanline consistency



Dynamic Programming without the Ordering Constraint

- Use GPU [Gong CVPR05]
  - Calculate 3-D matrix (x,y,d) of matching costs
  - Aggregate using shiftable 3x3 window
  - Find reliable matches along horizontal lines
  - Find reliable matches along vertical lines
  - Fill in holes
- Match reliability =
  cost of scanline passing through match – cost of scanline *not* passing through match

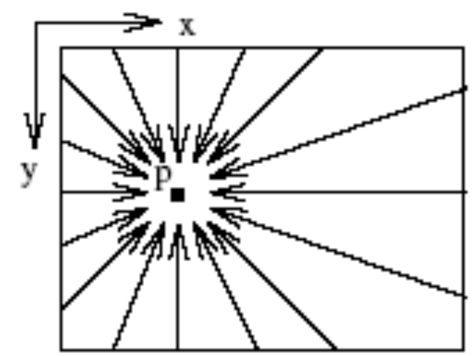
Near Real-time Results



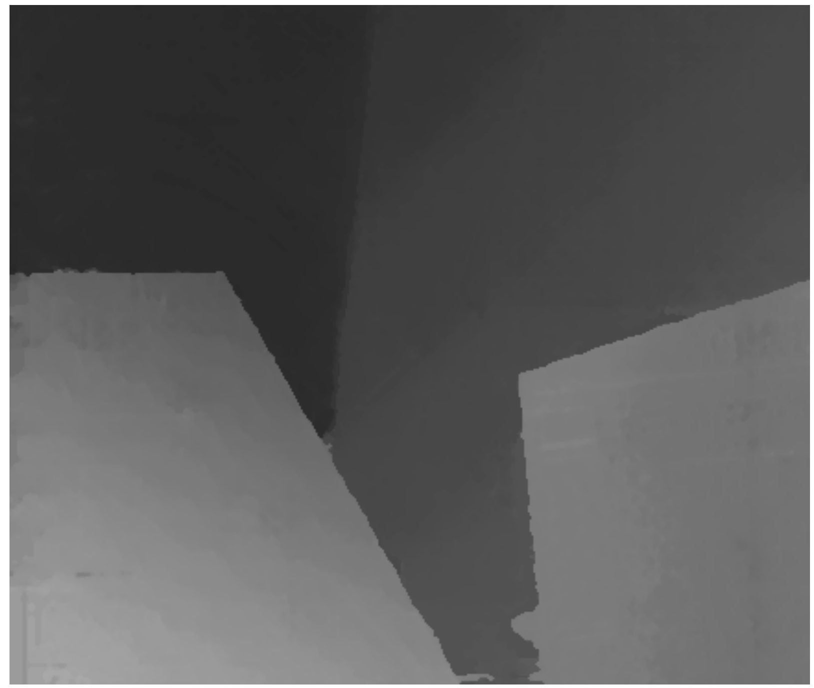
10-25 frames per second depending on image size and disparity range

Semi-global optimization

- Optimize: $E=E_{data}+E(|D_p-D_q|=1)+E(|D_p-D_q|>1)$ [Hirshmüller CVPR05]
  - Use mutual information as cost
- NP-hard using graph cuts or belief propagation (2-D optimization)
- Instead do dynamic programming along many directions
  - Don't use visibility or ordering constraints
  - Enforce uniqueness
  - Add costs



