Journal of Experimental Psychology:
Human Perception and Performance
1990, Vol. 16, No. 4, 693–705

0096-1523/90/$00.75

Image Statistics and the Perception of Apparent Motion

David L. Gilden
Vanderbilt University

Bennett I. Bertenthal
University of Virginia

Sinan Othman
Department of Electrical Engineering
University of Virginia

The short- and long-range apparent motion processes are discussed in terms of the statistical properties of images. It is argued that the short-range process, exemplified by the random-dot kinematogram, is primarily sensitive to the dipole statistics, whereas the long-range process, exemplified by illusory occlusion, is treated by the visual system primarily in terms of the tripole and higher statistical correlation functions. The studies incorporate the balanced dot, which is a unique stimulus element that permits high pass filtering while preserving detailed positional information. Low spatial frequencies are shown to be critical for texture segregation in random-dot kinematograms, independent of the grain size or number density of texture elements. Illusory path perception in the long-range process is shown not to require low spatial frequencies, but is sensitive rather to global temporal phase coherency. These results are interpreted in terms of the respective roles of the power and phase spectra in perceptual organization. The construction of balanced dots is discussed in detail.

The notion that visual processing commences with the identification of primitives and the construction of a primal sketch (Marr, 1982) has led to considerable effort into determining the nature of these primitives. Spatial frequency is a natural candidate in this regard. Electrophysiological and psychophysical studies (De Valois & De Valois, 1980; Graham, 1981; Shapley & Lennie, 1985; Wilson & Bergen, 1979) reveal that a crude Fourier analysis occurs in early vision, and that low and high spatial frequencies are differentiated as channels throughout the visual system.

The observation that the visual system performs a crude Fourier analysis does not in itself provide an obvious set of perceptual primitives. The Fourier transform of an image is isomorphic to the image; there is exactly as much information and complexity in the transform as there is in the image. Identification of primitives must follow a partitioning of the information in the transformed image into those elements that can be demonstrated to be essential for perception. The Fourier transform has the convenient mathematical property of breaking into two parts, the power and phase spectra, and this natural splitting has set the foundation for inquiries into what primitives might be discovered in the transform domain. Historically, the search for perceptual primitives has been confined to isolating features within the power spectrum. The feature that has received the greatest attention is the power at the low spatial frequencies.

In this article we are concerned with the role that spatial frequency analysis plays in perceptual organization. We make four key distinctions in this analysis. The first is between space and time as two separate domains where organization can occur. We study apparent motion because it is in this regime where organization is simultaneous in both domains. The second is between the power and phase spectra as components of the Fourier transform through which spatial organization can be mediated. We attempt to clarify the range of percepts for which the power spectrum and, in particular, low spatial frequencies are relevant. The third distinction is between the short- and long-range apparent motion processes. Our empirical work addresses the relation between the two components of the Fourier transform and the type of organization underlying the two motion processes. The final distinction, which is elaborated throughout these studies, is between random textures and images that contain figural information of some sort. This distinction was articulated by Anstis (1980), who used it to distinguish between the long- and short-range apparent motion processes; in the long-range process figural organization precedes the perception of motion, whereas the reverse is true in the motion of random textures in random-dot kinematograms. We begin with a brief theoretical discussion of the image information contained in the power and phase spectra. This is followed by a review of the experimental literature that relates to the spectral analysis of static and moving images.

The Image Information in the Power and Phase Spectra

The information in an image can be summarized by the hierarchy of its *n*-point correlation functions. These functions are straightforward generalizations of the autocorrelation

This research was supported by a National Research Service Award Postdoctoral Fellowship HD-07036 from the National Institute for Child and Human Development (NICHD) to David L. Gilden and by NICHD Grant HD-16195 and NICHD Career Development Award HD-00678 to Bennett I. Bertenthal.

The displays used in the apparent motion studies were programmed by Tom Breedon. Lenses and frames were provided by Arnold B. Popkin. Portions of this article were previously presented at the May 1988 meeting of the Association for Research in Vision and Ophthalmology held at Sarasota, Florida.

