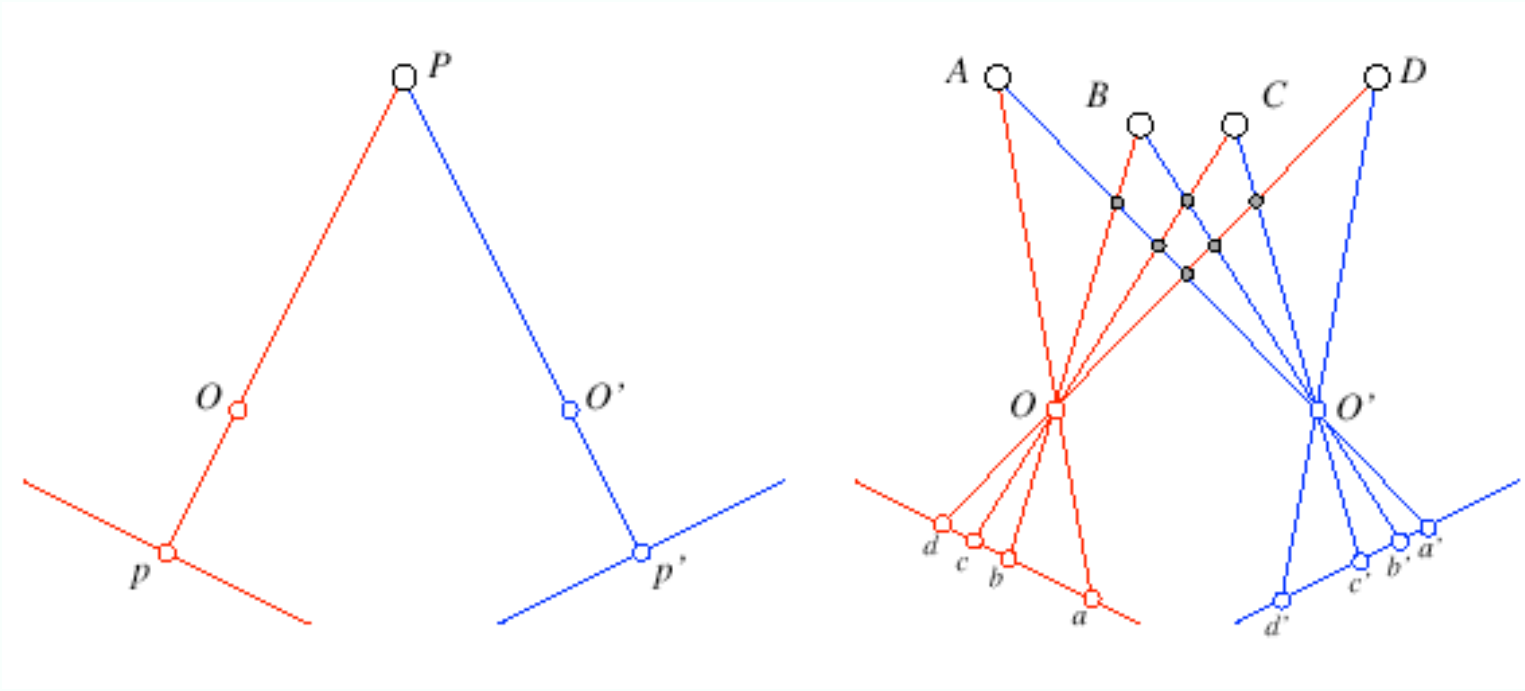
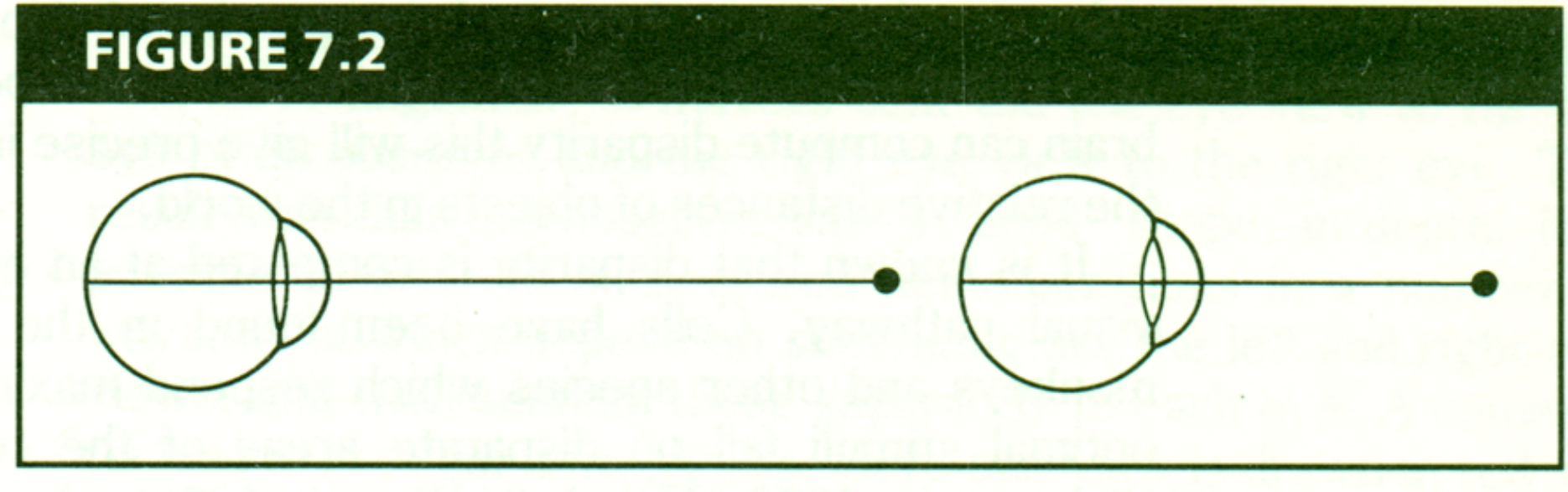
Two views

- Depth cues include
  - vergence
  - accomodation
  - stereopsis
  - motion
- Issues
  - what geometric information is available?
  - what matches are available? are correct?



