Computer Vision

10. Rectification and Depth

I-Chen Lin

College of Computer Science
National Yang Ming Chiao Tung University

Objective

- Image rectification
- Depth reconstruction

Related books:

• David A. Forsyth and Jean Ponce, Computer Vision: A Modern Approach, Prentice Hall, New Jersey, (1st Ed. 2003, 2nd Ed. 2012).

• R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision 2nd Ed., Cambridge University Press, 2004.

Plenty of slides are modified from the reference lecture notes or project pages:

• Prof. J. Rehg, Computer Vision, Georgia Inst. of Tech.

• Dr. Ng Teck Khim, Computer Vision and Graphics for Special Effects lecture notes.

• Prof. D.A. Forsyth, Computer Vision, UIUC.

• Prof. D. Lowe, Computer Vision, UBC, CA.

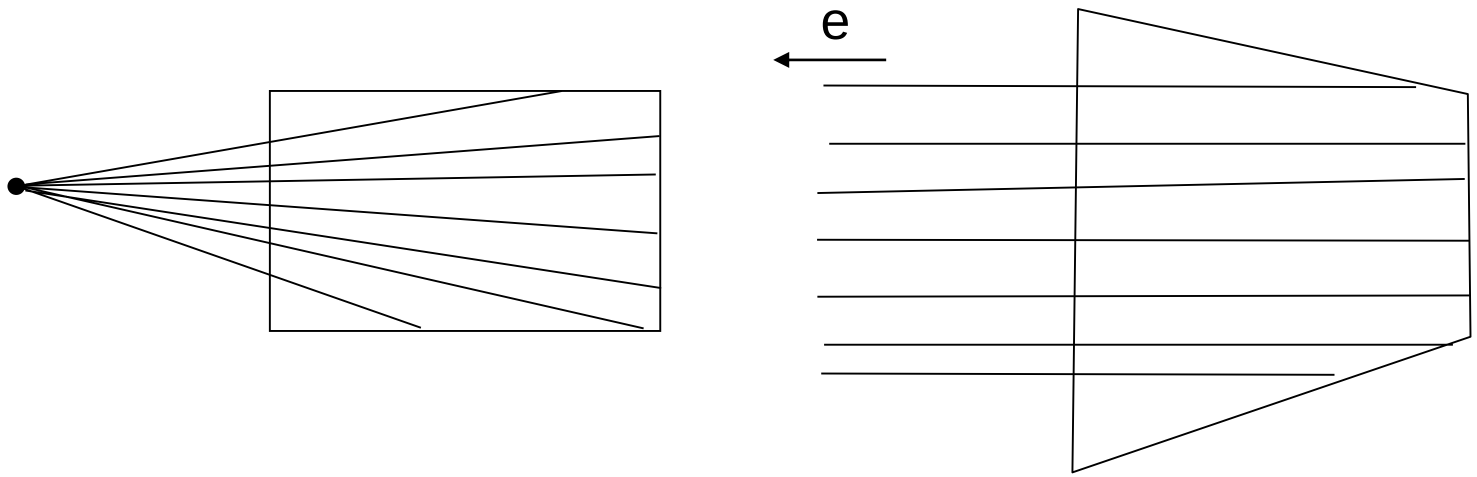
• Prof. T. Darrell, Computer Vision and Applications, MIT.

• Prof. M. Pollefeys, Computer Vision, UNC.

• S. M. Seitz, C. R. Dyer, "View Morphing", Proc. SIGGRAPH'96, pp. 21-30.

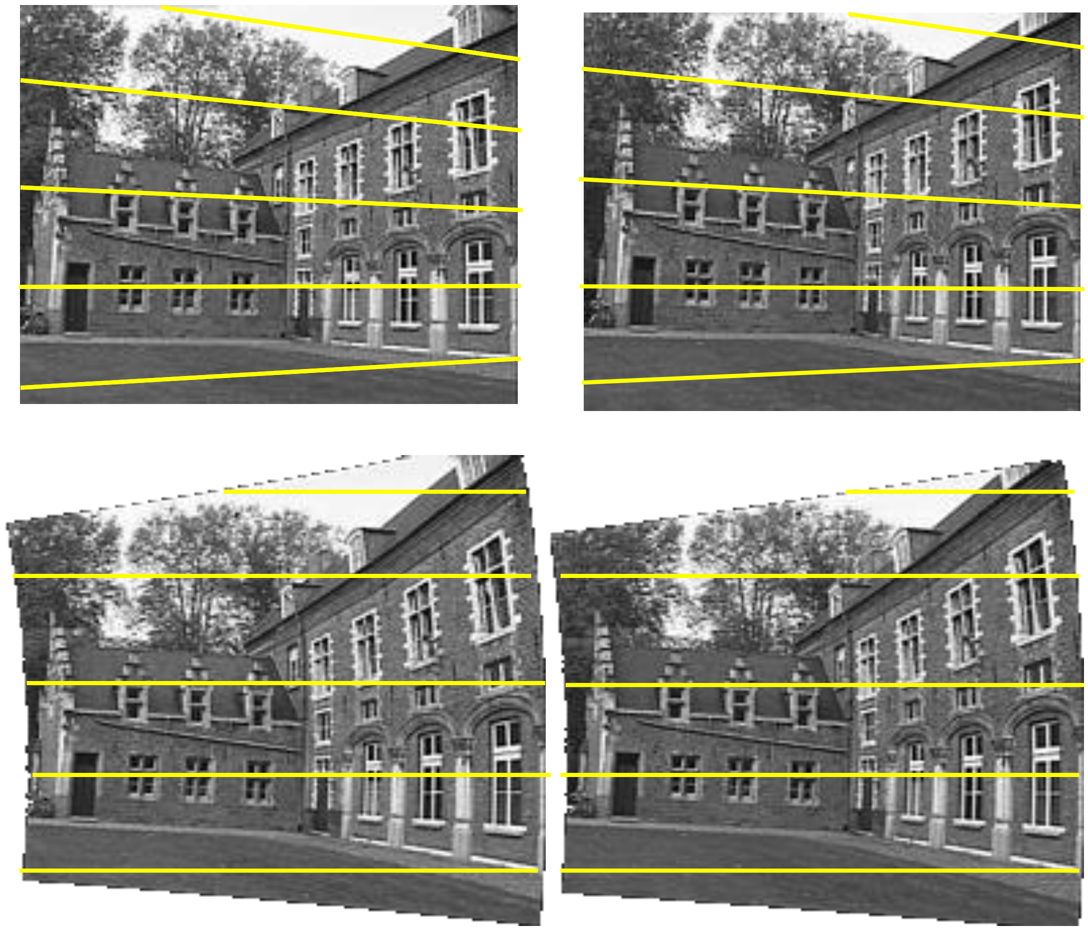
Image pair rectification

- Apply projective transformation so that epipolar lines correspond to horizontal scanlines.
- Simplify stereo matching by warping the images.
  - map epipole e to (1,0,0).



Questions

- Can we apply warping for image rectification problem?

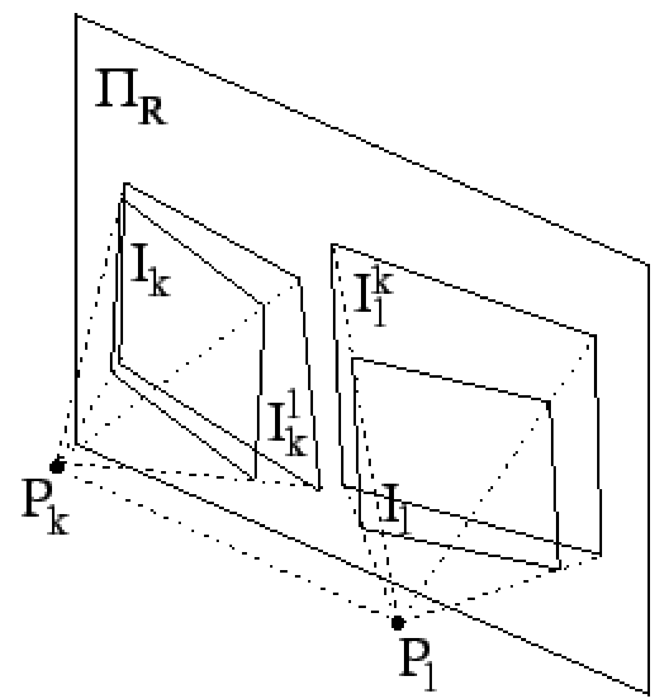


Estimating the matrix H

- From non-parallel → parallel.
  - The transformation is not unique.

  - Planar rectification in computer vision.
    - Rotate the epipole.

$$[1 \quad 0 \quad 0]^T = He, \text{ where e is the epipole}$$



Estimating the matrix H

- Hint: Rotate the "view direction" and make epipoles perpendicular to optic axis.

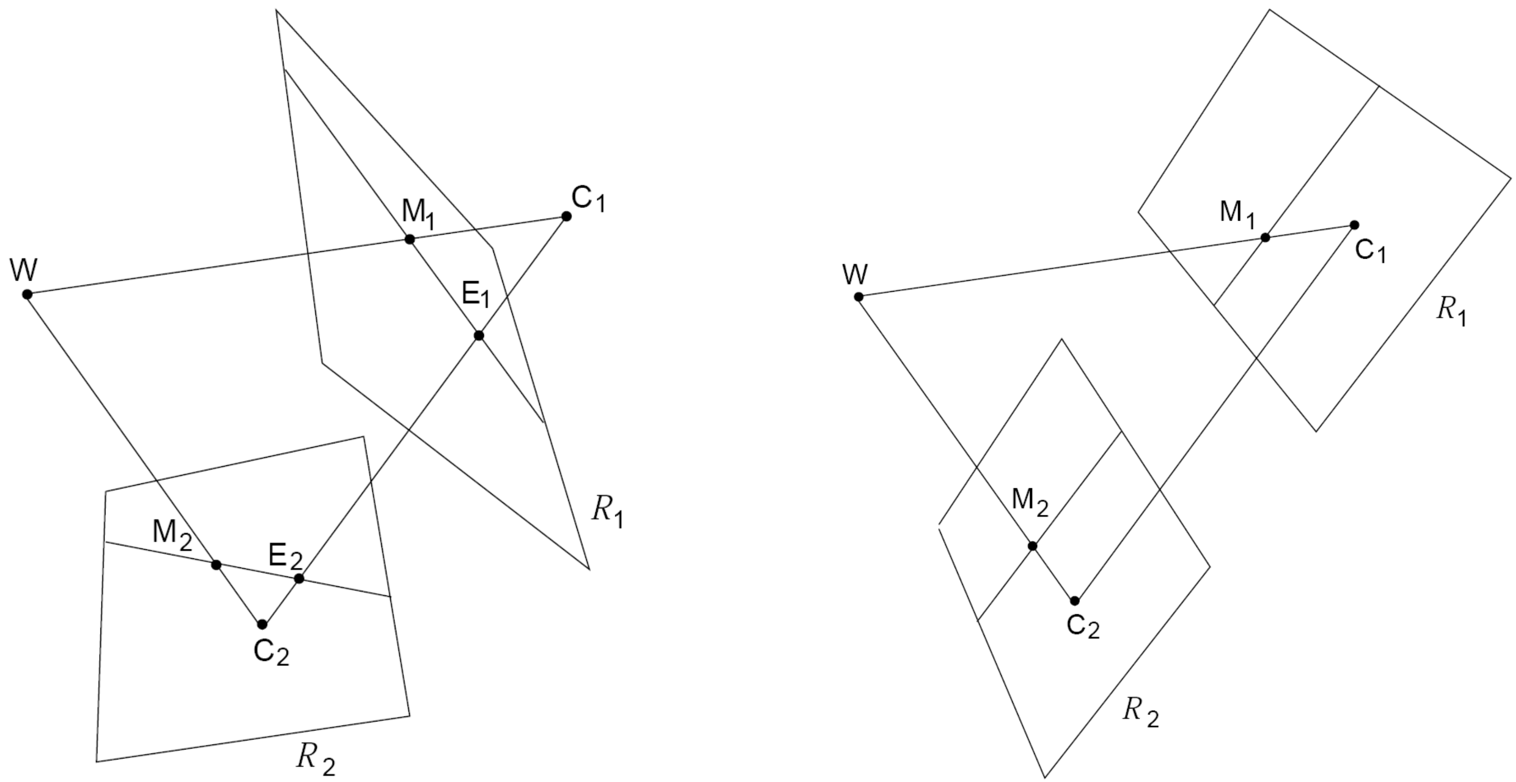


Figure from A. Fusiello, et al., A compact algorithm for rectification of stereo pairs. Machine Vision and Applications, 12(1):16-22, 2000.