Correspondence concerning this article should be addressed to David L. Gilden, Department of Psychology, Vanderbilt University, Nashville, Tennessee 37240.

function, although they are more difficult to evaluate and do not always result in a simple interpretation. An exceptionally clear exposition of this hierarchy is due to Julesz (Julesz, 1975; Julesz, 1981; Julesz, Gilbert, Shepp, & Frisch, 1973), who expressed the n-point correlation functions in terms of randomly thrown n-gons. Julesz introduced the notion of n-gon statistics into the psychological literature through the analysis of random-dot patterns. Although these ideas generalize in a completely straightforward manner to any image, random-dot patterns provide an intuitive vehicle for evaluating the information captured in the power and phase spectra.

The n-gon statistics of a random-dot image are developed by considering the ensemble of all n-sided polygons. The collection of 2-gons is the set of all line segments, the collection of 3-gons is the set of all triangles, and so on. Consider for a given k one member of the set of k-gons. Associated with this object is the probability that a random toss of the k-gon onto the image will have all k vertices simultaneously connecting random dots. The collection of k-gons creates a probability distribution; the k-gon image statistics.

The first few moments of the n-gon hierarchy are related to familiar distributions. The l-gon statistic is the probability that a randomly thrown point will land on any of the dots. This probability is proportional to the total number of dots in the collection or to the averaged pattern contrast. The 2-gon statistics give the probability that a randomly thrown dipole will have both ends land on dots. This probability is proportional to the autocorrelation function of the pattern. The various lags are just the different size dipoles that generate the probability distribution. The 3-gon statistics are less familiar and considerably more complicated, but have been useful in analyses of clustering (see, e.g., Peebles, 1980).

The complexity of the distribution functions of randomly thrown n-gons increases rapidly with n. Randomly thrown dipoles have 2 degrees of freedom: dipole length and orientation. Randomly thrown triangles have 4 degrees of freedom: any of the three combinations of side lengths and angular measure used to prove triangle congruency and overall orientation. In general, the distribution function of n-gons has dimension $2n - 2$. For $n > 3$, the distribution functions become increasingly intractable and are not useful for mathematical analysis. The structural information in the power and phase spectra can be evaluated by noting that the dipole, or 2-gon, statistics establish the bottom of the structure hierarchy; there is no structural information in the point, or 1-gon, statistics. The dipole statistics and the power spectrum are related through the Wiener–Khinchin theorem, which states that the autocorrelation function and power spectrum are Fourier transform pairs. Thus, the power spectrum contains exactly that subset of image information in the dipole statistics, and the phase spectrum contains all of the structural information in the n-gon hierarchy for $n > 2$. The structural information is contained primarily in the phase spectrum.

Visual Processing of Static Images

Figures

Perceptual achievements that require some sort of figural organization are exceedingly robust to transformations on the power spectrum. Piotrowski and Campbell (1982) illustrated the irrelevance of the power spectrum in picture perception by combining the power spectrum of one picture with the phase spectrum of another. Invariably, the combined picture is immediately recognized as the one from which the phase spectrum was drawn. In a more rigorous psychophysical testing of the same issue, Palmer, Kube, and Kurschke (1987) showed that the perception of pointing triangles is not affected by reducing the entire power spectrum to a single value nor by replacing the power spectrum by random numbers. Kleiner and Banks (1987) extended these ideas by showing that infants over the age of 2 months ignore the power spectrum in favor of the phase spectrum in the recognition of faces. Although these latter studies may have methodological problems (Badcock, 1989), these observations make it unlikely that perceptual organization (Gestalt) is related to the power spectrum, although there are claims to the contrary (Ginsburg, 1986).