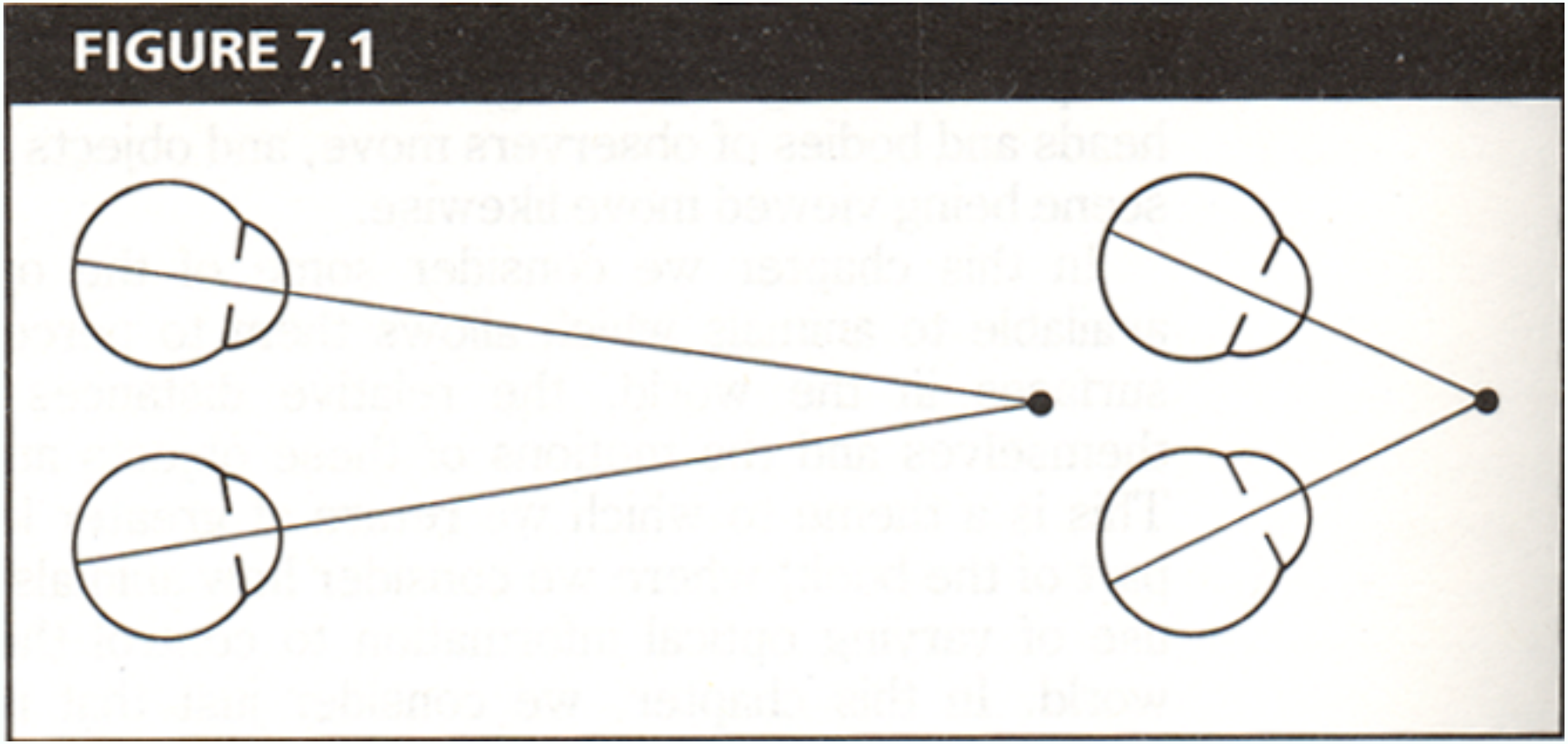
Correspondence errors = depth errors

Accomodation and focus



From Bruce and Green, Visual Perception,
Physiology, Psychology and Ecology

Convergence



From Bruce and Green, Visual Perception,
Physiology, Psychology and Ecology

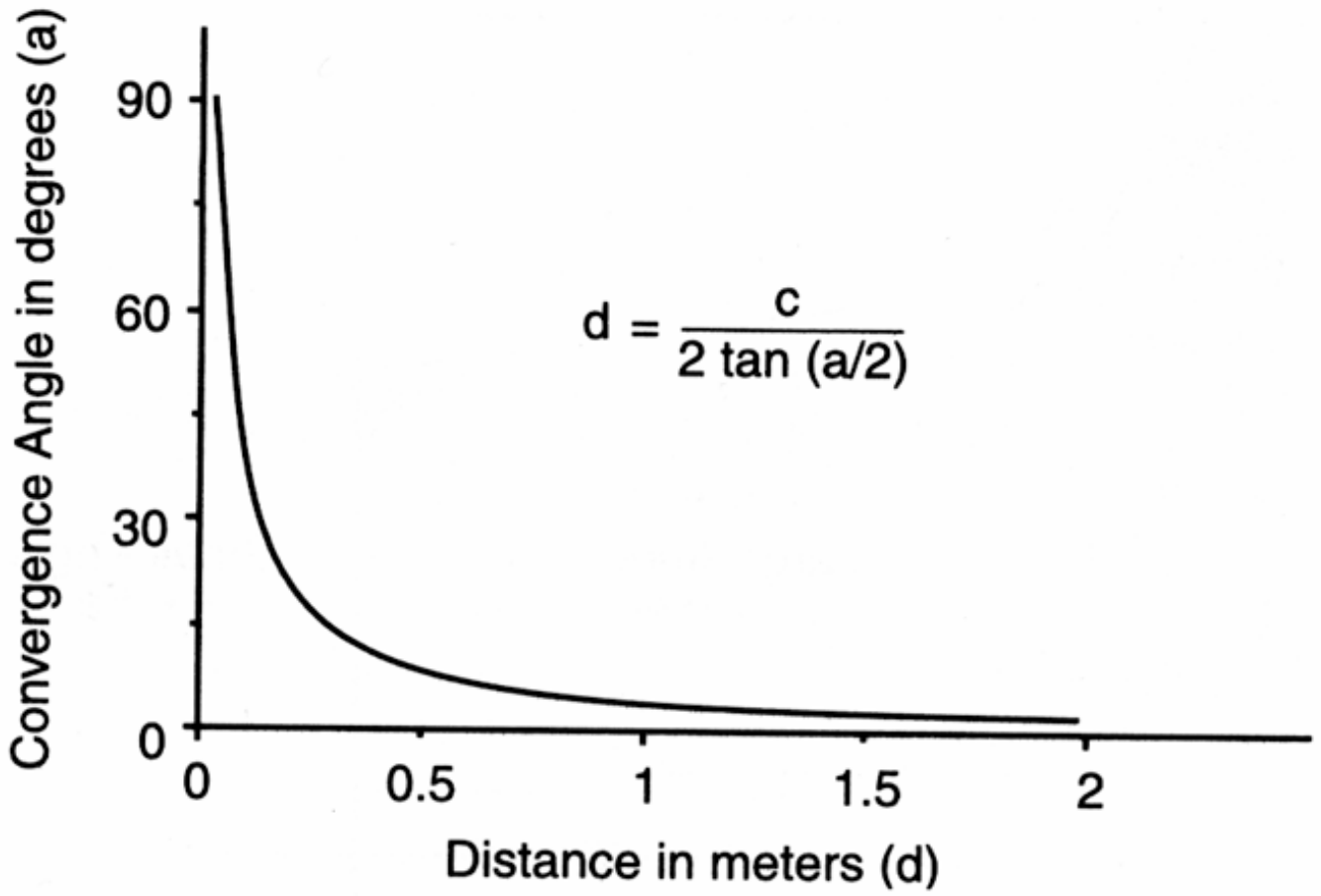
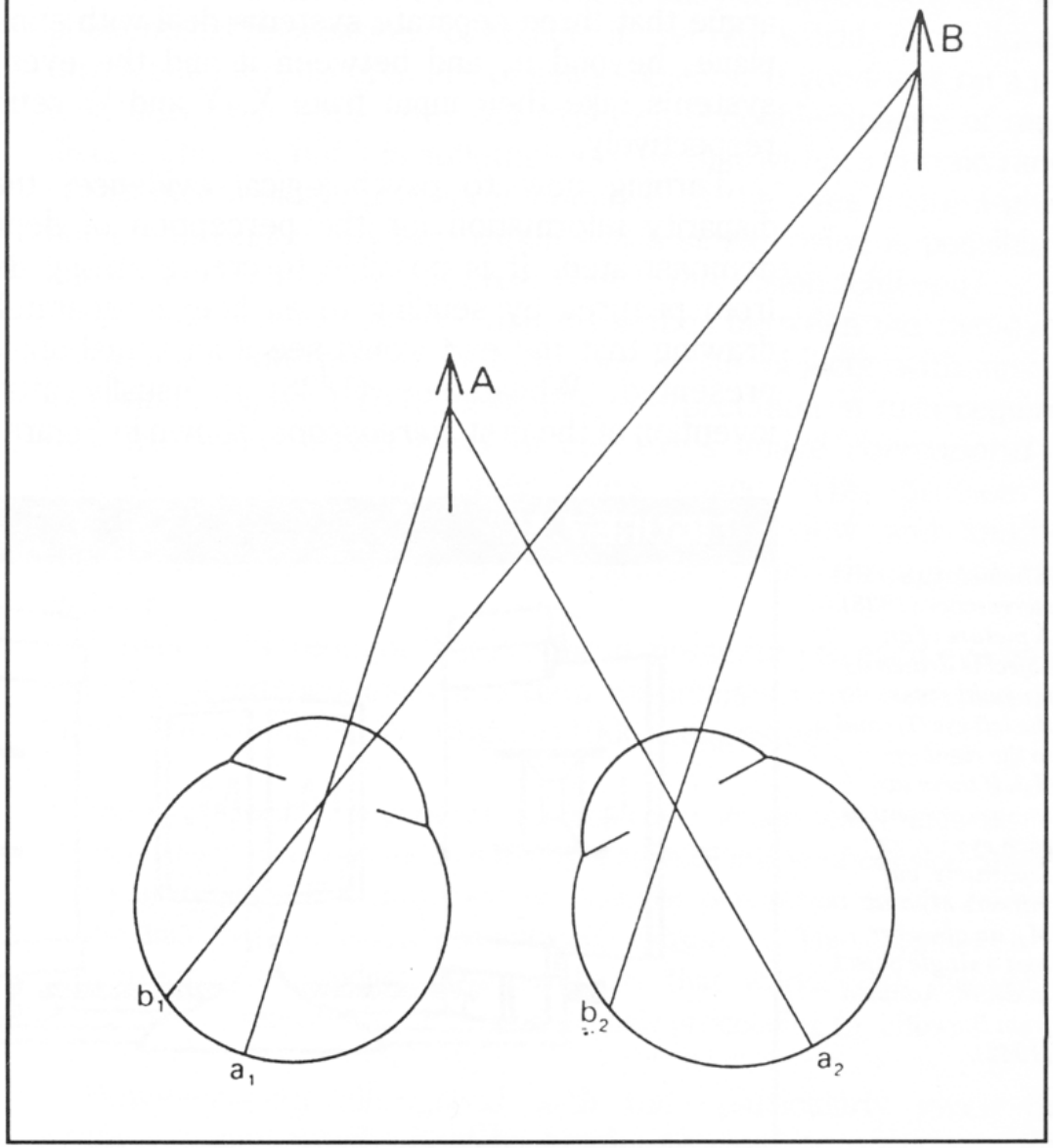


Figure 5.2.3 Convergence as a function of distance. The angle of convergence changes rapidly with distances up to a meter or two but very little after that.

From Palmer, "Vision Science", MIT Press

FIGURE 7.3



From Bruce and Green, Visual Perception, Physiology, Psychology and Ecology

Disparity occurs when
Eyes verge on one object;
Others appear at different
Visual angles

