Space From Motion

Joe Lappin

31 March 2004

Theme: **Motion is a principal source of visual information about spatial structure.**

1) Grouping & image segregation: grouping spatially connected features by common motion ("common fate"), & segregating those with unrelated motions. (Implicit question: What is "common" motion?)

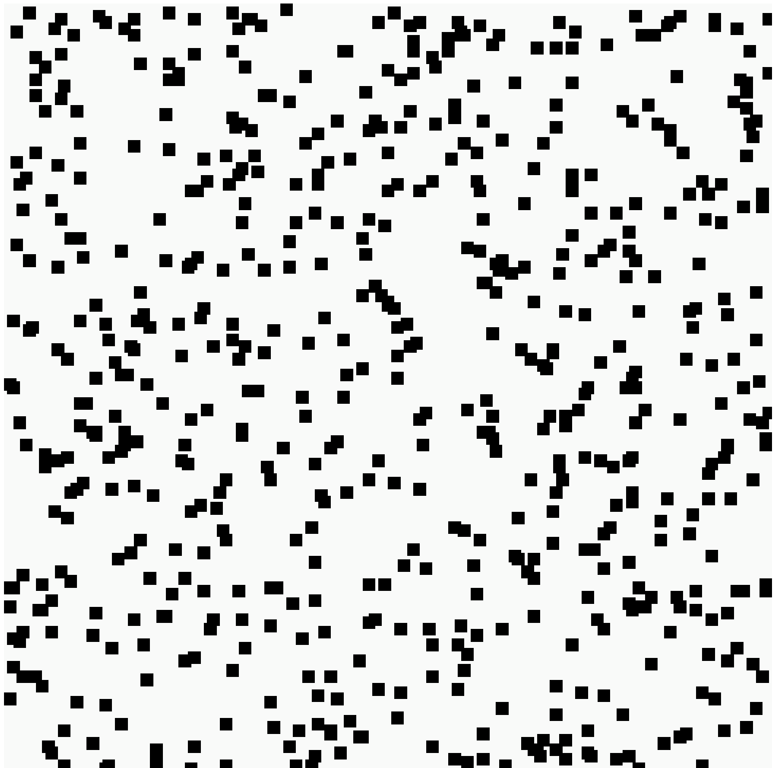
2) Motion-defined form, and figure-ground segregation:

3) 3D structure from motion — from image deformations produced by rotation in depth, views from different neighboring perspectives.

4) Optical information — from objects to images to percepts

Reading:

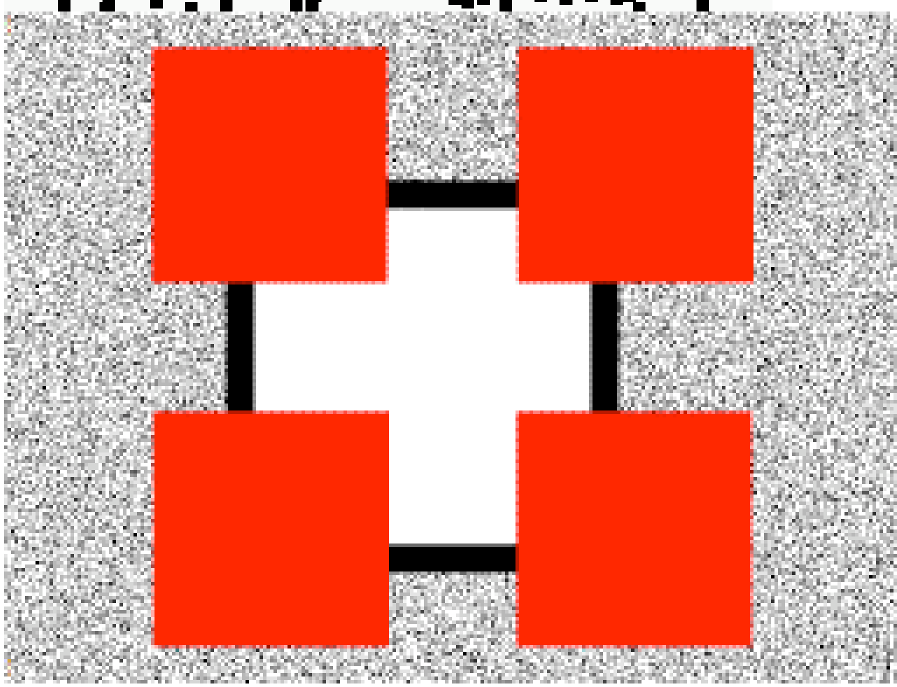
• Lappin, J.S., Doner, J.F., & Kottas, B.L. (1980). Minimal conditions for the detection of structure and motion in three dimensions. *Science*, **209**, 717-719.