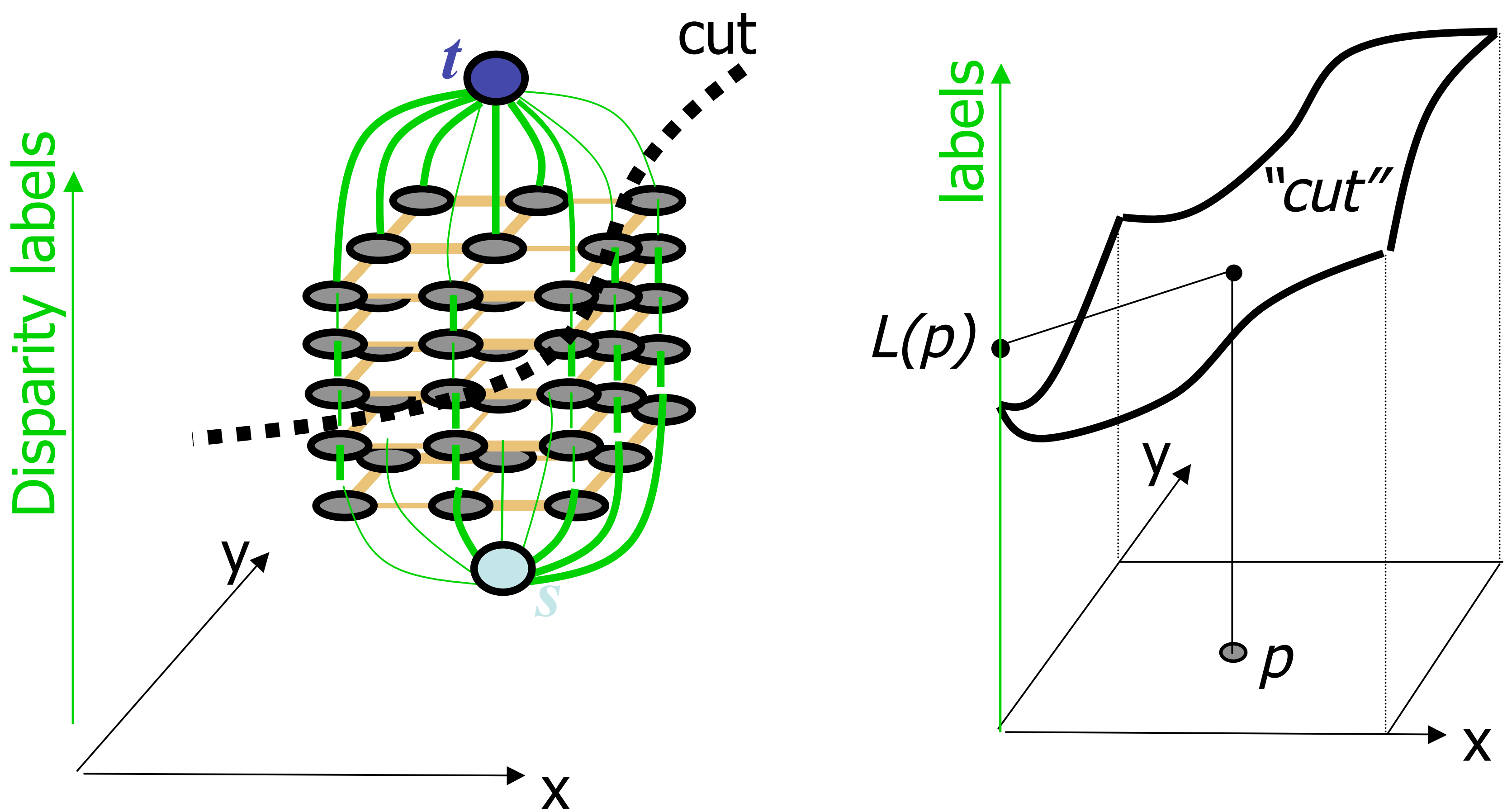
Results of Semi-global optimization



Results of Semi-global optimization

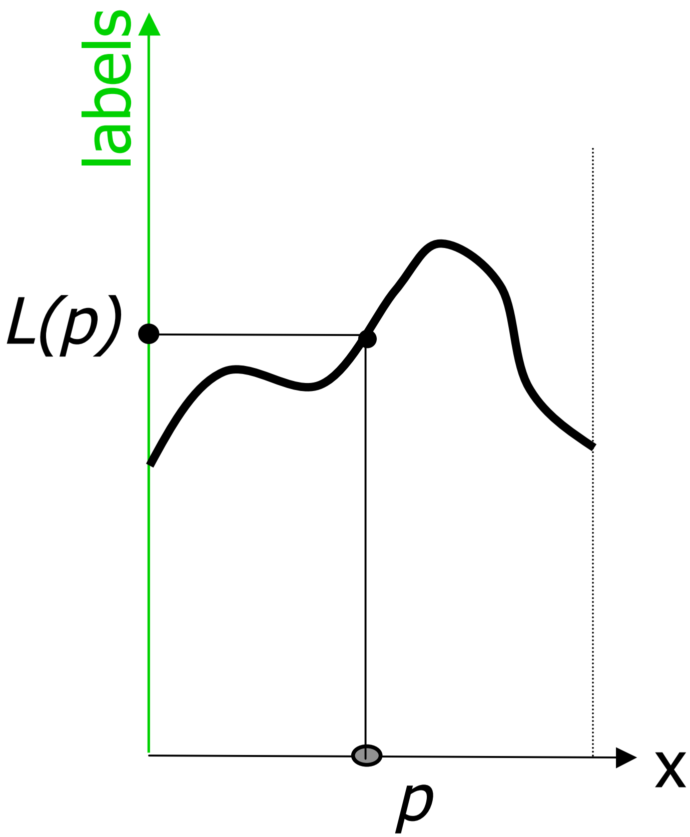
No. 1 overall in Middlebury evaluation
(at 0.5 error threshold as of Sep. 2006)

