CS4495/6495
Introduction to Computer Vision

3B-L3 *Stereo correspondence*

For now assume parallel image planes

- Assume parallel (co-planar) image planes…
- Assume same focal lengths…
- Assume epipolar lines are horizontal…
- Assume epipolar lines are at the same y location in the image…

- That’s a lot of assuming, but it allows us to move to the correspondence problem – which you will be solving!

Correspondence problem

Multiple match hypotheses satisfy *epipolar constraint*, but which is correct?

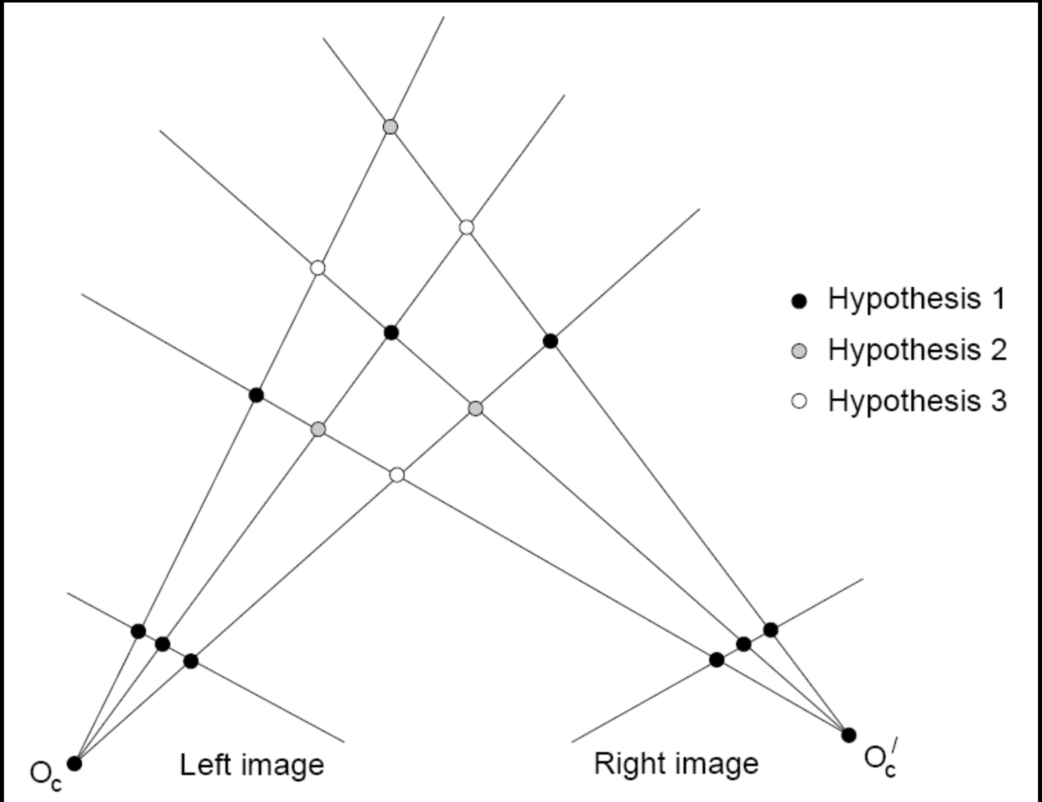


Figure from Gee & Cipolla 1999

Correspondence problem

Beyond the hard constraint of epipolar geometry, there are “soft” constraints to help identify corresponding points

- Similarity
- Uniqueness
- Ordering
- Disparity gradient is limited

Correspondence problem

Beyond the hard constraint of epipolar geometry, there are "soft" constraints to help identify corresponding points

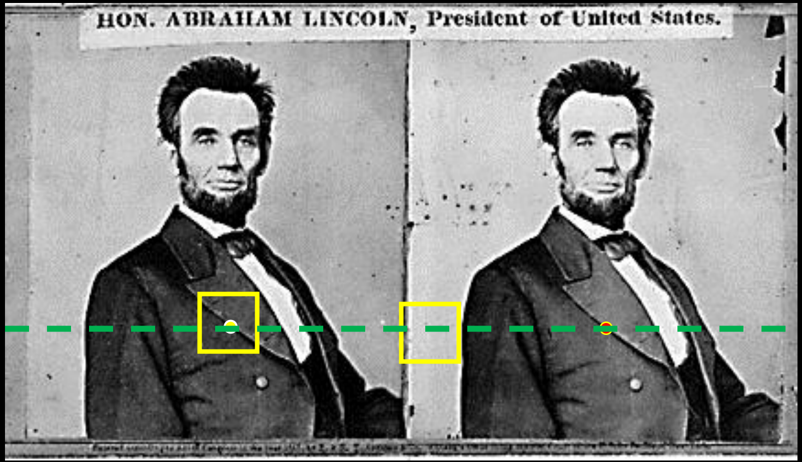
- **Similarity**
- Uniqueness
- Ordering
- Disparity gradient is limited

Correspondence problem

To find matches in the image pair, we will assume

- Most scene points visible from both views
- Image regions for the matches are similar in appearance

