

## Minimal Conditions for the Visual Detection of Structure and Motion in Three Dimensions

**Abstract.** Human observers detected the global three-dimensional organization of visual patterns consisting of only two successive frames of randomly positioned dots, corresponding to projections of a rotating sphere. A perfectly coherent sphere yielded a stable perceptual organization that was detected more accurately than other slightly less organized patterns.

One of the impressive achievements of the visual nervous system is the perception of stable structures moving in three-dimensional space. Even though the metric properties of the stimulus patterns projected onto the two-dimensional surface of photoreceptors are continually changed by movements of the observer and movements of environmental objects, the perceived structure of three-dimensional space remains invariant under these projective transformations.

Psychophysical research on the "kinetic depth effect" (1, 2) has suggested that perceptual stability of three-dimensional space arises directly from the optical projections of dynamic patterns without requiring the intervention of cognitive processes (3). A small number of lines or spots that form no recognizable pattern in a stationary two-dimensional projection yield a compelling impression of a solid object when the projection is moved as if rotating in three-dimensional space. Most of this evidence, however, has relied upon subjective judgments; objective and quantitative determinations of observers' sensitivities to structural details have been lacking. Additionally, most experiments have used patterns of relatively small numbers of elements, usually displayed in motion over several seconds (4). As a result, the dependence of this perceptual sensitivity on the complexity of the pattern and the duration of observation remains unclear.

Our experiments provide objective evidence for the perception of global three-dimensional structure and motion in stimulus patterns consisting of only two successive frames of several hundred randomly positioned dots. We used a temporal variant of Julesz's "cyclopean" method (5), in which the geometric information for this perception was provided only by the space-time

relationship between the two frames of dots rather than by either of the individual frames. Apparently, the organization of such patterns was detected by a visual process that determined the three-dimensional structure and motion capable of maximizing the correspondence between successive frames. Furthermore, this self-organizing process was nonlinearly stable.

The stimulus patterns consisted of 512 dots randomly distributed with uniform probability density over the surface of a transparent sphere (Fig. 1A) and displayed by polar projection (6) on a computer-controlled cathode-ray tube (Tektronix 604, with fast-decay P-15 phosphor). The second frame of the pattern followed with no interstimulus interval (7) and contained the same dots displaced as if the sphere were rotated 5.6° around a vertical axis through its center. We manipulated the spherical organization of this pattern by varying the proportion of dots that maintained the same spherical positions in the two successive frames.

The specific question in these experiments was whether observers could discriminate differences in spherical organization. We had previously found that most observers—about two-thirds of the 80 or so persons who have seen these patterns—spontaneously reported an impression of a sphere rotating in depth. Now we sought more rigorous objective evidence that the global three-dimensional organization was indeed detected. Could observers discriminate patterns that differed with respect to the correlations of dot positions in the two successive frames?

The result was that a completely coherent spherical pattern, in which the three-dimensional dot positions were perfectly correlated in the two frames, could be reliably discriminated from oth-

er less correlated patterns. Discriminations between patterns of lower degrees of correlation, however, were not very accurate. Thus, the perfectly correlated spherical pattern yielded a stable perceptual organization that collapsed under slight perturbations.

A two-alternative forced-choice task was used to measure the discriminability of two different degrees of correlation. In a given block of trials, the correlation of the dot positions in the two frames of one pattern was set at some constant value—1.0, .969, or .938—while the correlation of the other pattern was reduced by

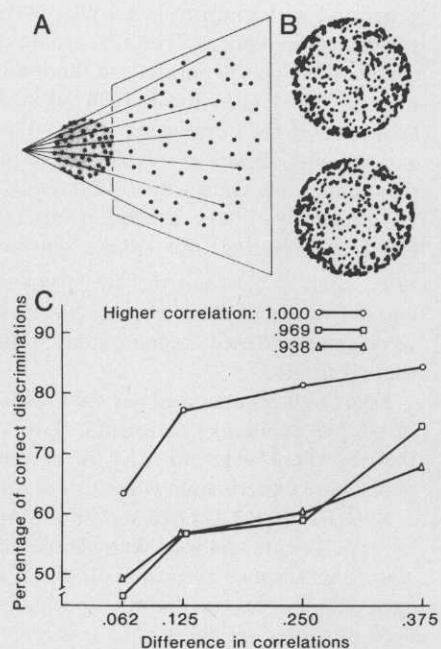


Fig. 1. (A) Schematic illustration of the two-dimensional projection of a spherical pattern. The patterns were binocularly viewed in a dimly lit room from a distance of 85 cm, subtending 3.15° of visual angle. (B) Photographic reproductions of two successive frames. (The original displays were more sharply focused than in these reproductions, with each spot having a uniform size and brightness.) The positions of dots in the second frame have been displaced as if the sphere were rotated 5.6° around a vertical axis. Each frame was displayed for 200 msec, with each point refreshed every 10 msec, and no interval between the two successive frames. (C) Average accuracy (for four observers) in discriminating between patterns with different correlations in a two-alternative forced-choice task. The two patterns differed with respect to the correlation of dot positions in the two successive frames that composed each pattern. The parameter distinguishing the separate curves is the value of the higher of the two correlations.

some constant amount—.062, .125, .250, or .375—in separate blocks of trials (8). Patterns of both degrees of correlation were presented in separate temporal intervals (separated by 0.5 second) on each trial, and the observer's task was to identify which of the two alternatives was more highly correlated. The correlations were determined by varying the proportions of dots that maintained corresponding positions on the spherical surface in the two successive frames, with the positions of the remaining dots independently generated in the two frames. At the beginning of each block of trials, 1024 possible dot positions were generated with uniform density over the surface of the sphere. The first frame of each pattern was generated by randomly placing 512 dots in these 1024 possible locations. If the correlation was  $p$ , then a randomly chosen  $p \times 1024$  locations maintained the same value (containing a dot or empty) in the second frame (although, of course, the sphere was rotated slightly between the two frames), and the values of the remaining locations were independently regenerated in the second frame.