2-D Optimization

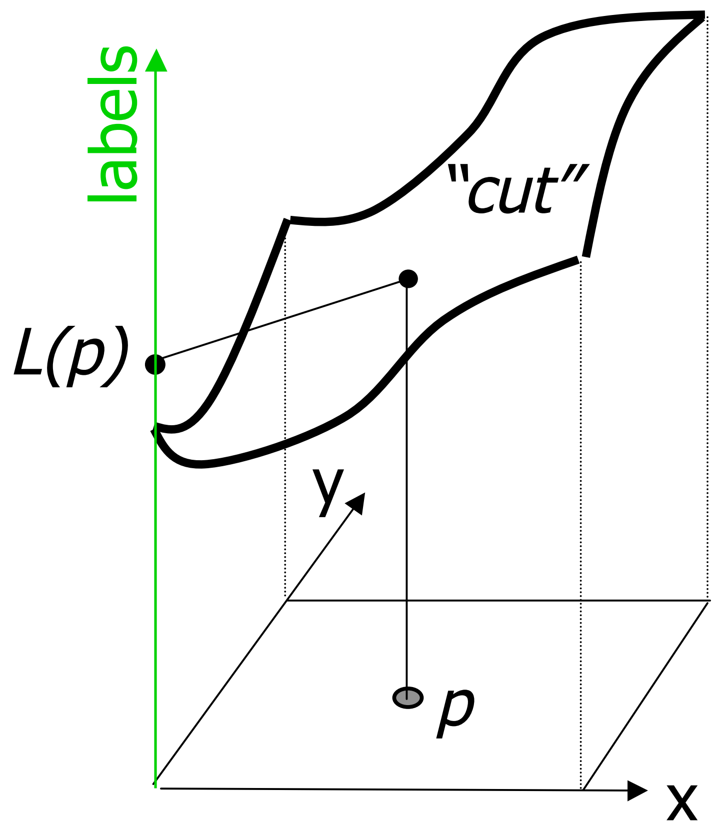


- Energy: Data Term + Regularization
- Find minimum cost cut that separates source and target

Scanline vs.
Multi-scanline optimization



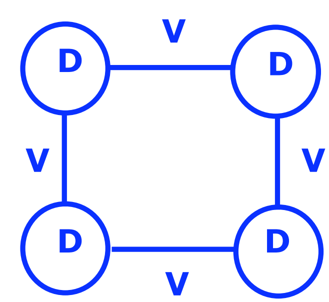
Dynamic Programming
(single scan line optimization)



s-t Graph Cuts
(multi-scan-line optimization)

Graph-cuts

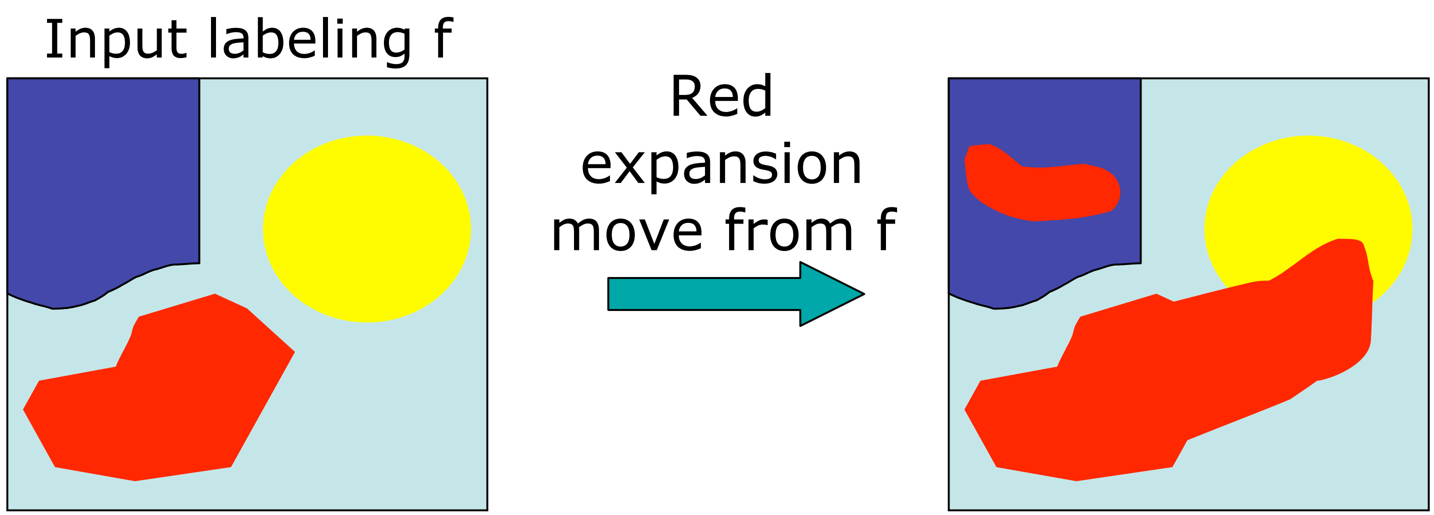
- MRF Formulation
  - In general suffers from multiple local minima
- Combinatorial optimization: minimize cost $\sum_{i \in \mathbf{S}} D_i(f_i) + \sum_{(i,j) \in \mathbf{N}} V(f_i, f_j)$ over <u>discrete</u> space of possible labelings f