Dense correspondence search

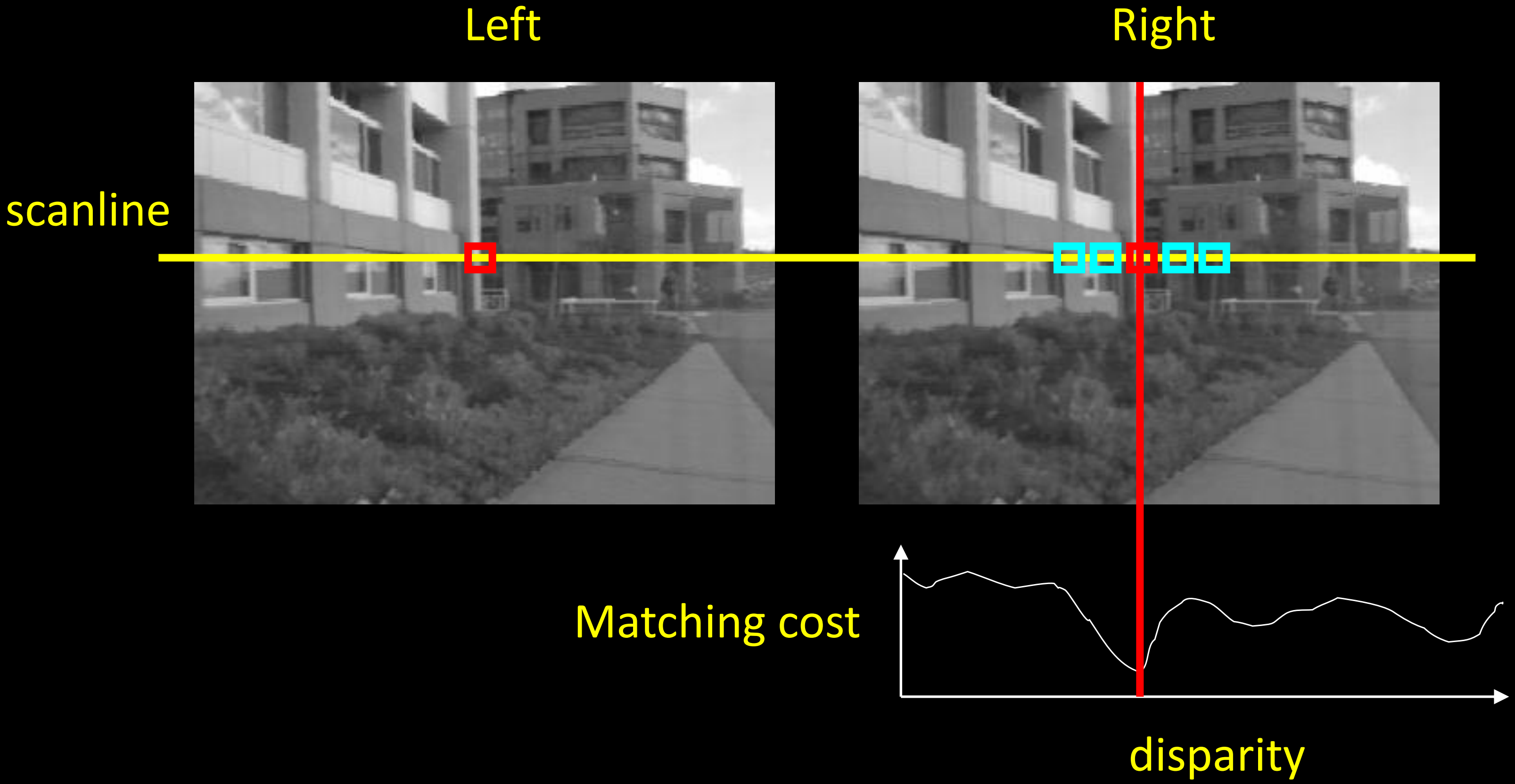


For each pixel / window in the left image

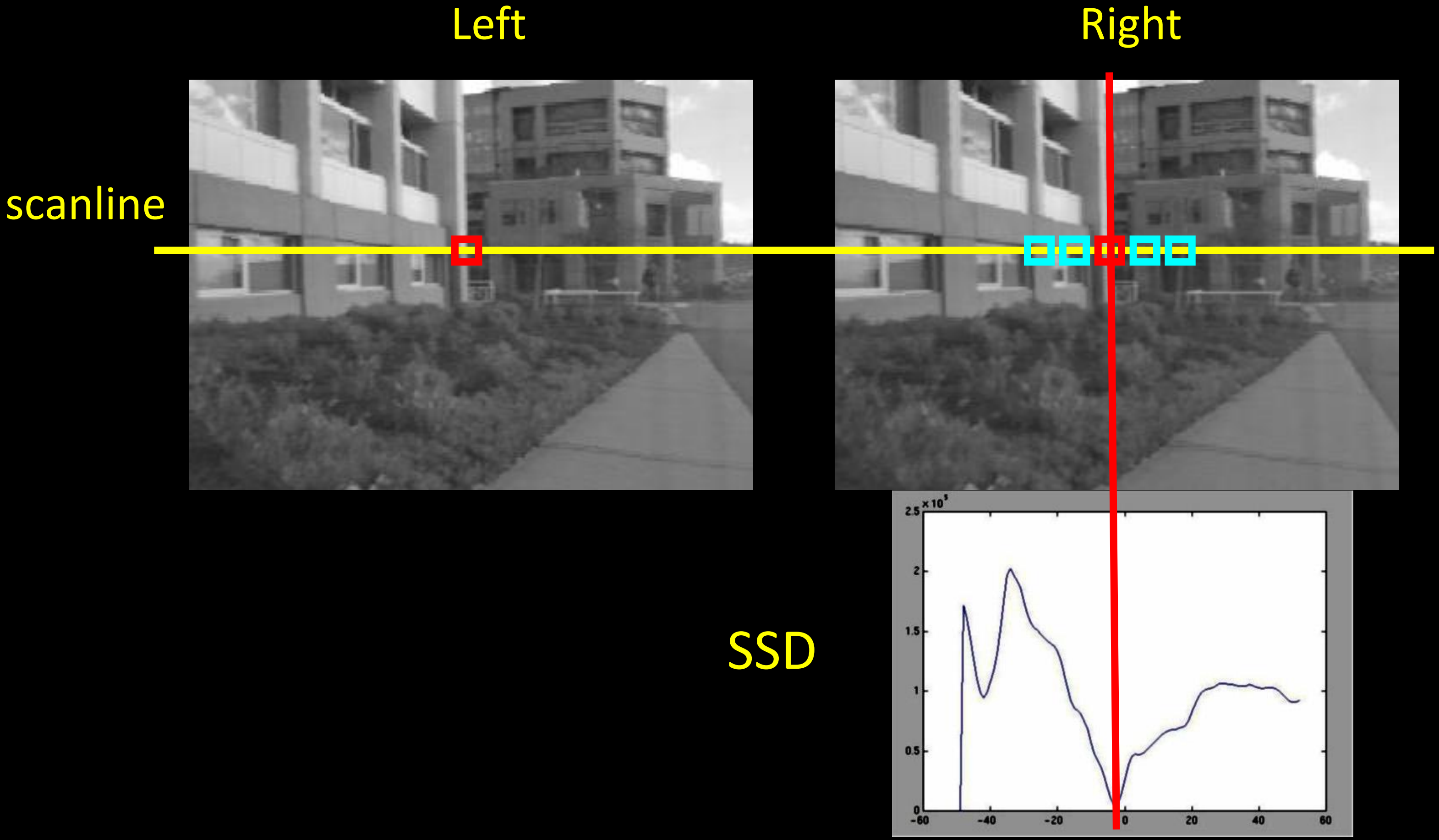
- Compare with every pixel / window on same epipolar line in right image
- Pick position with minimum match cost (e.g., SSD, normalized correlation)

