COS 429: COMPUTER VISON

STEREO (1 lecture)

- Stereo Reconstruction
- The Stereo Fusion Problem
- Random Dot Stereograms
- Binocular Fusion Algorithms

- Reading: Chapters 11

Many of the slides in this lecture are courtesy to Prof. J. Ponce

Stereopsis

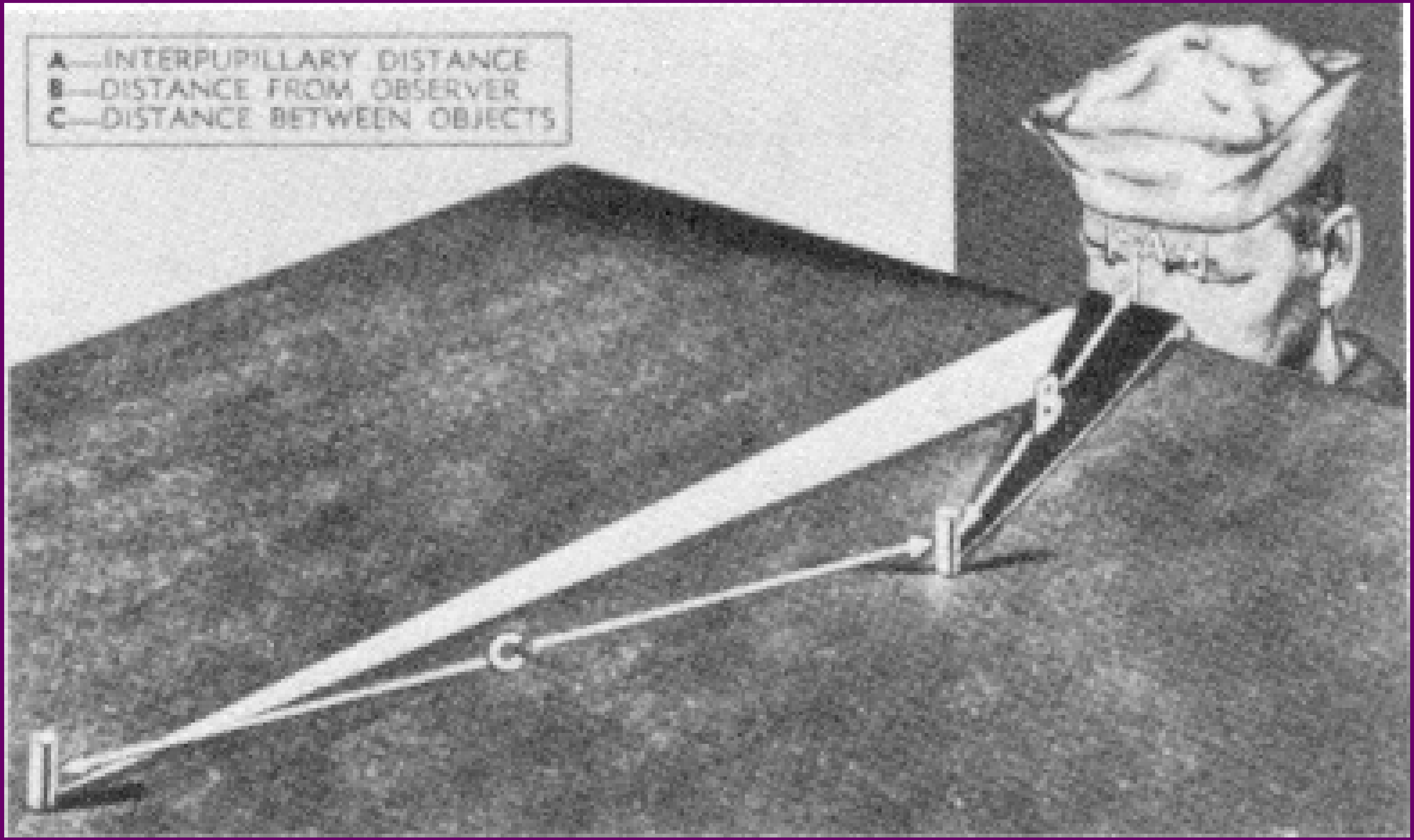
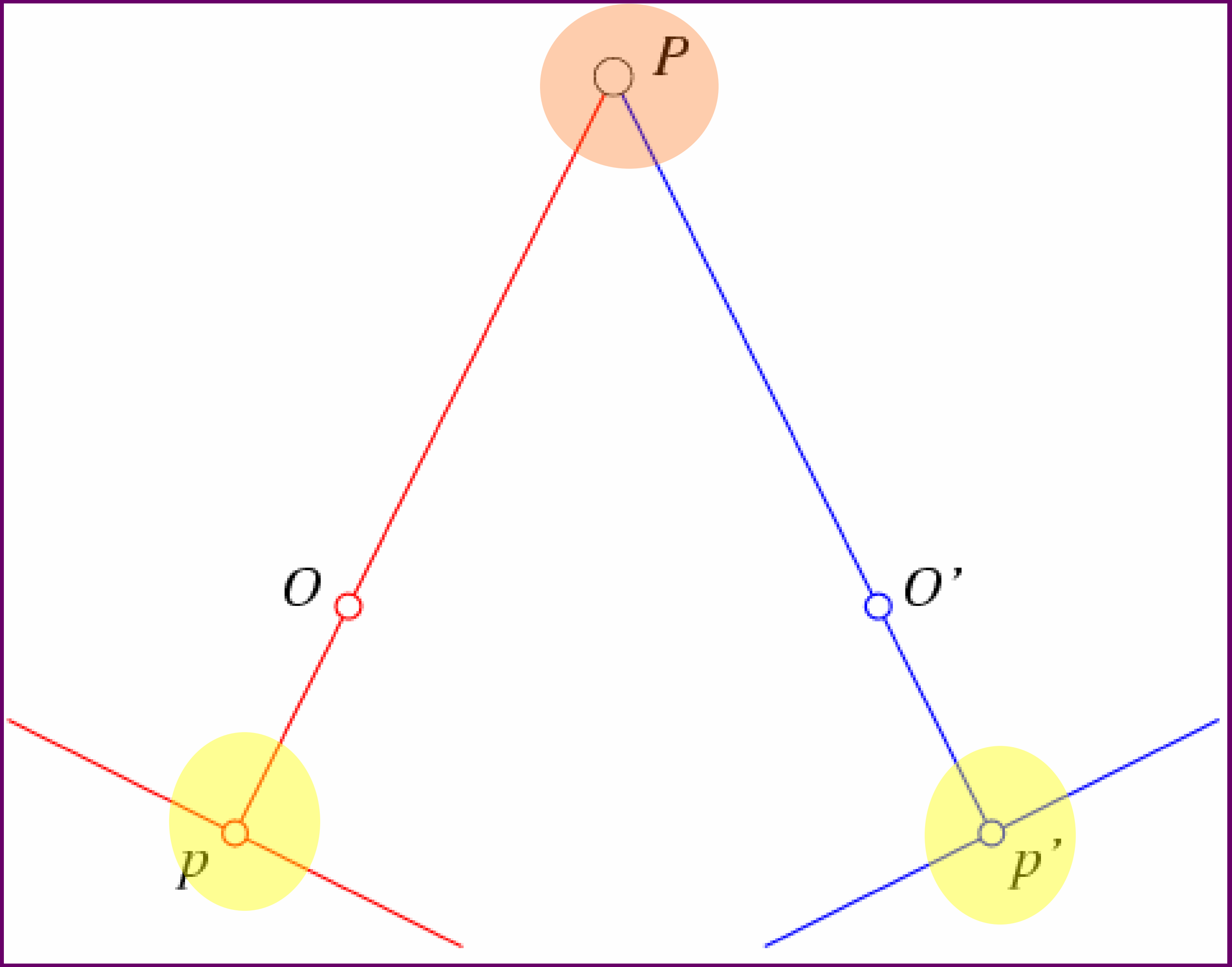


Figure from US Navy Manual of Basic Optics and Optical Instruments, prepared by Bureau of Naval Personnel. Reprinted by Dover Publications, Inc., 1969.

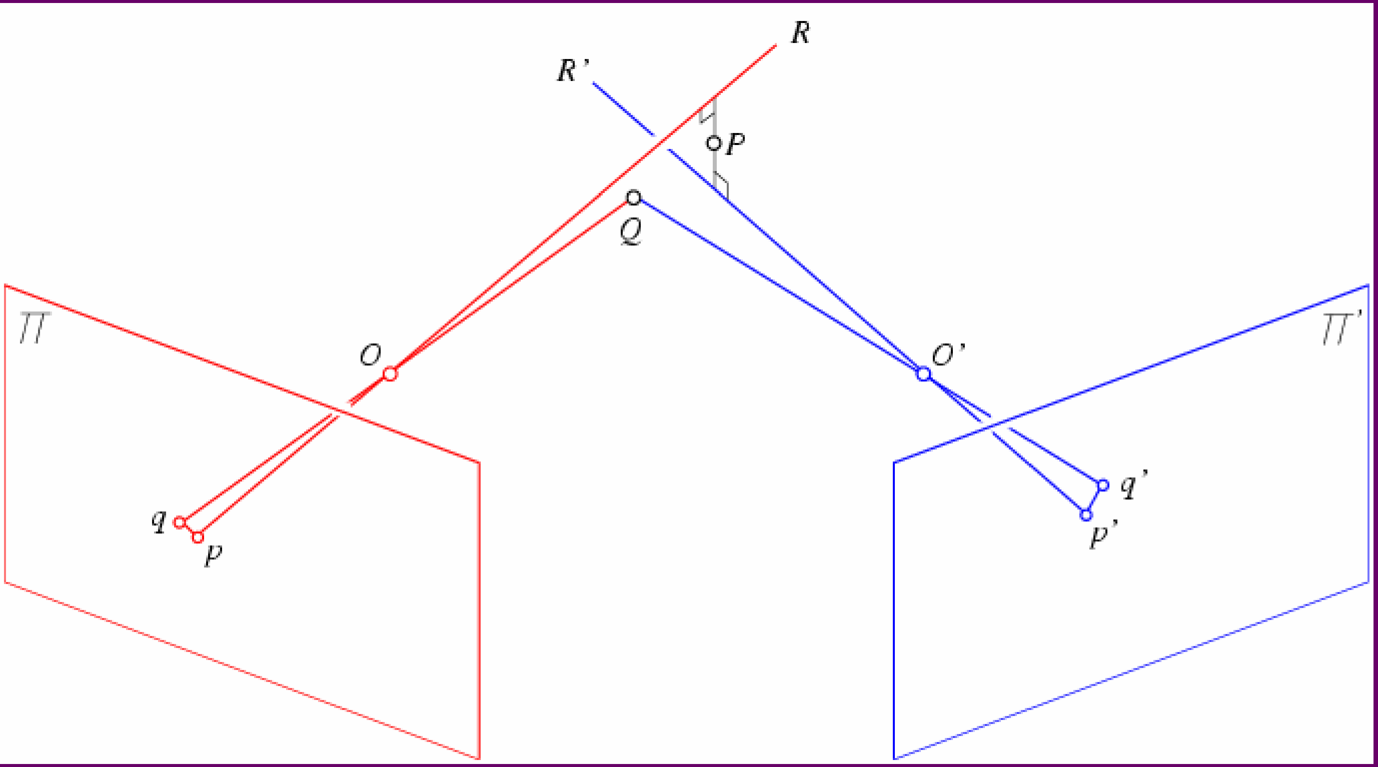
A spider

The INRIA Mobile Robot, 1990.