Four well-practiced observers served in several related experiments. Two of the observers were naïve as to the purpose of the experiments whereas the other two (J.D. and J.L.) had served in similar experiments and were knowledgeable about the purpose and rationale of these experiments. In the main experiment, each observer served for six sessions of four blocks of 100 trials—two blocks for each of the 12 conditions.

The principal result (Fig. 1C) was that discrimination accuracy was consistently higher when one of the two alternative patterns was a perfectly correlated sphere than when it was reduced to only .969 or .938. (When the two replications for each subject in each condition were averaged, the perfectly correlated pattern was more accurately detected than either the .969 or .938 pattern in all 32 comparisons.) There was no reliable difference between the detections of the patterns with .969 or .938 correlation. The data in Fig. 1C have been averaged across the four observers, since all four produced essentially the same pattern of results.

We were surprised by the magnitude of the effect of a small reduction in the coherence of any part of the sphere and by the small effects of further reductions. To further examine this phenomenon, we collected additional data for conditions in which the higher of the two alternative correlations was either .75 or .50. The correlation of the other alternative was reduced by either .25, .375, or .50 in separate blocks of trials. The result was no consistent difference in detectability of patterns with correlations of .75 versus .50 and little or no consistent difference in the detectability of these and the other partially correlated spherical patterns. The superior detectability of the perfectly correlated sphere indicated a nonlinear stability in the global perceptual organization (9).

The stability of this global organization is inconsistent with potential computational procedures that are either linear or local. Autocorrelation (defined on the group of projective transformations), for example, is merely a summation of measures defined on relations between pairs of points. Ullman (2) has recently proposed a more powerful computational procedure for determining three-dimensional structure from perspective transformations, but the output of this procedure is also based on a summation of results from strictly local mappings between neighboring points in successive frames.

As a control experiment to verify that this superior discriminability of the coherent spherical pattern was associated with its three-dimensionality, we also examined performance under an analogous set of conditions with two unconnected but superimposed rectilinear plane patterns. The pattern of results for these planar patterns was very different, with no superiority in the detectability of the perfectly correlated pattern, no effect from small reductions in the correlation, and a competitive interference rather than global organization between two planes displaced in opposite directions.

Thus, a single discrete projective transformation provides sufficient information for the detection of structure and motion in three dimensions. The underlying visual process is self-organizing, yielding a nonlinear stability sensitive to

the global coherence of the changing optical pattern.

JOSEPH S. LAPPIN  
JONATHAN F. DONER

Department of Psychology,  
Vanderbilt University,  
Nashville, Tennessee 37240

BRIAN L. KOTTAS  
Army Research Institute,  
Fort Knox, Kentucky 40121

#### References and Notes

1. H. Wallach and D. N. O'Connell *J. Exp. Psychol.* **45**, 205 (1953); M. L. Braunstein, *Depth Perception Through Motion* (Academic Press, New York, 1976); G. Johansson, *Sci. Am.* **232**, 76 (June 1975).
2. S. Ullman, *The Interpretation of Visual Motion* (MIT Press, Cambridge, Mass. 1979).
3. J. J. Gibson, *The Senses Considered As Perceptual Systems* (Houghton Mifflin, Boston, 1966).
4. G. Johansson, *Psychol. Res.* **38**, 379 (1976); J. E. Cutting and L. T. Kozlowski, *Bull. Psychonom. Soc.* **9**, 353 (1977).
5. B. Julesz, *Foundations of Cyclopean Perception* (Univ. of Chicago Press, Chicago, 1971). Our patterns, however, provided no binocular disparity. Instead, parallax was defined by the displacement between successive views to the same eye.
6. The perspective of the projective mapping correspond to positioning the convergence point one sphere-diameter away from the center of the sphere. This corresponds to a 3:1 ratio of the distance of the farthest to nearest points, and a 3:1 ratio in the extent of projected displacements of points on the nearest front and farthest rear surfaces when the sphere was rotated. While this amount of perspective was greater than would have corresponded to the viewing distance between display and observer, we have found that such perspective is necessary for detecting the three-dimensional structure in these displays.
7. Detectability of the organized structure of these patterns deteriorates with interstimulus intervals as brief as 50 msec.
8. The correlations of dots on the front and rear surfaces of the sphere were separately manipulated. The data have been averaged over these two conditions because there was no consistent difference between them. This was confirmed in a supplementary experiment specifically devoted to the comparison of these two conditions. The similarity of performance in the two conditions is remarkable because the projected spatial density and extent of displacement under motion was as much as three times greater on the front surface. Moreover, the observers reported not even seeing the motion of dots on the front surface. Perturbations of the dot positions were equally disruptive on either surface, however, indicating that the global coherence of the entire sphere was detected.
9. The special stability of the fully correlated pattern contrasts slightly with results recently reported by J. T. Petersik [*Percept. Psychophys.* **25**, 328 (1979)], who found that subjective judgments of the coherence and depth of a continually rotating sphere could be maintained even under large amounts of visual noise.
10. Supported by NSF research grant BNS 78-05857. We thank C. S. Harris for helpful discussions of these experiments and R. Fox for valuable comments on a preliminary draft of this report.