Adapted from Li Zhang

Correspondence search: (dis)similarity constraint



Correspondence search: (dis)similarity constraint

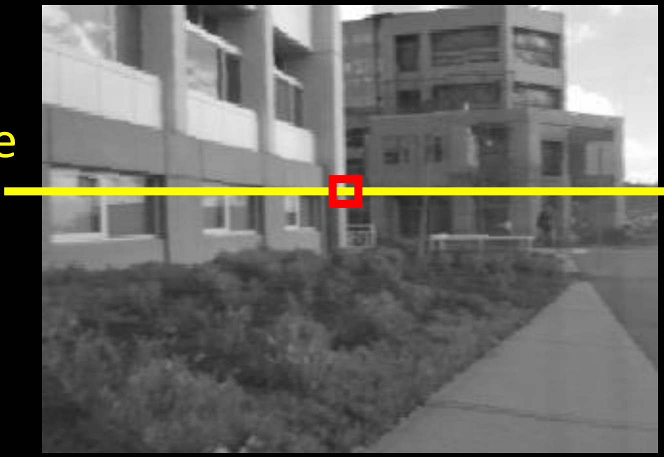


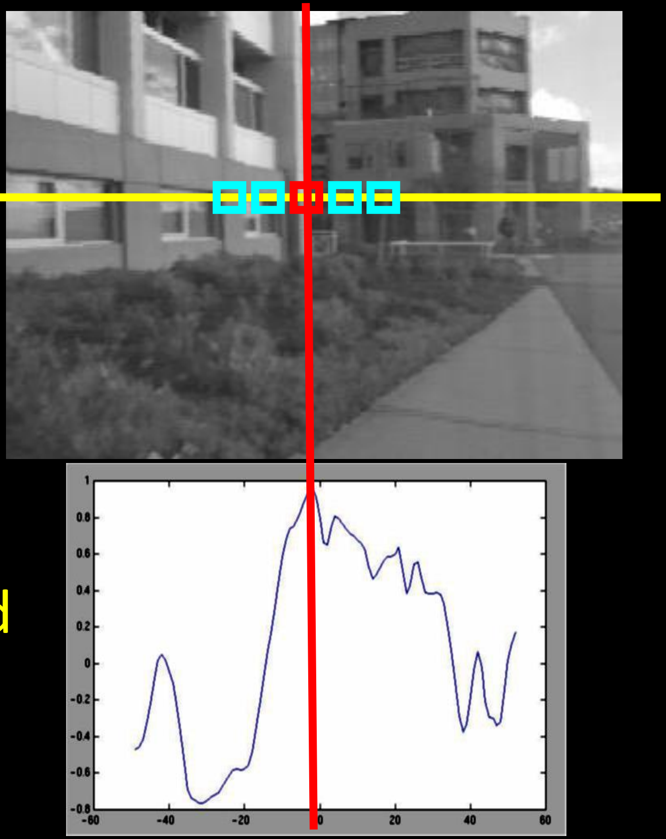
Correspondence search: similarity constraint