  - Exponential search space $O(k^n)$
  - NP hard in most cases for grid graph
  - Approximate practical solution [Boykov PAMI01]

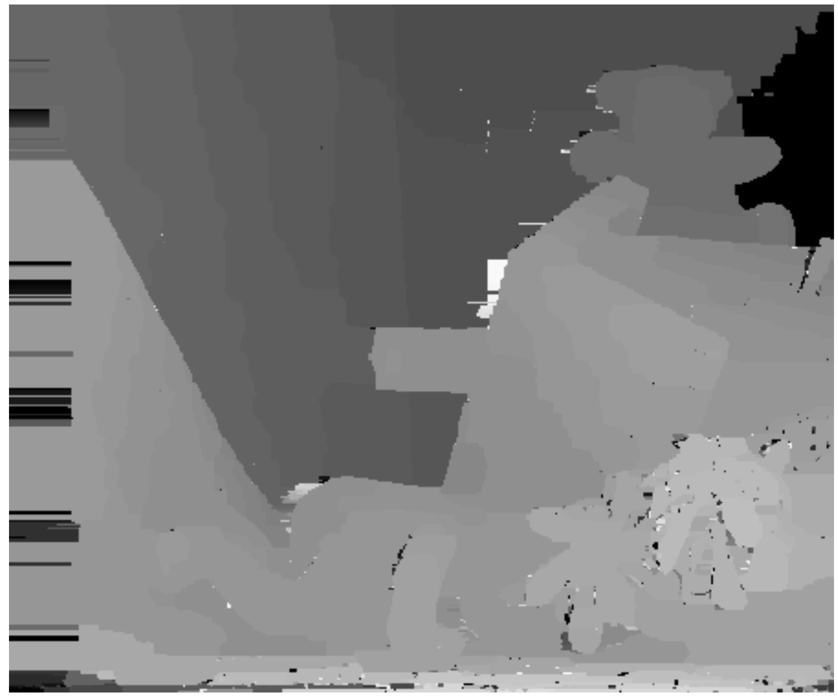
Alpha Expansion Technique

- Use min-cut to efficiently solve a special two label problem
  - Labels “stay the same” or “replace with α”
- Iterate over possible values of α
  - Each rules out exponentially many labelings



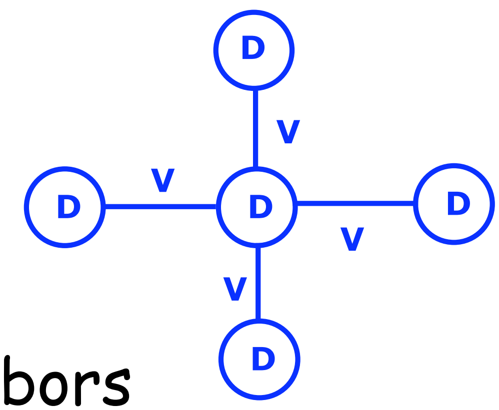
Results using Graph-Cuts

- Include occlusion term in energy [Kolmogorov ICCV01]



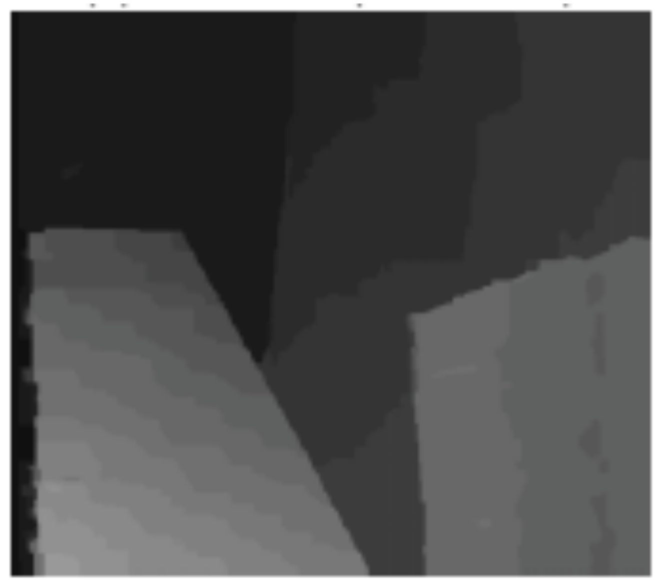
Belief Propagation

- Local message passing scheme in graph
  - Every site (pixel) in parallel computes a belief
    - *pdf* of local estimates of label costs
  - Observation: data term (fixed)
  - Messages: *pdf's* from node to neighbors
- Exact solution for trees, good approximation for graphs with cycles