Illusions

Visual illusions are one domain for which a specific region of the power spectrum, the low spatial frequencies, has been claimed to significantly contribute to perception. Several classic illusions (Muller–Lyer, Ponzo, and Poggendorf) involve configurations of lines where length, curvature, or displacement is not veridically perceived. It has been suggested that low spatial frequencies in the cortical Fourier-transformed image are responsible for these illusions (Ginsburg, 1975; Ginsburg & Evans, 1979). Carlson, Moeller, and Anderson (1984) provided a critical test of these claims by constructing the illusions from arrays of dots in which the *DC* point (the power at zero spatial frequency) was removed. This dot is referred to as a balanced dot (see Appendix for a detailed discussion of the power spectrum of a balanced dot). Estimates of the perceived length ratio for the Muller–Lyer illusion showed that there was at most a 4% reduction in the perceived length ratio in figures composed of balanced dots compared with figures composed of normal, unfiltered dots. Carrasco, Figueroa, and Willen (1986), using an adaptation procedure, showed that the perceived length ratio dropped by about 5% after adaptation to a low frequency grating (2 cycles/degree), whereas there was no decrement following adaptation to a high-frequency grating (11 cycles/degree). It is, thus, evident that low spatial frequencies play only a minor role in determining the perceived length of the Muller–Lyer lines. Other line-drawn illusions are not expected to show anything different. Casual inspection of the Poggendorf illusion (Carlson et al., 1984) reveals that the illusion is still present following high-pass filtration.

Another class of illusions is produced by arrangements of elements that yield illusory contours. Ginsburg (1975) claimed that illusory contours, in particular those in the Kanizsa triangle, are physically present in the low spatial frequency components of the inducing elements. Tyler (1977) criticized this claim on several grounds, but most compellingly by simply noting that the illusory contour is still present in a Kanizsa triangle that has been high-pass filtered. Further analysis of the filtration technique used by Ginsburg (1975), conducted by Becker and Knopp (1978), showed that high frequencies had been reintroduced by digital-processing artifacts. They made the further plausible observation that illusory contours are well defined and quite sharp, and therefore could not possibly be produced by a low-pass filter.

It is evident that neither the recognition of scenes nor illusions containing figural information are organized by information that exists in the power spectrum. There may be illusions involving textures that are sensitive to the presence of low spatial frequencies, but the appropriate tests have not been conducted. In evaluating these results, it must be remembered that the phase spectrum in fact contains virtually all of the image information, and apparently virtually all of the information required for spatial perceptual organization.

Random Textures

The role of the power spectrum in random textures has been studied by Julesz and collaborators (Julesz, 1975; Julesz, 1981; Julesz et al., 1973) in terms of texture segregation in random-dot patterns. Julesz showed that, in general, the visual system cannot immediately (preattentively) distinguish between textures that are equated on their point and dipole statistics, but that differ in their higher order statistics. The counterexamples to this general rule, the textons, are limited in number, and are not an issue in completely random patterns. There is immediate segregation, however, if the patterns differ in their dipole or point statistics. Another way of stating this result is that the visual system cannot immediately compute the phase spectrum for random-dot patterns, but the computation of the power spectrum is quite rapid.

Thus, it is apparent that the role of the phase and power spectrum are quite different in the perception of figures and random textures. Figural organization is mediated primarily through information in the phase spectrum—that part of the *n*-point hierarchy commencing with the 3-point statistics. However, texture segregation of random textures is primarily limited to the information in the power spectrum—the 1- and 2-point statistics. In the following section, we see that this distinction may be preserved in apparent motion.

Visual Processing of Apparent Motion

The Long-Range Process: Figural Motion

A common tool used to study perceptual primitives is the competition paradigm in long-range apparent motion. In this paradigm, the visual system must resolve one or more ambiguities in the apparent motion path. Typically, there is one element in the first frame and two elements in the second. The dominating path yields a set of stimulus attributes that are matched under an apparent motion transformation. These attributes are interpreted as constituting primitive features for figural identity. An exemplary experiment using the competition paradigm was conducted by Ramachadran, Ginsburg, and Anstis (1983) in an attempt to elucidate the role of low spatial frequencies as a matching attribute. They showed that filled squares moved to filled circles rather than to outline squares, with the interpretation that the match was made on the basis of the low spatial frequency content of the filled figures. Prazdny (1986) brought this interpretation into question by showing that identical results could be obtained even when there is no luminance domain information distinguishing the competing alternatives. In these experiments, the competing alternatives were presented as segregating textures in random-dot kinematograms. Matched attributes in these stimuli must derive from the perceived form of the segregating surface. These forms only exist in a perceptual system that can correlate over time, and any statistical analysis that the visual system brings to such forms is subsequent to an analysis of the light distribution in individual frames.