(Appendix) Estimating the matrix H (cont.)

- Solutions suggested by the authors of view morphing.
  1. Rotate the epipole around an axis d_i. $R_{\theta_i}^{d_i}$
     - Make image planes parallel.
  2. Rotate the image plane R_{ϕ_i}
     - Make epipolar lines parallel.
  3. Align the epipolar line T

$$\boxed{W_0 = R_{\phi_0} R_{\theta_0}^{d_0}} \qquad \boxed{W_1 = T R_{\phi_1} R_{\theta_1}^{d_1}}$$

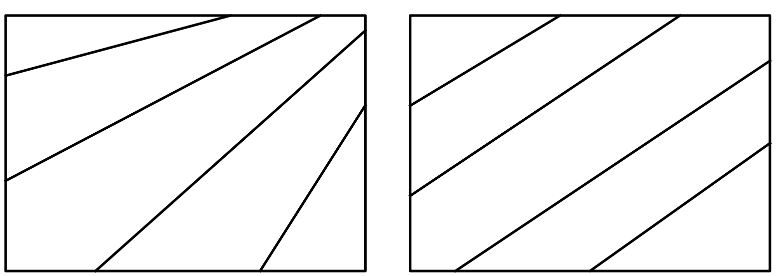
(Appendix) Estimating the matrix H (I)

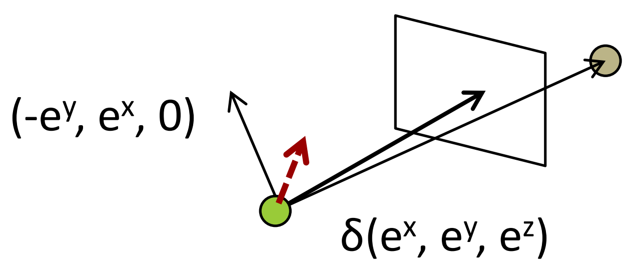
- Rotate around axis $d_0 = \begin{bmatrix} d_0^x & d_0^y & 0 \end{bmatrix}^T$ E.g. $d_0 = \begin{bmatrix} -e_0^y & e_0^x & 0 \end{bmatrix}^T$

$$e_0 = \begin{bmatrix} e_0^x & e_0^y & e_0^z \end{bmatrix}^T$$

e_0 is the unit eigenvector of F corresponding to eigenvalue 0.

$$R_{\theta_0}^{d_0} = \begin{bmatrix} \left(d_0^x\right)^2 + \left(1 - \left(d_0^x\right)^2\right)\cos\theta_0 & d_0^x d_0^y \left(1 - \cos\theta_0\right) & d_0^y \sin\theta_0 \\ d_0^x d_0^y \left(1 - \cos\theta_0\right) & \left(d_0^y\right)^2 + \left(1 - \left(d_0^y\right)^2\right)\cos\theta_0 & -d_0^x \sin\theta_0 \\ -d_0^y \sin\theta_0 & d_0^x \sin\theta_0 & \cos\theta_0 \end{bmatrix}$$



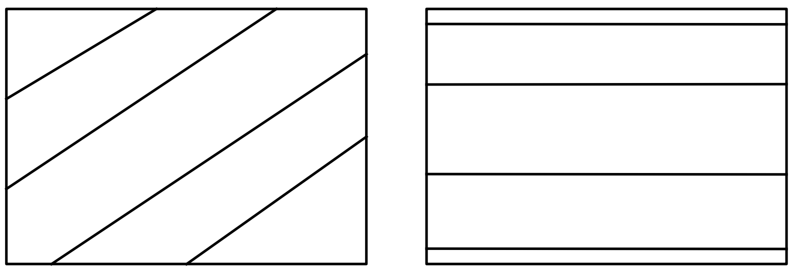


(Appendix) Estimating the matrix H (II)

- The new epipoles are $\begin{bmatrix}\tilde{e}_i^{\,x} & \tilde{e}_i^{\,y} & 0\end{bmatrix}^T = R_{\theta_i}^{d_i} e_i$

- Rotate the image plane so that epipolar lines are horizontal.

$$R_{\phi_i} = \begin{bmatrix} \cos\phi_i & -\sin\phi_i & 0 \\ \sin\phi_i & \cos\phi_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \phi_i = -\tan^{-1}\left(\frac{\tilde{e}_i^{\,y}}{\tilde{e}_i^{\,x}}\right)$$



(Appendix) Estimating the matrix H (III)

$$\tilde{F} = \left(\left(R_{\phi_1} R_{\theta_1}^{d_1}\right)^{-1}\right)^{T} F \left(R_{\phi_0} R_{\theta_0}^{d_0}\right)^{-1}$$

$$= R_{\phi_1} R_{\theta_1}^{d_1} F R_{-\theta_0}^{d_0} R_{-\phi_0} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & a \\ 0 & b & c \end{bmatrix}$$

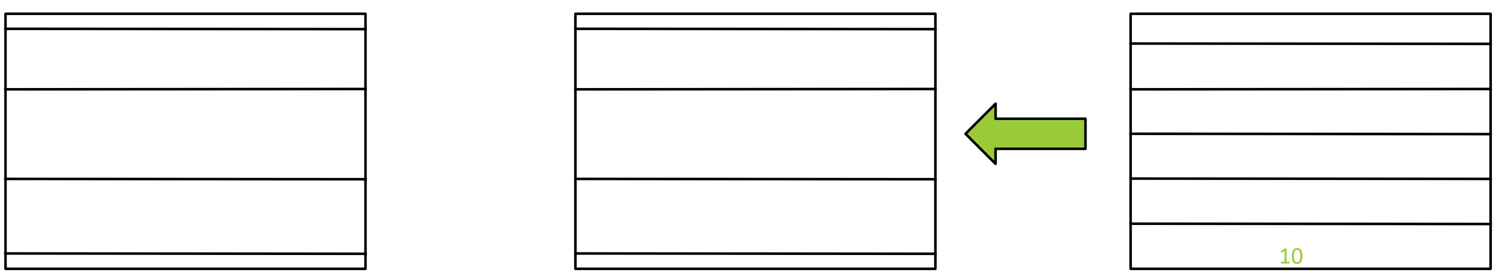
$$\boxed{W_0 = R_{\phi_0} R_{\theta_0}^{d_0}}$$

$$\boxed{W_1 = T R_{\phi_1} R_{\theta_1}^{d_1}}$$

(*The T used in view morphing*)

Our target → $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$

Multiply : $T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -a & -c \\ 0 & 0 & b \end{bmatrix}$



If we would like to keep the scene geometry invariant, ….

(Appendix) Estimating the matrix H (IV)

$$\hat{F} = \left(\left(TR_{\phi_1} R_{\theta_1}^{d_1}\right)^{-1}\right)^T F \left(R_{\phi_0} R_{\theta_0}^{d_0}\right)^{-1}$$

$$= (T^{-1})^T R_{\phi_1} R_{\theta_1}^{d_1} F R_{-\theta_0}^{d_0} R_{-\phi_0} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

The ^F and T used in view morphing

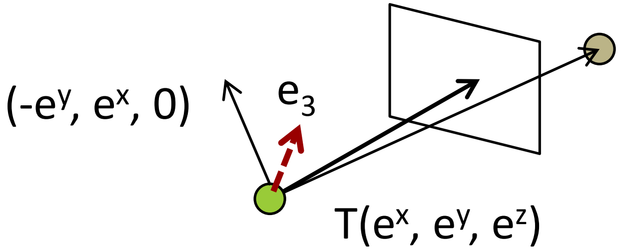
H by coordinate transformation

- We can use coordinate transformation and calculate only one rotation matrix.
- Assume e_1 is the vector pointing to the epipole, and e_1 , e_2 , e_3 are orthogonal to each other.

$$\boldsymbol{e_1} = \frac{\boldsymbol{T}}{\|\boldsymbol{T}\|} \qquad e_2 = \frac{1}{\sqrt{T_x{}^2 + T_y{}^2}}[-T_y, T_x, 0]^T \qquad e_3 = e_1 \times e_2$$

- Since we would like $e_1 \rightarrow x'$, $e_2 \rightarrow y'$, $e_3 \rightarrow z'$