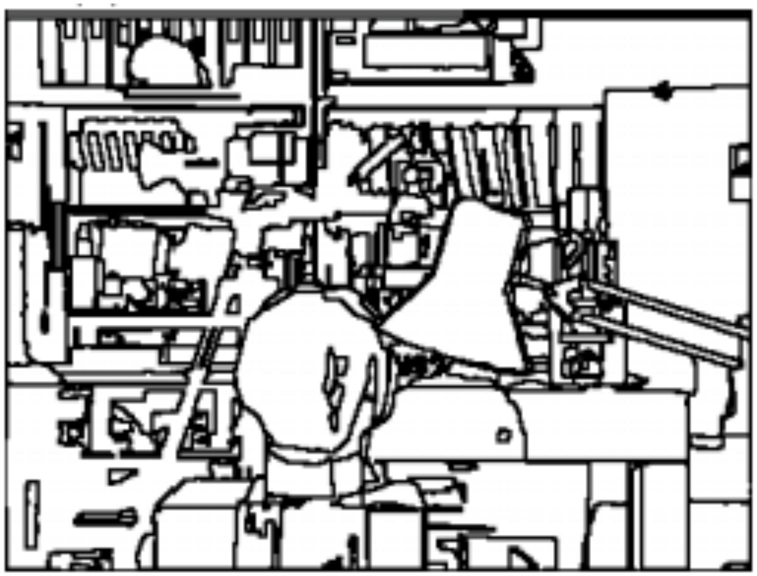


Belief Propagation for Stereo

- Minimize energy that considers matching cost, depth discontinuities and occlusion [Sun ECCV02, PAMI03]

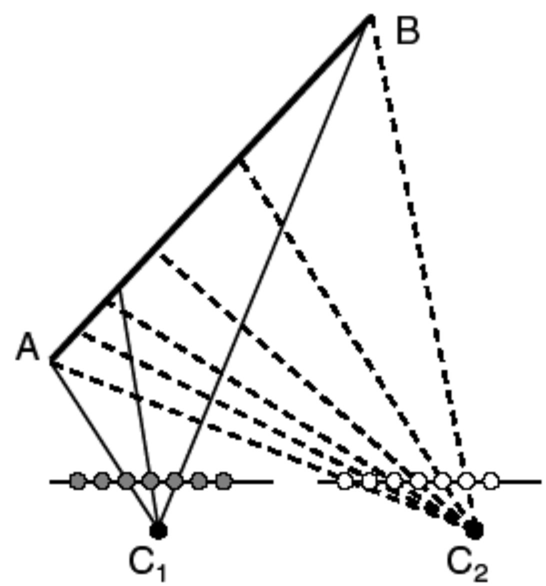


Belief Propagation and Segmentation



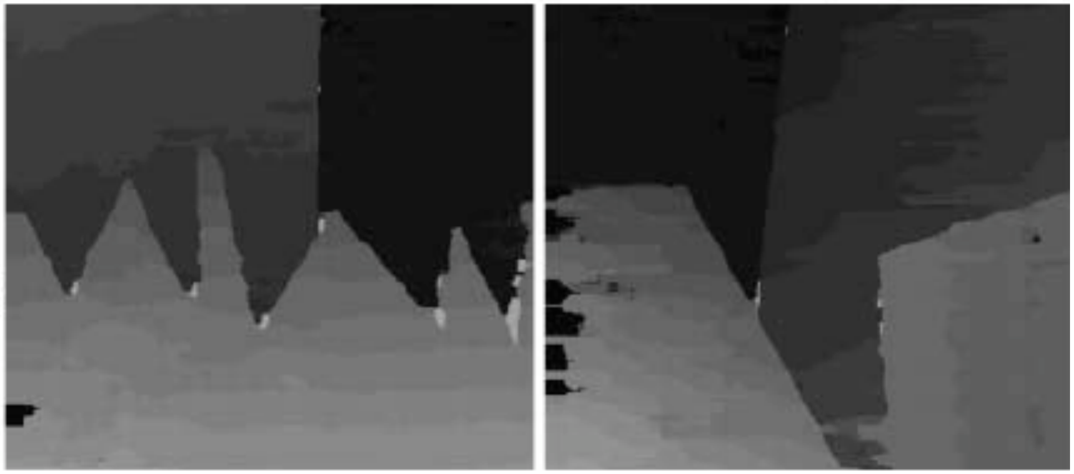
Uniqueness Constraint

- Each pixel can have exactly one or no match in the other image
  - Used in most of the above methods
- Unfortunately, surfaces do not project to the same number of pixels in both images [Ogale CVPR04]