There are in fact many properties that filled figures share that are not shared by outline figures. The most obvious shared figural property is that filled figures do not have holes in them. Having or not having a hole is a distinguishing topological property that cannot be represented merely in the power spectrum; all levels of the *n*-point hierarchy change when a hole is introduced into a filled figure. Chen (1985) argued that it is not the matching in spatial frequency that is prescribing apparent motion paths, but rather that the matched attributes are fundamental topological properties of the stimuli (although see Rubin & Kanwisher, 1985, for dissenting comments on these studies).

Green (1986a, 1986b), using Gaussian modulated gratings (Gabors) to represent stimuli with varying spatial frequency, did find a bias for matching to occur between elements of similar spatial frequency. These latter experiments are not subject to a topological reinterpretation because all Gabors are topologically equivalent. Thus, it does appear as if spatial frequency is used as a primitive in solving the correspondence problem, as long as no other information is supplied in the stimulus. It is important to emphasize that Green's (1986a, 1986b) studies do not satisfactorily address the necessity of spatial frequency matching because, as we have already discussed, apparent motion paths are observed between figures that are quite different in detail (Chen, 1985).

The Short-Range Process: Random Texture Motion

The short-range process has been studied intensively in the context of random-dot kinematograms, and it was within this experimental paradigm that Braddick (1974) suggested it as an independent process. A random-dot kinematogram is a two-frame version of a random-dot stereogram, where the two frames are displayed in alternation. Correlated element motion within a delimited region of the array produces the percept of a randomly textured shape moving over a background composed of randomly moving elements. The short-range process is so named because the displacement of the coherent region is a critical variable and was thought to be limited to about 15 arc min (Braddick, 1974, 1980), although it was later shown that there is no absolute displacement limit (Chang & Julesz, 1983; Lappin & Bell, 1976). Subsequent studies showed spatial frequency to be a critical variable; the maximum displacement tolerable for the perception of a segregating surface ($d_{\max}$) decreases upon high-pass filtering and increases upon low-pass filtering (Chang & Julesz, 1983).

The studies outlined in the previous sections suggest that the reliance on the phase spectrum in the perception of static figures is preserved in long-range apparent motion; the perception of long-range apparent motion is one of a figure undergoing a transformation. The short-range process, however, appears to be related to the perception of static random texture and so is limited by the visual system's inability to calculate beyond the second-order statistics. The figural information that arises in texture segregation is subsequent to a process that is essentially blind to the statistics that comprise the phase spectrum.

Overview of the Studies

In this article we provide critical tests for the conjecture that low spatial frequencies are relevant for global field organization in the short- and long-range processes. These tests introduce the balanced dot into apparent motion displays. The logic that we use is similar to that of Carlson et al. (1984); if low spatial frequencies are critical for global field organization, then this organization should be disrupted when the stimuli are composed of balanced dots. Our work is divided into three sets of experiments. In the first (Experiment 1), we construct balanced dots and provide a psychophysical test that verifies that the dots have attenuated power at low spatial frequencies. In the second (Experiment 2), we reexamine the role of low spatial frequencies in the short-range process by using balanced dots. In the third set (Experiments 3 and 4), we study the relation between temporal asynchrony and spatial frequency in the perception of illusory occlusion and in the determination of $d_{\max}$ in random-dot kinematograms.

Experiment 1: Construction of Balanced Dots

We desired to create a stimulus that psychophysically does not blur (perceptually does not get fuzzy; i.e., has a minimum of power at low spatial frequencies). The balanced dot, introduced by Carlson et al. (1984), is an example of such a stimulus. In constructing balanced dots, it is necessary to distinguish the mathematical conditions for balancing in the luminance domain from the psychophysical objective, which was to create a nonblurring stimulus. This distinction arises because of compressive nonlinearities in the visual system at superthreshold luminance (Hood & Finkelstein, 1986). In the Appendix, we review the mathematical development for balancing in the luminance domain. Here, we describe the procedures followed in psychophysically validating our technique of construction.

The principal difference between an unfiltered dot (simply the result of turning on a pixel) and a balanced dot is that unfiltered dots can be blurred. Balanced dots cannot blur because they have no overall contrast; their *DC* point has been canceled. Balanced dots disappear instead of blurring when viewed through a lens of sufficiently large diopter or when viewed from a distance. Similarly, a pattern composed of spatially contiguous unfiltered dots may still be visible even when the individual dots are not resolved. However, a pattern composed of balanced dots will disappear before becoming fuzzy because it cannot blur at the dot level.