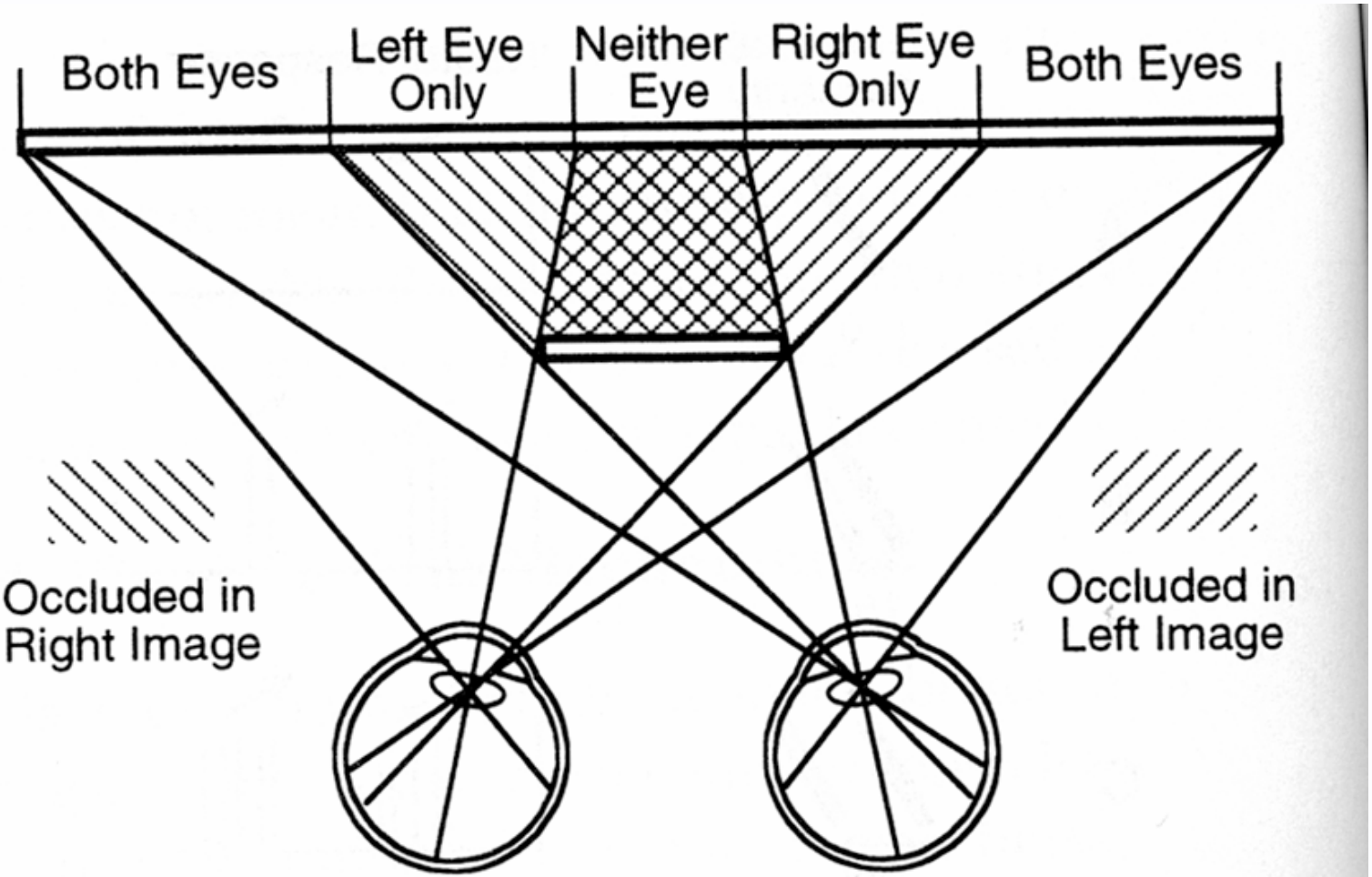
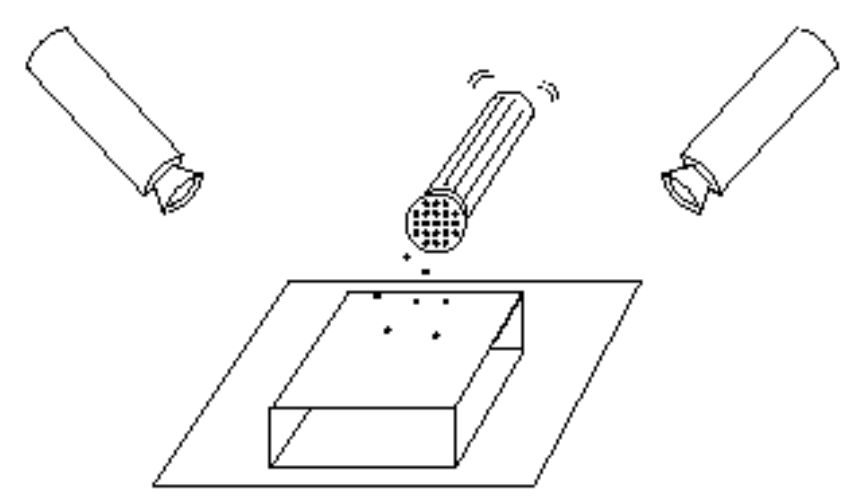
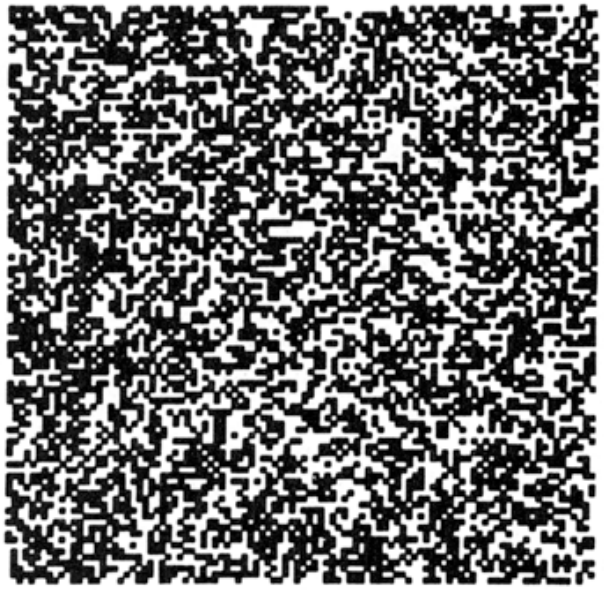


Figure 5.3.23 Da Vinci stereopsis. Depth information also arises from the fact that certain parts of one retinal image have no corresponding parts in the other image. (See text for details.)

From Palmer, "Vision Science", MIT Press

Random Dot Stereograms





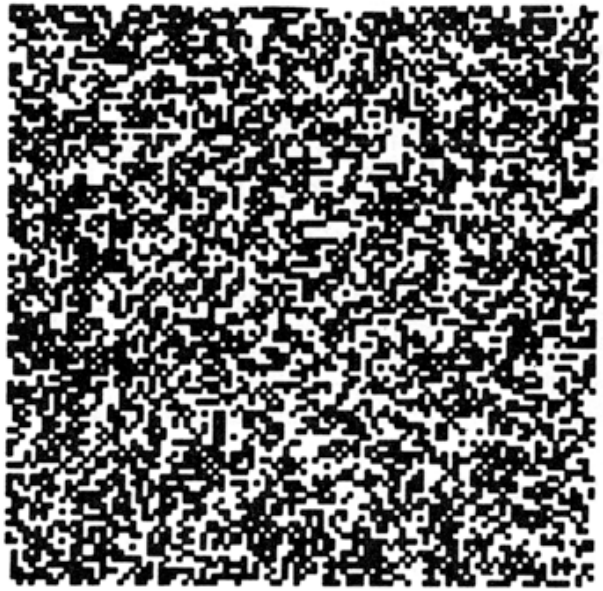
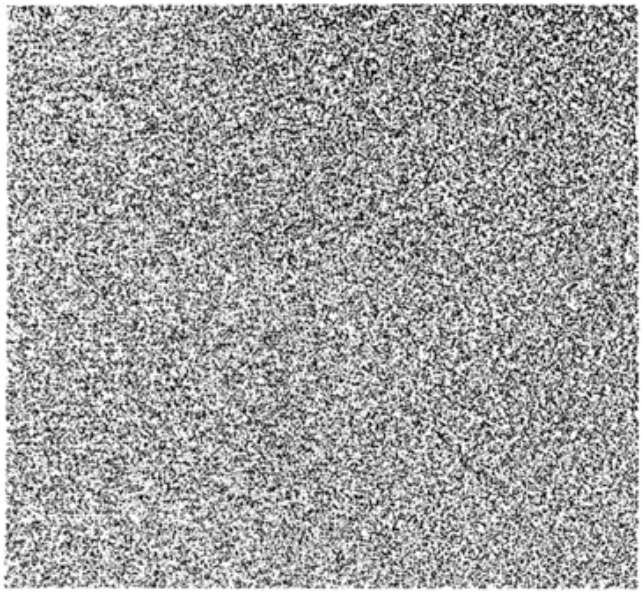


Figure 5.3.8 A random dot stereogram. These two images are derived from a single array of randomly placed squares by laterally displacing a region of them as described in the text. When they are viewed with crossed disparity (by crossing the eyes) so that the right eye's view of the left image is combined with the left eye's view of the right image, a square will be perceived to float above the page. (See pages 210–211 for instructions on fusing stereograms.)

From Palmer, "Vision Science", MIT Press



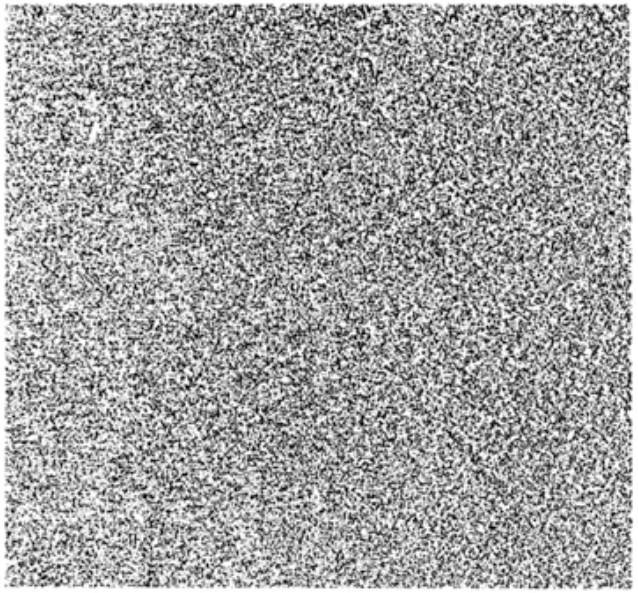


Figure 5.3.9 A random dot stereogram of a spiral surface. If these two images are fused with crossed convergence (see text on pages 210–211 for instructions), they can be perceived as a spiral ramp coming out of the page toward your face. This perception arises from the small lateral displacements of thousands of tiny dots. (From Julesz, 1971.)

From Palmer, “Vision Science”, MIT Press

Homogenous coordinates refresher

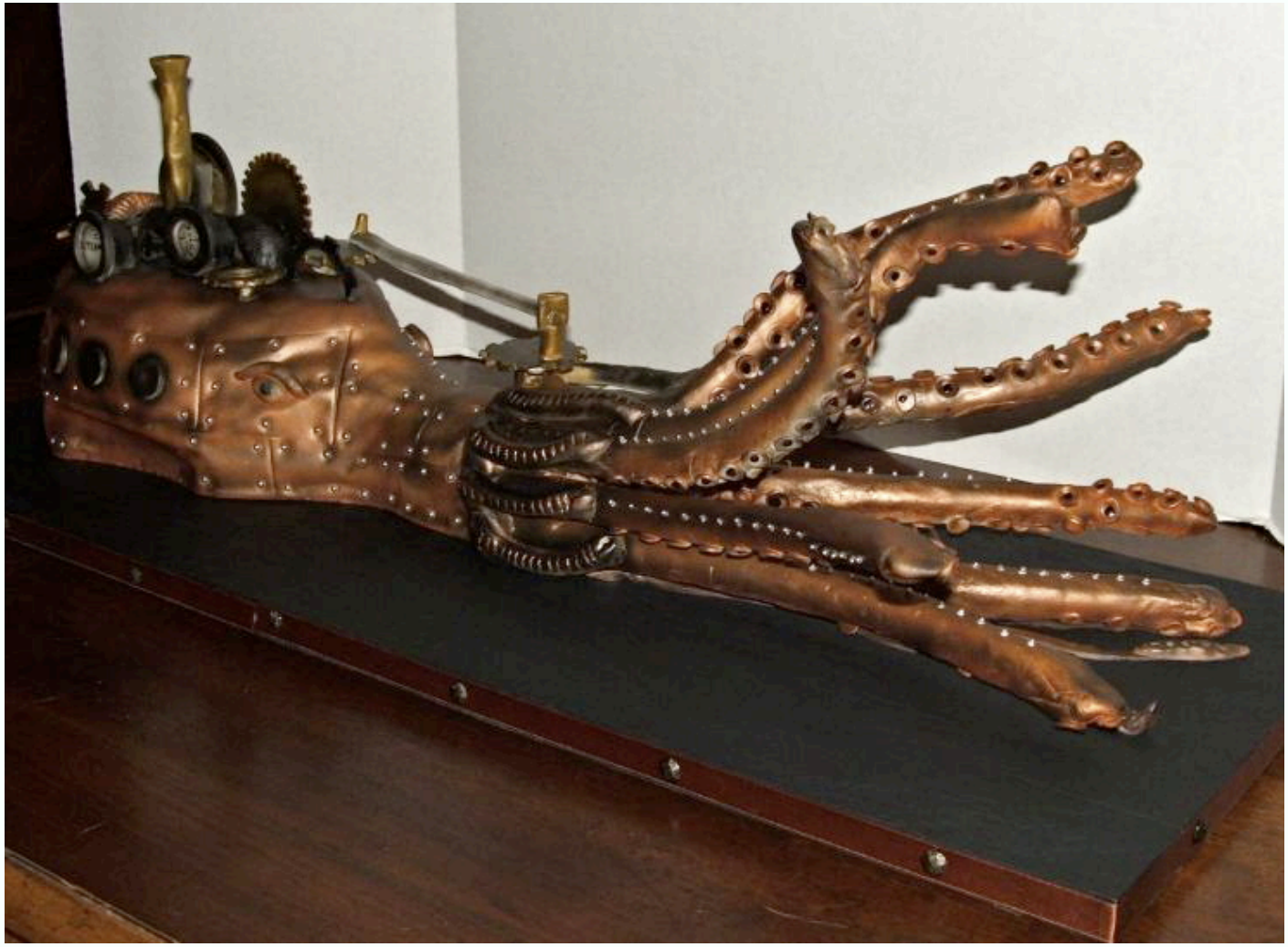
- Remember:
  - 3 coordinates in plane
  - 4 in 3D
  - equivalence relation --- two points are the same if one is parallel to other
- Lines on the plane
  - can be described using homogenous coords
- Planes in 3D
  - can be described using homogenous coords