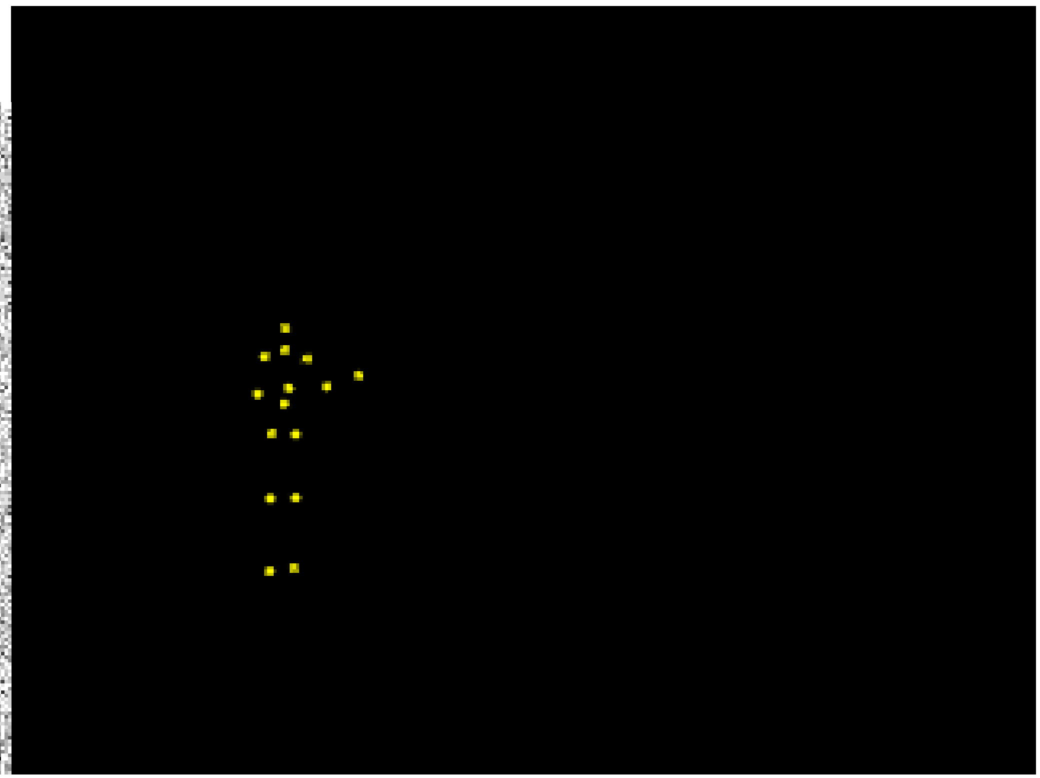


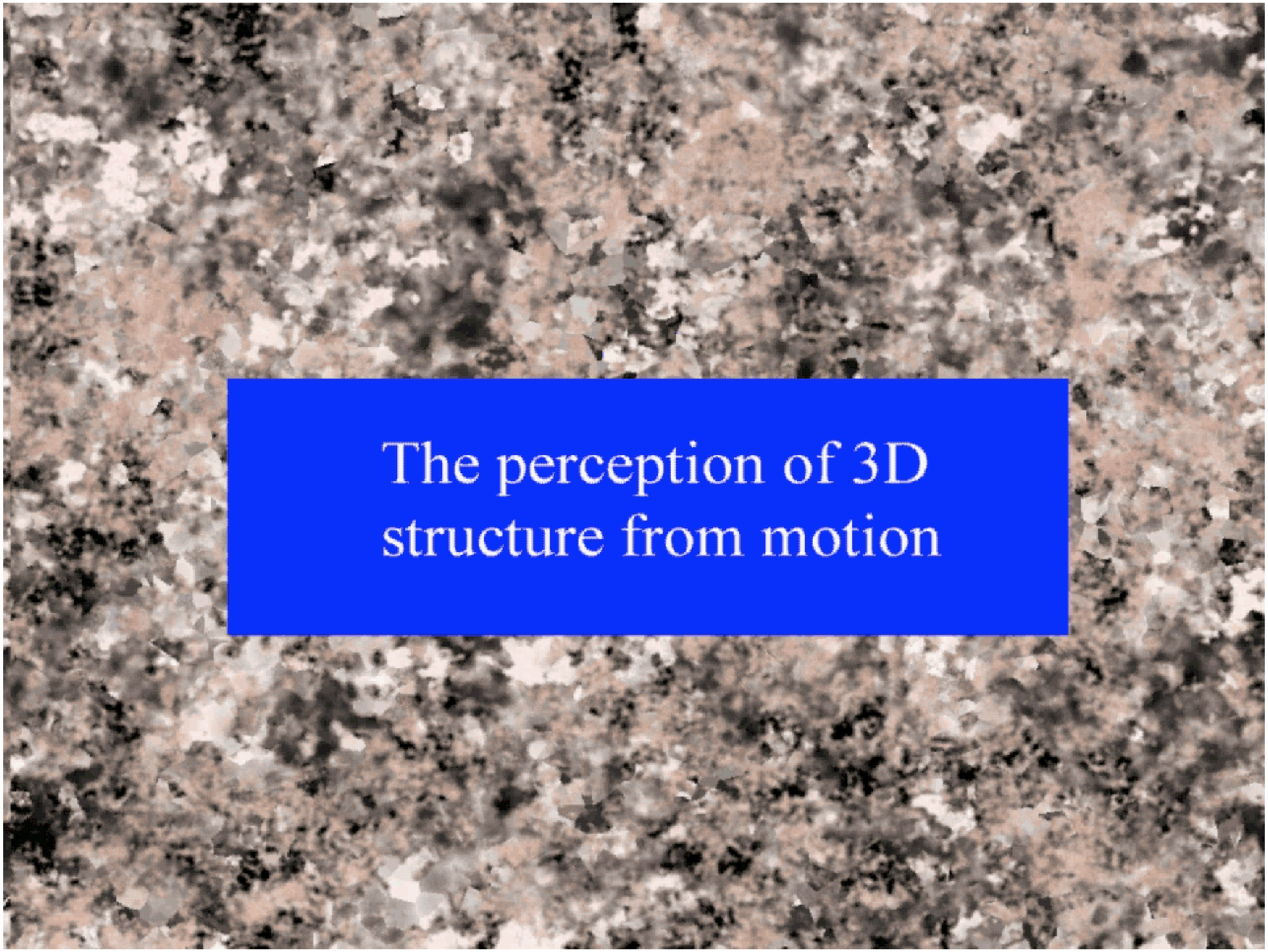
* *Upper left*: 2D form defined by motion; figure-ground segregation