$$R_{rect} = \begin{pmatrix} e_1{}^T \\ e_2{}^T \\ e_3{}^T \end{pmatrix}$$

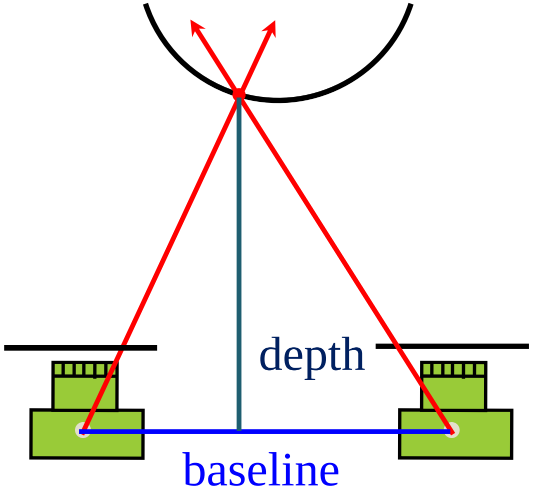


Dense depth reconstruction

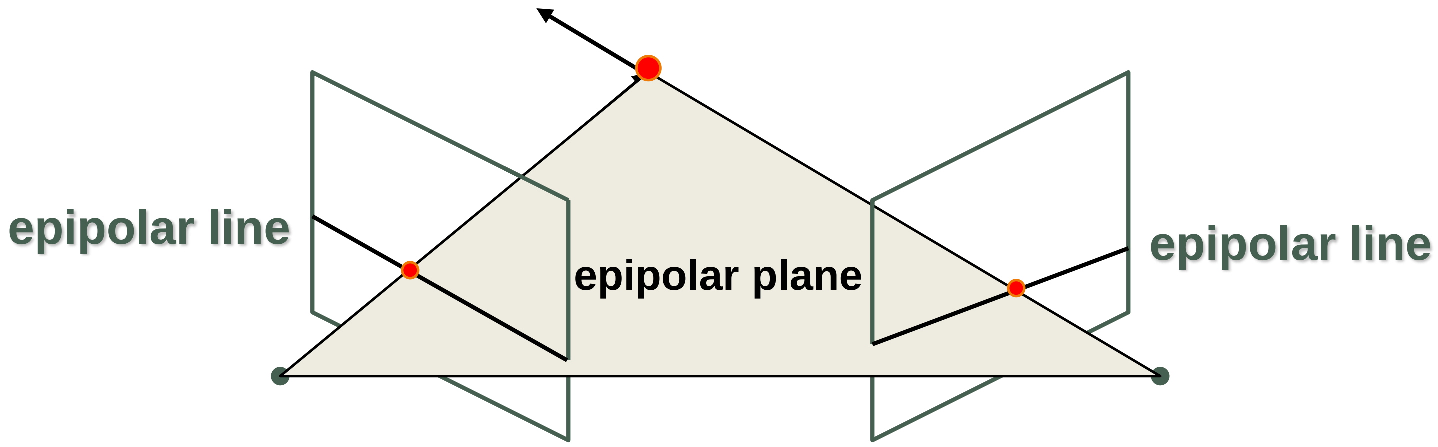
- Given sparse and reliable point correspondences, we can evaluate the fundamental matrix (or essential matrix) and the extrinsic parameters.

- Our next step is to reconstruct surface or dense depth map from images.





The epipolar constraint



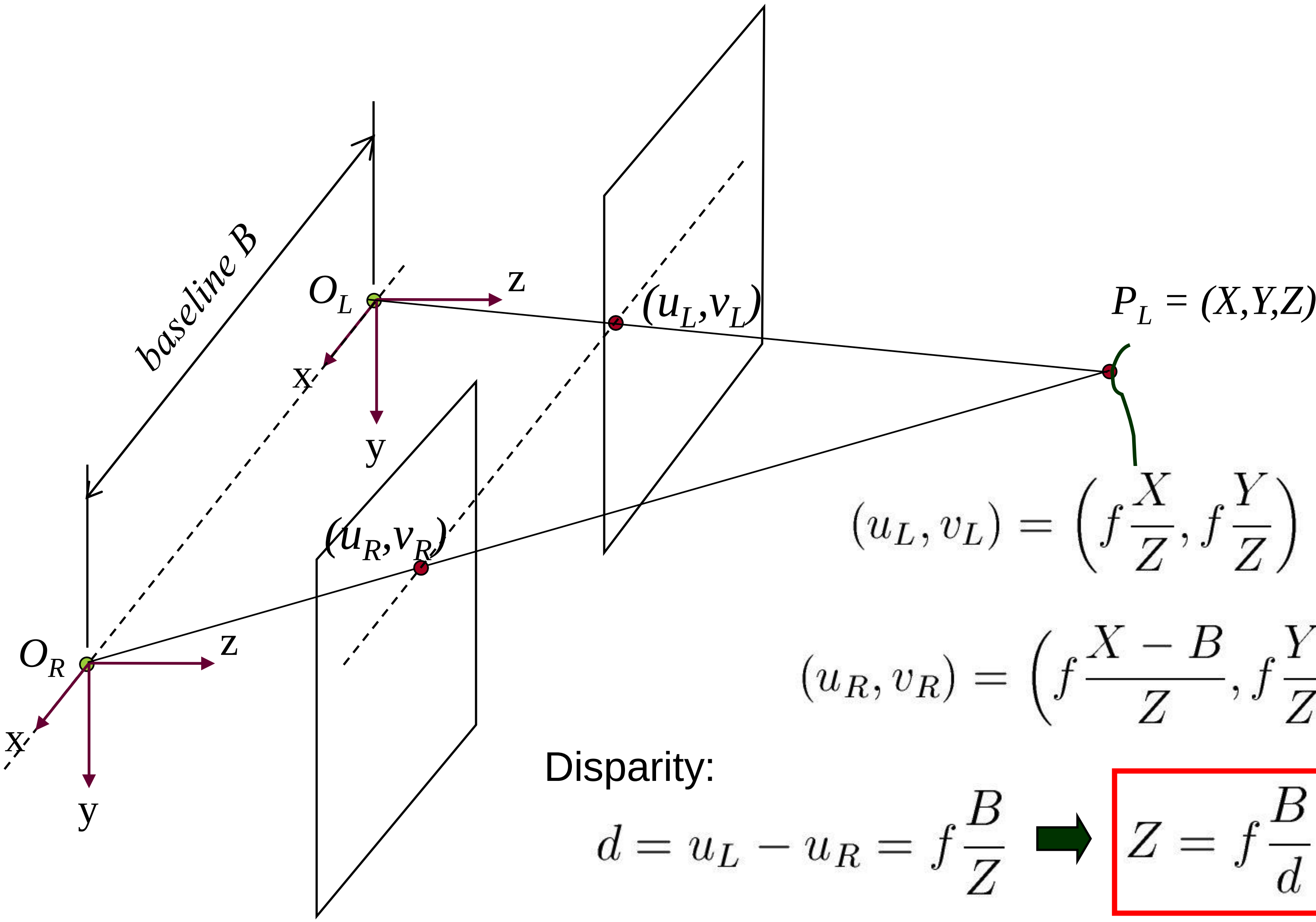
- Epipolar Constraint
  - Matching points lie along corresponding epipolar lines.
  - Reduces correspondence problem to 1D search along *conjugate epipolar lines*
  - Greatly reduces cost and ambiguity of matching

Slide credit: Steve Seitz

Simplest case: rectified images

- Image planes of cameras are parallel.
- Focal points are at same height.
- Focal lengths are the same.
- Then, epipolar lines fall along the horizontal scan lines of the images

Simplest case: rectified images



$$(u_L, v_L) = \left(f\frac{X}{Z}, f\frac{Y}{Z}\right)$$

$$(u_R, v_R) = \left(f\frac{X-B}{Z}, f\frac{Y}{Z}\right)$$

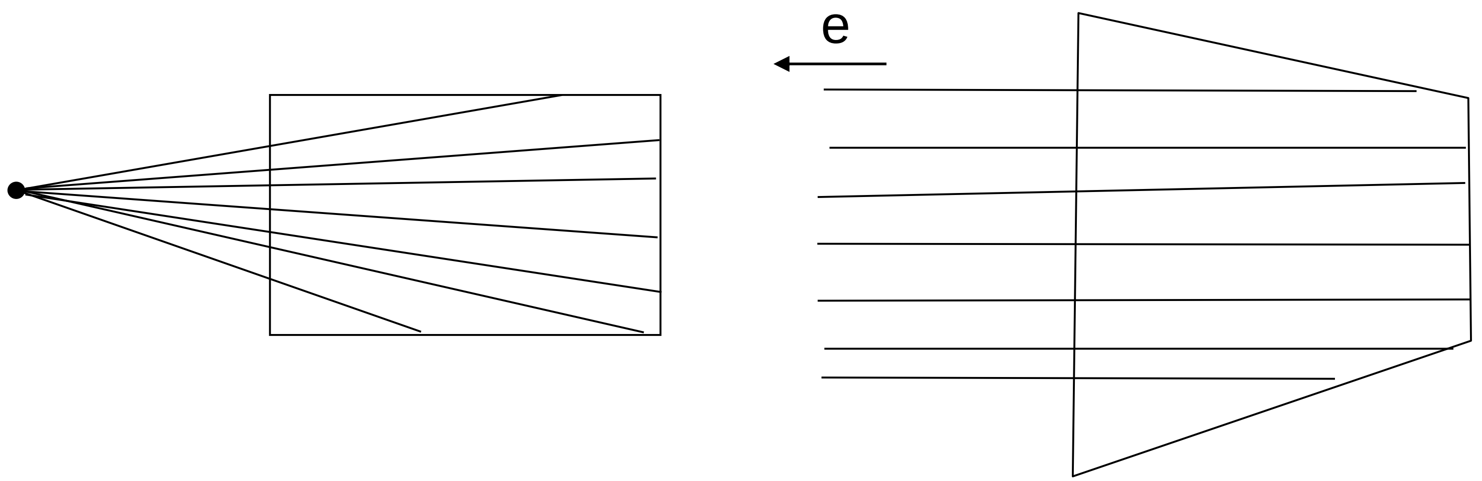
Disparity:

$$d = u_L - u_R = f\frac{B}{Z} \Rightarrow \boxed{Z = f\frac{B}{d}}$$

Prof. D. Lowe, Computer Vision, UBC, CA.

Image pair rectification

- Apply projective transformation so that epipolar lines correspond to horizontal scanlines.
- Simplify stereo matching by warping the images.
  - map epipole e to (1,0,0).
  - It will be introduced later.



Correspondence: epipolar constraint

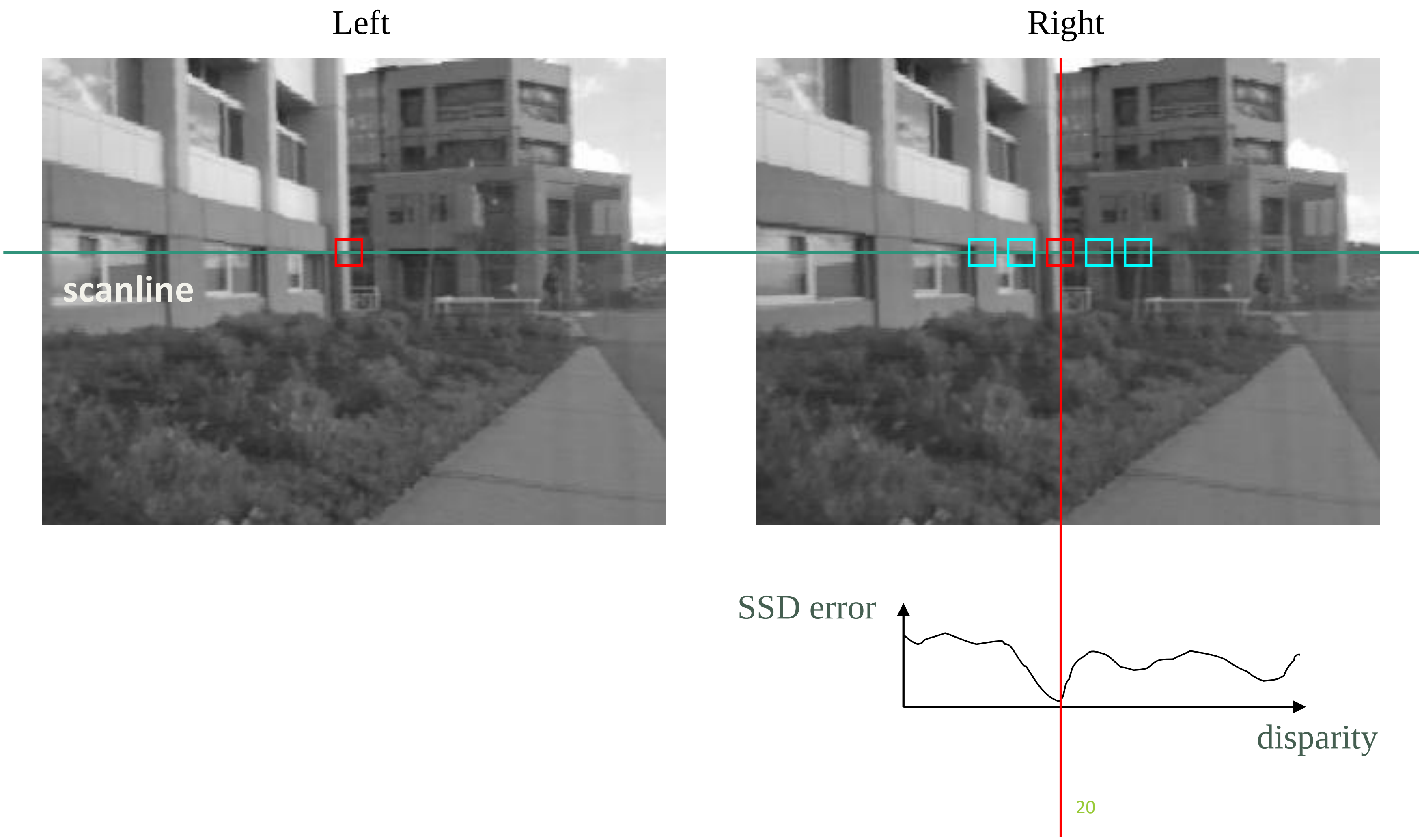
- The epipolar constraint helps, but much ambiguity remains.