Stereo vision: Fusion and Reconstruction



Reconstruction



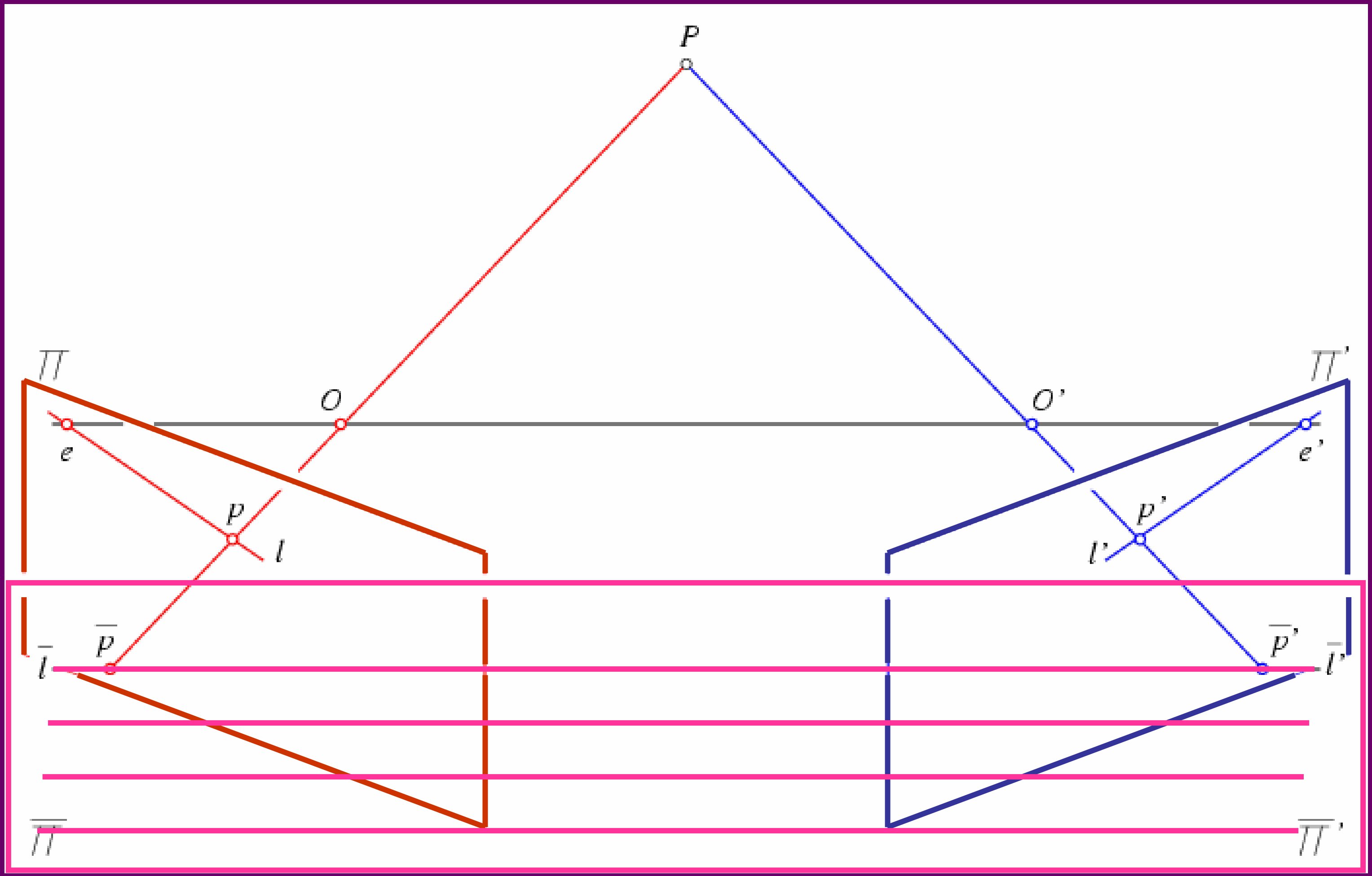
- Linear Method: find P such that

$$\begin{cases} \boldsymbol{p} \times \mathcal{M}\boldsymbol{P} = 0 \\ \boldsymbol{p}' \times \mathcal{M}'\boldsymbol{P} = 0 \end{cases} \Longleftrightarrow \begin{pmatrix} [\boldsymbol{p}_\times]\mathcal{M} \\ [\boldsymbol{p}'_\times]\mathcal{M}' \end{pmatrix} \boldsymbol{P} = 0$$

- Non-Linear Method: find Q minimizing

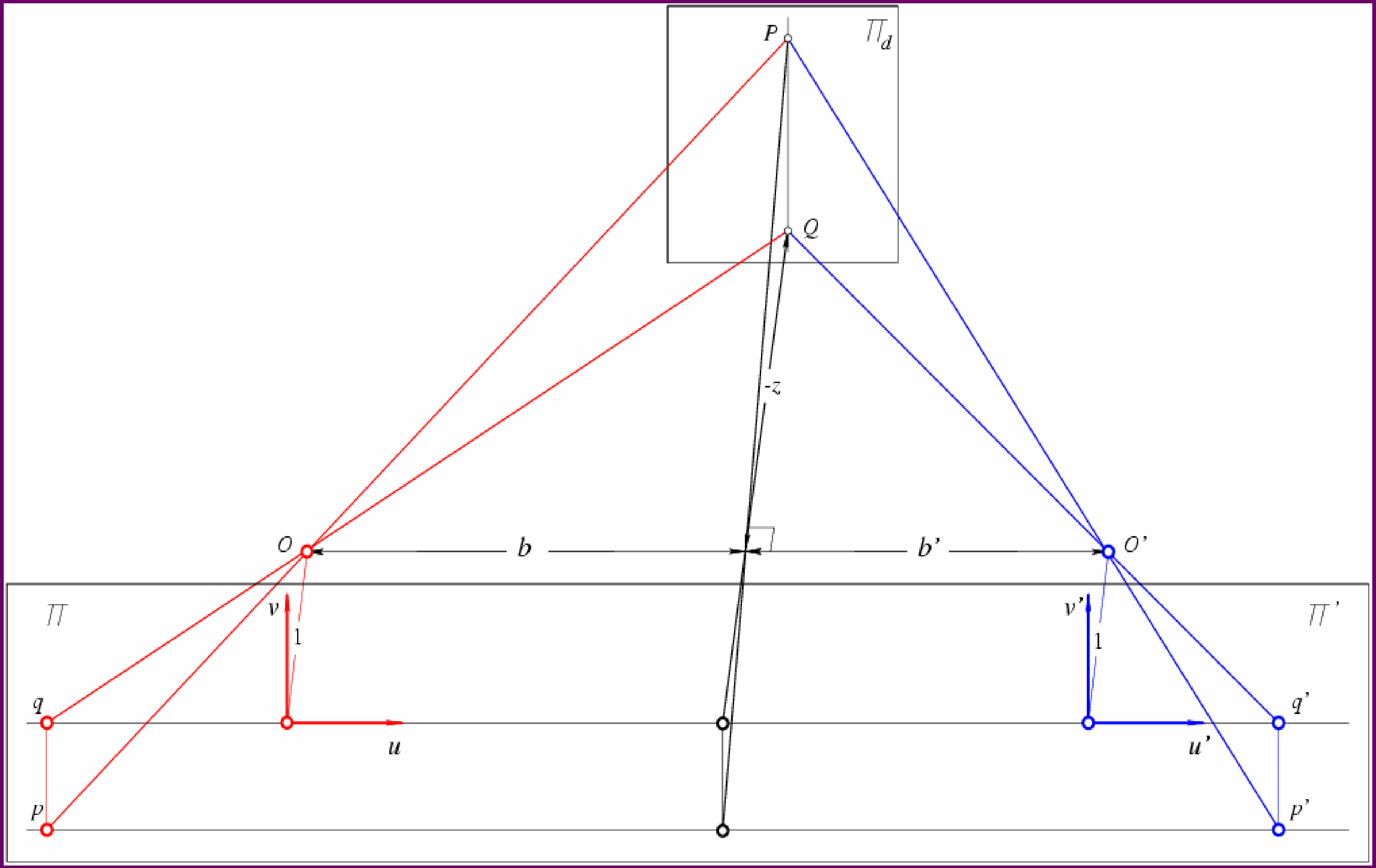
$$d^2(p, q) + d^2(p', q')$$

Rectification



All epipolar lines are parallel in the rectified image plane.

Reconstruction from Rectified Images



Disparity: $d=u'-u$. ⇒ Depth: $z = -B/d$.