Useful geometric construction

- Equation of line through p_1, p_2
  - det(p_1, p_2, x)=0
- Equation of plane through P_1, P_2, P_3
  - det(P_1, P_2, P_3, x)=0

The fundamental matrix

- A point in view one can lie on a line in view two
  - not anywhere IMPORTANT
  - only on epipolar line
- Each point corresponds to a line
  - the coefficients of the line depend linearly on the point's coefficients
- The family of lines passes through a point
  - the epipole



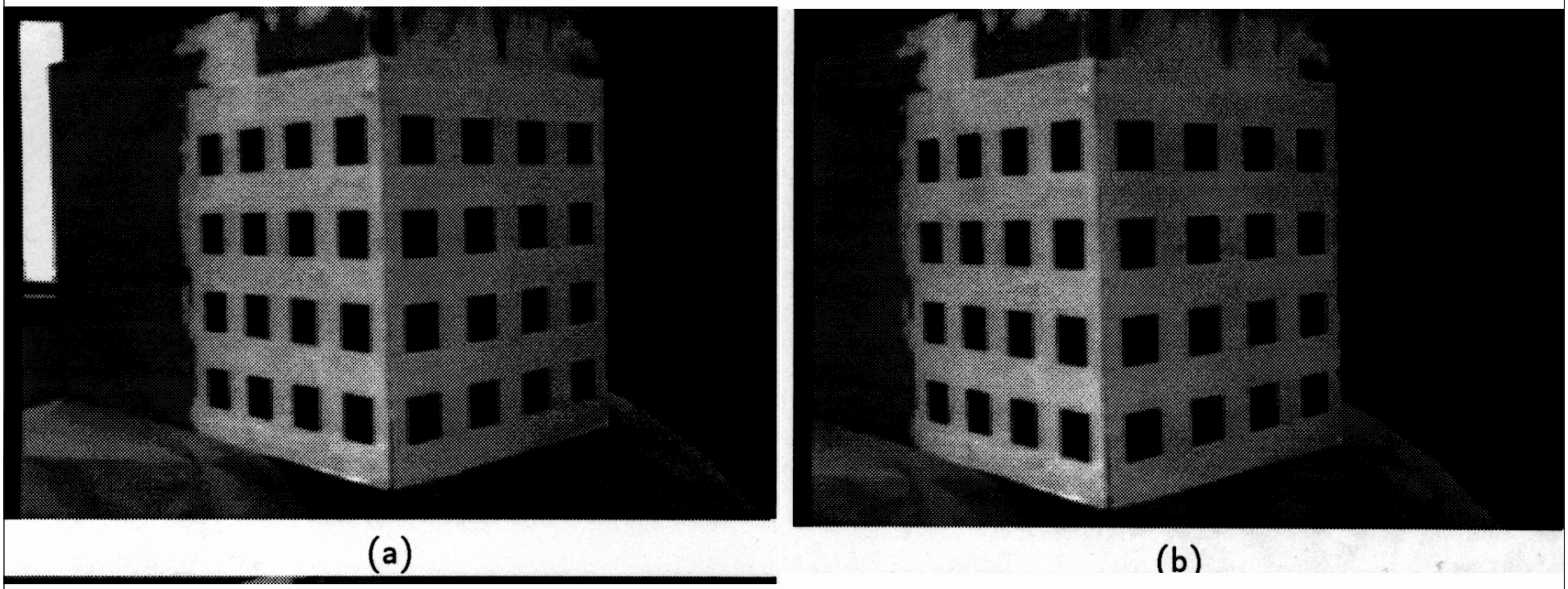
What do we know about matches?

- Geometry:
  - We work with points and lines in HC's
  - A point in left image corresponds to a line in right image
    - the coefficients of the line depend linearly on the point's coefficients
  - A 2D family of points in left gives a 1D family of lines in right
    - also, right->left
- All this means
  - there is a Fundamental matrix
    - which has determinant zero

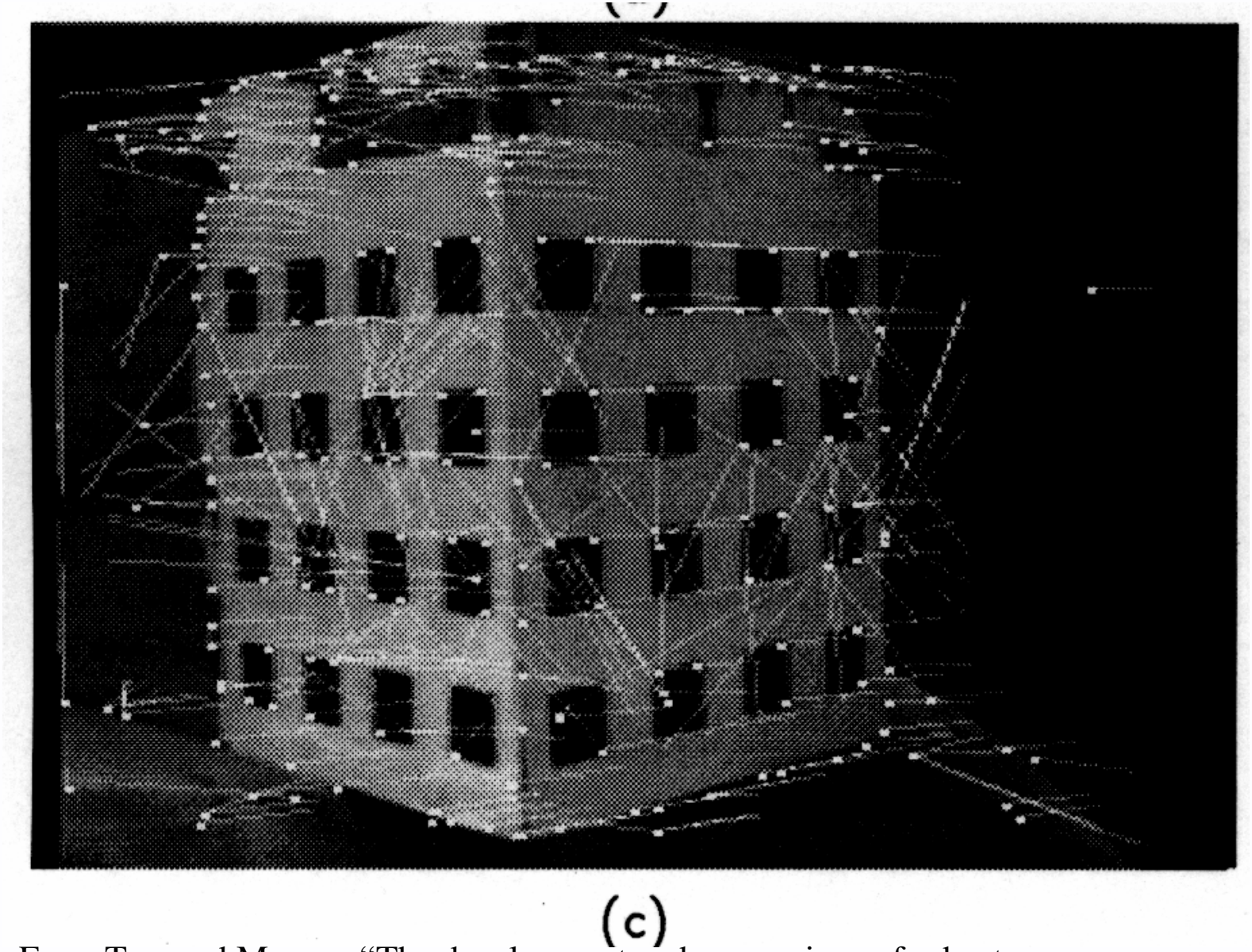
$$\mathbf{x}^T \mathcal{F} \mathbf{x} = 0$$