Correspondence: photometric constraint

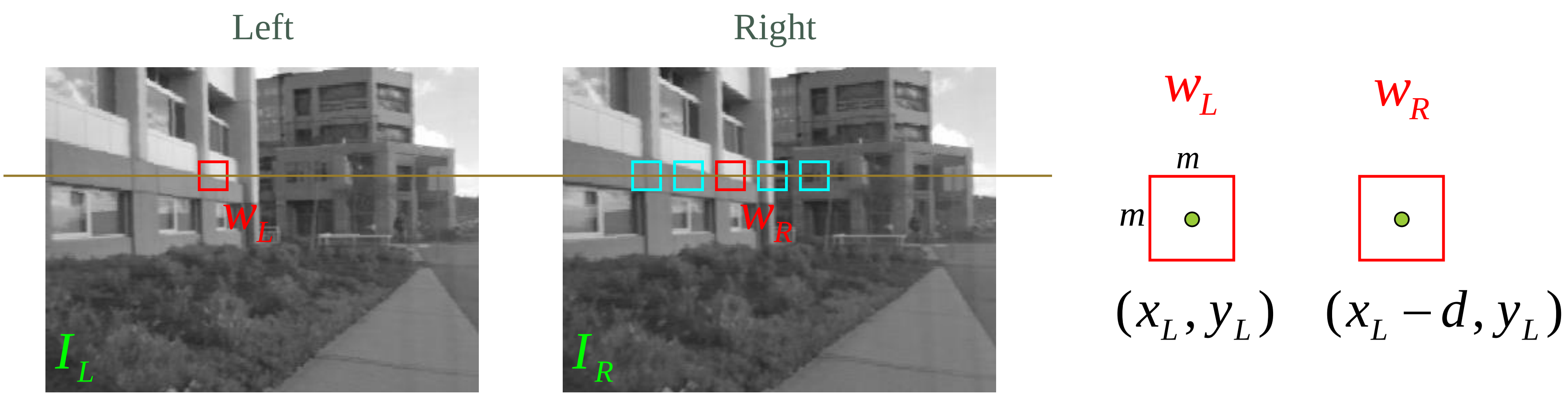
- Corresponding world point has same intensity in both images.
- True for Lambertian surfaces.
    - A Lambertian surface has a brightness independent of viewing angle.

- However,
    - Noise
    - Specularity
    - Non-Lambertian materials
    - …

Correspondence by correlation



Sum of squared differences



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

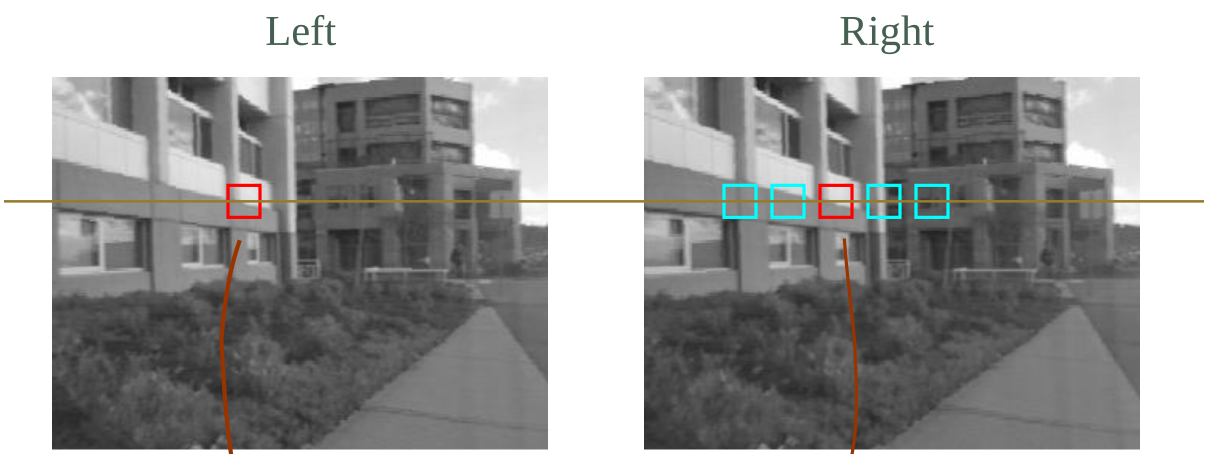
We define the window function :

$$W_m(x, y) = \{u, v \mid x - \tfrac{m}{2} \le u \le x + \tfrac{m}{2}, y - \tfrac{m}{2} \le v \le y + \tfrac{m}{2}\}$$

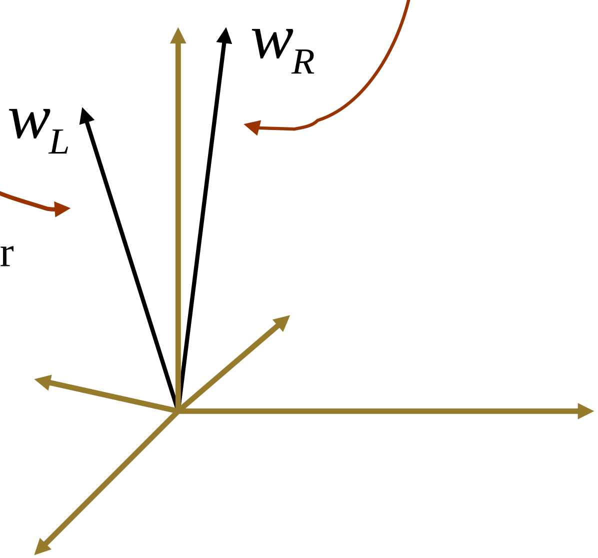
The SSD cost measures the intensity difference as a function of disparity :

$$C_r(x, y, d) = \sum_{(u,v) \in W_m(x,y)} [I_L(u, v) - I_R(u - d, v)]^2$$

Images as vectors



“Unwrap” an image to form a vector, using raster scan order



Each window is a vector in an m^2 dimensional vector space. Normalization makes them unit length.

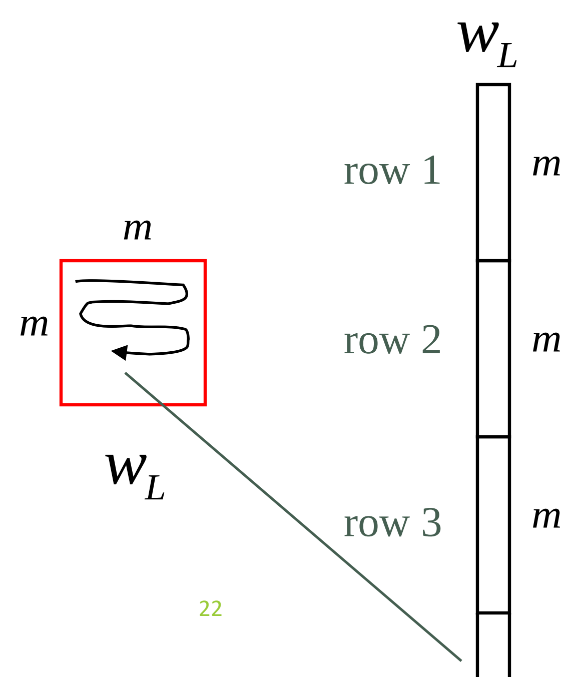
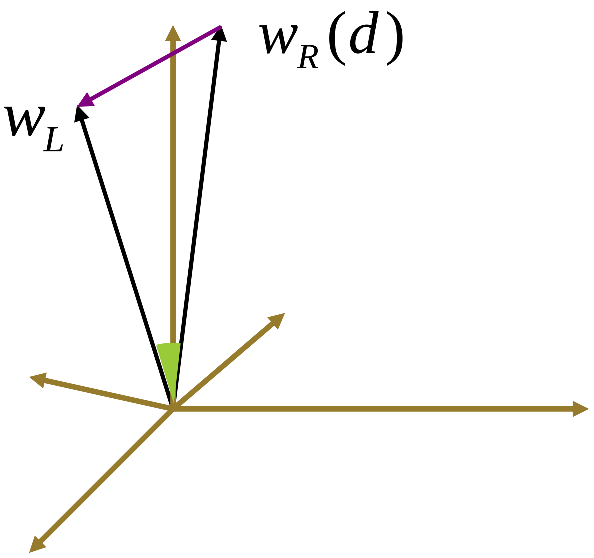


Image metrics



(Normalized) Sum of Squared Differences

$$C_{\mathrm{SSD}}(d) = \sum_{(u,v)\in W_m(x,y)} [\hat{I}_L(u,v) - \hat{I}_R(u-d,v)]^2$$

$$= \left\| w_L - w_R(d) \right\|^2$$

Normalized Correlation

$$C_{\mathrm{NC}}(d) = \sum_{(u,v)\in W_m(x,y)} \hat{I}_L(u,v)\hat{I}_R(u-d,v)$$

$$= w_L \cdot w_R(d) = \cos\theta$$

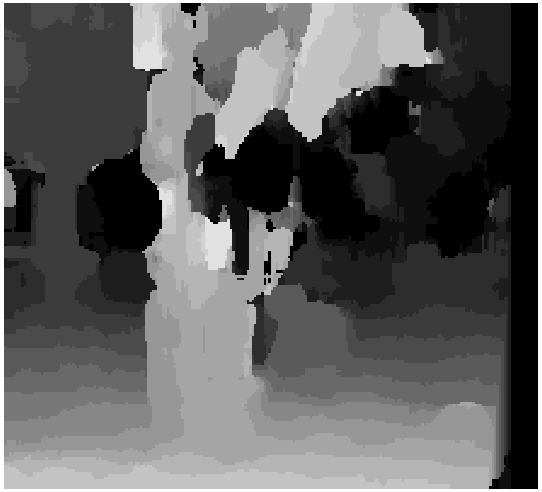
$$d^* = \arg\min_d \left\| w_L - w_R(d) \right\|^2 = \arg\max_d w_L \cdot w_R(d)$$