Continuous Approach

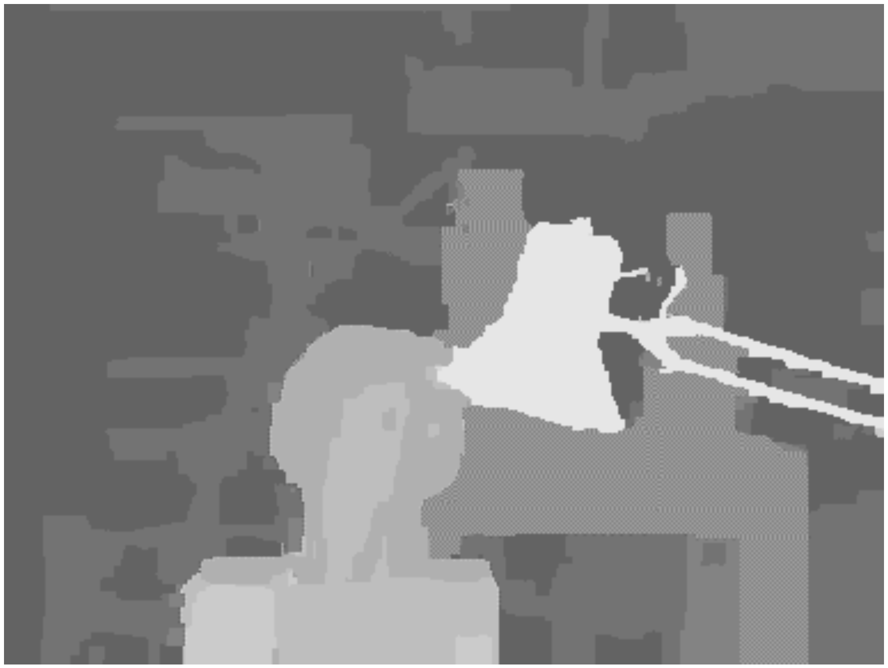
- Treat intervals on scanlines as continuous entities and not as discrete sets of pixels
- Assign disparity to beginning and end of each interval
- Optimize each scanline
  - Would rank 8,7 and 2 for images without horizontal slant
  - Ranks 22 for Venus !!!



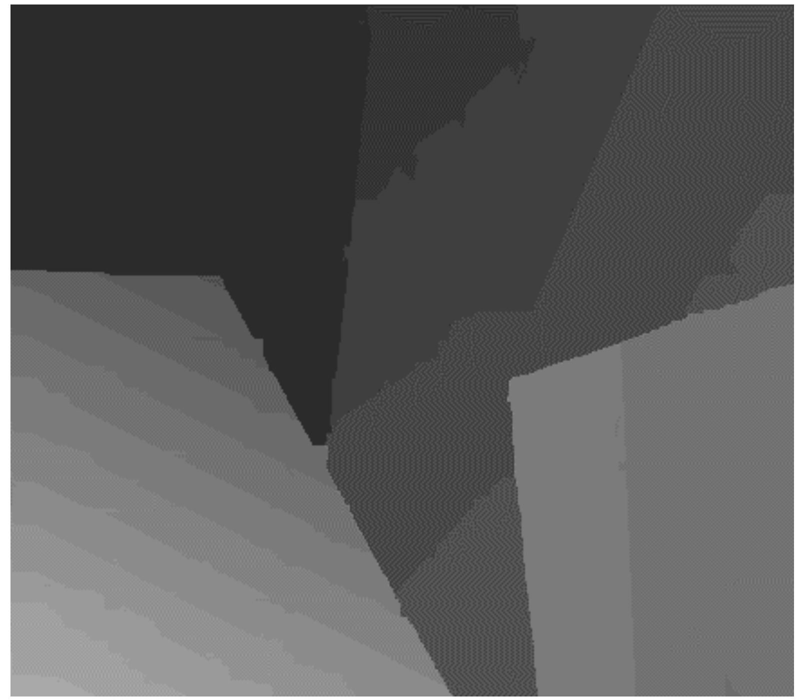
Visibility Constraint

- Each pixel is either occluded or can have one disparity value (possibly subpixel) associated with it [Sun CVPR05]
  - Allows for many-to-one correspondence
- Symmetric treatment of images
  - Compute both disparity and occlusion maps
  - Left occlusion derived from right disparity and right occlusion from left disparity
- Optimize using Belief Propagation
  - Iterate between disparity and occlusion maps
- Segmentation as a soft constraint