Stereopsis

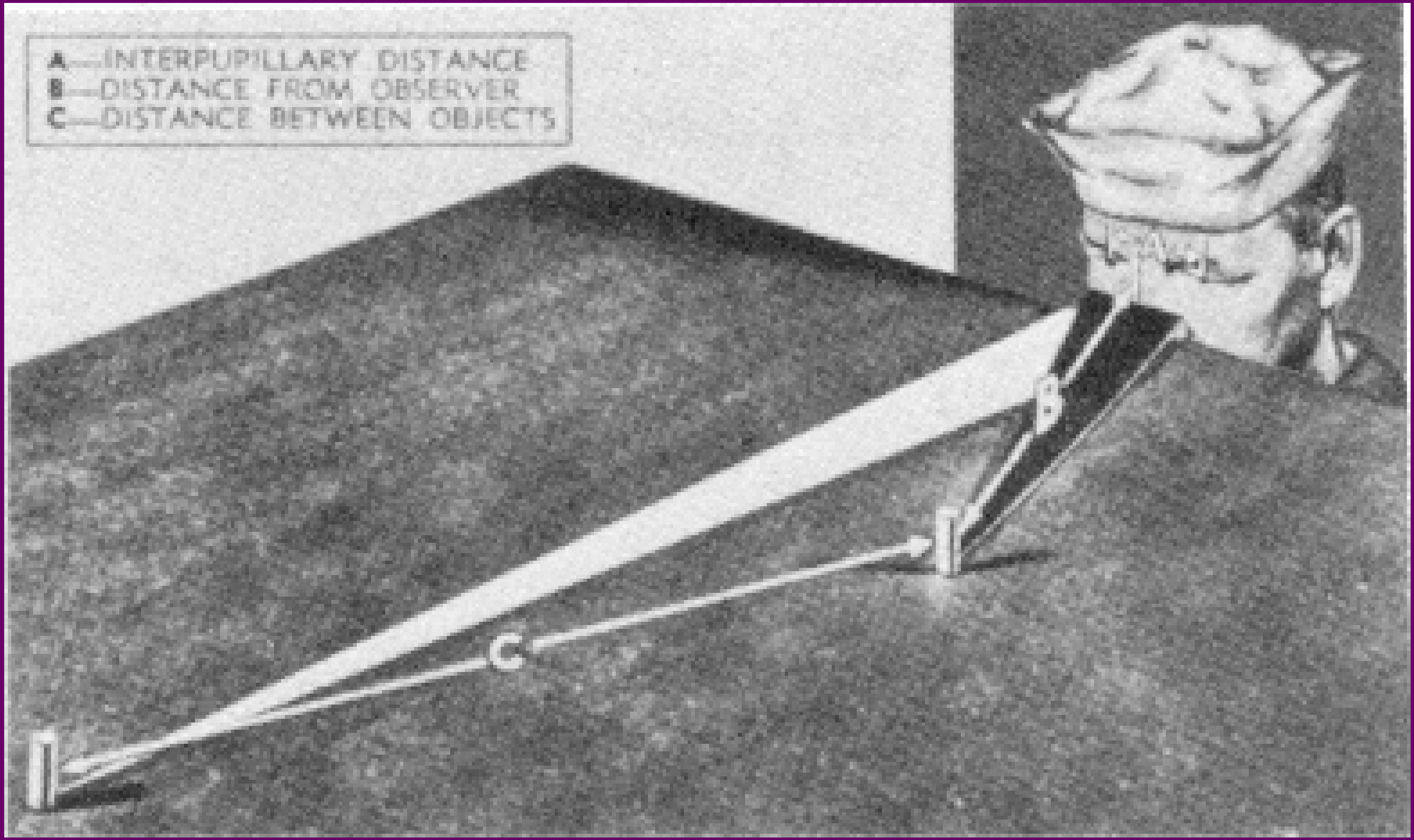
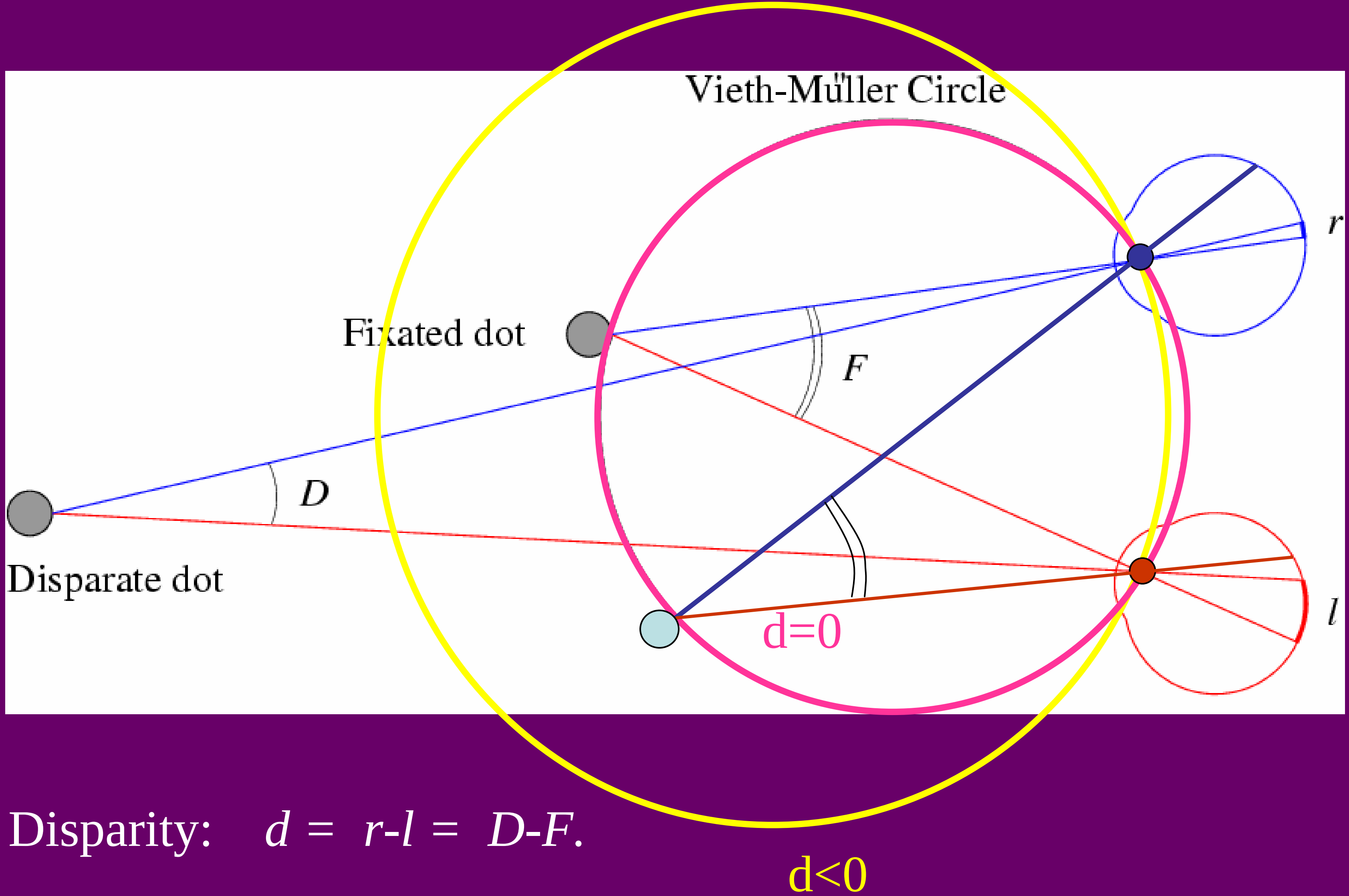


Figure from US Navy Manual of Basic Optics and Optical Instruments, prepared by Bureau of Naval Personnel. Reprinted by Dover Publications, Inc., 1969.

Human Stereopsis: Reconstruction



Disparity: $d = r\text{-}l = D\text{-}F$.

In 3D, the horopter.

Human Stereopsis: Reconstruction

What if *F* is not known?

Helmoltz (1909):

- There is evidence showing the vergence angles cannot be measured precisely.

- Humans get fooled by bas-relief sculptures.

Human Stereopsis: Reconstruction

What if *F* is not known?

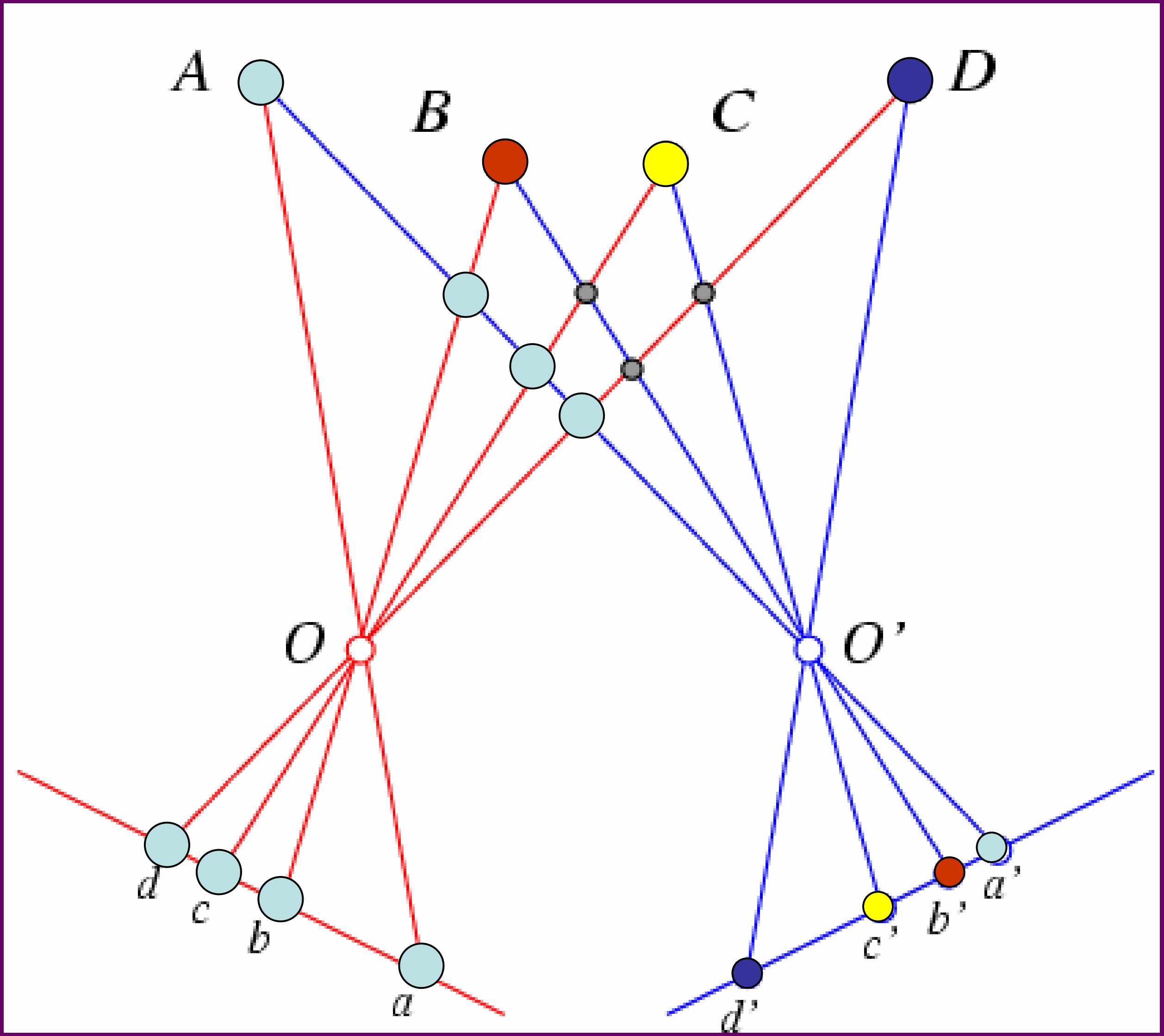
Helmoltz (1909):

- There is evidence showing the vergence angles cannot be measured precisely.