* *Lower left:* Spatial connections & motion directions can depend on 3D organization.

* *Lower right:* “biological motion” (Johansson, 1973) — bending; hierarchies of spatial connections & motions; meaningful 3D objects & events



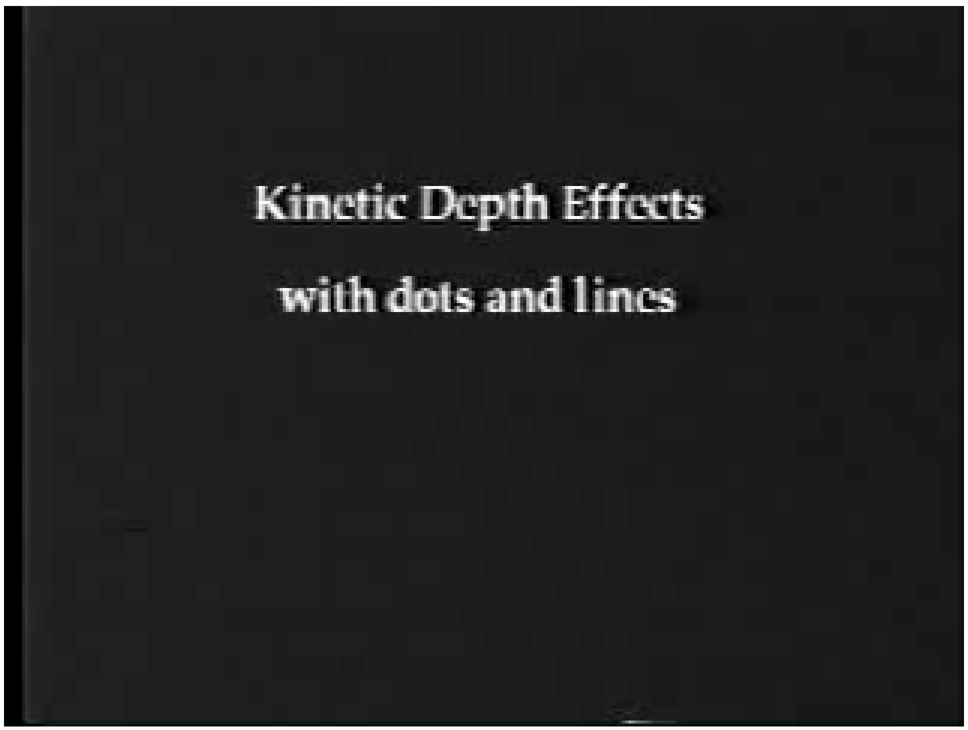




(Thanks to James Todd, Ohio State University for this illustration.)

KDE & 3D Structure from Motion

- Motion yields immediate and compelling perceptions of 3D structure and motion — even for images that have no visible depth when stationary. The illustration below is similar to one reported by Wallach & O'Connell (1953), who named this the "kinetic depth effect" (KDE). In the 1970s, 80s, and 90s, this probably was the most widely cited article on perceiving 3D structure from motion.