Stereo results

Window size

- Effect of window size.
  - Some approaches have been developed to use an adaptive window size (try multiple sizes and select best match)

W = 3

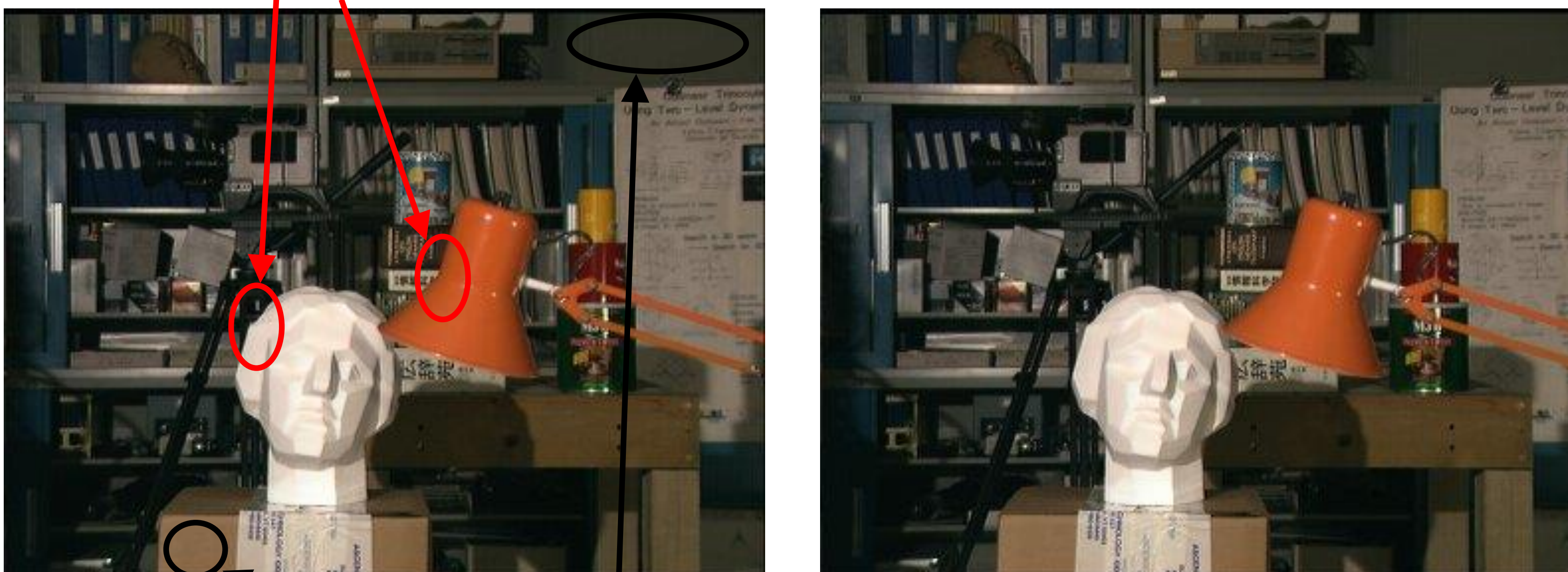


W = 20

Two major roadblocks

- Textureless regions create ambiguities
- Occlusions result in missing data

Occluded regions

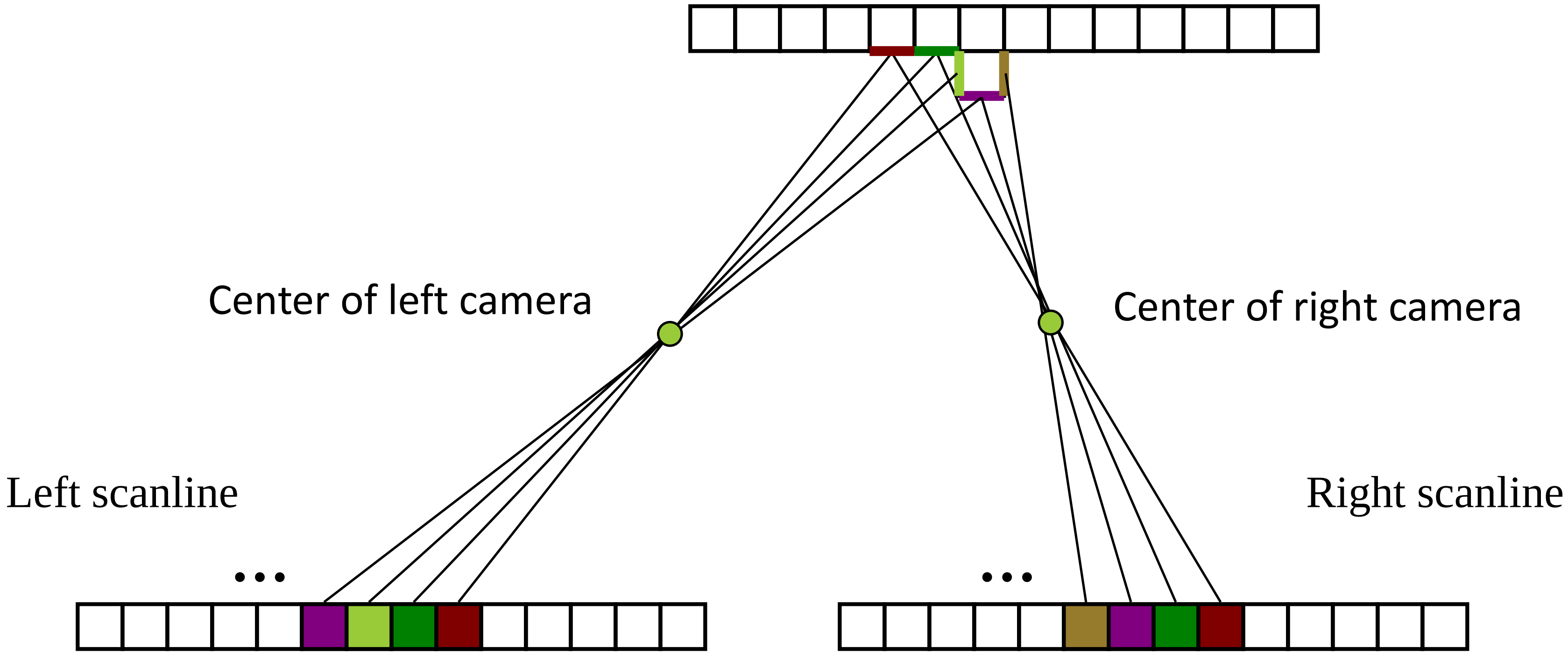


Textureless regions

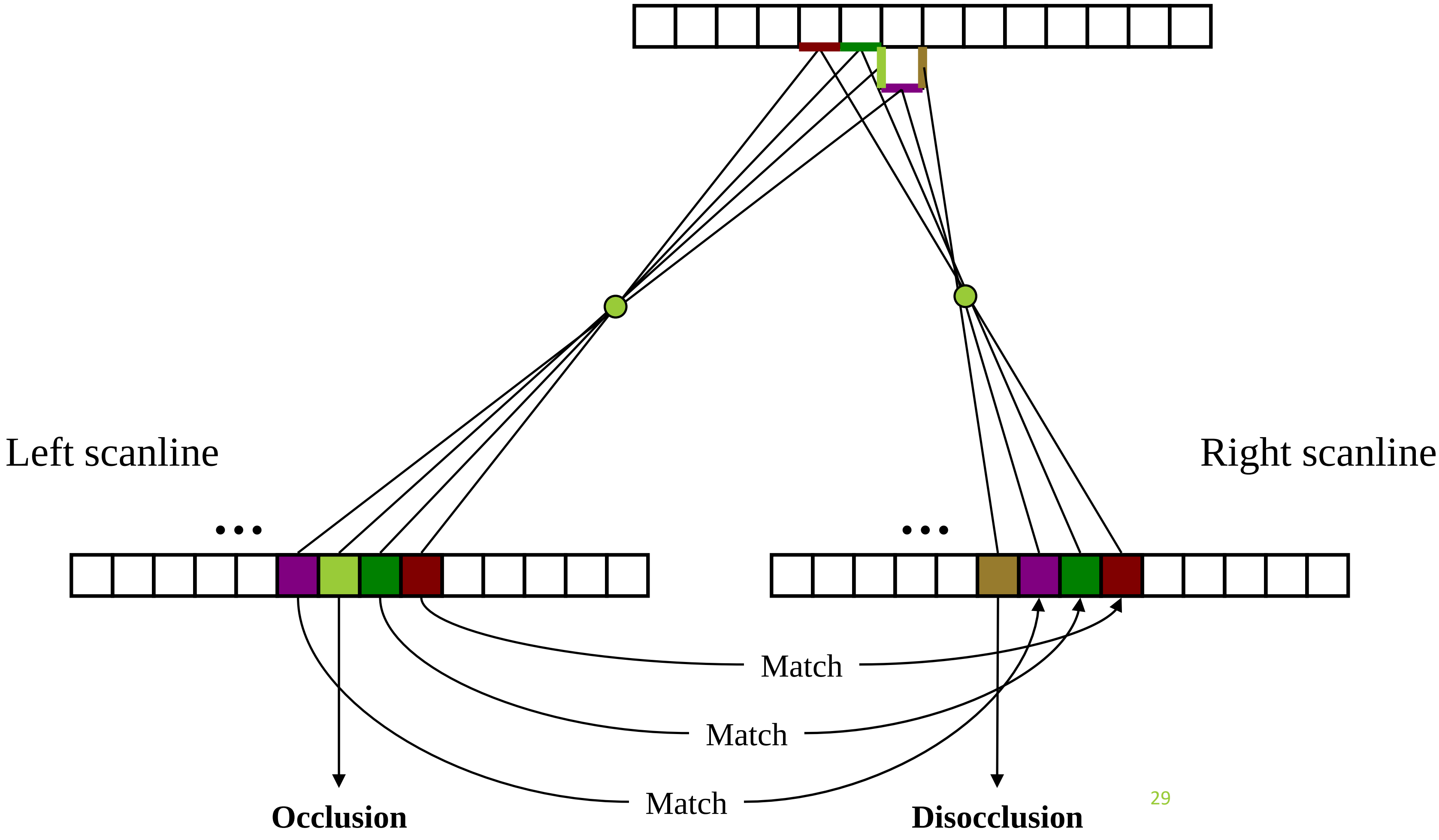
Dealing with ambiguities and occlusion

- Ordering constraint:
  - Impose same matching order along scanlines.
- Uniqueness constraint:
  - Each pixel in one image maps to unique pixel in other.
- Can encode these constraints easily in dynamic programming.