Estimating the fundamental matrix

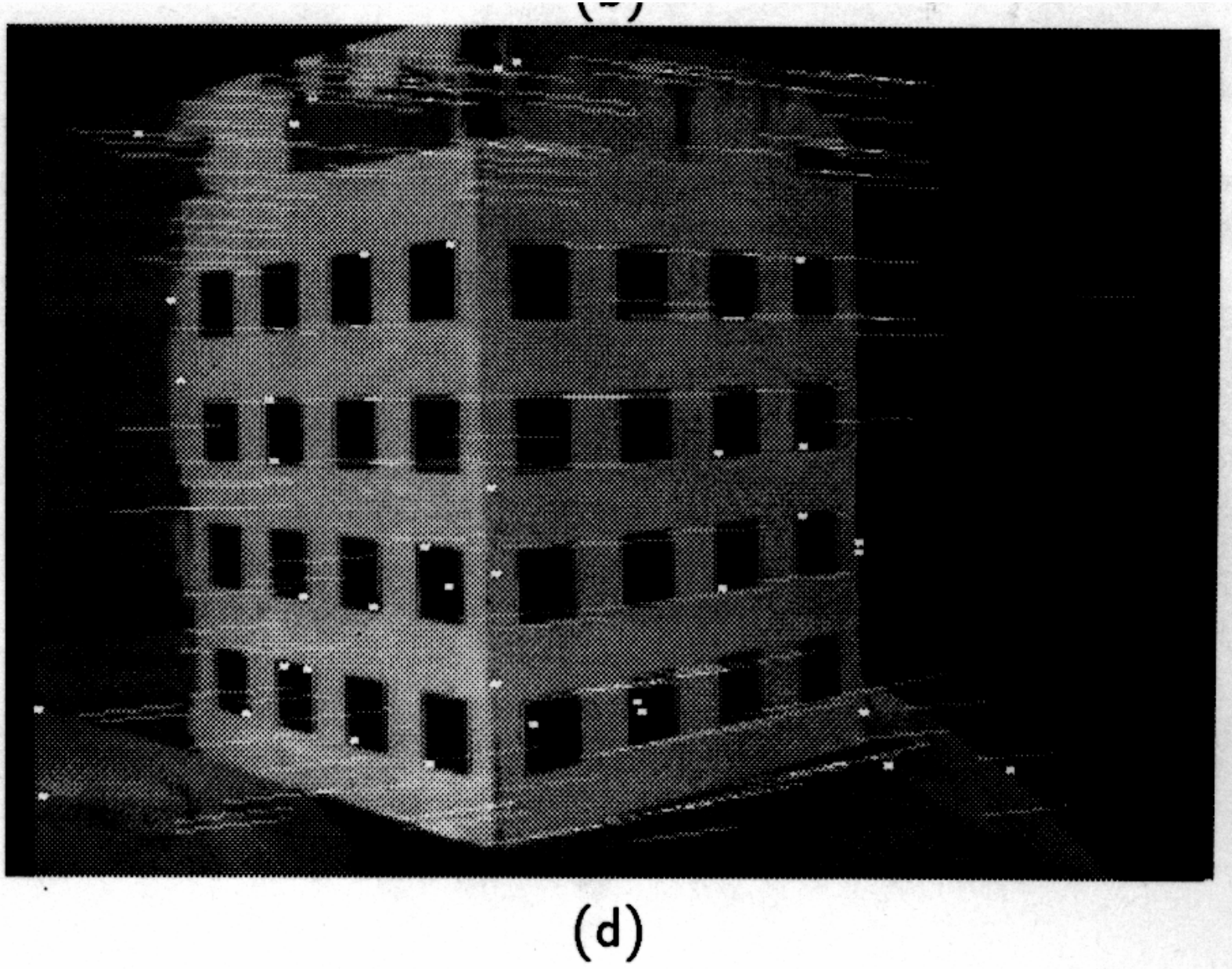
- We need to estimate 7 degrees of freedom
- Algorithm 1:
  - Take 8 point correspondences
  - Estimate linearly
- Algorithm 2 (better):
  - Take 7 point correspondences
  - Estimate linear family
  - Solve cubic
    - Check roots with 8'th point if three real



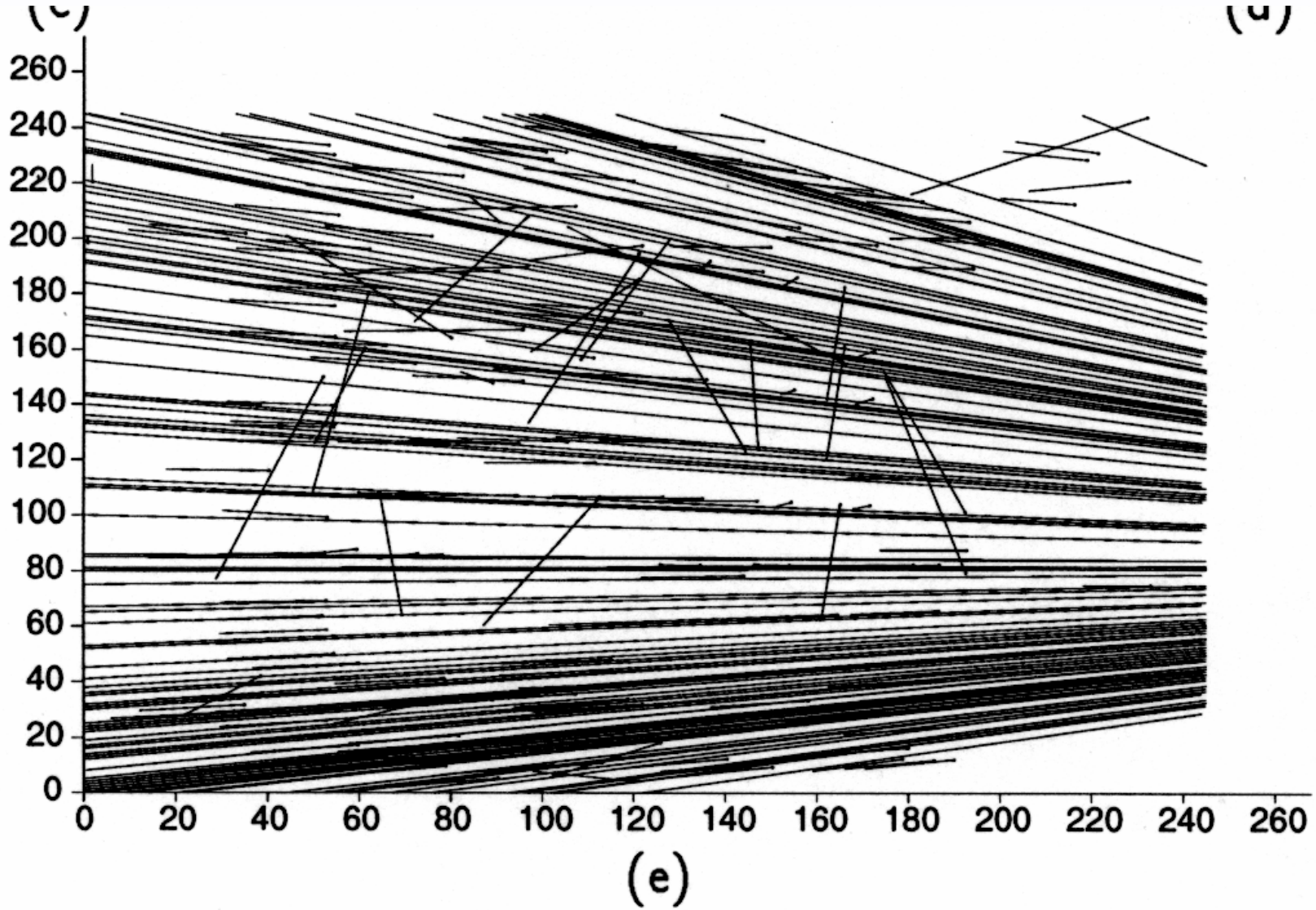
From Torr and Murray, “The development and comparison of robust methods for estimating the fundamental matrix”



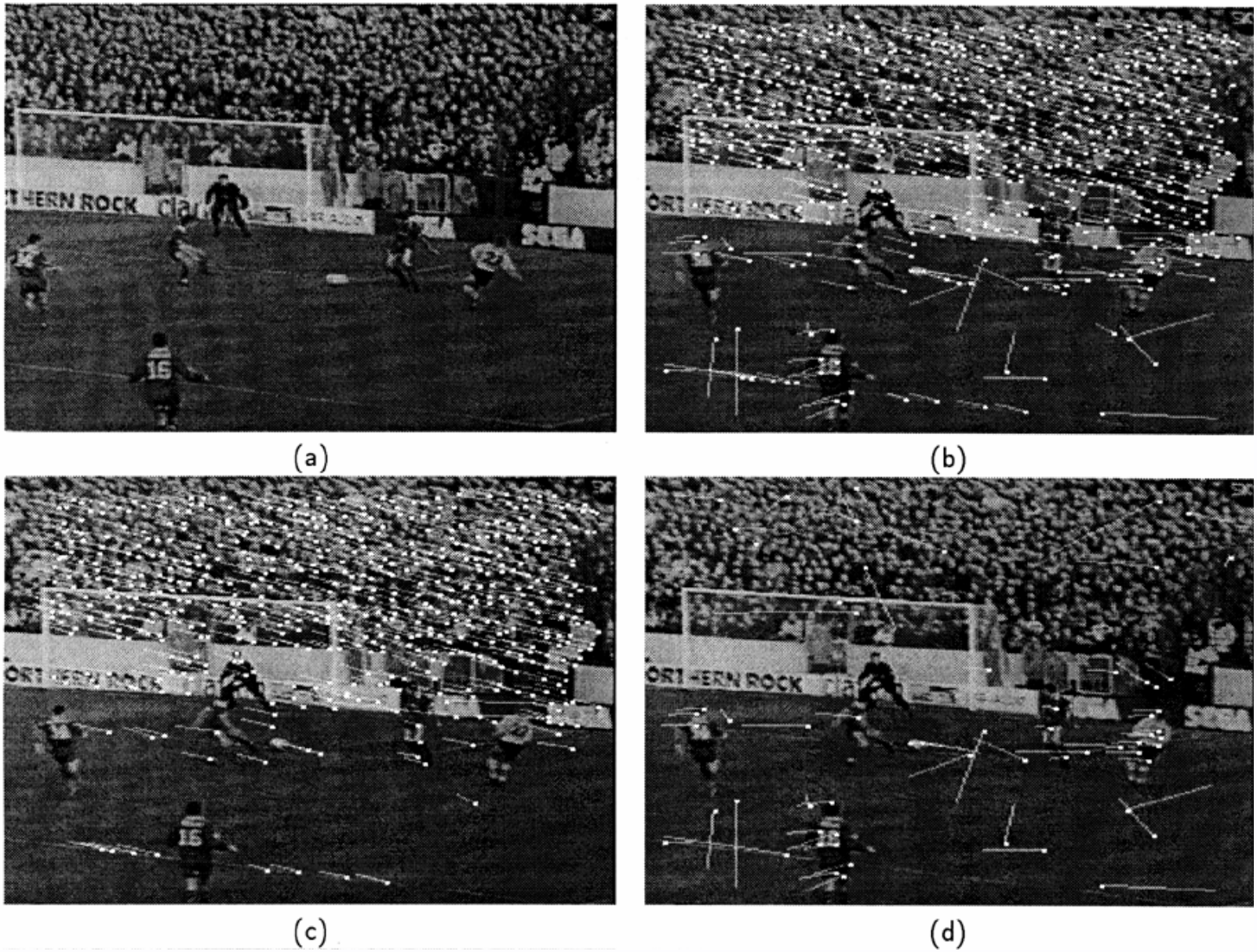
From Torr and Murray, “The development and comparison of robust methods for estimating the fundamental matrix”