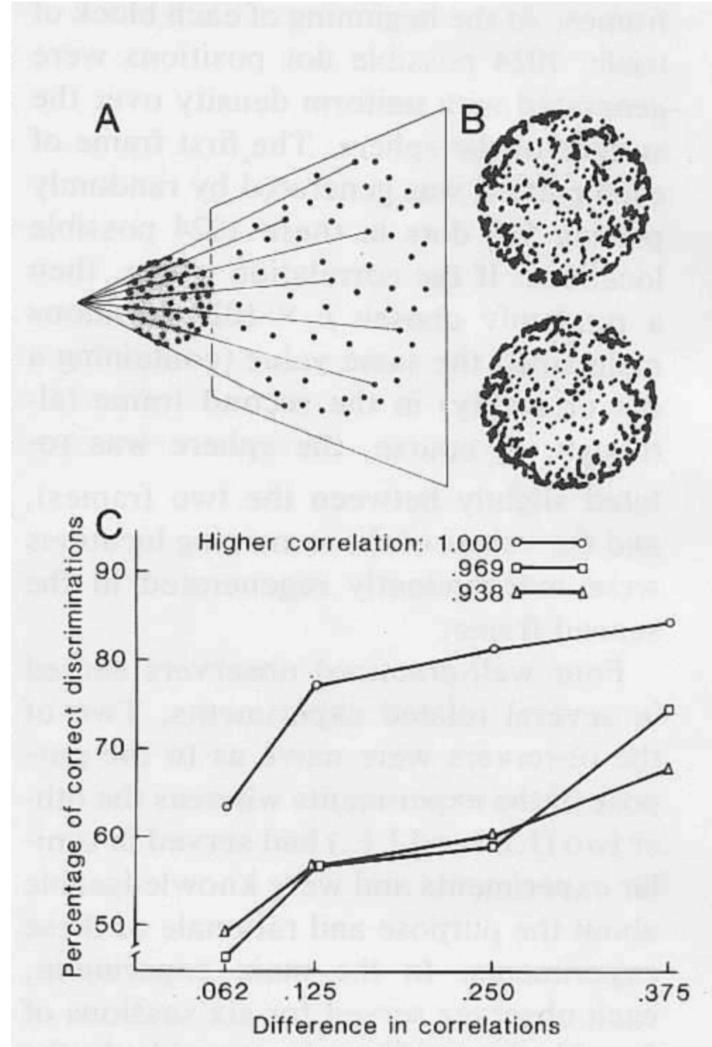
- Notice the similarity with stereoscopic vision, where a given object is seen from multiple perspectives — simultaneously with the two eyes in stereopsis, and sequentially in time in the KDE.

- **Basic issue:** What is the optical information?

(Thanks to Farley Norman, Western Kentucky Univ., for this computer-generated demonstration.

The illustration at the right is from Lappin, Doner, & Kottas (1980). (A) shows a schematic illustration of the 2D projection of a spherical pattern containing 512 dots randomly scattered over the surface of the sphere. (B) shows photographs of two images of the same pattern related by a rotation of 5.6˚ around the central vertical axis. (C) shows the discrimination accuracies for observers who discriminated between two patterns with different correlations between the dot positions in <u>just two frames</u>.

Observers were more accurate when one of the two alternative patterns was perfectly correlated. Accuracy was much lower when the higher of the two correlations was even a little less than 100%. Thus, perception of this coherent 3D structure and motion seems to involve a nonlinear stability, which seems to be fragile under these minimal conditions.