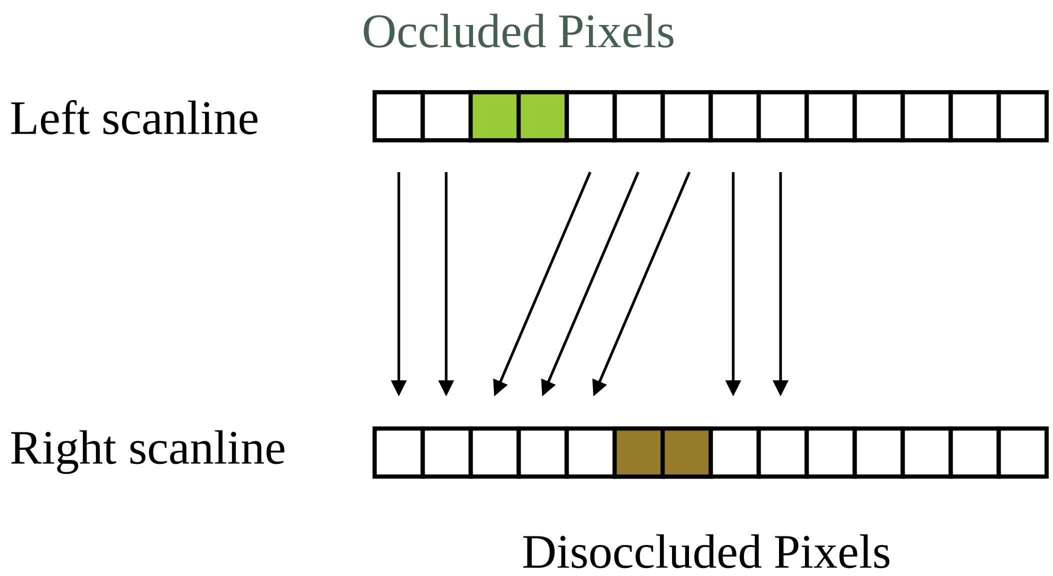
Pixel-based stereo



Stereo correspondences



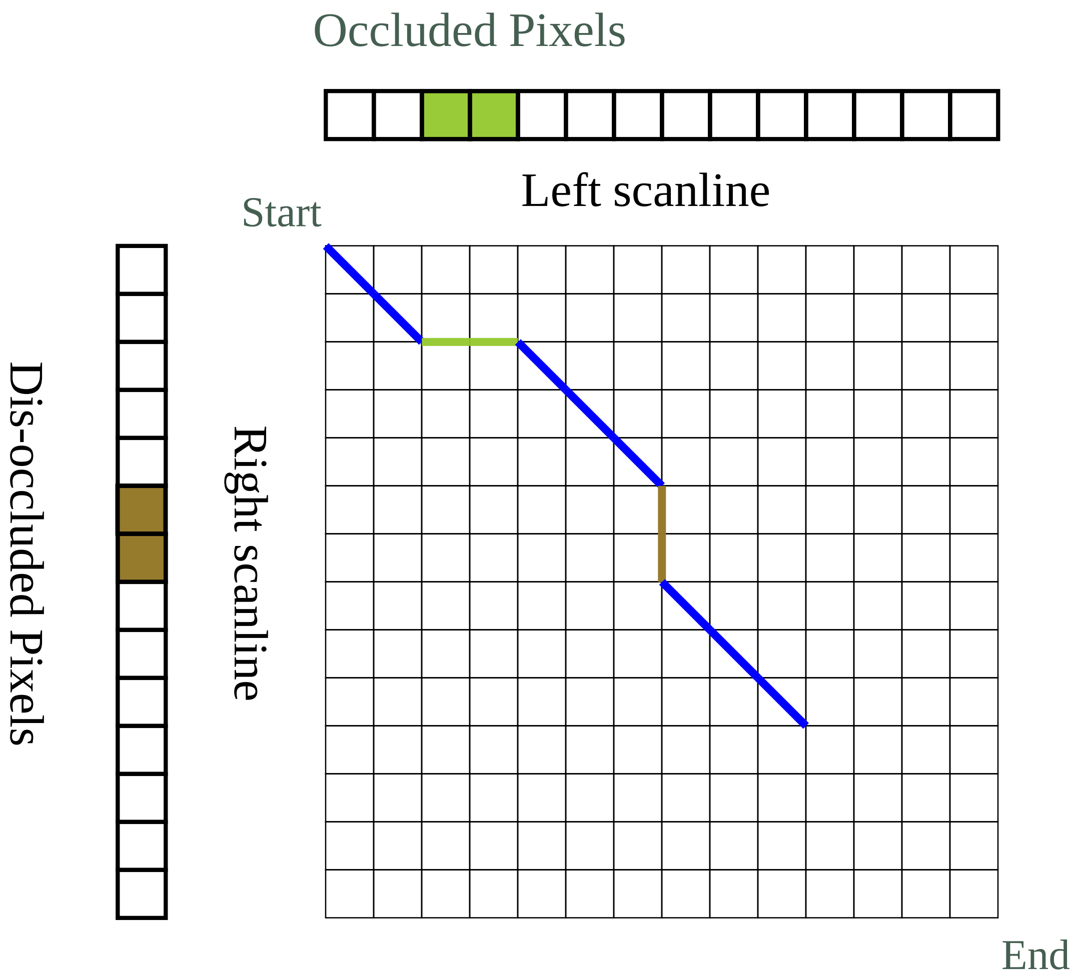
Search over correspondences



Three cases:

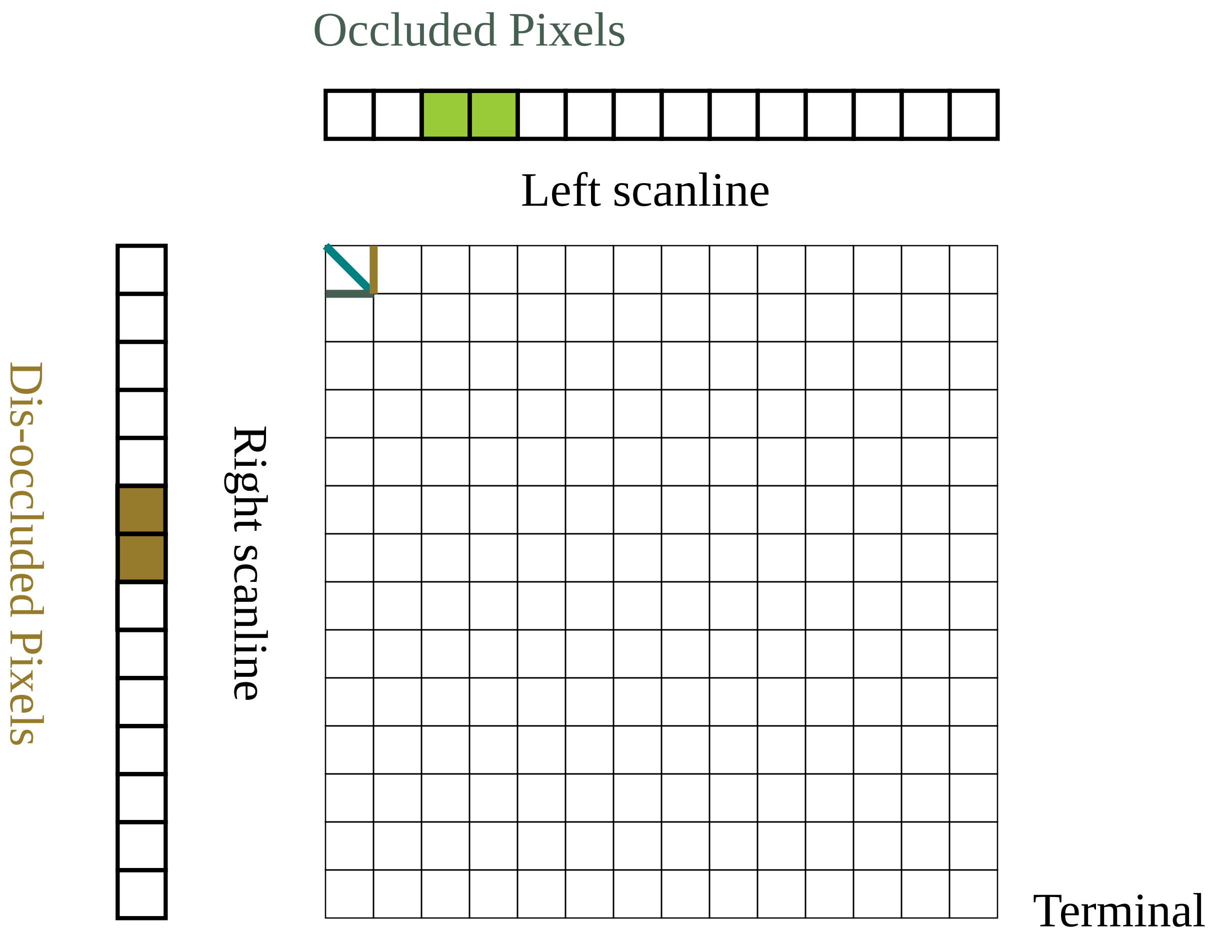
- Sequential – cost of match
- Occluded – cost of no match
- Disoccluded – cost of no match