Results using Symmetric Belief Propagation

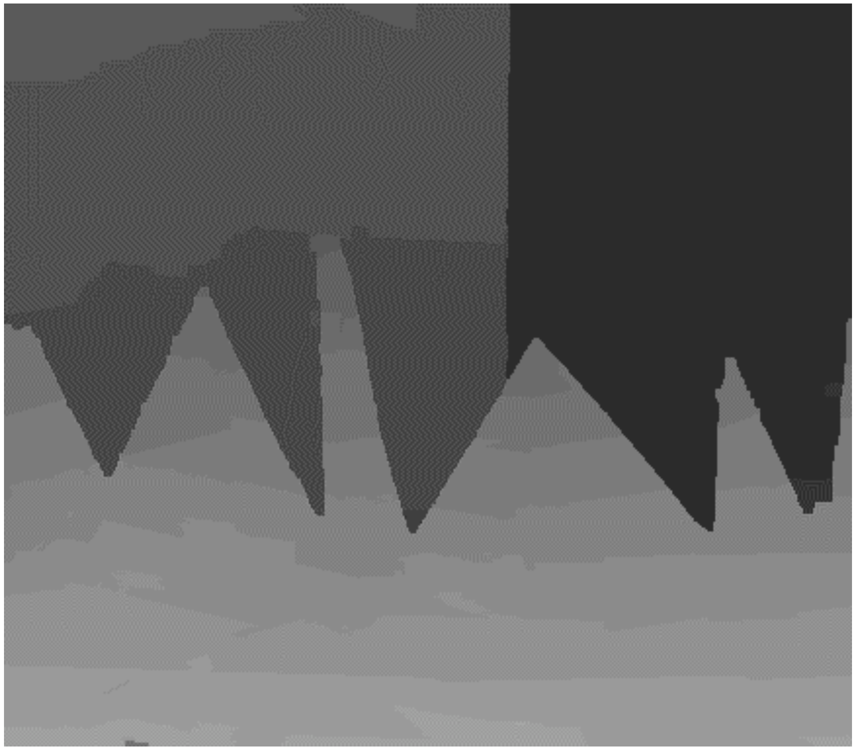


No. 1 in Middlebury evaluation (June 2005)

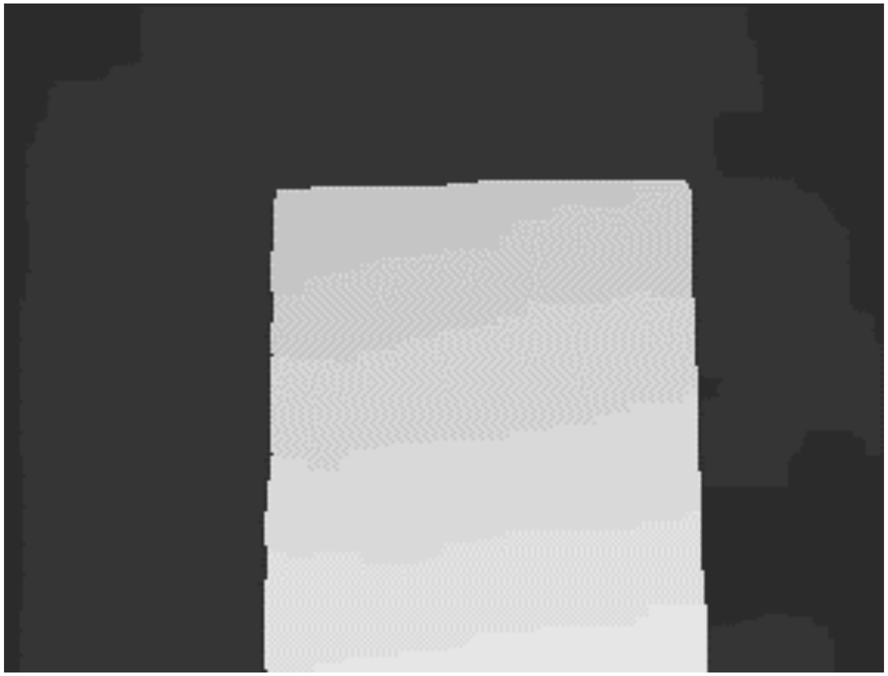


No. 3 in Middlebury evaluation (No. 1 in New Middlebury evaluation) (June 2005)