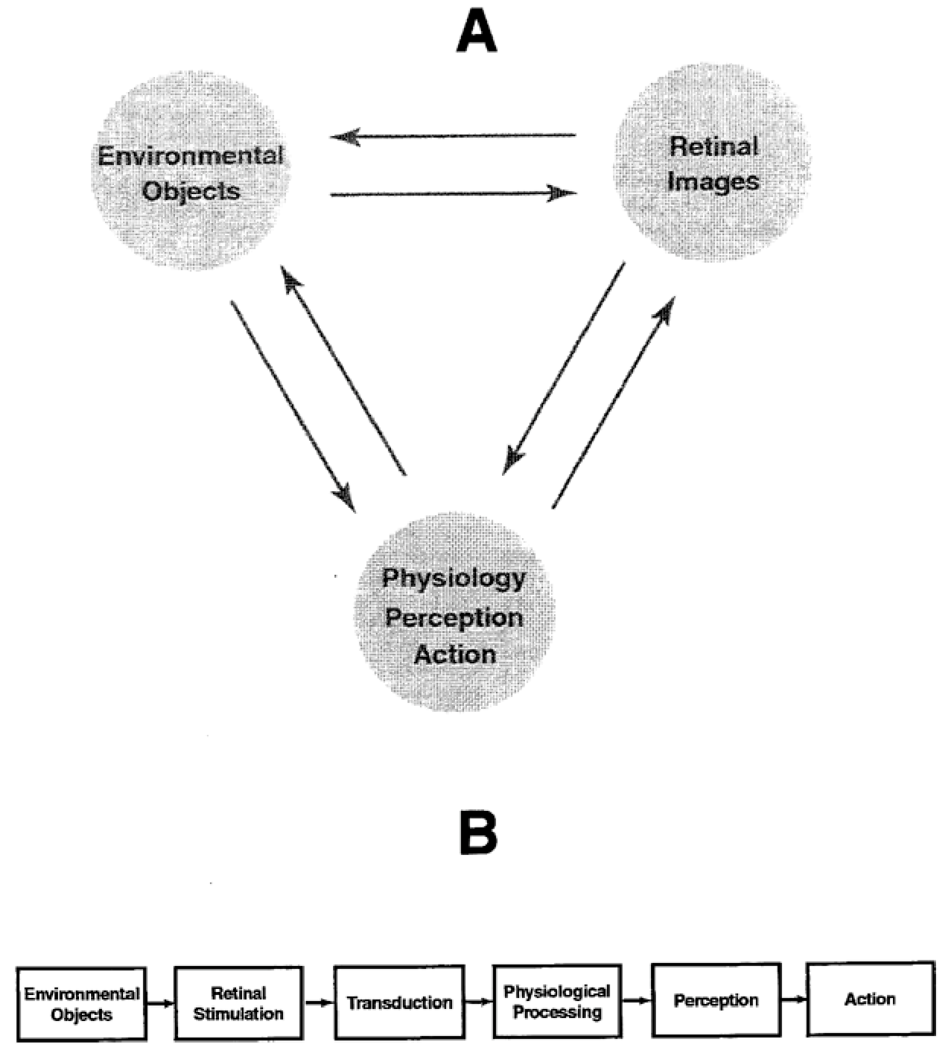
This phenomenon might suggest a visual process that roughly resembles *<u>autocorrelation</u>*,

$$A(M) = \sum\sum \{f(x, y) \cdot M[f(x, y)]\}$$

where *M* is a rotation in depth. The present 3D discriminations, however, were nonlinear.

Information ⇔ Structural Correspondence (Isomorphism)

Lappin & Craft (2000): Spatial information consists of an (approximate) isomorphism of spatial structure in two or more physical domains.



OK, so what specific image structure specifies what specific object structure?

Koenderink & van Doorn (1992), Lappin & Craft (2000):

1) Retinal images are images of surfaces.

2) Correspondences between surfaces and their images can be described by differential geometry.

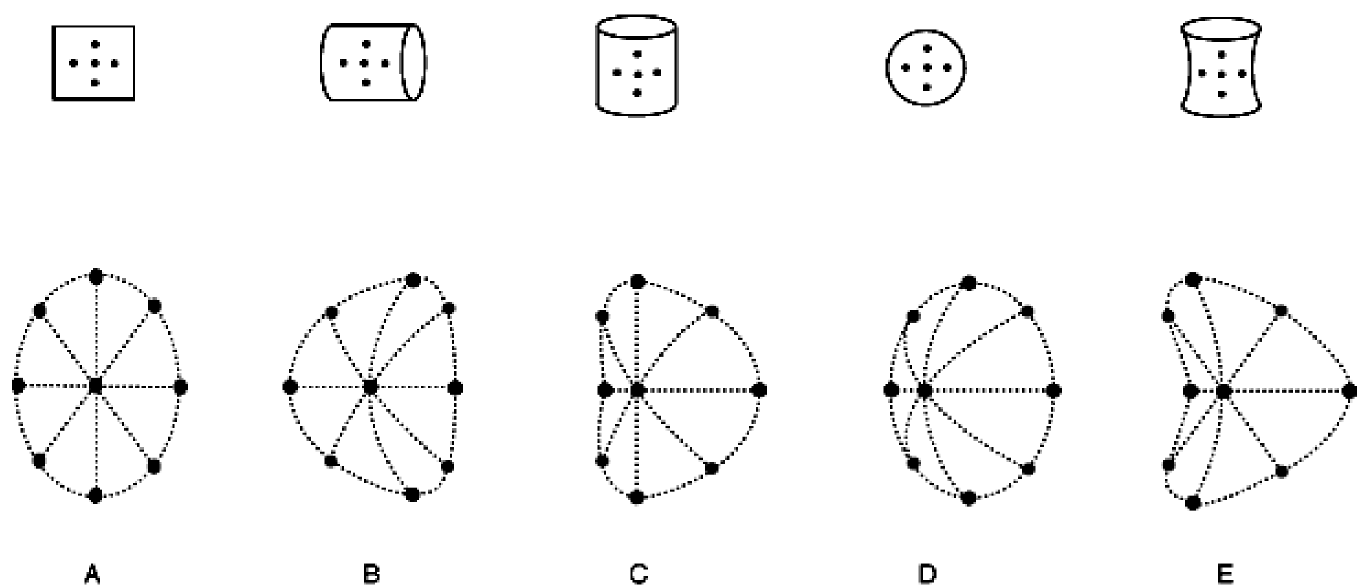
3) Information about local surface shape is given by 2-D 2nd-order differential structure of images of objects rotating in depth.