- Humans get fooled by bas-relief sculptures.

- Relative depth can be judged accurately.

(Binocular) Fusion

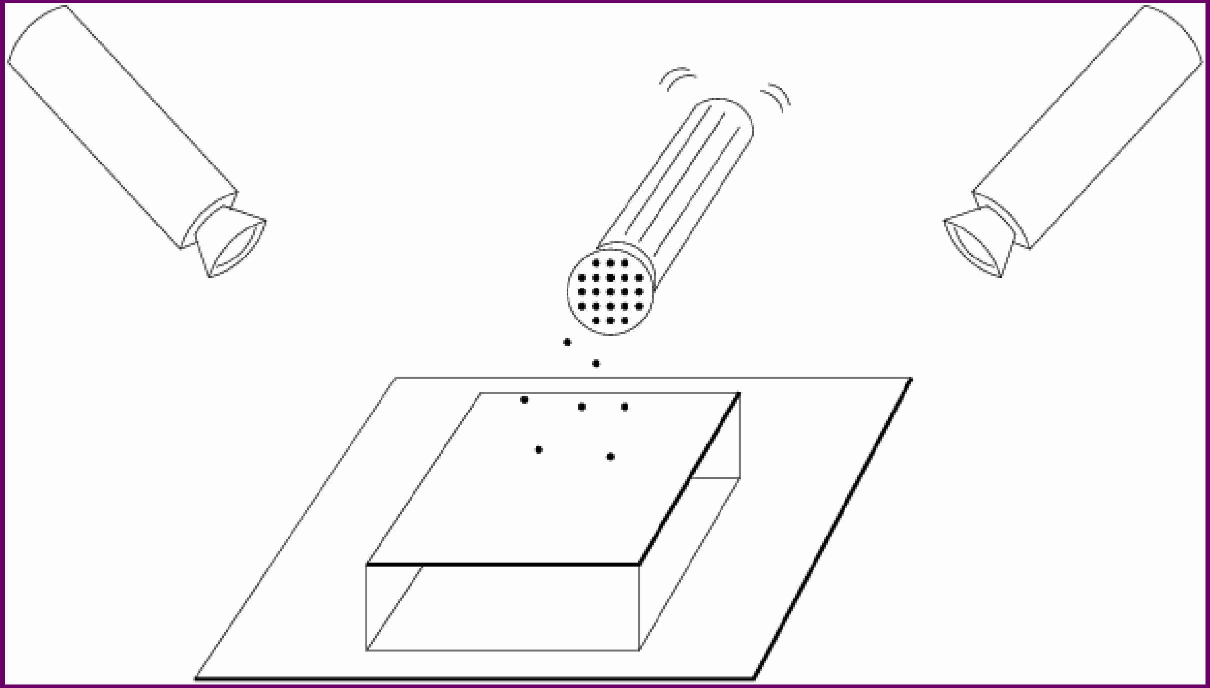


Human Stereopsis: Binocular Fusion

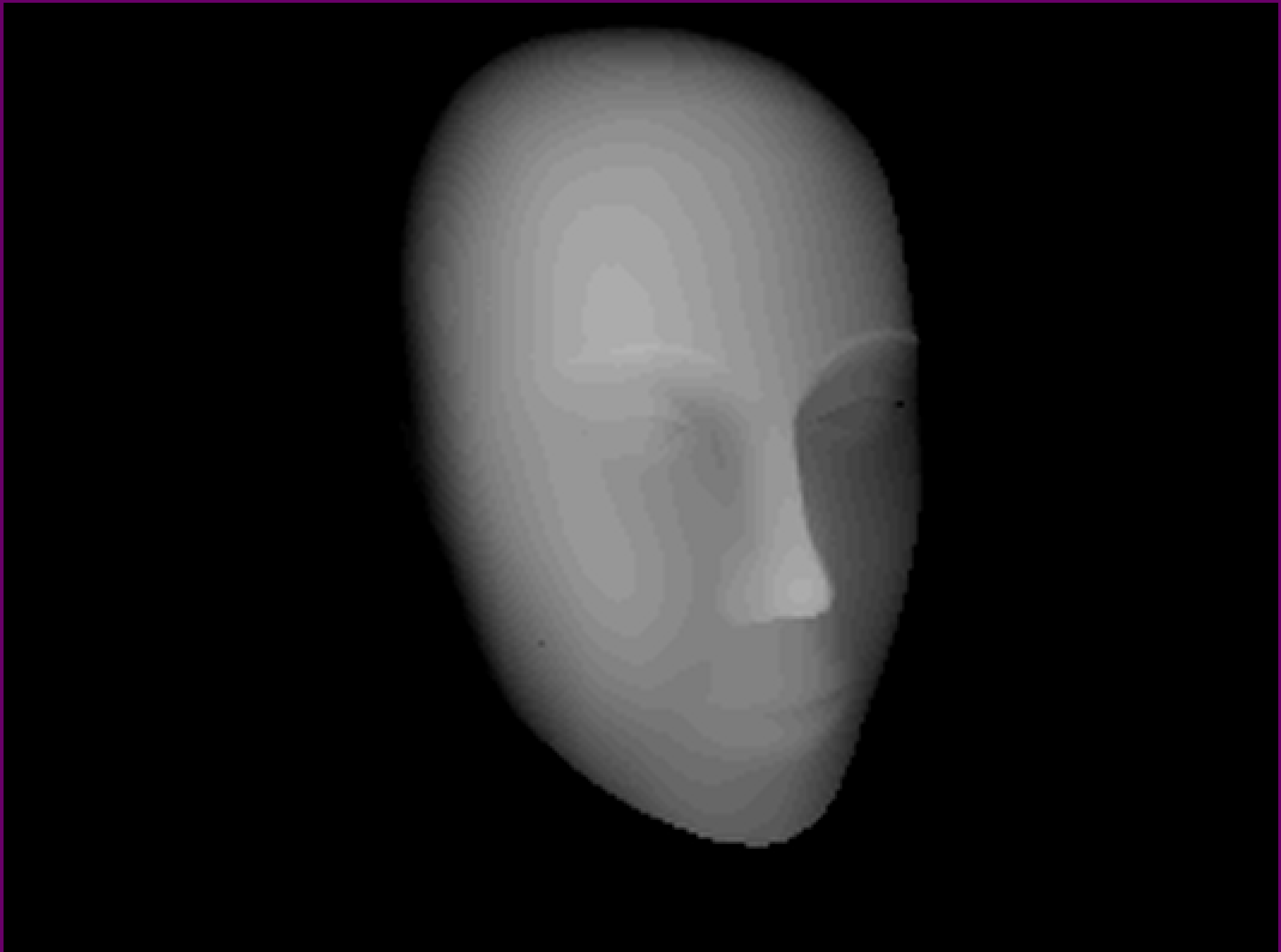
How are the correspondences established?

Julesz (1971): Is the mechanism for binocular fusion a monocular process or a binocular one??

- There is anecdotal evidence for the latter (camouflage).



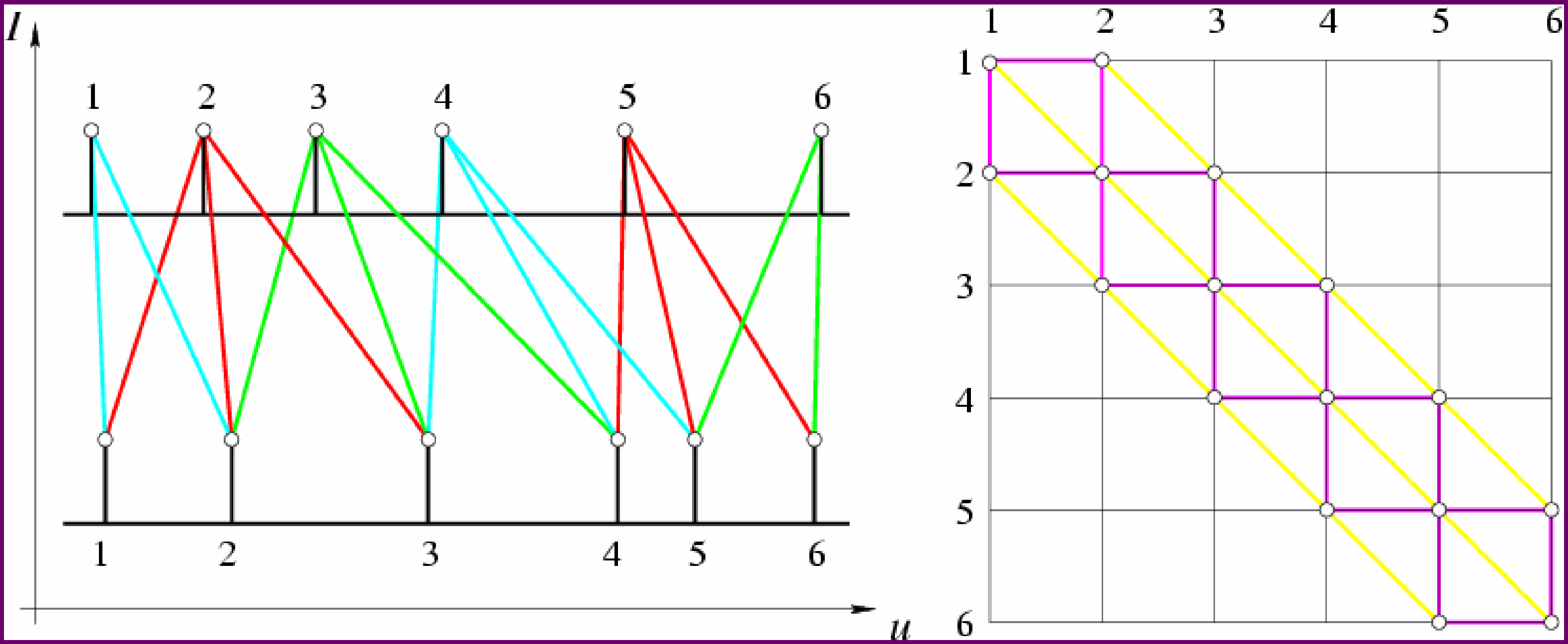
- Random dot stereograms provide an objective answer