In this experiment, subjects attempted to identify geometric patterns (hereinafter called forms) composed either of unfiltered dots or balanced dots. The forms were viewed through positive diopter lenses. Each form was displayed in a range of sizes. Confirmation that the low spatial frequencies had been removed from the balanced dots would be evidenced by (a) an inability to recognize forms composed of balanced dots at diopters where forms composed of unfiltered dots were easily recognized, and (b) the independence of performance in recognizing balanced dot forms and form size, where form size is an important determinant in the recognition performance of forms composed of unfiltered dots.

Method

Subjects. Four subjects participated in the study. Two were naive, but all subjects were experienced psychophysical observers.

Stimuli. All stimuli were displayed on a Gould image-processing system, with a screen resolution of 512 × 512 pixels and 256 gray levels. A photometric analysis of the screen image was used to derive the specific nonlinear relationship between the quantized gray level (0–255) and the actual screen luminance. This analysis allowed for an explicit comparison between psychophysically determined balanced dots and dots balanced in the luminance domain.

Dot definition. Carlson et al. (1984) did not discuss the construction of balanced dots, except to remark that the balancing condition must be satisfied to within 1% accuracy, and that this is readily achieved electronically. It is, in fact, generally obvious when a balanced dot is constructed. We used the following method for determining the luminance values required for balancing.

The balanced dots used in this study were formed from a 4 × 4 square grid of pixels, the inner 2 × 2 grid forming the bright center and the outer 12 pixels forming a dark annulus. The entire balanced dot subtended 7.0 arc min. The background gray level and center brightness were chosen first. The luminance of the annulus was varied until the dots either disappeared or were maximally indistinct. Whether a balanced dot actually disappears depends on the observer's distance from the screen and the brightness chosen for the dot center. This procedure was repeated at least four times before the beginning of any experimental session. In all cases, the gray value chosen by a given subject for the annulus never varied by more than 1%. Different subjects also agreed on the appropriate gray level with differences never exceeding 1%. The unfiltered dots consisted of a 2 × 2 matrix of uniformly illuminated pixels. The luminance level was always chosen to be that of the center of the balanced dot. The unfiltered dots subtended 3.5 arc min.

This informal method of determining the balancing condition is apparently the procedure adopted by Carlson et al. (1984) in producing balanced dots. In fact, it is sufficient, but we desired a straightforward psychophysical experiment that would validate this procedure. For this reason we developed the form identification task.

Form definition. The forms used in this experiment are illustrated in Figure 1. This figure shows schematically how the forms were constructed from individual balanced dots, and how the balanced dots appeared on the display monitor. A balanced dot phenomenally consists of a bright core with a dark surround, all embedded in a uniform gray field. Forms constructed from unfiltered dots were constructed by simply setting the luminance of the dark surround of each balanced dot to the value of the ambient background. The forms came in four sizes, subtending 3.5°, 5.2°, 7.0°, and 8.8°. The dot density for all sizes was constant. If the form was composed of balanced dots, there was a 4-pixel space (1 dot width) between each dot. All forms were displayed for 2 s.

Procedure. Subjects dark-adapted for about 10 min in scotopic room conditions. They then viewed a block of forms monocularly through a randomly chosen lens. The lens strength varied from 1 to