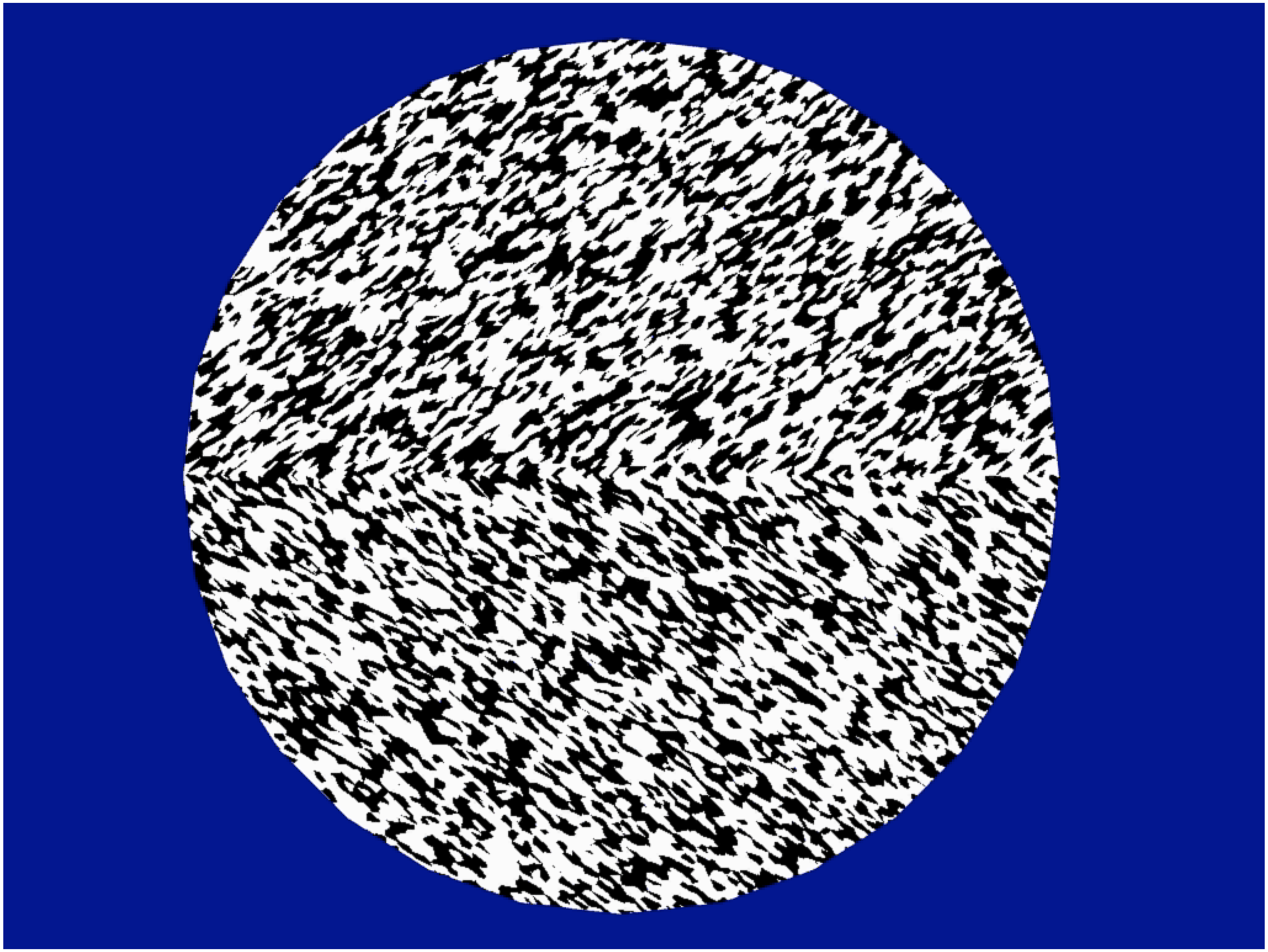
Local surface shape:

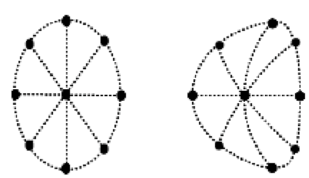
- The qualitative local surface shape at any point on a smooth surface is one of (only) 4 alternative types.