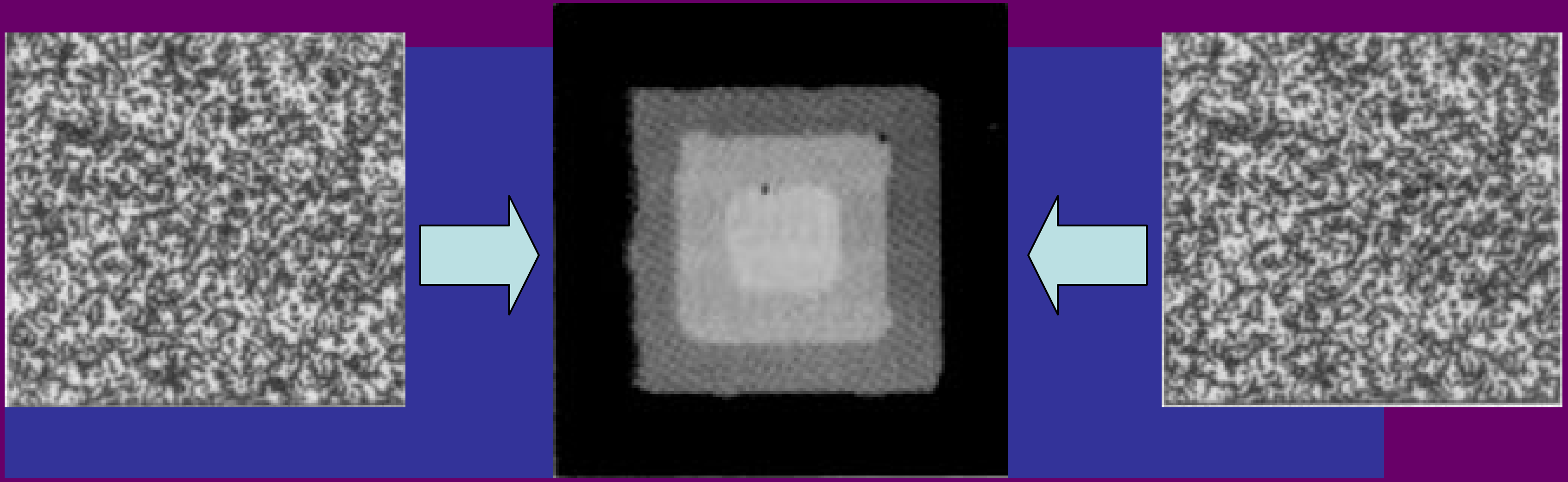


Some fun for yourself

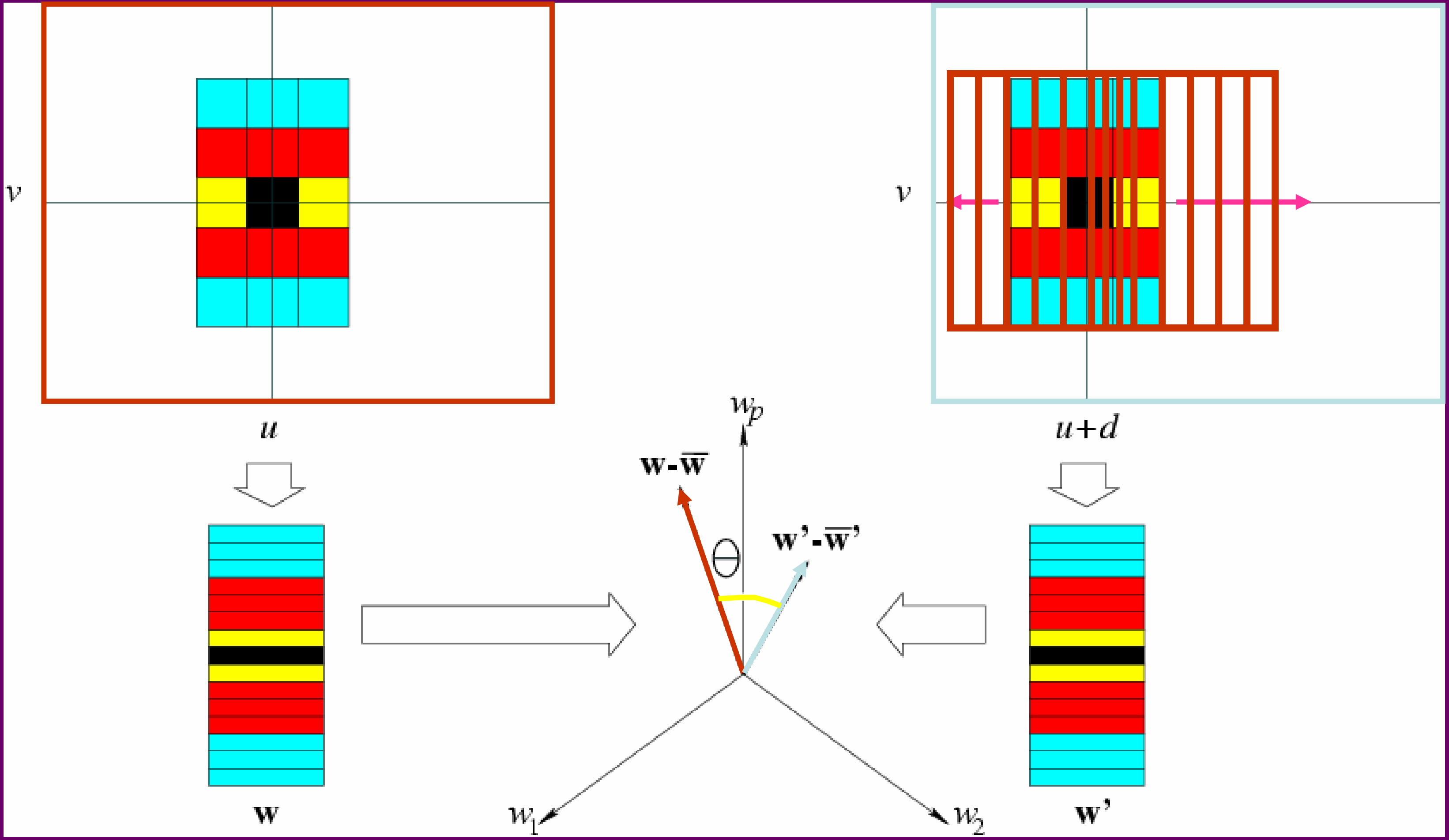
- Google “random dot stereograms” for some fun examples
- For example:
  - http://www.tanos.co.uk/gallery/randomdot/

A Cooperative Model (Marr and Poggio, 1976)





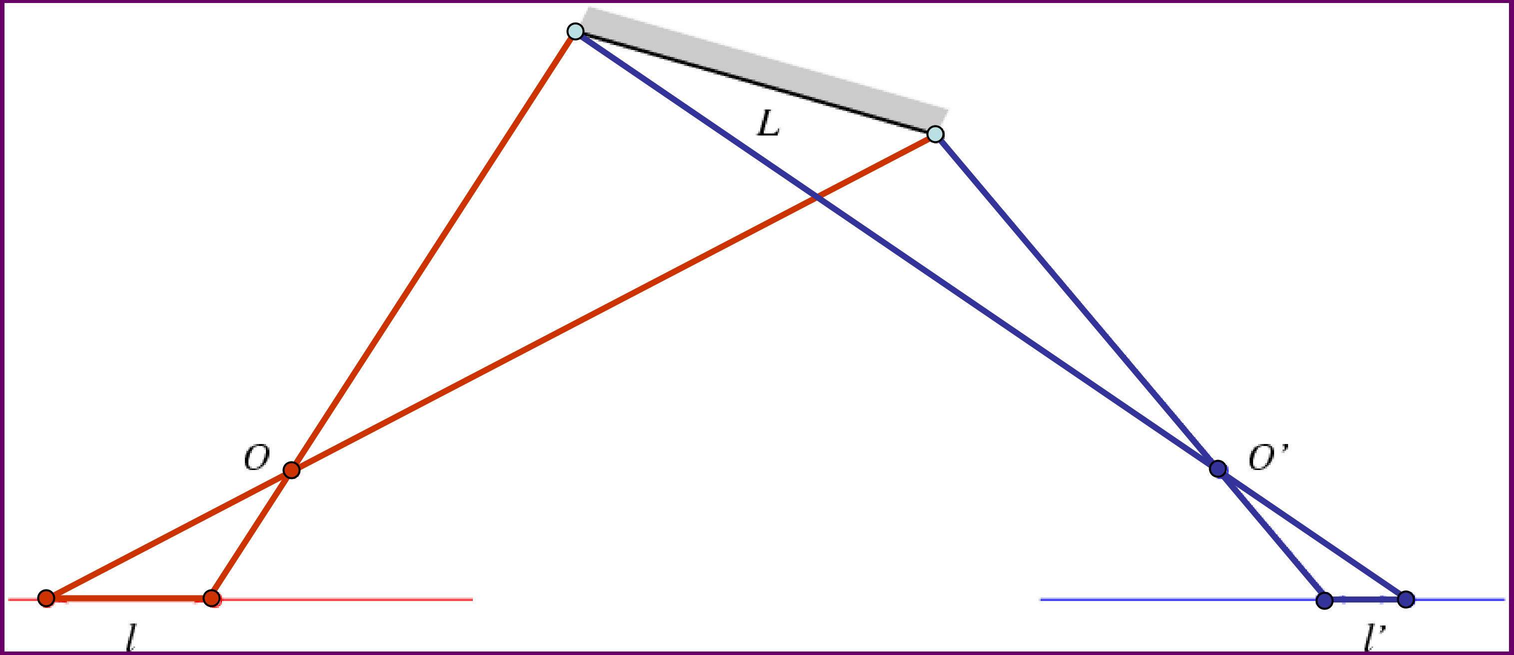
Correlation Methods (1970--)