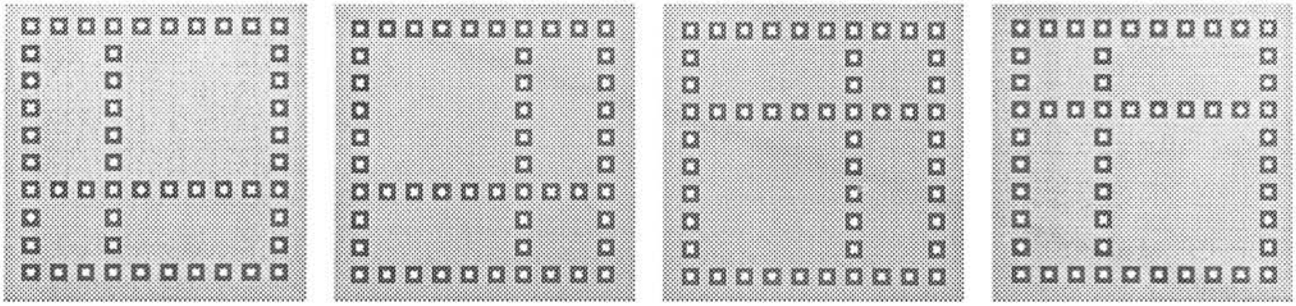

Figure 1. Forms used in the identification experiment. The forms displayed here are schematically represented as being constructed from balanced dots. Note that each dot is constructed from a bright core and dark surround. A dot is psychophysically balanced when the perceived average luminance of the dot is equal to the perceived luminance of the background. The dot density was one dot every eight pixels or every other balanced-dot width. These forms cannot be distinguished on the basis of the total flux from the screen.

7 diopters in 1-diopter increments. Viewing distance was 105 cm. Each block consisted of viewing all four forms, in each of four sizes, composed of both balanced and unfiltered dots. Stimuli were randomized within blocks. Identifications were made verbally and were recorded by an assistant. The entire experiment consisted of the completion of seven blocks, one block for each diopter lens.

Results and Discussion

Form recognition performance was assessed using a repeated measures analysis of variance (ANOVA) with diopter (7), dot type (2), and form size (4) as the independent variables. There were significant main effects for diopter, $F(6, 18) = 120.8$, $p < .0001$, dot type, $F(1, 3) = 371.8$, $p < .0001$, and form size, $F(3, 9) = 25.1$, $p < .0001$. There were significant interactions of dot type by form size, $F(3, 9) = 24.0$, $p < .0001$, dot type by diopter, $F(6, 18) = 62.5$, $p < .0001$, form size by diopter, $F(18, 54) = 3.5$, $p < .0002$, and dot type by form size by diopter, $F(18, 54) = 3.0$, $p < .0008$.

The results are displayed in Figure 2. In the unfiltered-dot condition, the forms subtending 3.5°, 5.2°, 7.0°, and 8.8° are denoted by the functions labeled 1, 2, 3, and 4, respectively. Error bars in the balanced dot condition are shown for the forms subtending 8.8° because the variability was generally greatest at this form size. It is quite clear from this graph that form size is not a parameter in the identification of balanced-dot forms. All four functions move together. Conversely, the functions corresponding to the unfiltered dot forms are differentiated by size; the larger forms are more visible at every diopter. In addition, the balanced-dot forms are identified no better than chance (25% correct) beyond 3 diopters. These results show that our construction of balanced dots satisfy both of the stated criteria: Form size was a factor only for the unfiltered dots, and unfiltered dots were more easily identified throughout the range of diopter.

Experiment 2: Random-Balanced-Dot Kinematograms

Our first study with balanced dots investigated whether d_{max} in random-dot kinematograms varied as a function of the spatial frequency content of the random texture. This study resembles that of Chang and Julesz (1983), who showed that high-pass filtering decreased d_{max}, whereas low-pass filtering increased d_{max}. The critical difference here is in the nature of the bandpass filter. The filter used by Chang and Julesz was global; the Fourier transform of the entire screen image was bandpass filtered. In the balanced-dot technique, the filter is local; the Fourier transform of each individual texture element is filtered.

Global filtration achieves the desired tailoring of the spatial frequency content of the texture by changing its graininess—the scale of fluctuation in the random-dot texture. In the experiments of Chang and Julesz (1983), the filtering was subject to the constraint that the total power be roughly constant. In the luminance domain, this constraint requires that the total number of illuminated pixels be constant across