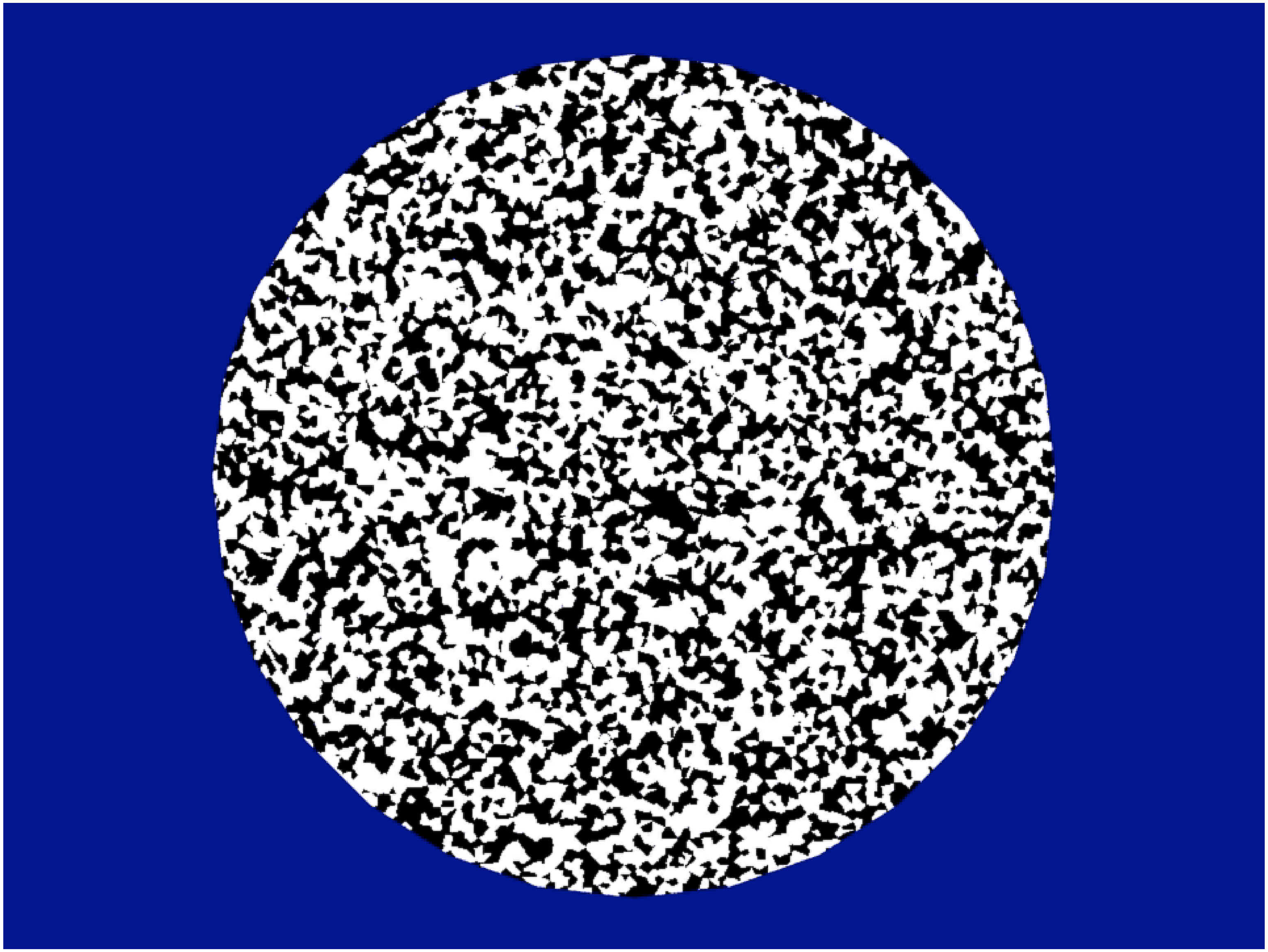
- These 4 surface shapes are specified by the 2D 2nd-order differential structures of both the surface and its image. (This isomorphism holds for images defined by motion, stereoscopic disparity, and texture.)

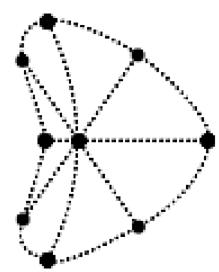
- This correspondence between the qualitative structures of local surface shapes and their images is invariant under motions in 3D.