Left

Right

scanline



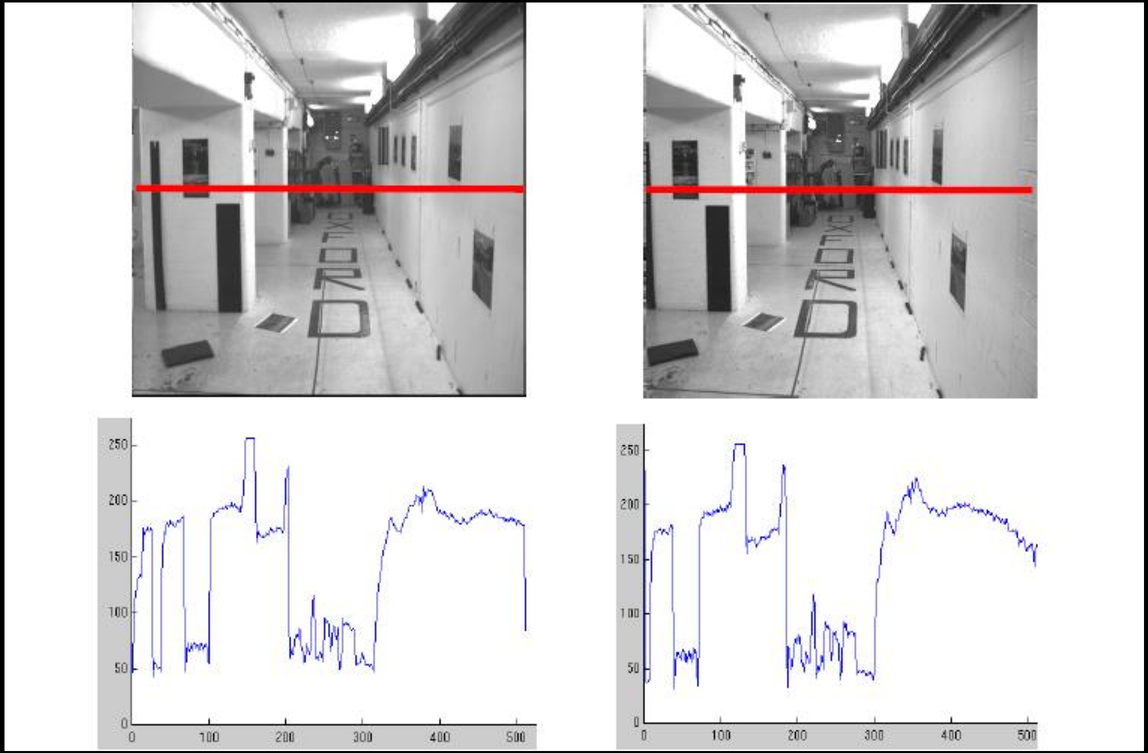


Normalized correlation

Correspondence problem

Source: Andrew Zisserman

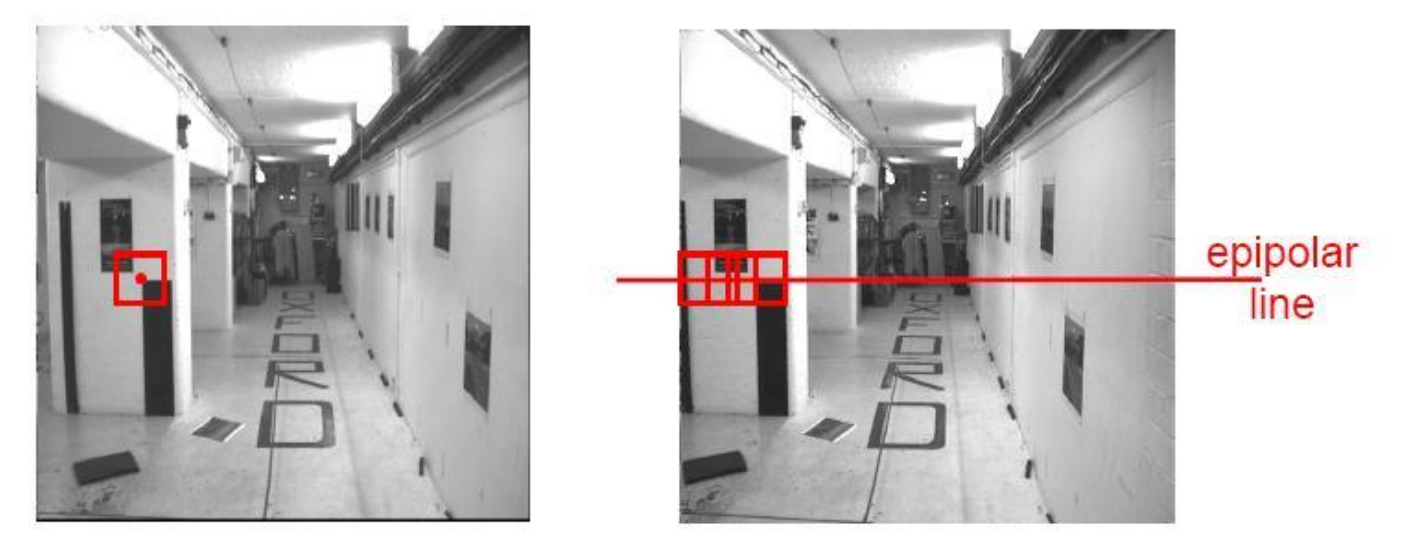
Correspondence problem



Intensity
profiles

Source: Andrew Zisserman

Correspondence problem



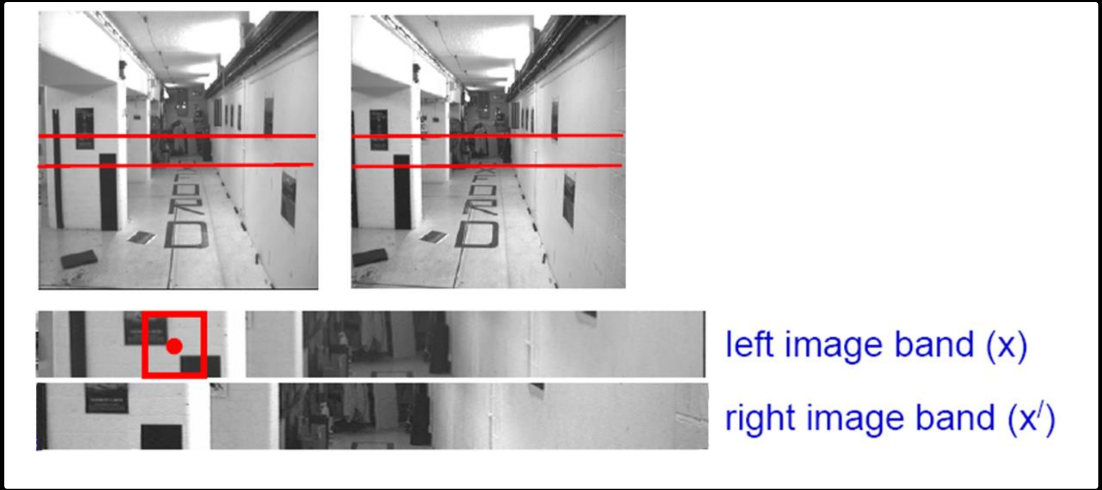
Neighborhoods of corresponding points are similar in intensity patterns.

Source: Andrew Zisserman

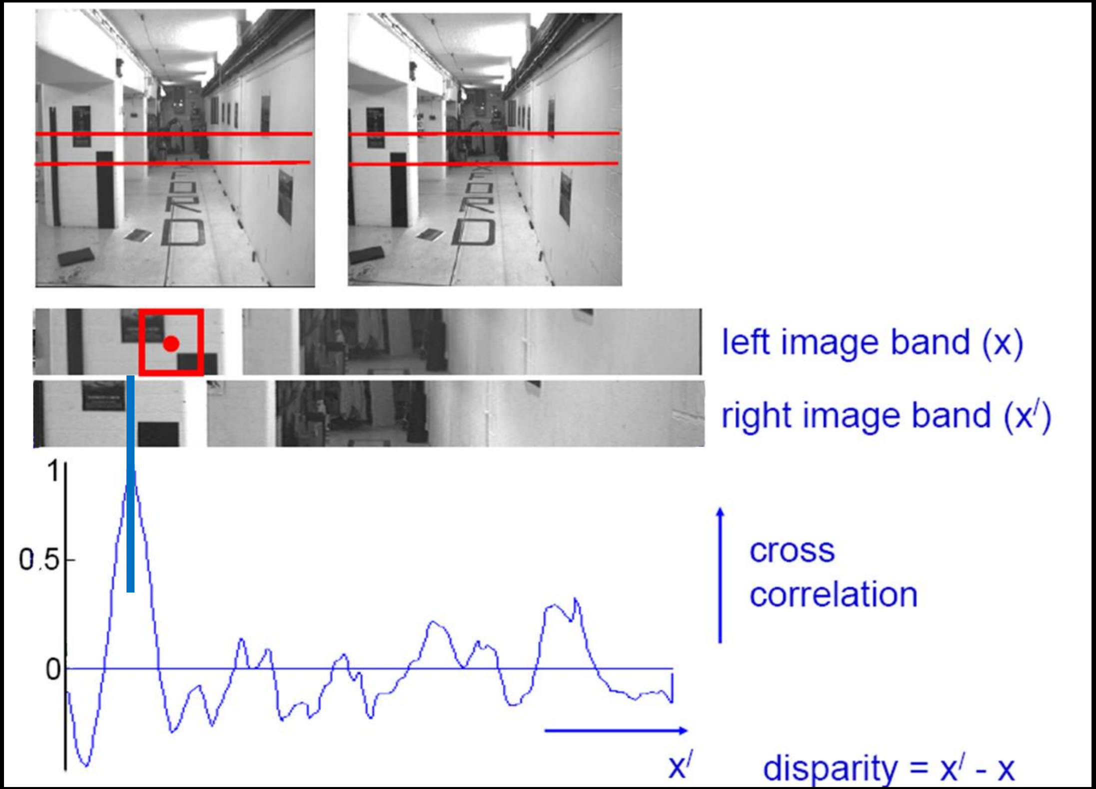
Correlation-based window matching

left image band (x)

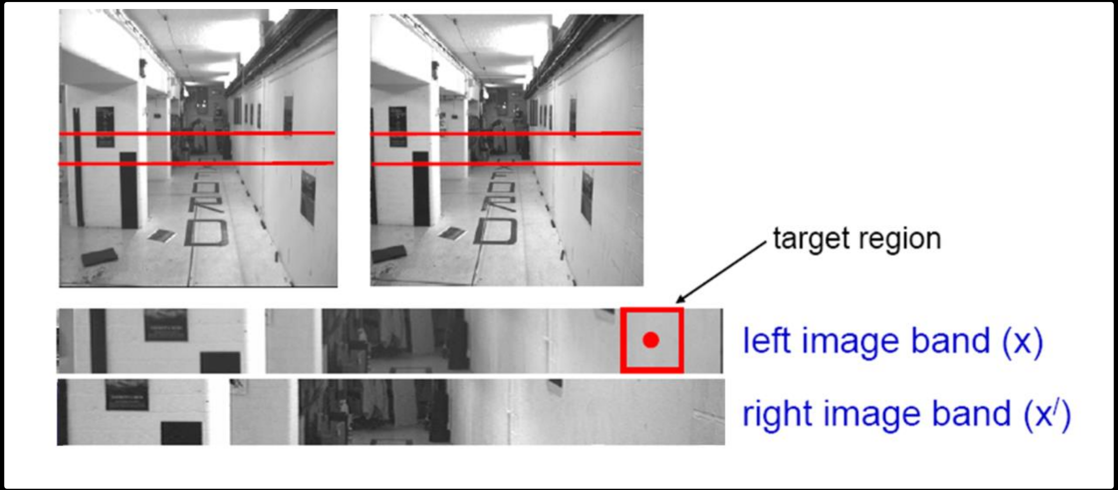
Correlation-based window matching



Correlation-based window matching



Correlation-based window matching



Correlation-based window matching

Effect of window size

Source: Andrew Zisserman

Effect of window size

W = 3

W = 20

Figures from Li Zhang

Correspondence problem

Beyond the hard constraint of epipolar geometry, there are “soft” constraints to help identify corresponding points