Matching with dynamic programming

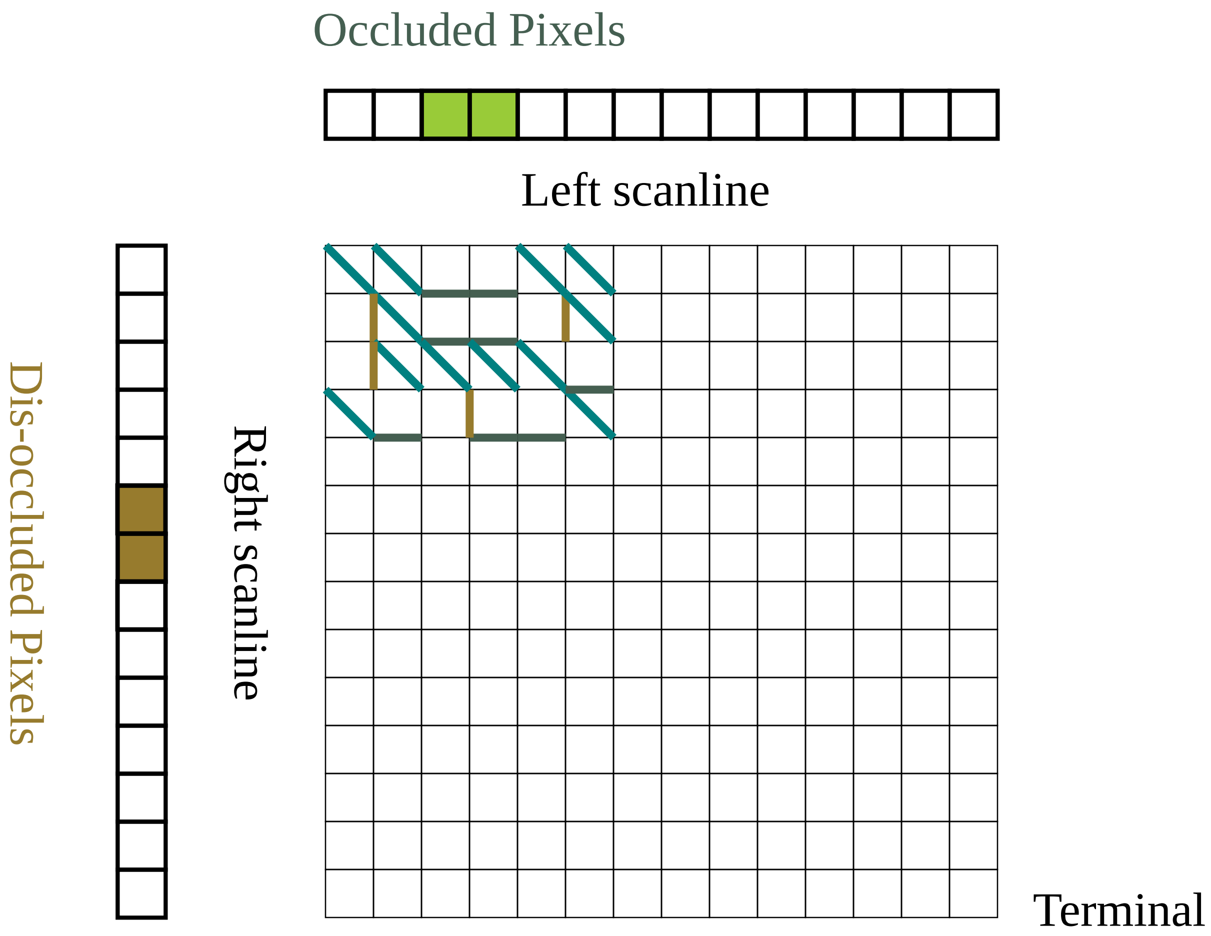


Dynamic programming yields the optimal path through grid. This is the best set of matches that satisfy the ordering constraint

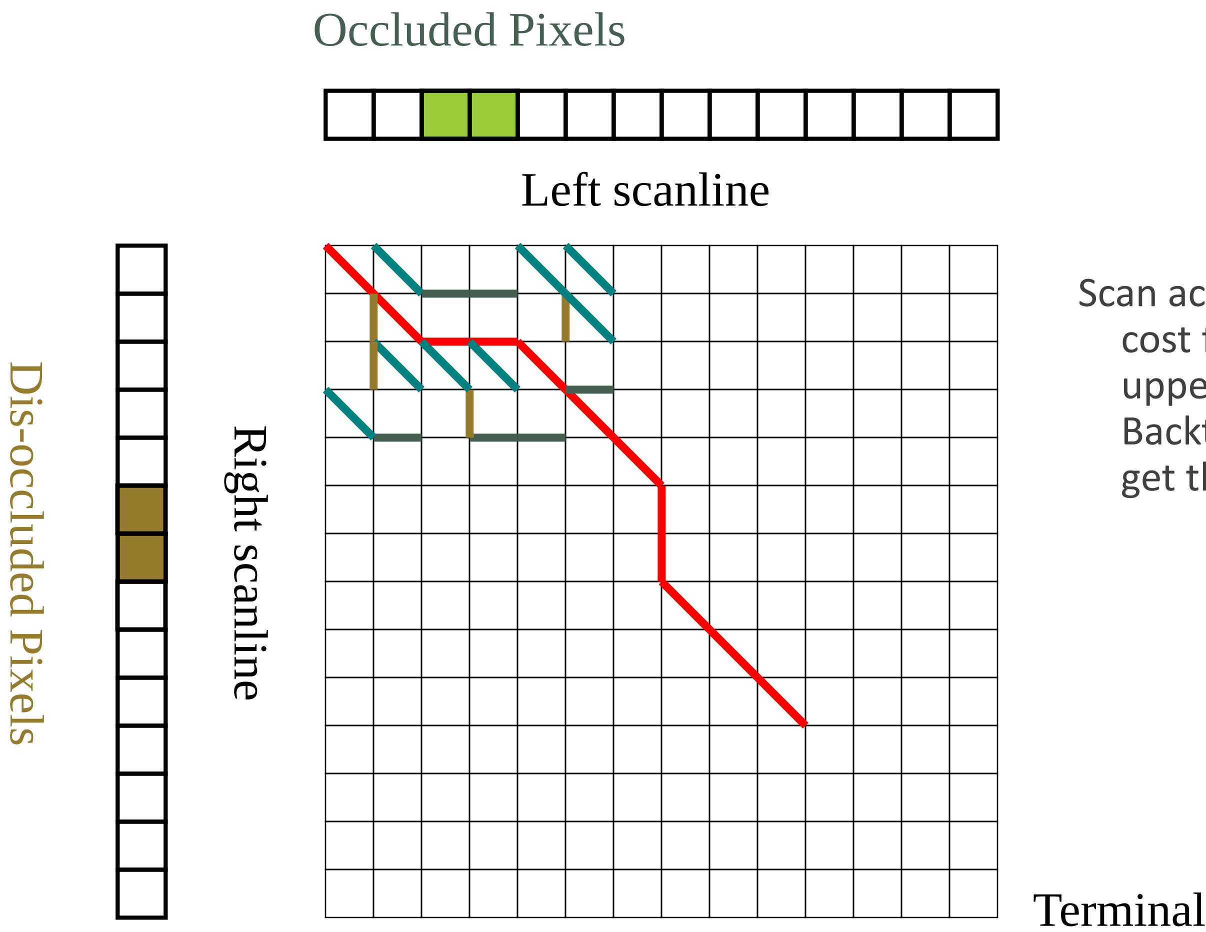
Matching with dynamic programming



Matching with dynamic programming

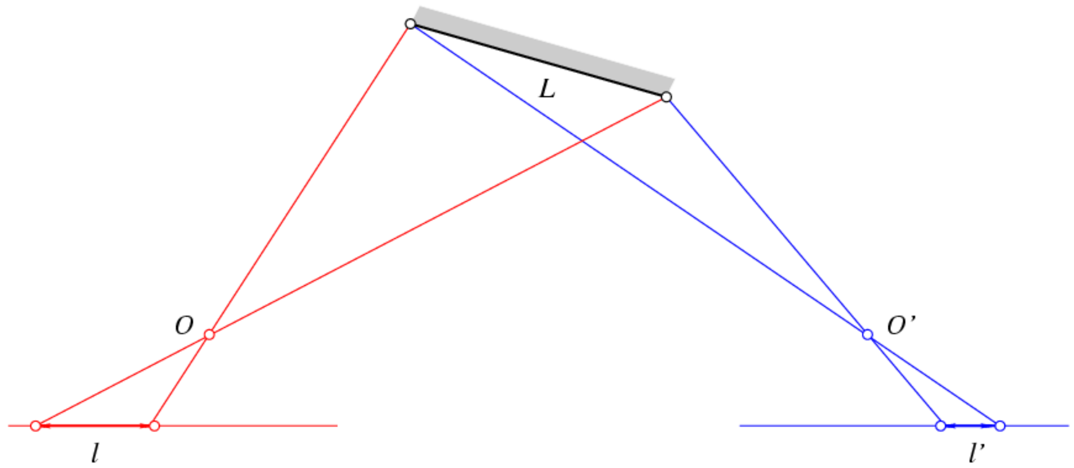


Matching with Dynamic Programming



Scan across grid computing optimal cost for each node given its upper-left neighbors. Backtrack from the terminal to get the optimal path.