Results using Symmetric Belief Propagation

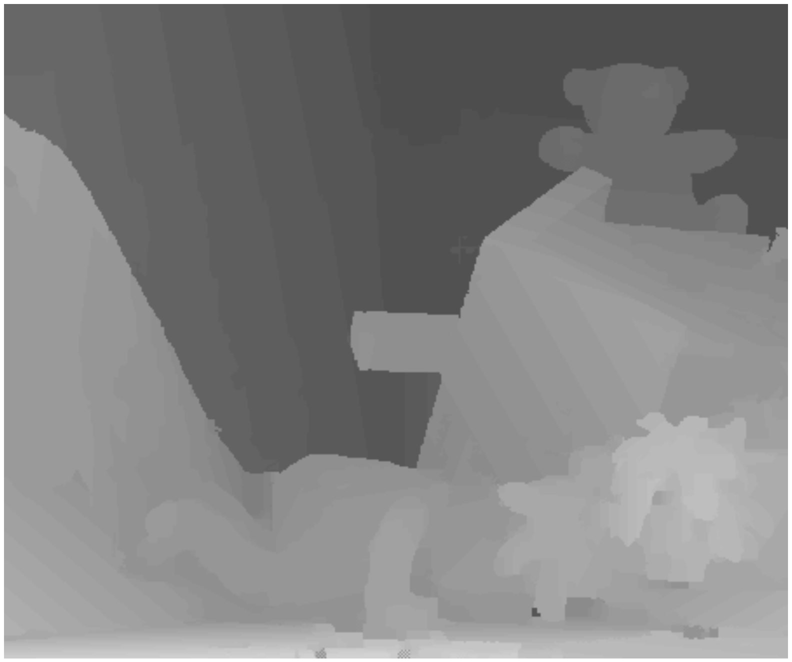


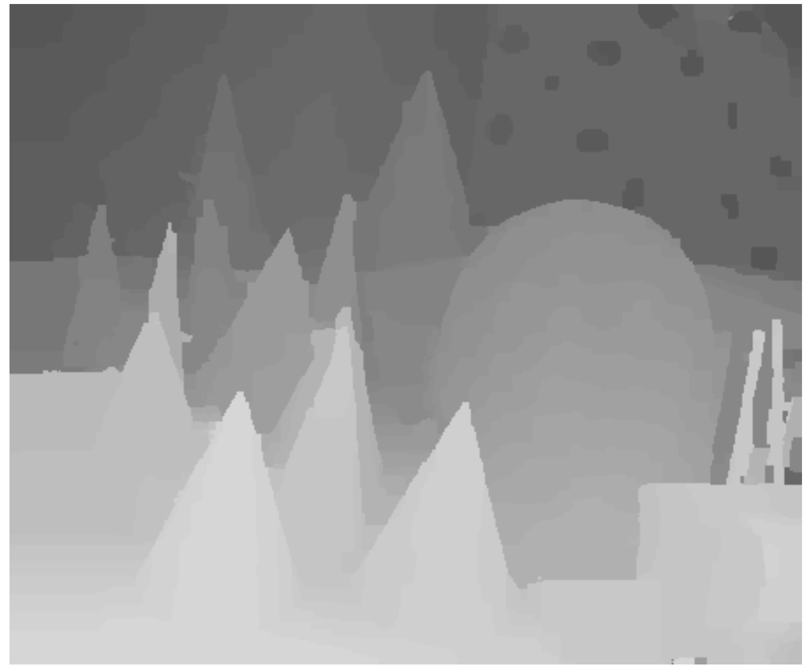
No. 1 in Middlebury evaluation
(June 2005)



No. 1 in Middlebury evaluation
(June 2005)

Results using Symmetric Belief Propagation





Bibliography

- D. Scharstein, Middlebury Stereo Evaluation Webpage, www.middlebury.edu/stereo
- D. Scharstein and R. Szeliski, A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms, IJCV 2002
- S. Birchfield and C. Tomasi, A pixel dissimilarity measure that is insensitive to image sampling, PAMI 1998
- Ramin Zabih and John Woodfill, Non-parametric Local Transforms for Computing Visual Correspondence, ECCV 1994
- M. Okutomi, Y. Katayama and S. Oka, A Simple Stereo Algorithm to Recover Precise Object Boundaries and Smooth Surfaces, IJCV 2002
- O. Veksler, Fast variable window for stereo correspondence using integral images, CVPR 2003

Bibliography

- J.C. Kim, K.M. Lee, B.T. Choi and S.U. Lee, A Dense Stereo Matching Using Two-Pass Dynamic Programming with Generalized Ground Control Points, CVPR 2005
- K.J. Yoon and I.S. Kweon, Locally Adaptive Support-Weight Approach for Visual Correspondence Search, CVPR 2005
- J. Sun, Y. Li, S.B. Kang and H.Y. Shum, Symmetric Stereo Matching for Occlusion Handling, CVPR 2005
- M. Gong and Y.H. Yang, Near Real-time Reliable Stereo Matching Using Programmable Graphics Hardware, CVPR 2005
- H. Hirschmüller, Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information, CVPR 2005

Bibliography

- Y. Boykov, O. Veksler and R. Zabih, Fast approximate energy minimization via graph cuts, PAMI 2001
- V. Kolmogorov and R. Zabih, Computing visual correspondence with occlusions via graph cuts, ICC 2001
- J. Sun, H.Y. Shum, and N.N. Zheng, Stereo matching using belief propagation, ECCV 2002
- J. Sun, H.Y. Shum, and N.N. Zheng, Stereo matching using belief propagation, PAMI 2003
- A. Ogale and Y. Aloimonos, Stereo correspondence with slanted surfaces: Critical implications of horizontal slant, CVPR 2004