Psychophysical evidence — from Warren Craft's Ph.D. thesis (Vanderbilt, 1997)

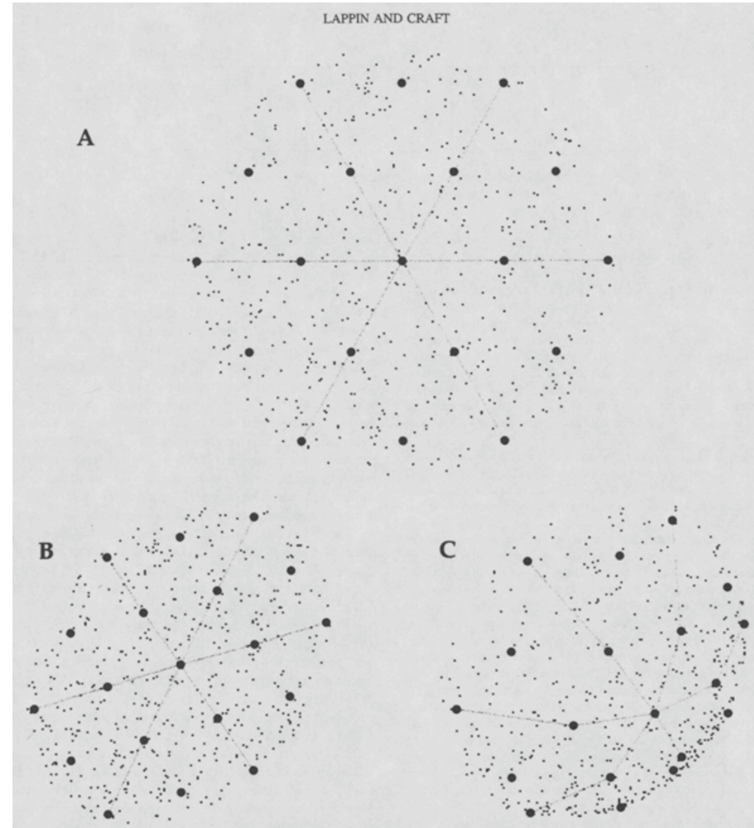
Lappin & Craft (2000) concluded:

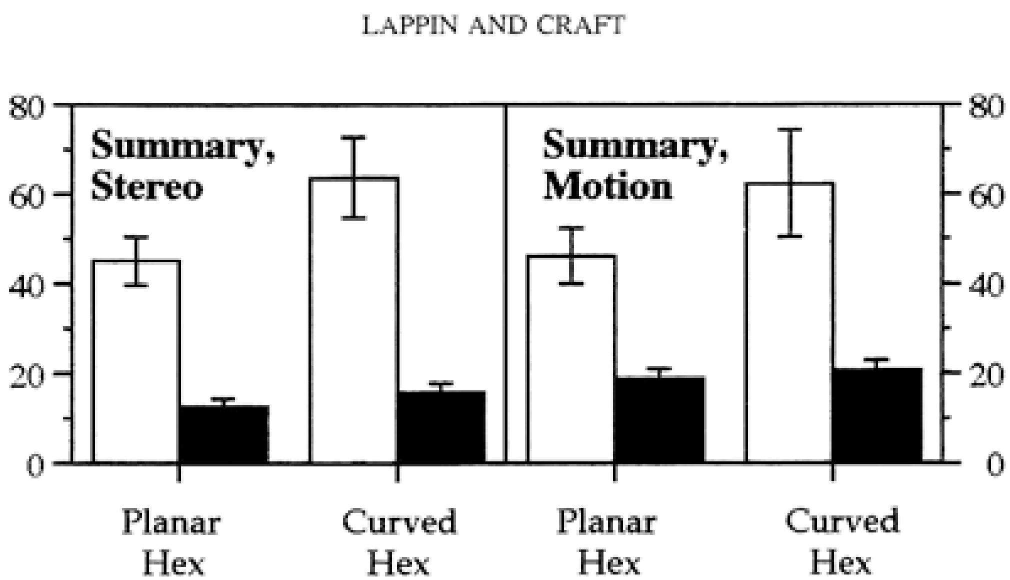
1) The 2D 2nd-order differential structure of moving and stereoscopic images constitutes a "spatial primitive" for seeing local surface shape.

2) This information must be specified at the retina.

— based on *hyperacuities* for detecting relative motion and binocular disparity, invariant under random perturbations of lower order structure.

Craft's (1997) experiments on acuities for relative position — (a) centered on the surface plane, and (b) coplanar with the surface — in patterns with 3D structure from stereo and from relative motion.





Results: Hyperacuities for detecting binocular disparity or relative depth on a sphere were as good as those for a plane, though the sphere requires 2D 2nd-order structure, but the plane requires only 1st-order structure.

Thank you for your attention!