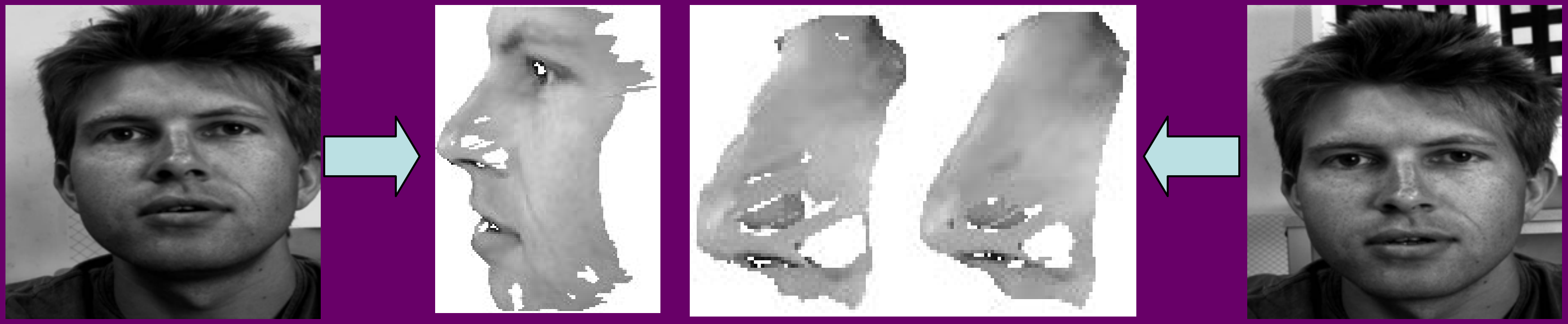
Slide the window along the epipolar line until *w.w'* is maximized.

Normalized Correlation: minimize θ instead. ⟺ Minimize $|w-w'|^2$.

Correlation Methods: Foreshortening Problems

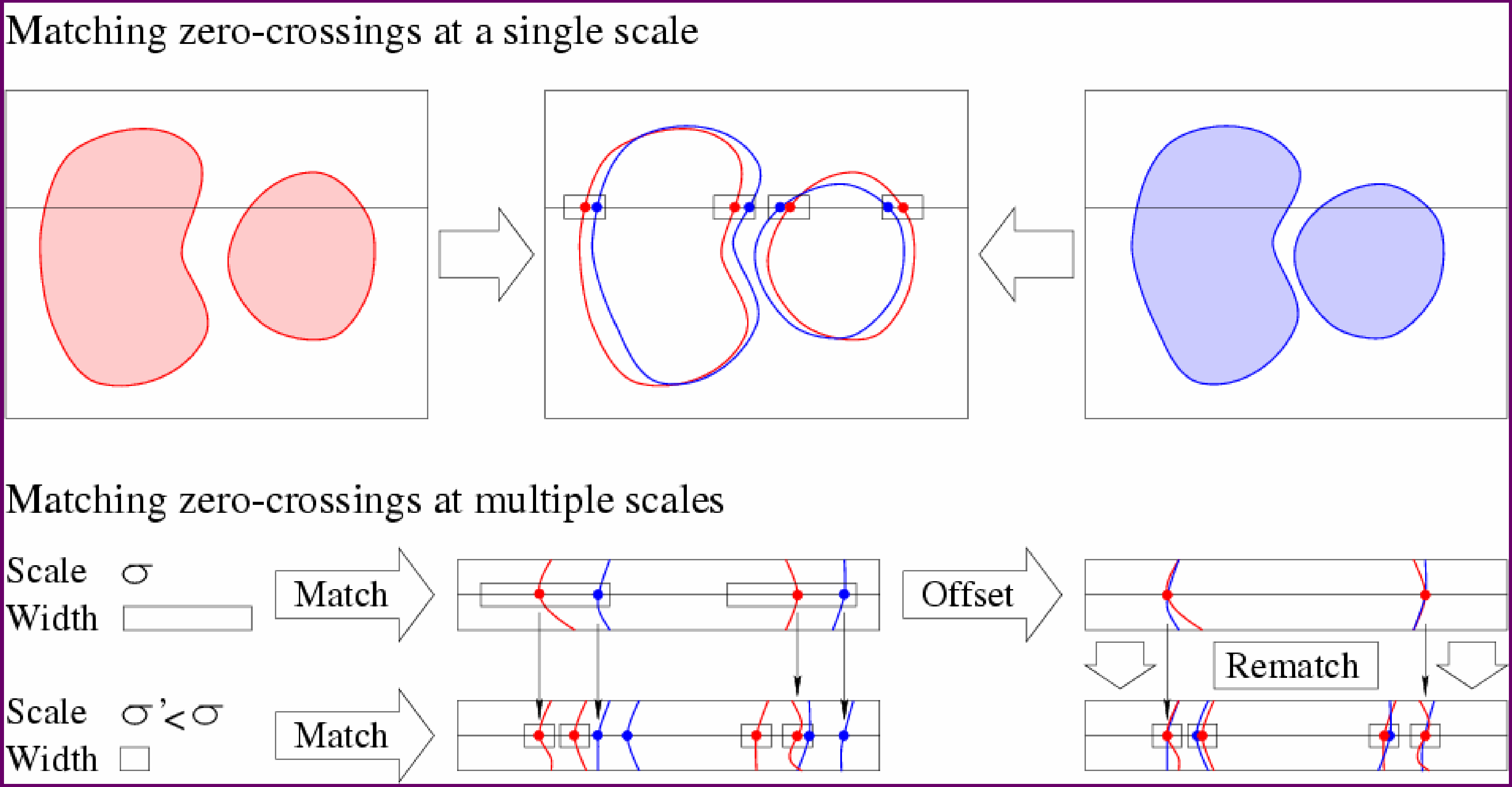


Solution: add a second pass using disparity estimates to warp the correlation windows, e.g. Devernay and Faugeras (1994).



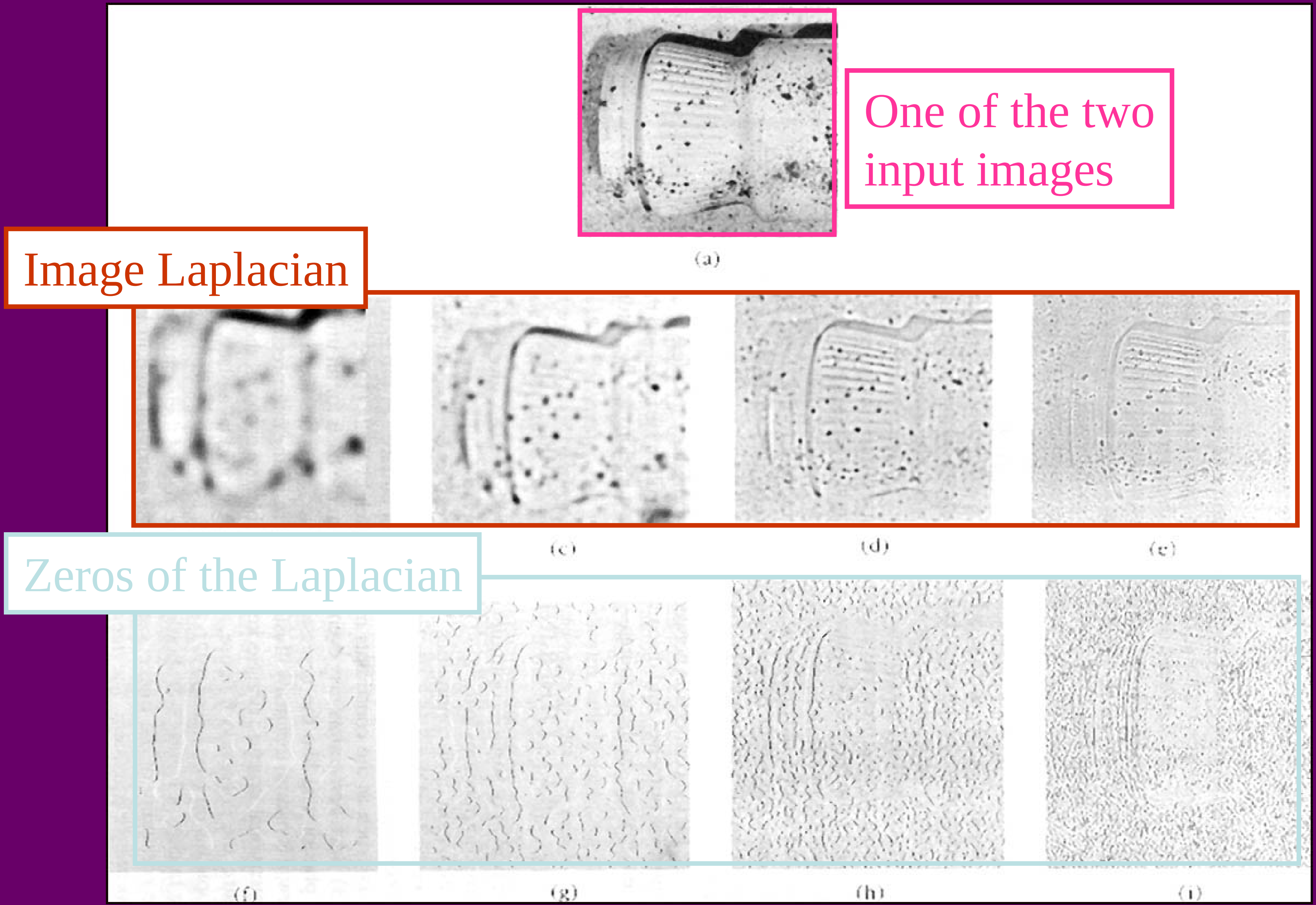
Reprinted from "Computing Differential Properties of 3D Shapes from Stereopsis without 3D Models," by F. Devernay and O. Faugeras, Proc. IEEE Conf. on Computer Vision and Pattern Recognition (1994).