From Torr and Murray, “The development and comparison of robust methods for estimating the fundamental matrix”

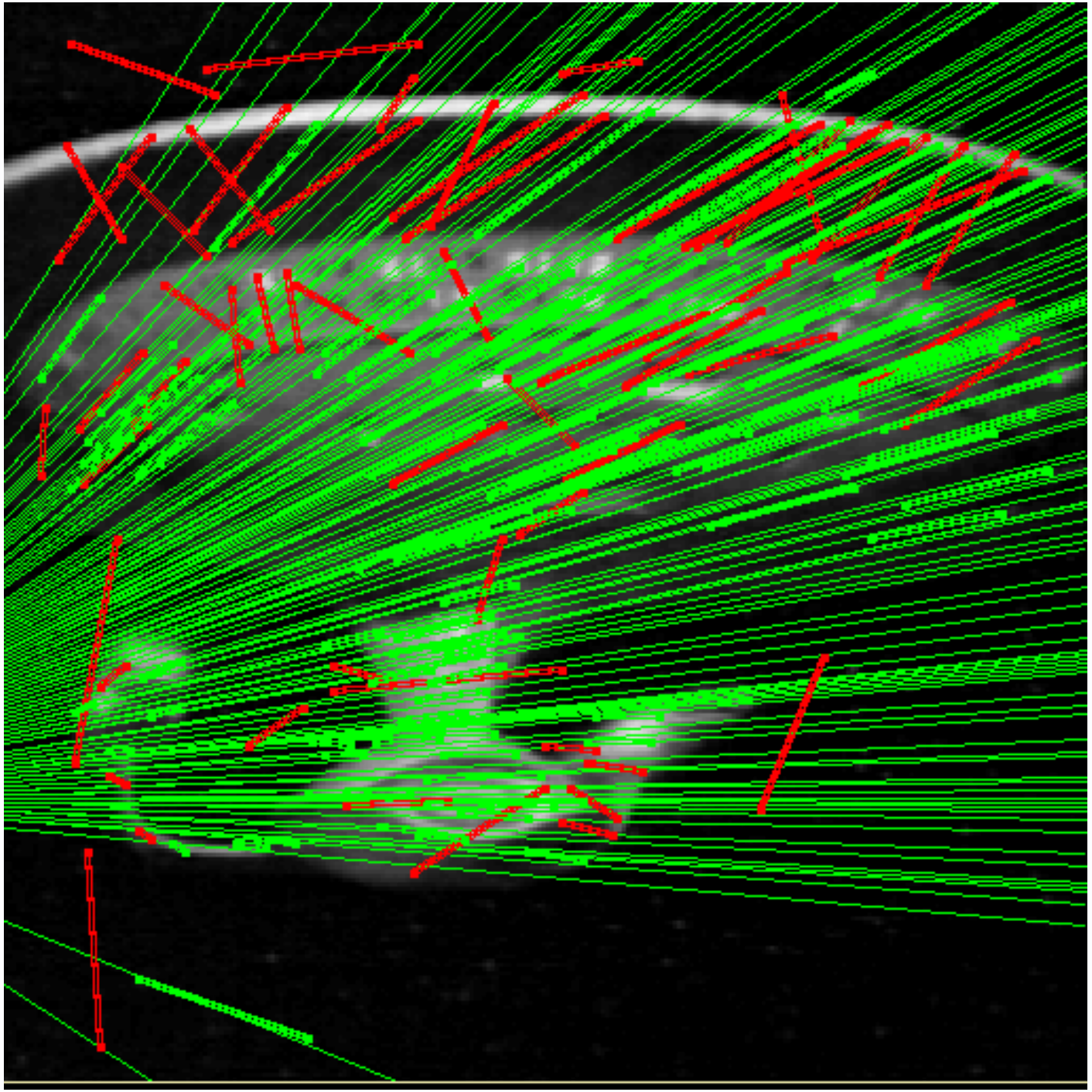


From Torr and Murray, “The development and comparison of robust methods for estimating the fundamental matrix”



From Torr and Murray, "The development and comparison of robust methods for estimating the fundamental matrix"

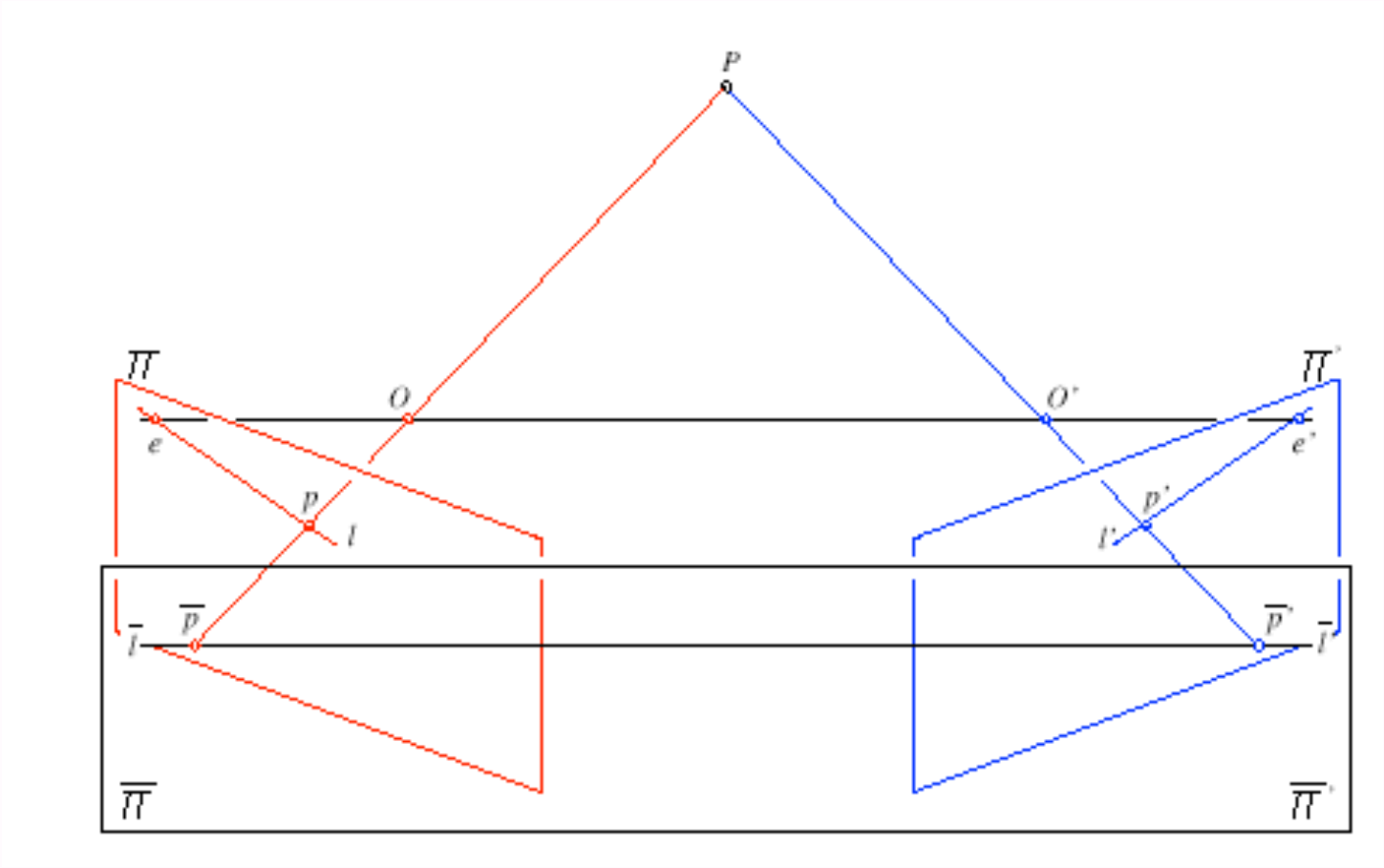
From Phil Torr's
Web page, at
Microsoft research



Pragmatics

- Simplify activities by assuming
  - That camera image planes are coplanar
  - That focal lengths are the same
  - That the separation is parallel to the scanlines
  - (all this used to be called the epipolar configuration)