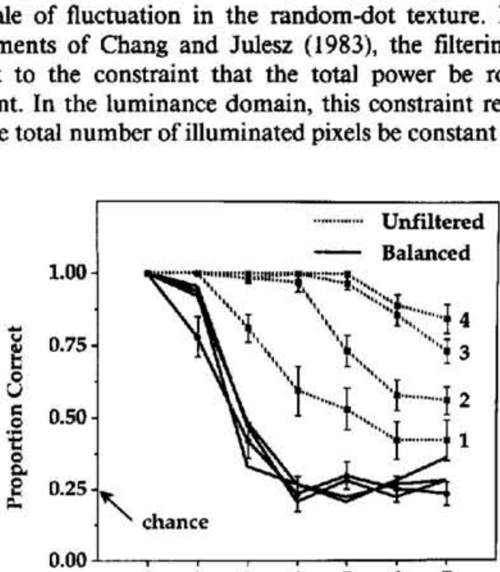

Figure 2. Results from the form identification study. The proportion correct is a function of the dot type, the diopter strength of the lens, and the form size. For unfiltered dots, the form sizes 3.5°, 5.2°, 7.0°, and 8.8° are associated with Curves 1, 2, 3, and 4, respectively. Form size is not differentiated for balanced dots. The pattern of results obtained shows that balanced dots do not blur. Error bars are shown in the balanced-dot condition for the forms subtending 8.8°, for which the standard errors were generally greatest.

filtering conditions. For textures that have a constant fraction of illuminated pixels, the fluctuation scale (grain area) is inversely proportional to the number of texture elements. Global high-pass filtering reduces the grain size and increases grain number, whereas global low-pass filtering merges regions and inflates the grain size. Because the screen size is constant in all conditions, textures that have more elements have a larger number density and a smaller average interelement separation. Textures with fewer elements have smaller number densities and larger spacings on average.

The number of grains in low-passed random-dot kinematogram displays can be on the order of 10, as inspection of the figures in Chang and Julesz (1983) illustrates. When there are so few distinct elements on the screen, even considering their irregular shapes and luminance distributions, it is uncertain that their motion is mediated by the short-range process. Small-number multielement displays are easily created that permit path perception over several degrees of arc. Illusory occlusion, described later, is a prime example. Insofar as small-number multielement apparent motion displays are generally regarded as being mediated by the long-range process, and given the general lack of definition of what constitutes the short- and long-range processes, it is not clear whether a low-passed display is in fact tapping the short-range process. Petersik (1989), in a review of the theoretical and experimental status of the two-process distinction, suggested that the increase in d_{max} in the low-passed displays may be a consequence of a transition to the long-range process because of the reduction in element number.

The local filtration technique used in the following experiment allows the spatial frequency content of a texture element to be adjusted independently of the interelement separation and the element number density. A display constructed of balanced dots looks essentially like a display of unfiltered dots, except that all the dots in the balanced display have dark rings around them. In this sense, any given unfiltered display induces a balanced isomorph that has exact positional correspondence. Using a local filtration technique in creating balanced dots, we were able to determine the dependence of d_{max} on spatial frequency without the confounds associated with changing grain size or number density.

Method

Subjects. Three experienced psychophysical observers, 1 naive as to the purpose of the study, served as subjects.

Stimuli. Two-frame random-dot kinematograms were created by randomly filling a screen of dimensions 15° × 15° with dots at 40% density. (See Braddick, 1974, for a detailed description of the construction of a random-dot kinematogram.) Both unfiltered-dot kinematograms and balanced-dot kinematograms were created. The coherent region that was rigidly displaced in the second frame was constructed to be a square with a missing quadrant. The odd quadrant was replaced with randomly positioned dots. The coherent square measured 7.0° × 7.0°. Four forms were defined in terms of which quadrant was missing from the coherent square. The short-range process was assessed by measuring the error rate in a form identification task as a function of the displacement of the coherent square. Five displacements were used: 3.5, 7, 14, 21, and 28 arc min. In terms of the size of the unfiltered dots, these displacements correspond to 1, 2, 4, 6, and 8 dot widths. Each display consisted of 10 presentations of both frames; each frame was shown for 67 ms. The interstimulus interval (ISI) was zero. Balanced dots were created as in Experiment 1.

Procedure. Subjects dark-adapted for about 10 min in scotopic room conditions. Each subject then viewed four blocks of stimuli. Viewing distance was 105 cm. A block consisted of each of the four forms displayed in each of the five element displacements; all forms and displacements were shown using both balanced and unfiltered dot kinematograms. Thus, each block consisted of 40 trials. Stimuli were randomized within blocks. Each display was viewed for 1.3 s, after which the subject would indicate which quadrant was missing. Responses were recorded by an assistant.