Multi-Scale Edge Matching (Marr, Poggio and Grimson, 1979-81)

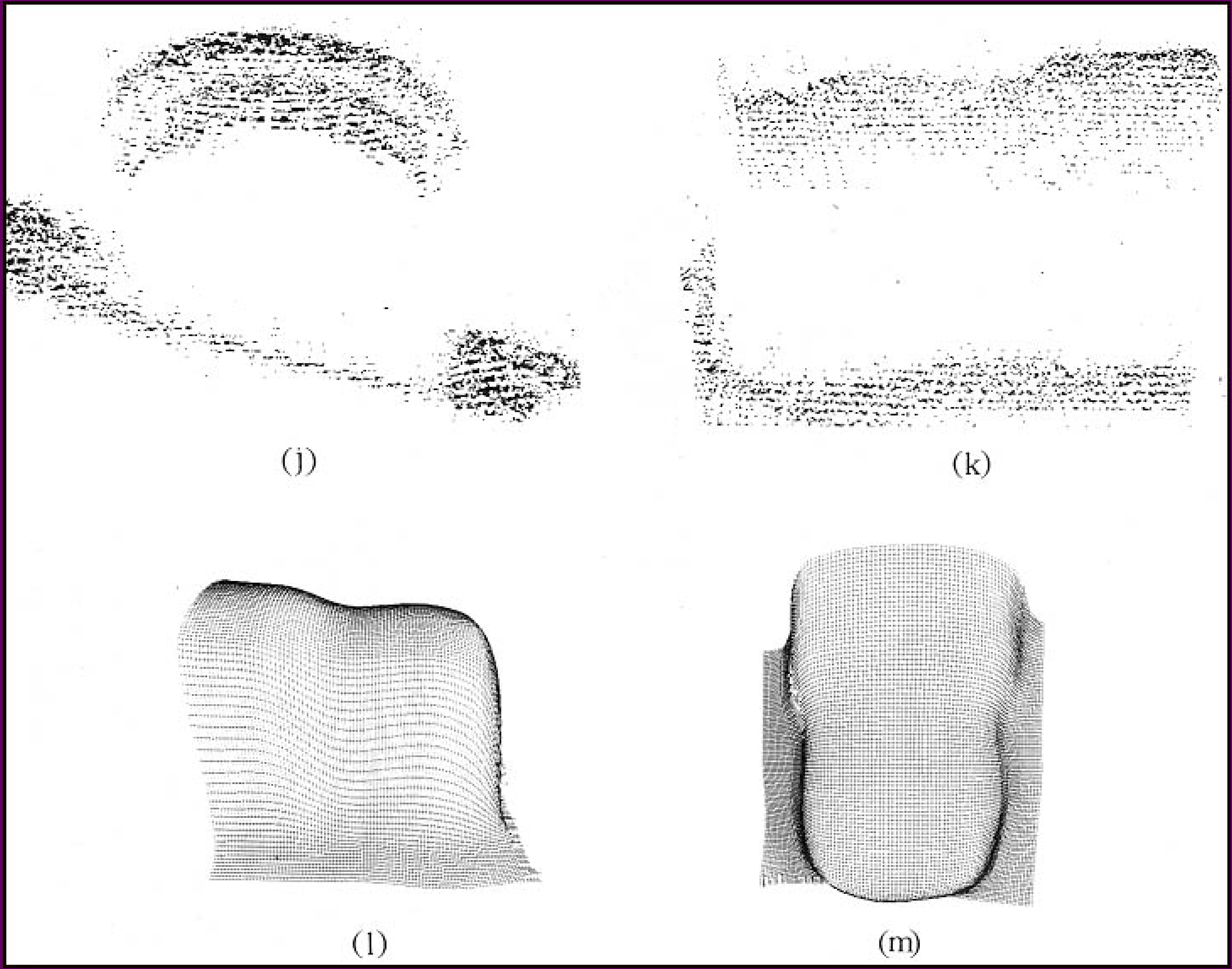


- Edges are found by repeatedly smoothing the image and detecting the zero crossings of the second derivative (Laplacian).
- Matches at coarse scales are used to offset the search for matches at fine scales (equivalent to eye movements).

Multi-Scale Edge Matching (Marr, Poggio and Grimson, 1979-81)

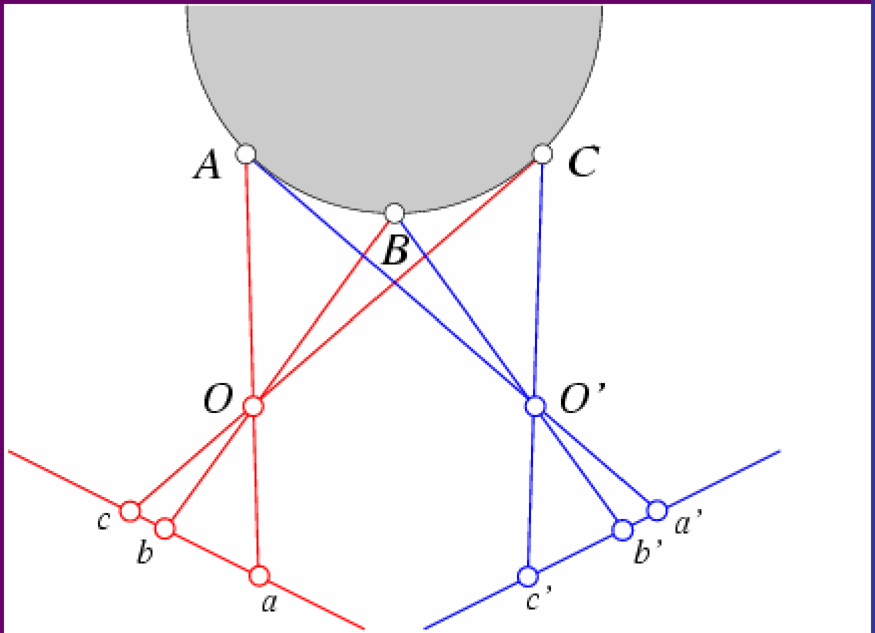


Multi-Scale Edge Matching (Marr, Poggio and Grimson, 1979-81)



(j) (k) (l) (m)

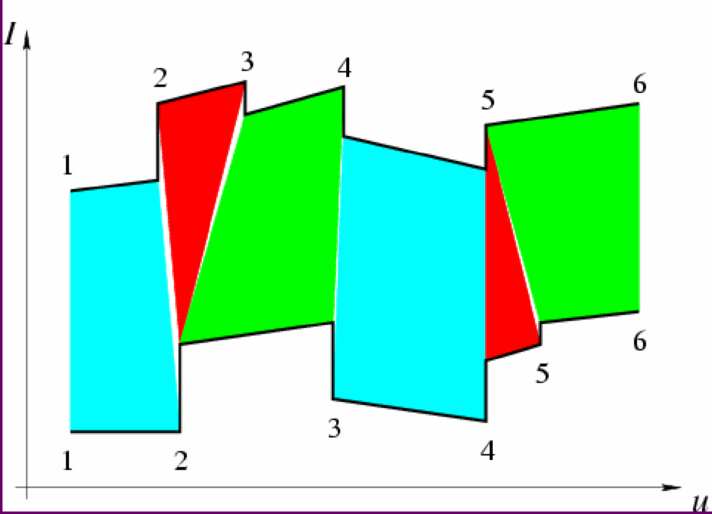
The Ordering Constraint

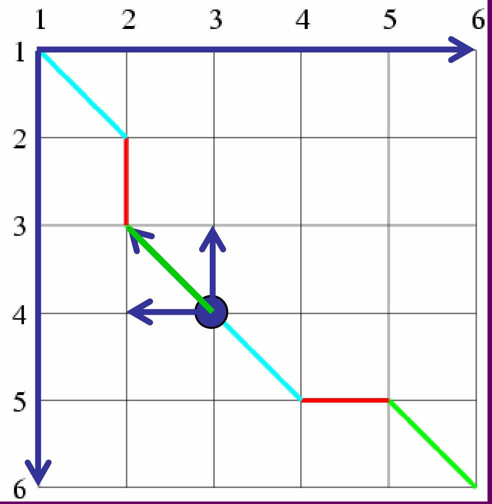


In general the points are in the same order on both epipolar lines.

But it is not always the case..

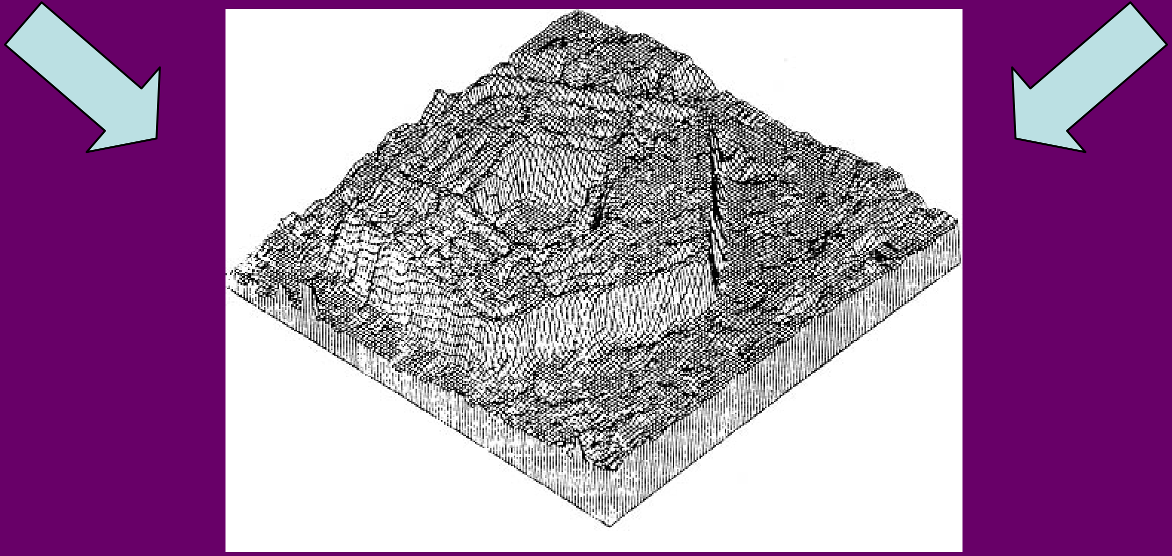
Dynamic Programming (Baker and Binford, 1981)