- Similarity
- Uniqueness
- Ordering
- Disparity gradient is limited

Uniqueness constraint

No more than one match in right image for every point in left image

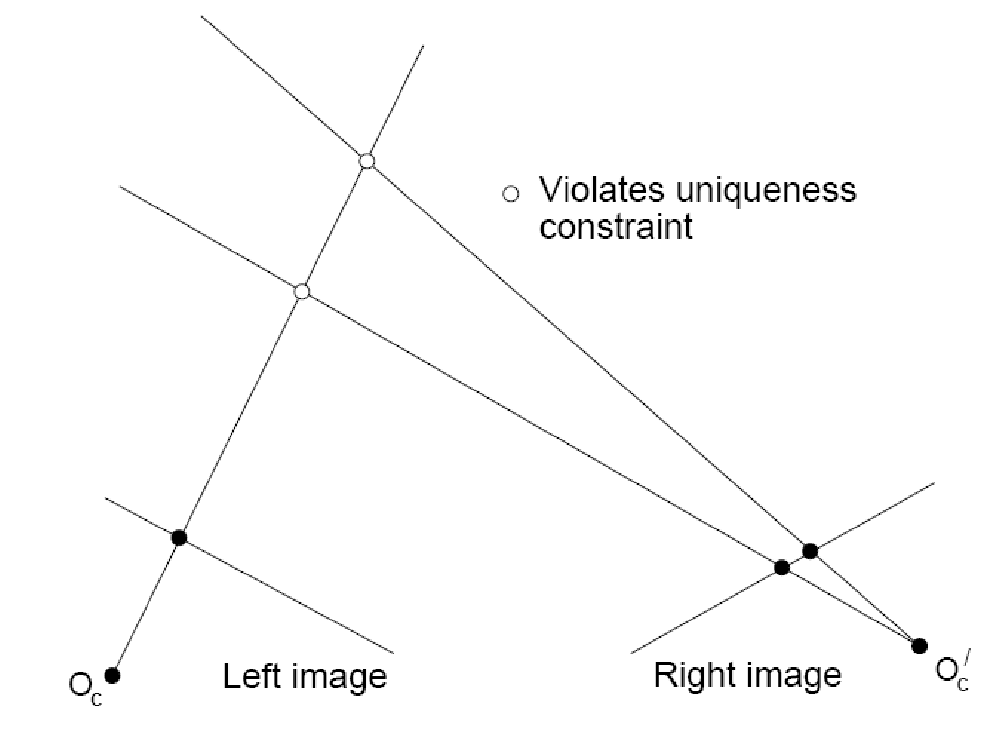
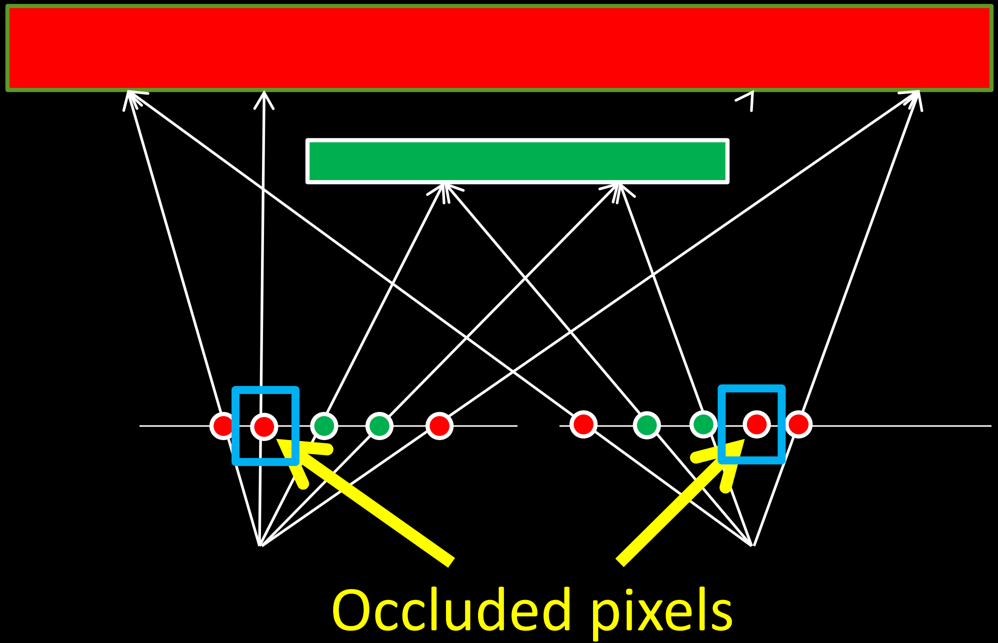


Figure from Gee & Cipolla 1999

Problem: Occlusion



Ordering constraint

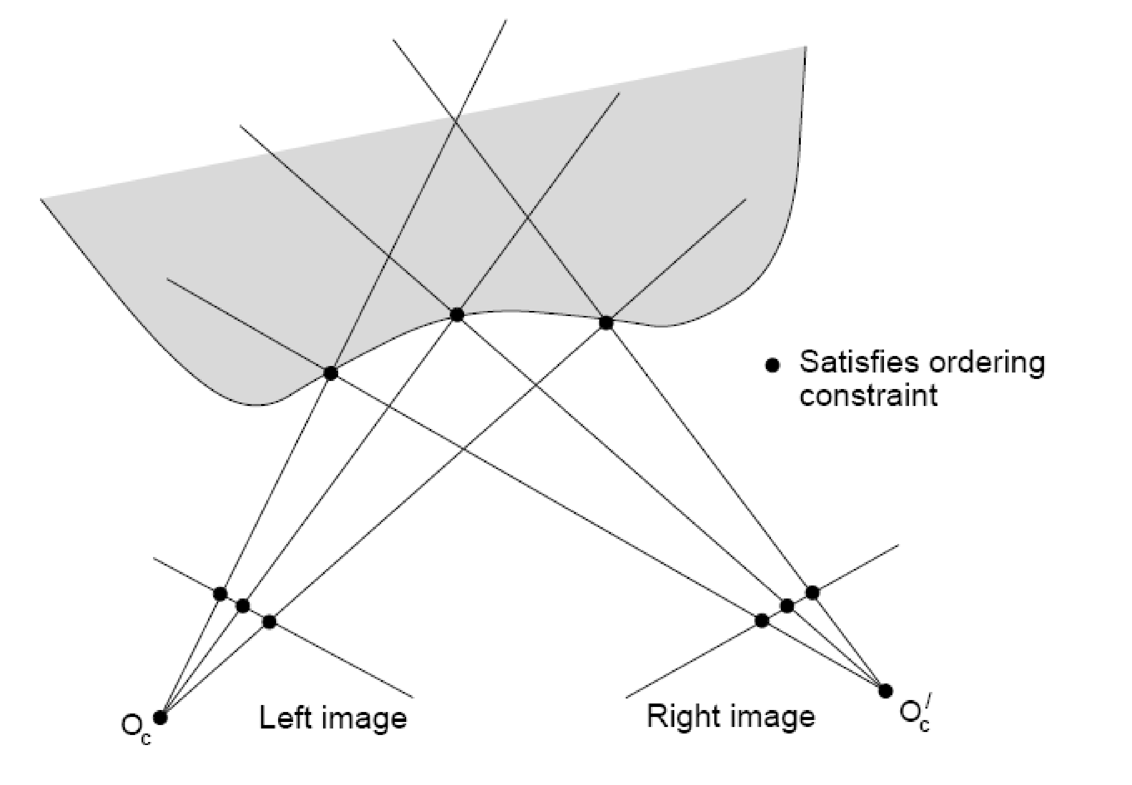
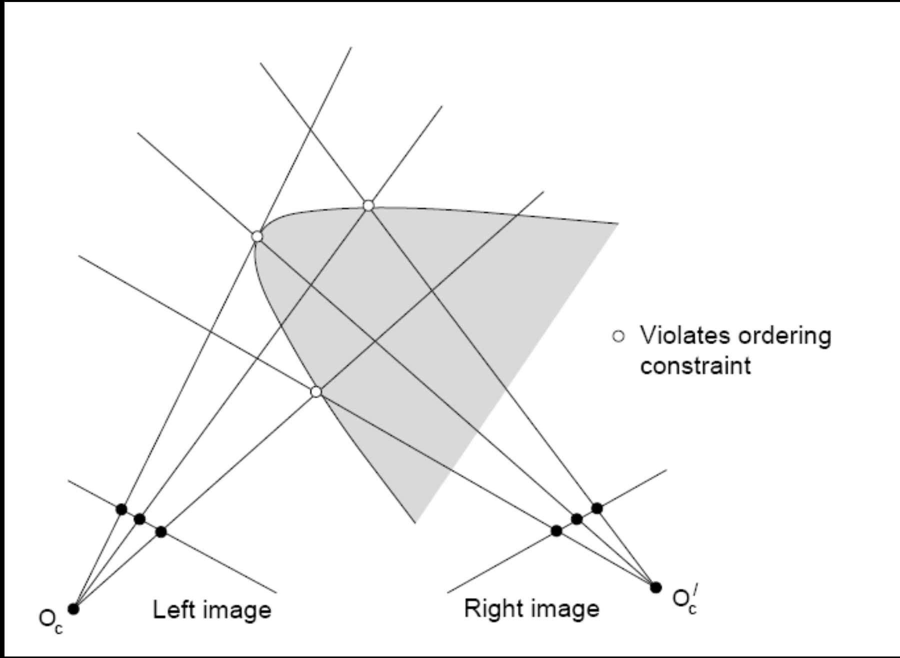