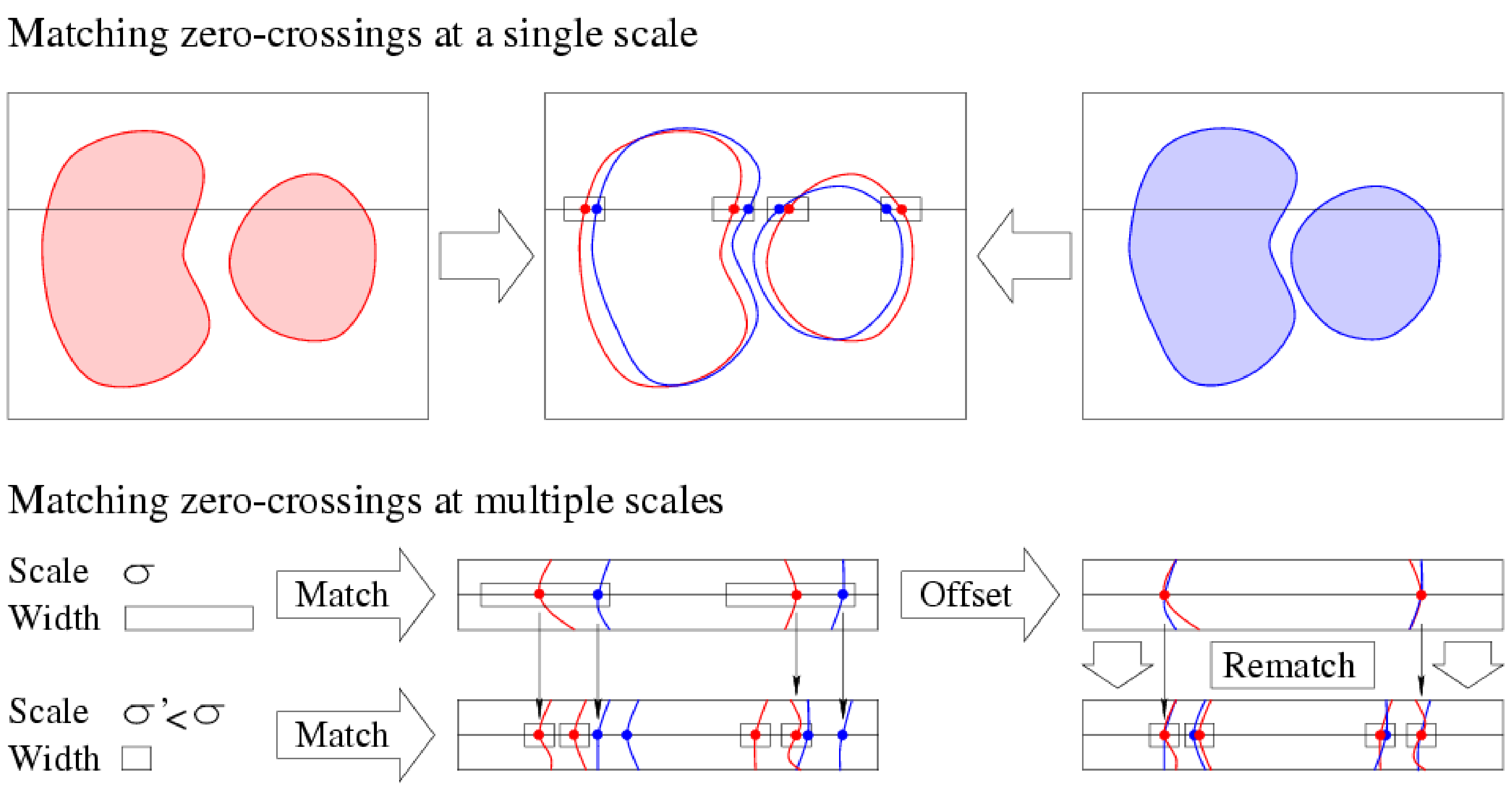
Foreshortening problems



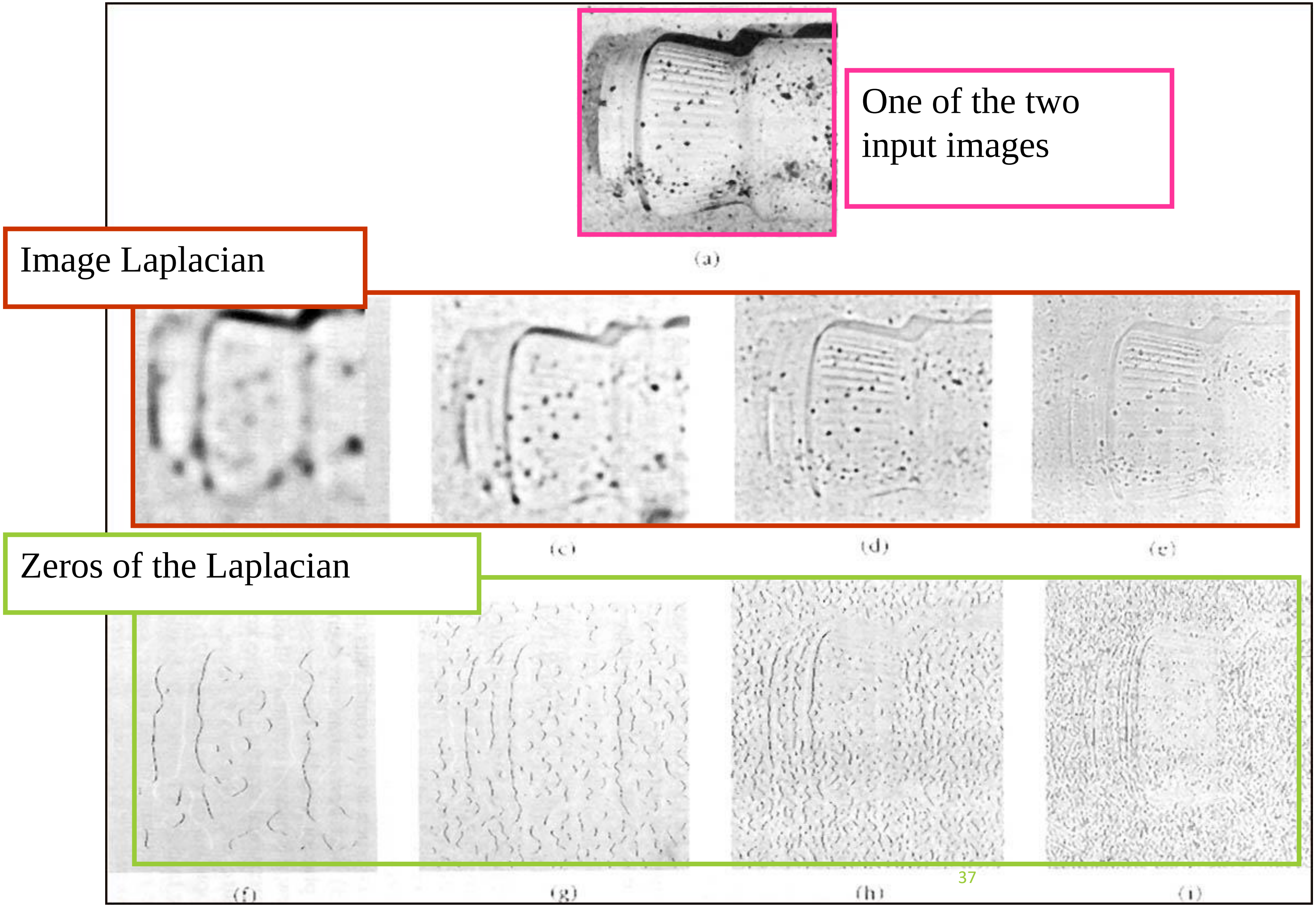
Solution: add a second pass using disparity estimates to warp the correlation windows, e.g. Devernay and Faugeras (1994).

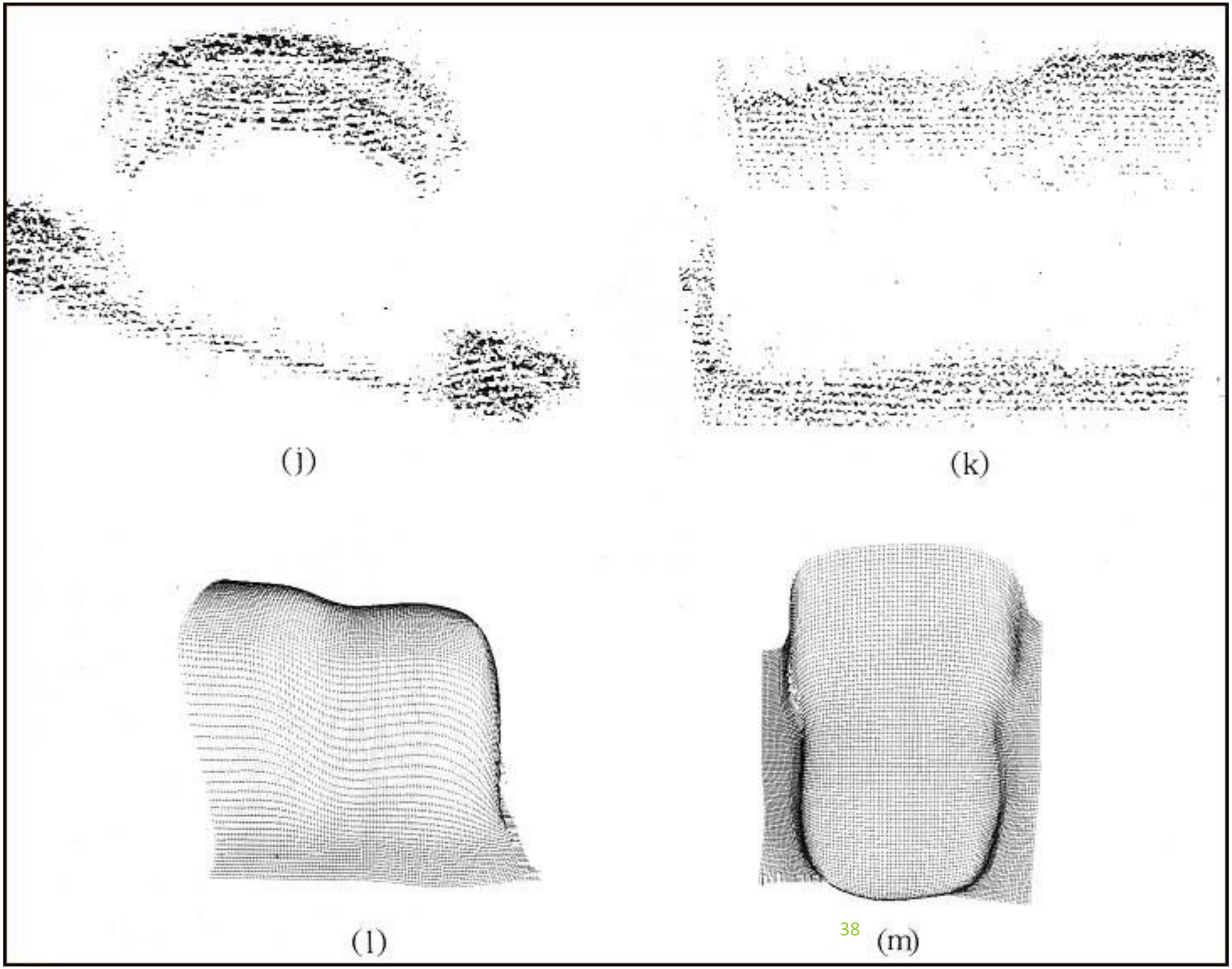
Reprinted from "Computing Differential Properties of 3D Shapes from Stereopsis without 3D Models," by F. Devernay and O. Faugeras, Proc. IEEE Conf. on Computer Vision and Pattern Recognition (1994).

Multi-Scale edge matching (Marr, Poggio and Grimson)

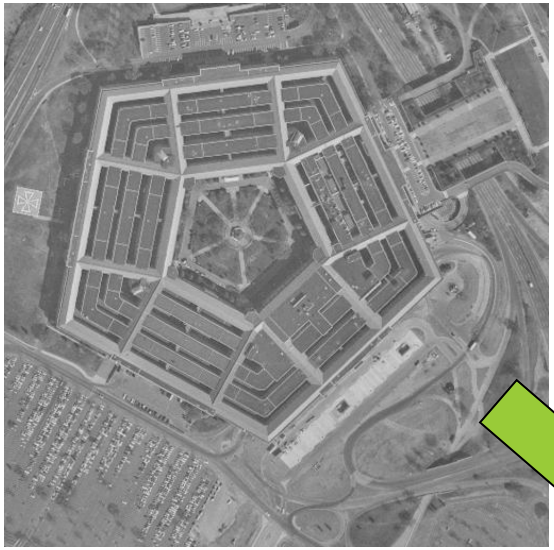


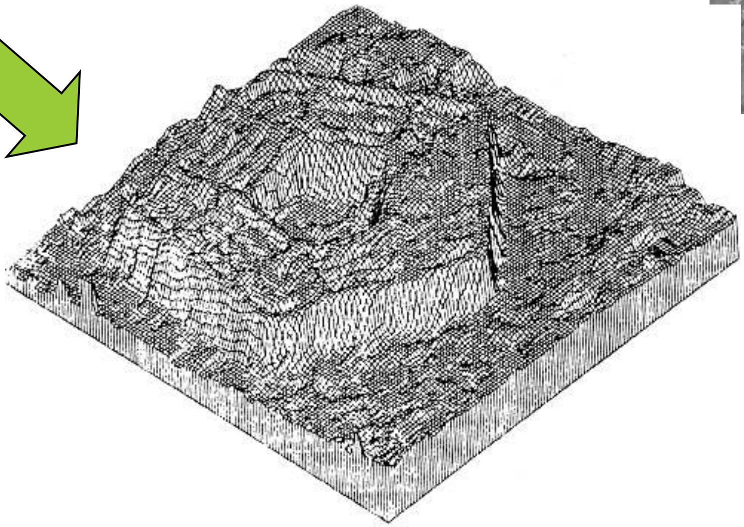
- Edges are found by repeatedly smoothing the image and detecting the zero crossings of the second derivative (Laplacian).
- Matches at coarse scales are used to offset the search for matches at fine scales (equivalent to eye movements).


One of the two input images
Image Laplacian
Zeros of the Laplacian
(a)
(c)
(d)
(e)
(f)
(g)
(h)
(i)



Dynamic programming (Ohta and Kanade)





Reprinted from “Stereo by Intra- and Intet-Scanline Search,” by Y. Ohta and T. Kanade, IEEE Trans. on Pattern Analysis and Machine Intelligence, 7(2):139-154 (1985). © 1985 IEEE.

Optimizing the disparity map

- What is the assumption and limitation of the above-mentioned DP-based method?

- More advanced optimization methods, e.g.
  - V. Kolmogorov et al., Computing Visual Correspondence with Occlusions via Graph Cuts, Proc. ICCV'01.
  - J. Sun et al, Stereo Matching Using Belief Propagation, IEEE T. PAMI, 2003.