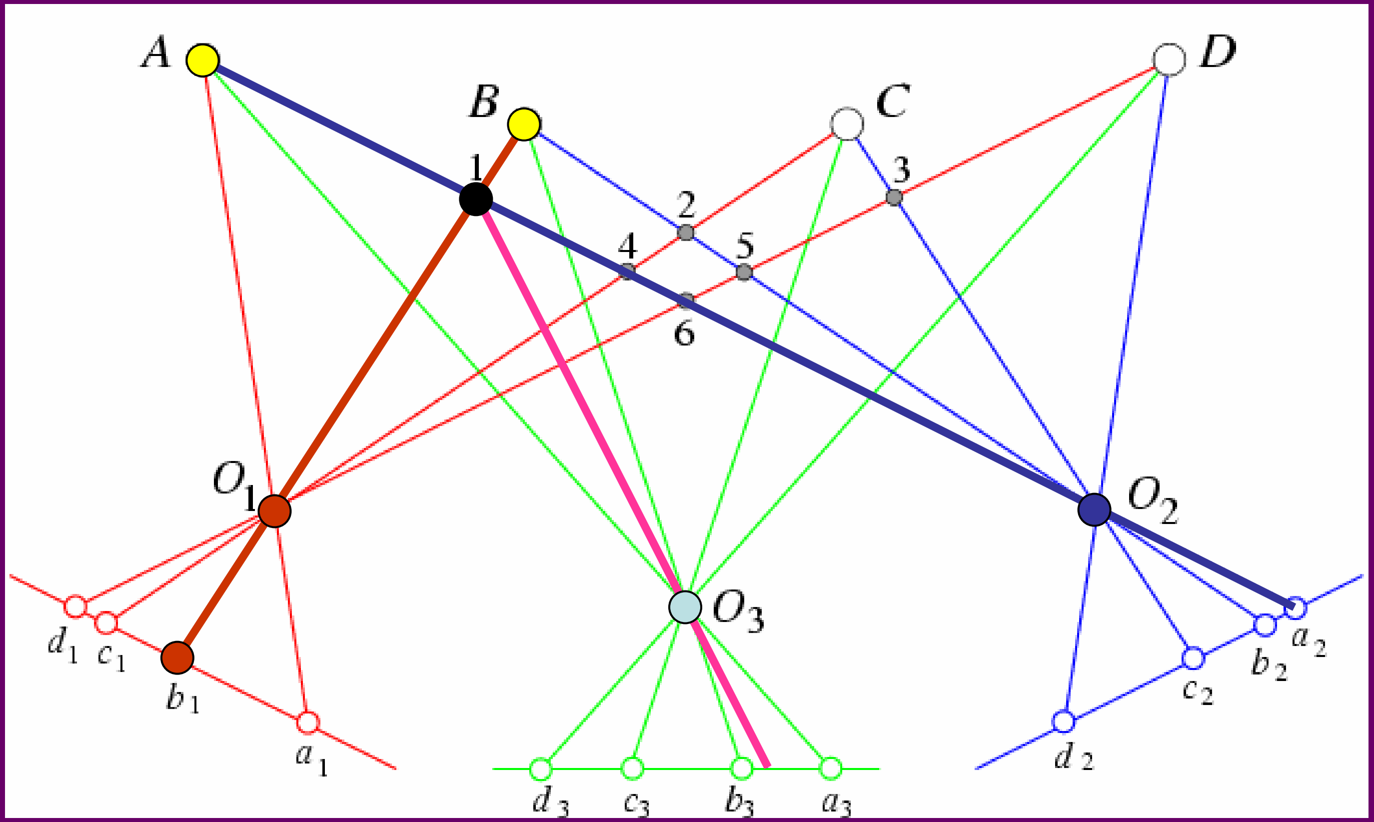
% Loop over all nodes (k, l) in ascending order.
for $k = 1$ to m do
  for $l = 1$ to n do
    % Initialize optimal cost $C(k, l)$ and backward pointer $B(k, l)$.
    $C(k, l) \leftarrow +\infty$; $B(k, l) \leftarrow$ nil;
    % Loop over all inferior neighbors (i, j) of (k, l).
    for $(i, j) \in$ Inferior − Neighbors(k, l) do
      % Compute new path cost and update backward pointer if necessary.
      $d \leftarrow C(i, j) +$ Arc − Cost(i, j, k, l);
      if $d < C(k, l)$ then $C(k, l) \leftarrow d$; $B(k, l) \leftarrow (i, j)$ endif;
      endfor;
    endfor;
  endfor;
% Construct optimal path by following backward pointers from (m, n).
$P \leftarrow \{(m, n)\}$; $(i, j) \leftarrow (m, n)$;
while $B(i, j) \neq$ nil do $(i, j) \leftarrow B(i, j)$; $P \leftarrow \{(i, j)\} \cup P$ endwhile.

Dynamic Programming (Ohta and Kanade, 1985)



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.

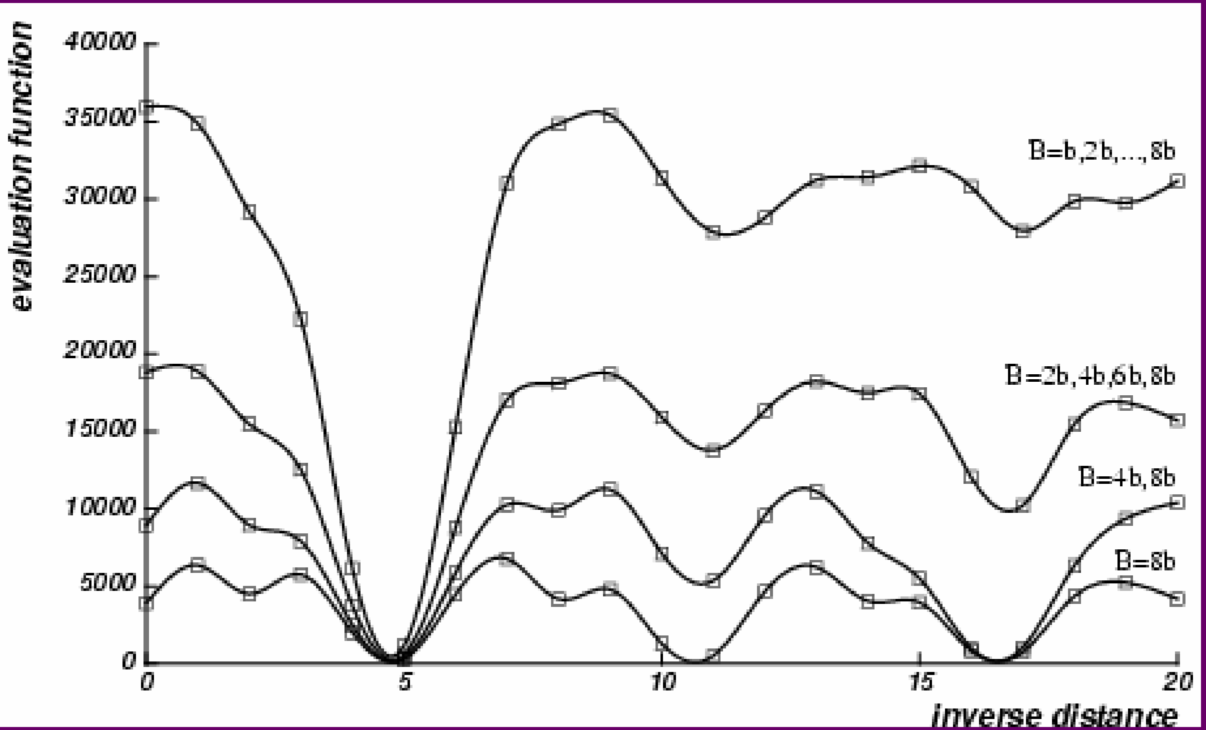
Three Views



The third eye can be used for verification..

More Views (Okutami and Kanade, 1993)

Pick a reference image, and slide the corresponding window along the corresponding epipolar lines of all other images, using inverse depth relative to the first image as the search parameter.

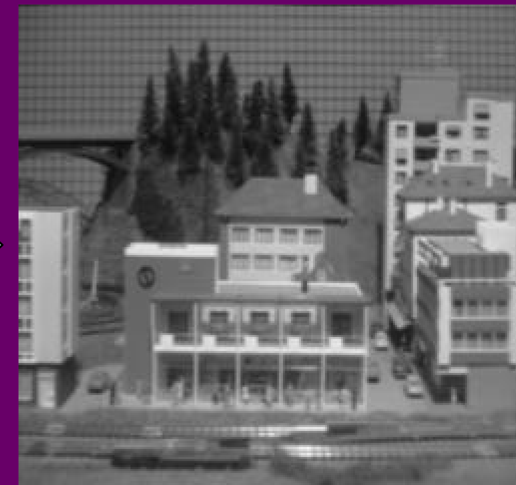


Reprinted from “A Multiple-Baseline Stereo System,” by M. Okutami and T. Kanade, IEEE Trans. on Pattern Analysis and Machine Intelligence, 15(4):353-363 (1993). \copyright 1993 IEEE.

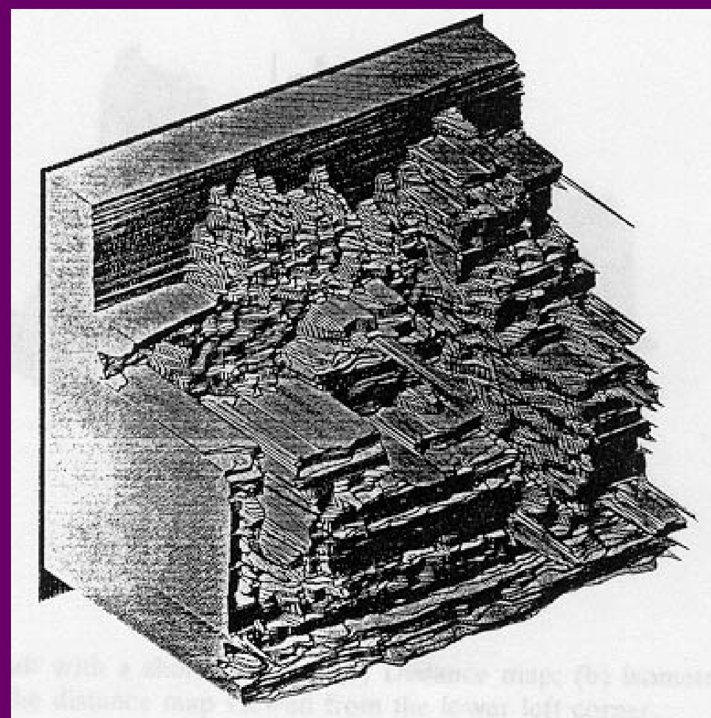
Use the sum of correlation scores to rank matches.

I1

I2



I10



Reprinted from “A Multiple-Baseline Stereo System,” by M. Okutami and T. Kanade, IEEE Trans. on Pattern Analysis and Machine Intelligence, 15(4):353-363 (1993). \copyright 1993 IEEE.