Figure from Gee & Cipolla 1999

Ordering constraint

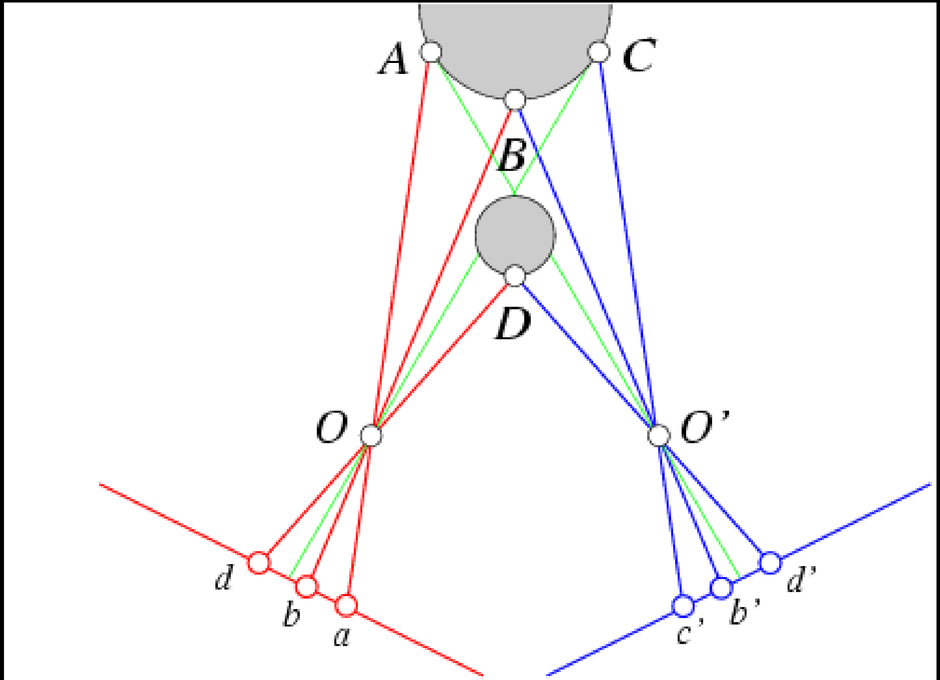
Won't always hold, e.g. consider transparent object...



Figures from Forsyth & Ponce

Ordering constraint

…or a narrow occluding surface

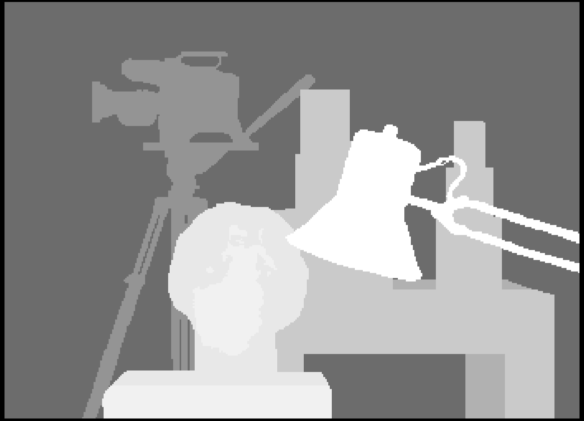


Figures from Forsyth & Ponce

Stereo results

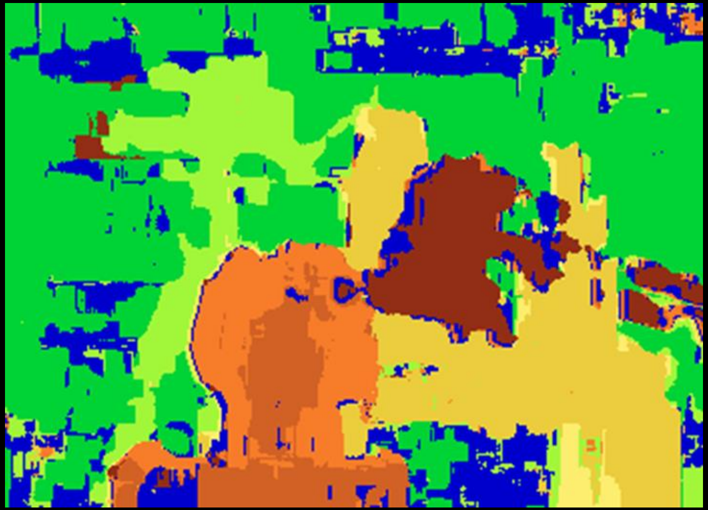
Image data from University of Tsukuba

Scene



Ground truth

Results with window search



Window-based matching
(best window size)

Ground truth

Better solutions

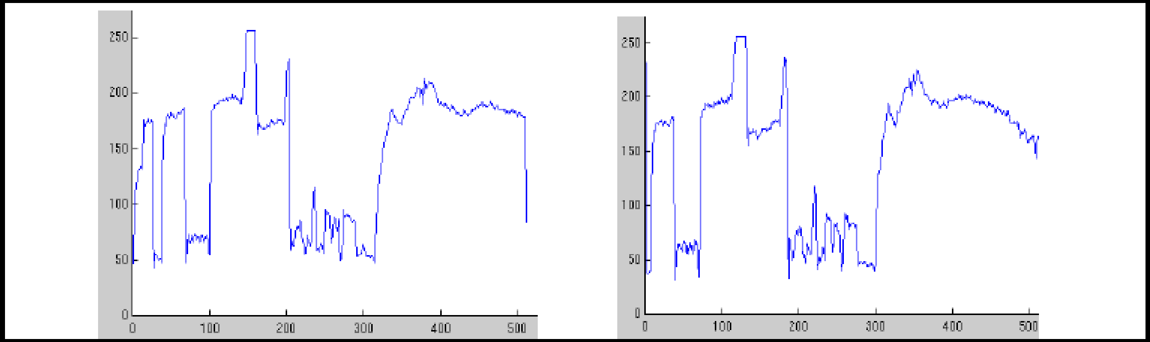
Beyond individual correspondences to estimate disparities:

- Optimize correspondence assignments jointly
  - Scanline at a time (DP)
  - Full 2D grid (graph cuts)

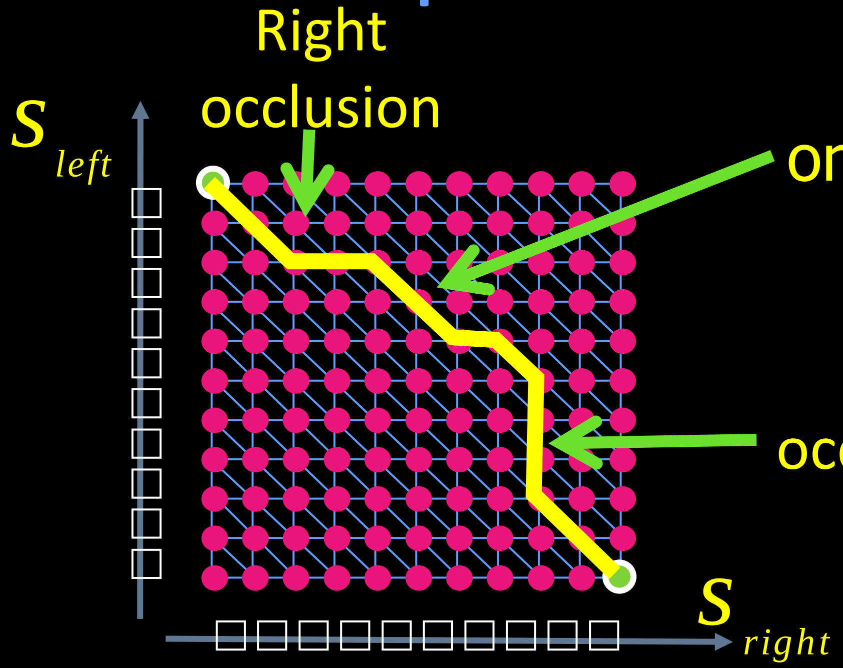
Scanline stereo

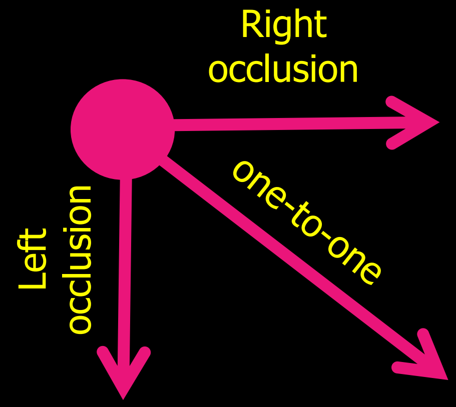
Coherently match pixels on the entire scanline

intensity



“Shortest paths” for scan-line stereo





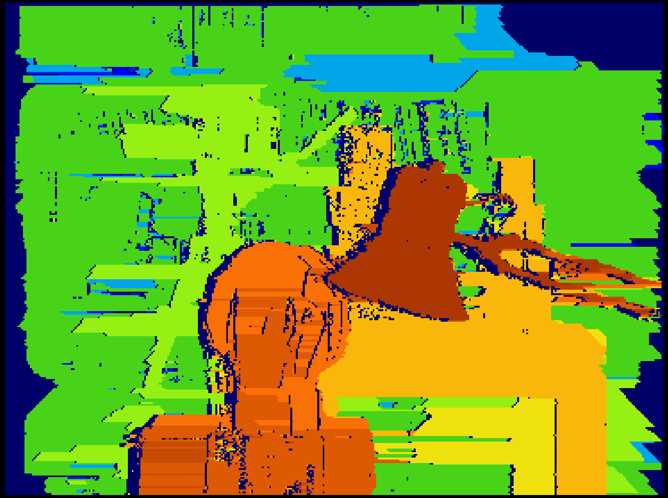
Can be implemented with dynamic programming

Ohta & Kanade ’85, Cox et al. ’96, Intille & Bobick, ‘01

Slide: Y. Boykov

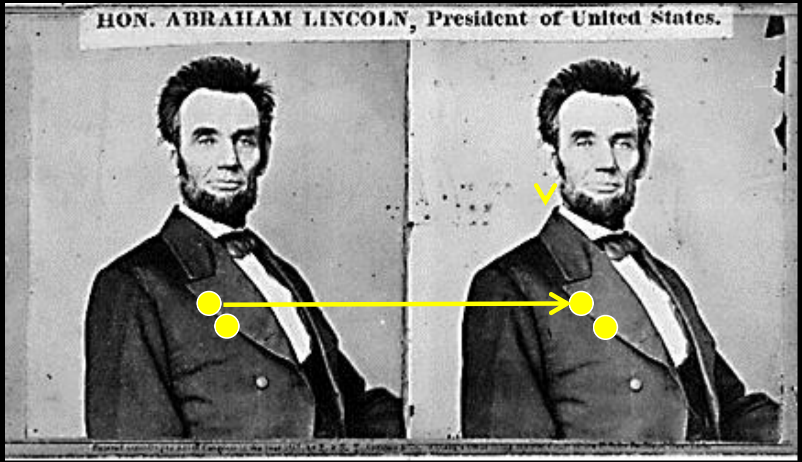
Coherent stereo on 2D grid

Scanline stereo generates streaking artifacts



Can’t use dynamic programming to find spatially coherent disparities/ correspondences on a 2D grid

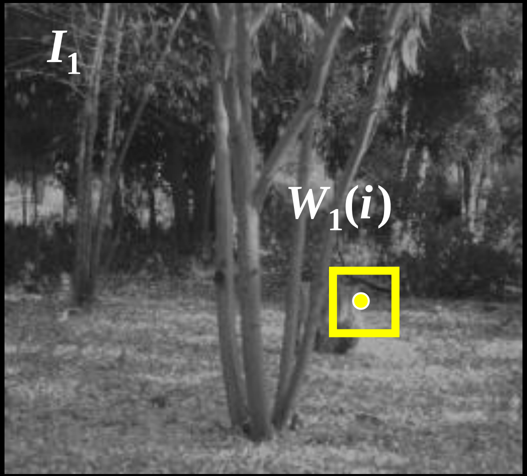
Stereo as energy minimization

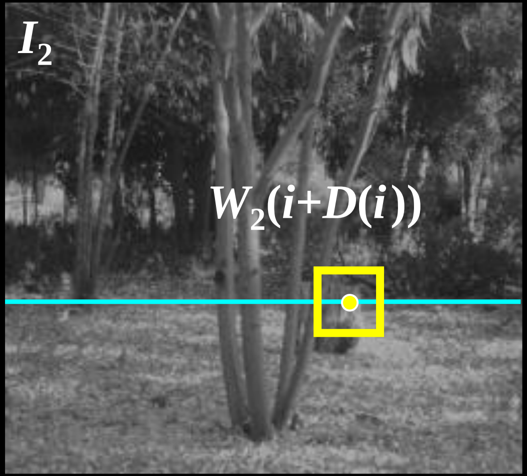


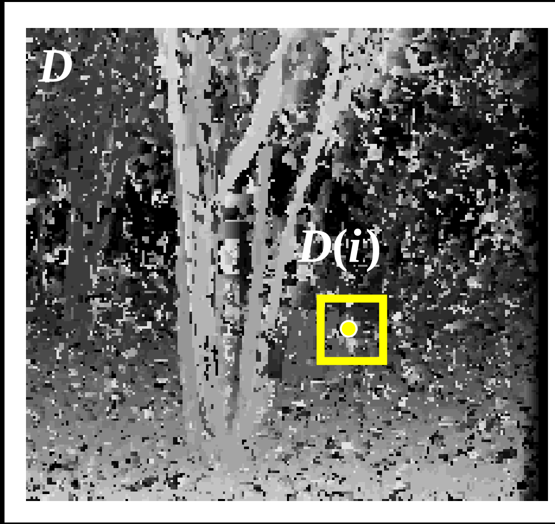
What defines a good stereo correspondence?