Rectification



Triangulation

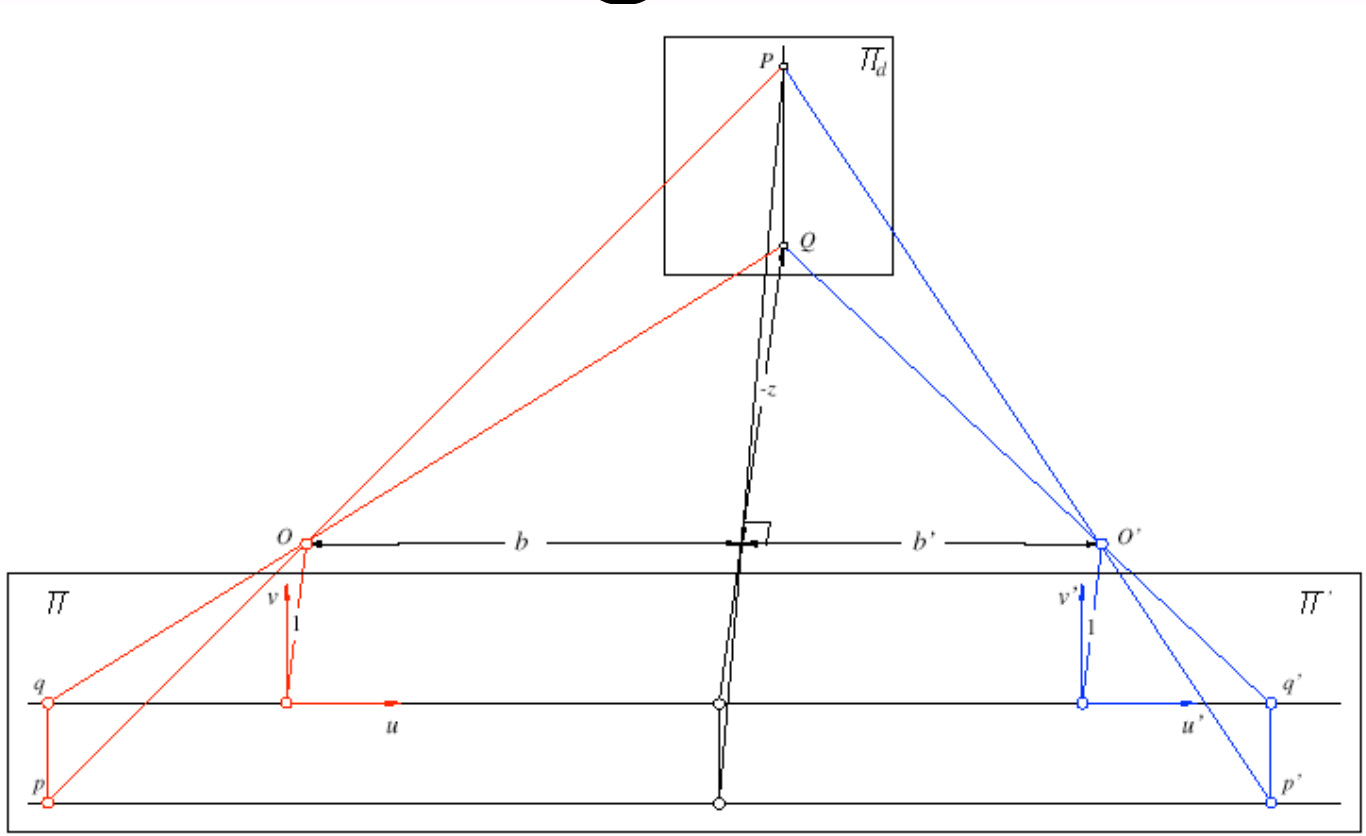
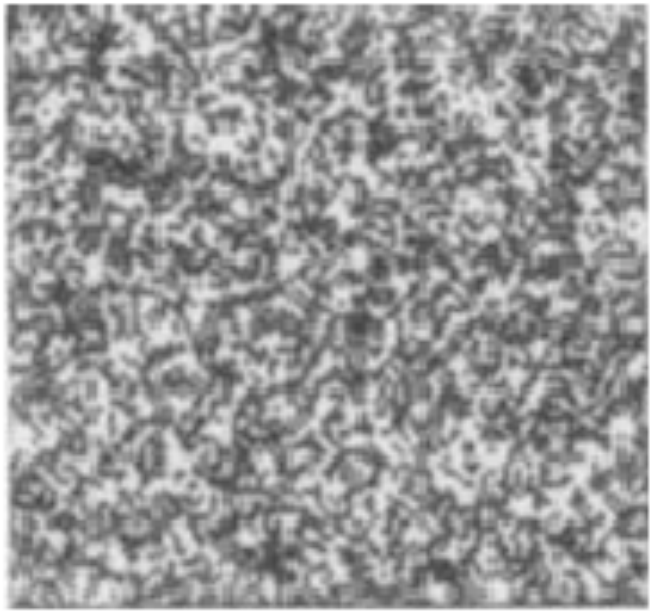
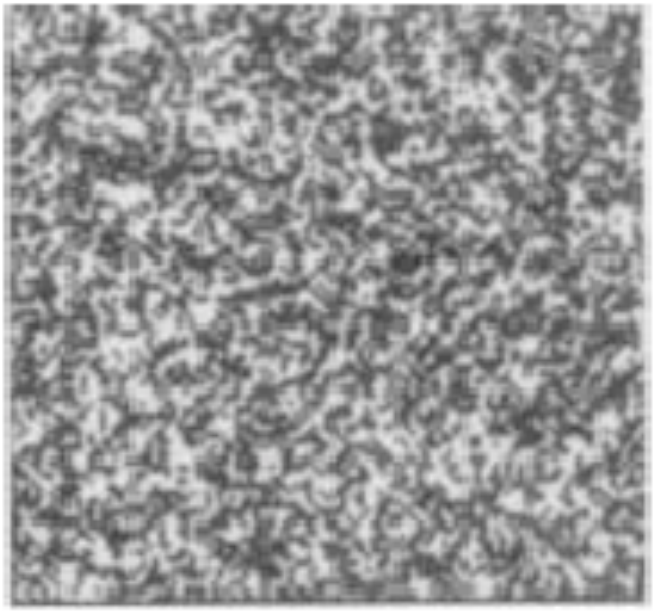
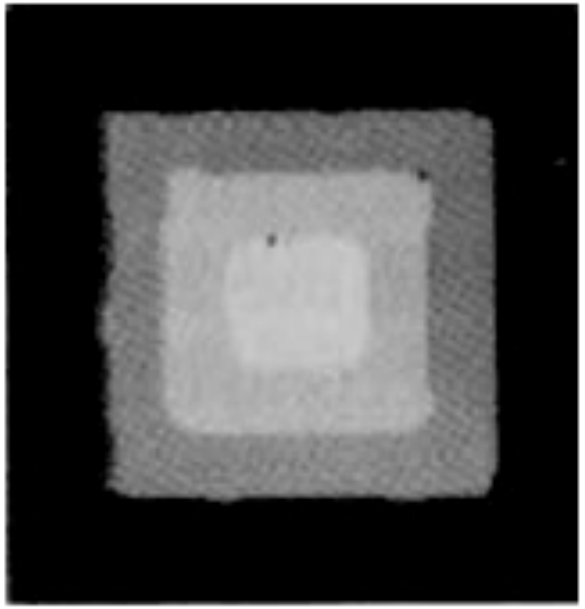
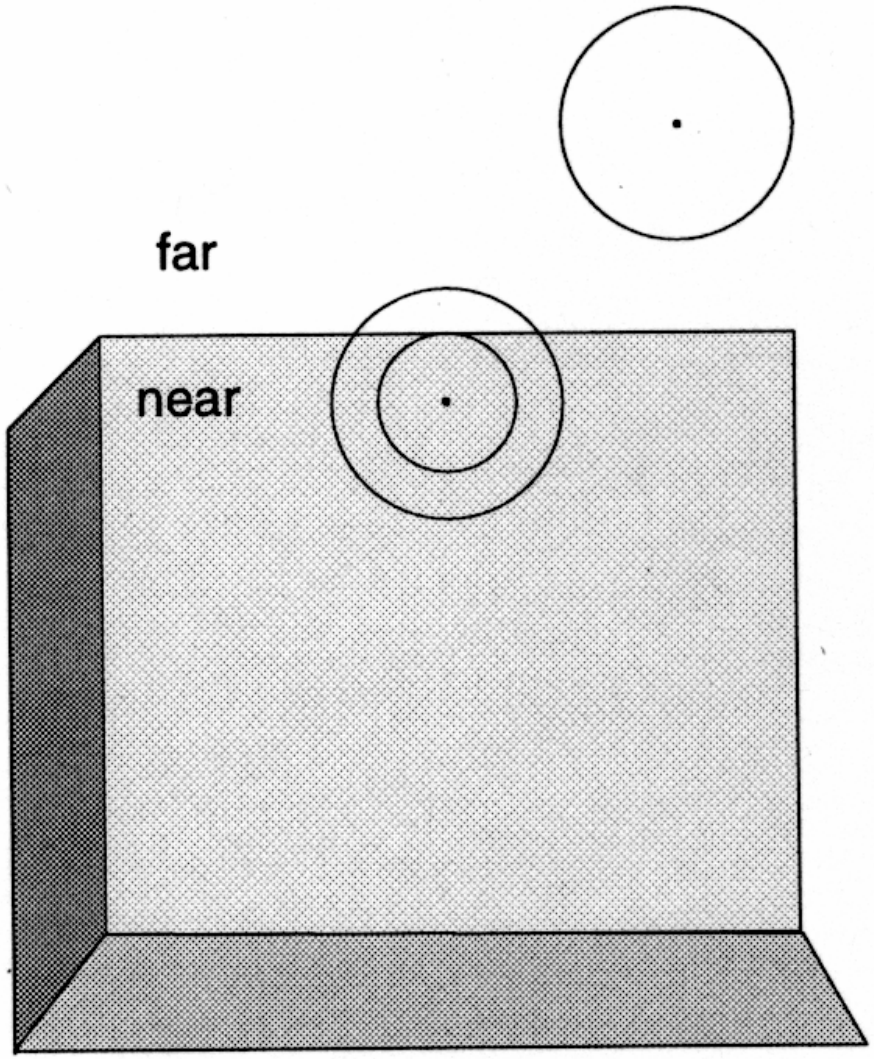


Figure 13.6. Triangulation for rectified images: the rays associated with two points p and p' on the same scanline are by construction guaranteed to intersect in some point P. As shown in the text, the depth of P relative to the coordinate system attached to the left camera is inversely proportional to the disparity $d = u' - u$. In particular, the preimage of all pairs of image points with constant disparity d is a *frontoparallel* plane Π_d (i.e., a plane parallel to the camera retinas).

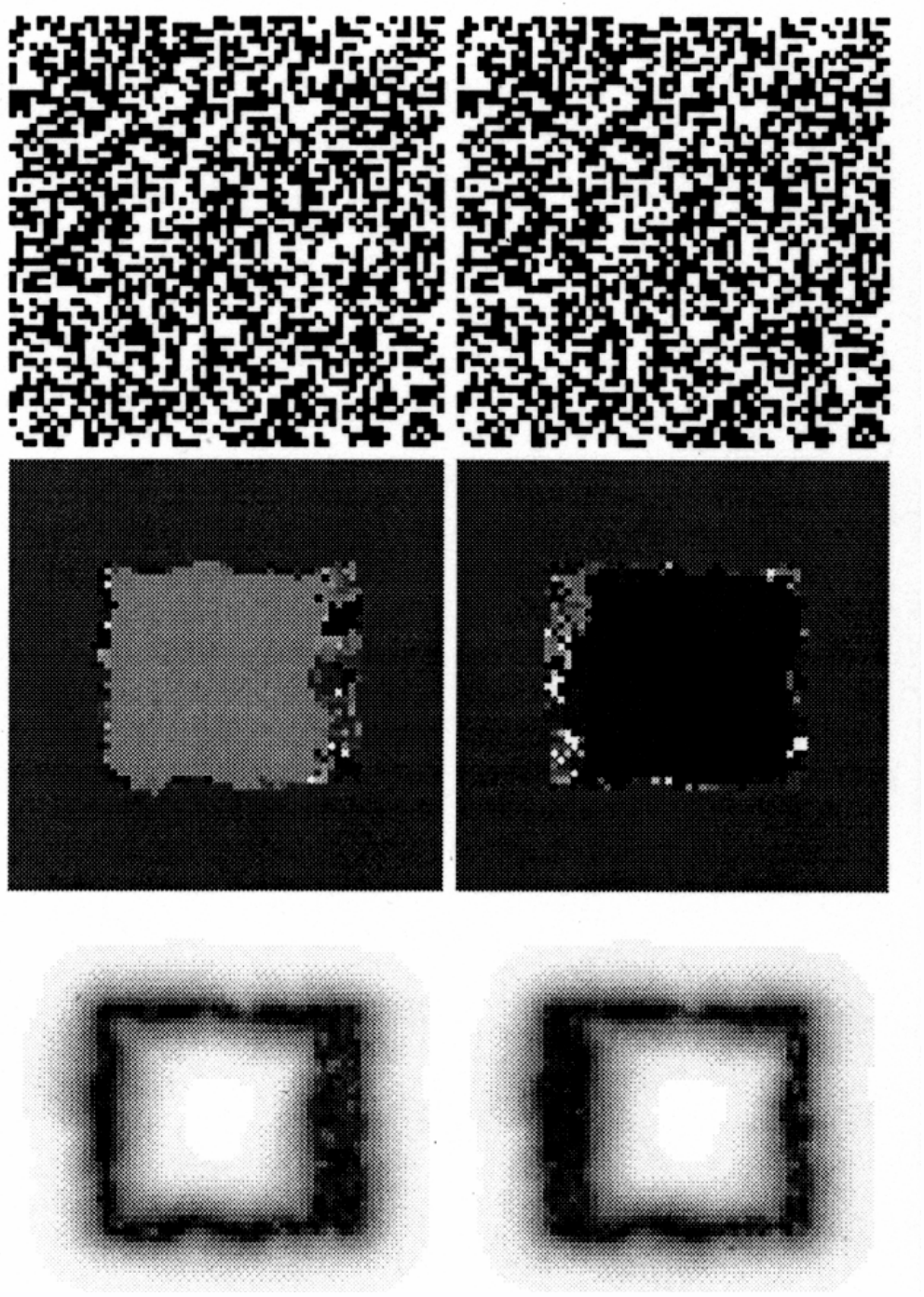
Pragmatics

- Issue
  - Match points
- Strategy
  - correspondences occur only along scanlines
  - represent points from coarse to fine
    - scale problems - some scales are misleading
- Issue
  - some points don’t have correspondences (occlusion)
- Match left to right, then right to left
  - if they don’t agree, break match

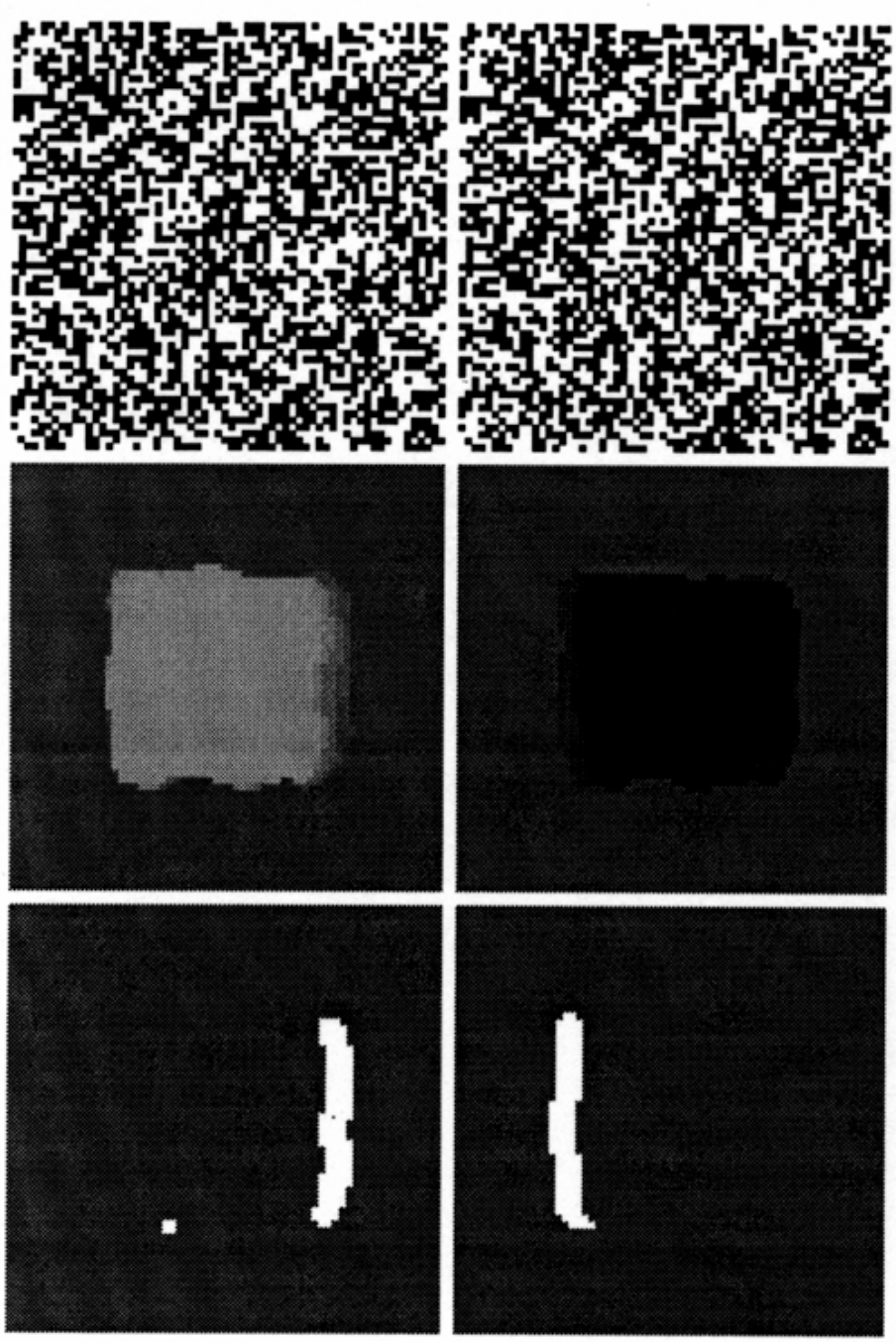






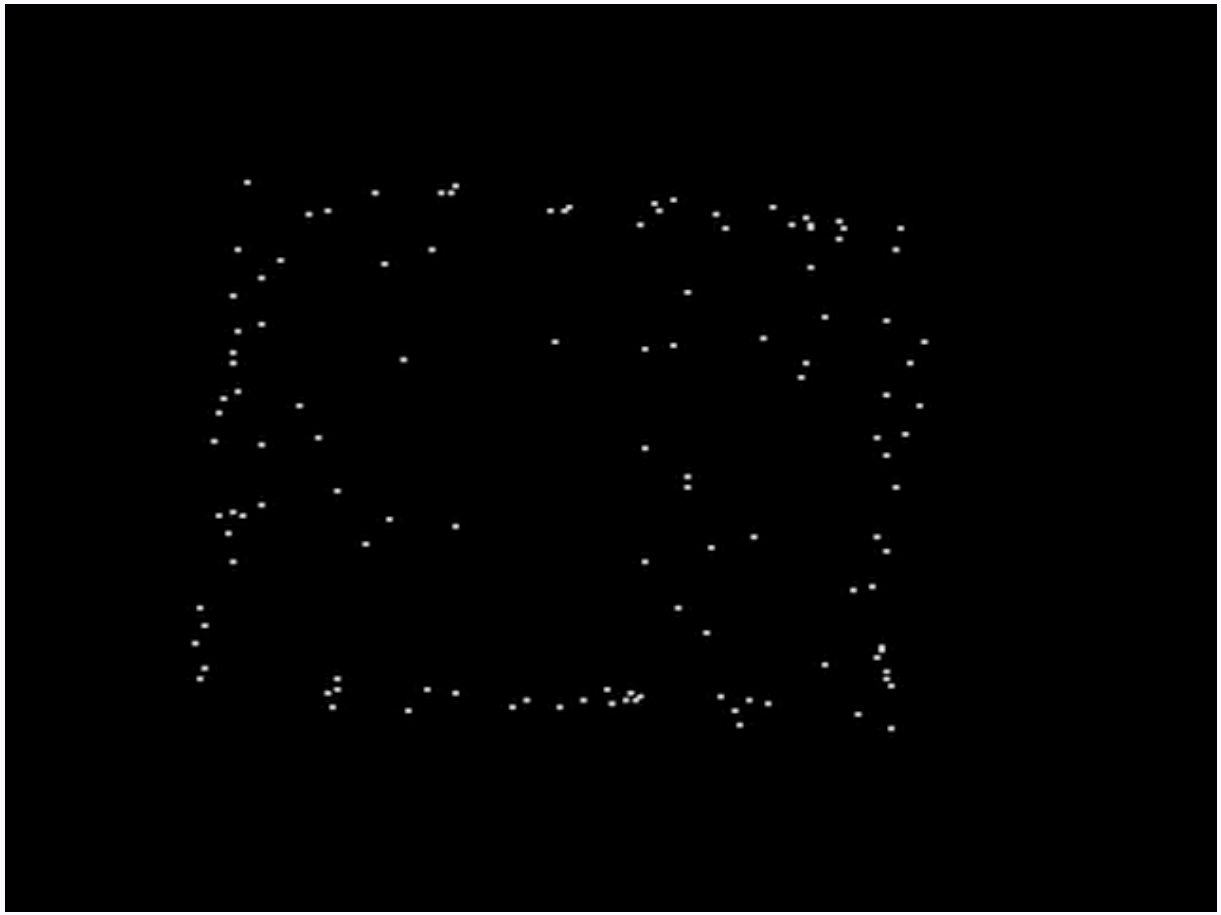
From Jones and Malik, "A computational framework for determining Stereo correspondences from a set of linear spatial filters



From Jones and Malik, “A computational framework for determining
Stereo correspondences from a set of linear spatial filters



From Jones and Malik, “A computational framework for determining
Stereo correspondences from a set of linear spatial filters



Reconstruction from multiple views

- From point matches in multiple views, we can get
  - camera positions and orientations
  - 3D point positions
  - using technology to be sketched in future lecture
- Issue:
  - dense reconstructions tend to be hard
    - because point matches are not as fine resolution as pixels
- Idea:
  - obtain point based reconstructions, then use brightness based stereo
  - multiple view stereo

Furukawa et al, Toward internet scale multi-view stereo, 2010

Image

Points

MVS surface

Dense 3D Motion Capture for Human Faces

Yasutaka Furukawa
Department of Computer Science and Engineering
University of Washington, Seattle, WA, USA

Jean Ponce
Willow Team, LIENS (CNRS/ENS/INRIA UMR 8548)
Ecole Normale Supérieure, Paris, France

Reconstructing Building Interiors from Images

Yasutaka Furukawa Brian Curless Steven M. Seitz
University of Washington, USA

Richard Szeliski
Microsoft Research, USA