1. ***Match quality*** - Want each pixel to find a good appearance match in the other image
2. ***Smoothness*** - of two pixels are adjacent, they should (usually) move about the same amount

Stereo matching as energy minimization



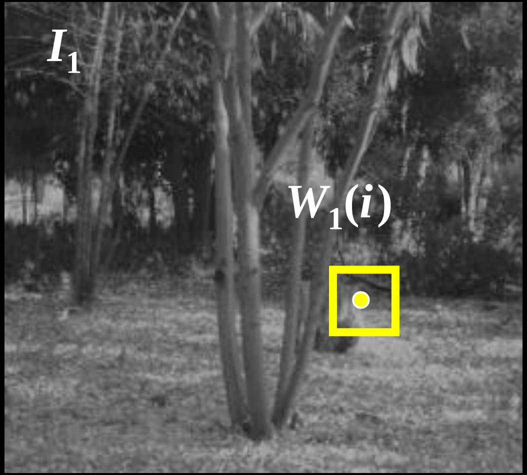


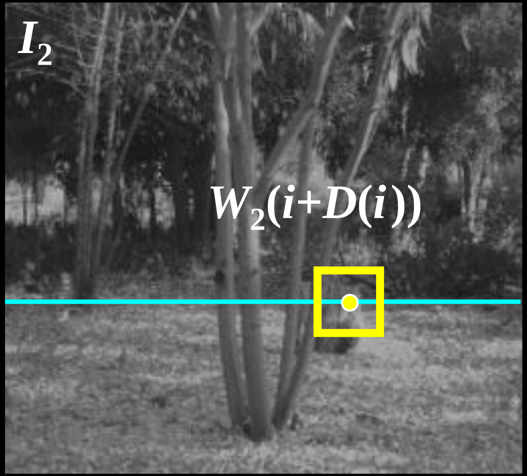


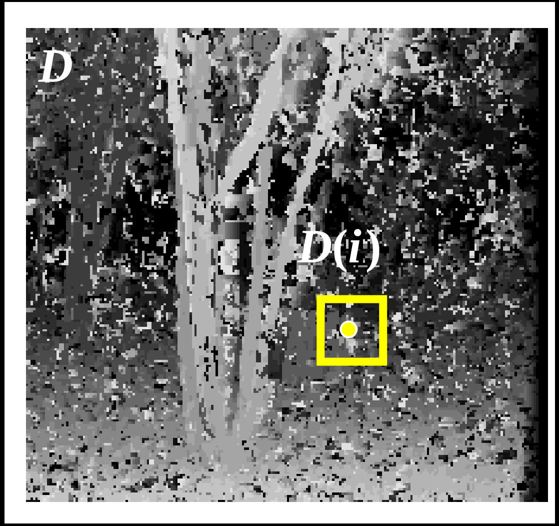
Data term: $\boxed{E_{\text{data}} = \sum_i \left(W_1(i) - W_2(i + D(i))\right)^2}$

Source: Steve Seitz

Stereo matching as energy minimization





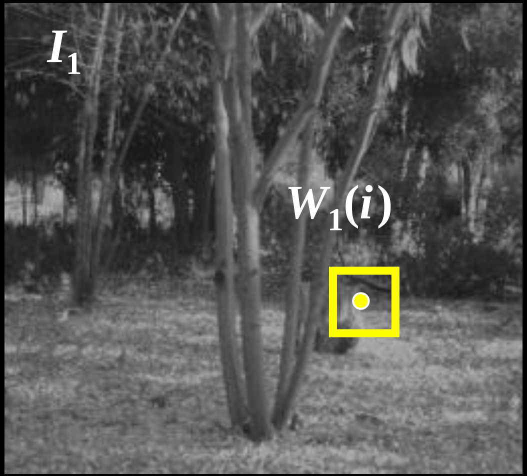


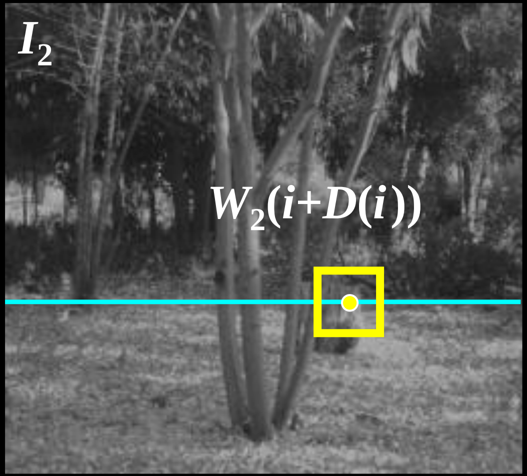
Smoothness term:

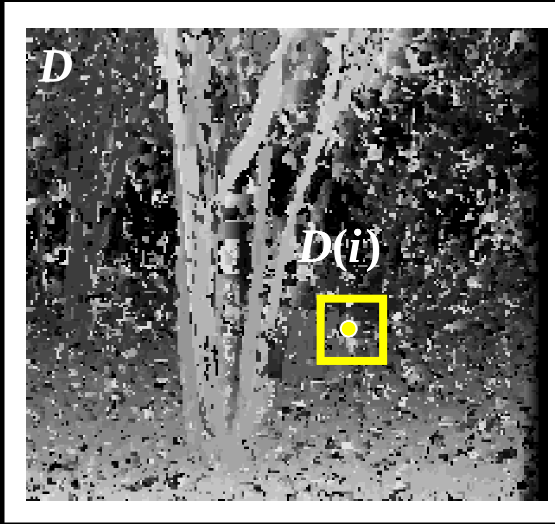
$$\boxed{E_{\text{smooth}} = \sum_{\text{neighbors } i,j} \rho\left(D(i) - D(j)\right)}$$

Source: Steve Seitz

Stereo matching as energy minimization







Total energy: $\boxed{E = \alpha E_{\text{data}}(I_1, I_2, D) + \beta E_{\text{smooth}}(D)}$

Source: Steve Seitz

Better results…

- Energy functions of this form can be minimized using *graph cuts*
- Y. Boykov, O. Veksler, and R. Zabih, Fast Approximate Energy Minimization via Graph Cuts, PAMI 2001

Better results…

State of the art method

Ground truth

For the latest and greatest: http://www.middlebury.edu/stereo

Challenges

- Low-contrast ; textureless image regions
- Occlusions
- Violations of brightness constancy (e.g., specular reflections)
- Really large baselines (foreshortening and appearance change)
- Camera calibration errors