Results and Discussion

The data were analyzed using a repeated measures ANOVA with dot type (2) and displacement size (5) as the independent variables. For reasons discussed later, we do not include results for balanced dots at the minimum displacement. Mean proportion correct in each condition is plotted in Figure 3 with the associated standard errors of the mean. Main effects for dot type, $F(1, 2) = 58.3$, $p < .017$, and displacement, $F(3, 6) = 6.84$, $p < .023$, were both significant. The Dot Type × Displacement interaction was marginal, $F(3,6) = 4.61$, $p < .053$, but we did not include larger displacements that would have brought the unfiltered condition to chance. Most important, the unfiltered dots did not show a decline in texture segregation for displacements less than 15 arc min, and discrimination was still above chance at a displacement of 25 arc min. Conversely, for balanced dots, discrimination was near chance by any displacement equal to or exceeding one balanced-dot width.

These results are consistent with earlier studies of high-pass filtering of random-dot kinematograms (Chang & Julesz, 1983) in demonstrating that d_{max} is reduced upon high-pass filtration. It is, however, difficult to compare quantitatively our results with those of Chang and Julesz because of the

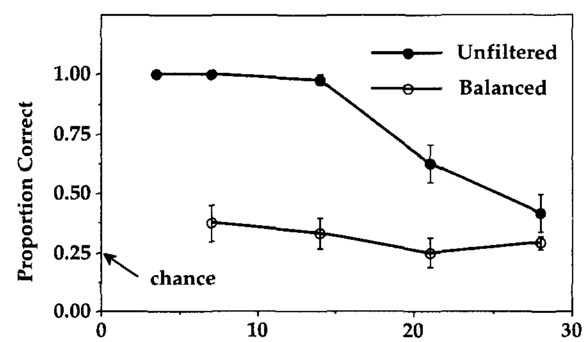

Figure 3. Results from the random-dot kinematogram study. The segregating textures in these studies are squares with a missing quadrant. The proportion correct in the identification of the location of the missing quadrant is a function of the dot type and the size of the between-frame displacement. Results are not plotted for balanced dots at the smallest displacement because this displacement was less than one balanced-dot width.

sensitivity of d_{max} to virtually every parameter used in the construction of the display, including dot size, target area, presentation time, as well as the specific task—figural identification as opposed to the identification of direction of motion of the segregating surface.

In this experiment, we have shown that removal of the *DC* point and suppression of low-spatial frequencies considerably decreases d_{max} independently of the number-density of elements and the interelement spacing. We can conclude therefore that the earlier results of Chang and Julesz (1983), that d_{max} decreases upon high-pass filtering, is not a result of the transition to the long-range process. Rather, the dependence of d_{max} on spatial frequency appears to be intimately related to the way in which random textures are treated in visual processing.

Further confirmation of the irrelevance of number-density of texture elements in the determination of d_{max} comes from experiments of Baker and Braddick (1982). In these experiments, the spatial frequency content of the individual dots was held constant and the number-density was varied. It was found that number-density had no effect on d_{max}. We have replicated these studies with balanced and unfiltered dots, varying number-density between 10% and 40% (50% is the maximum), and also found no effect on d_{max}. We can therefore make the stronger claim that not only does spatial frequency enter independently of number density, but that number-density is itself not a parameter.

Texture segregation was found in balanced-dot/random-dot kinematograms for a displacement of half a balanced-dot width (recall that the width of an unfiltered dot is half that of an unbalanced dot). At this displacement, there was perfect performance in shape identification. There are two reasons why discrimination was possible at this displacement. First, it may be that although the DC point is canceled in balanced dots the low-spatial frequencies are not entirely suppressed. There is some leakage in the quadratic portion (see Appendix and Figure A1) of the power spectrum near the *DC* point where suppression is complete. This leakage may permit texture segregation at sufficiently small displacements. Second, when the dot displacement is less than one dot width, there are random overlaps produced when the coherent region is shifted in the second frame. These overlaps do supply some partial contour information. Because we did not remove these overlaps, we do not stress this result, and have not included the minimum displacement for balanced dots in our analyses.

Experiment 3: Global Field Organization in Illusory Occlusion

There is a class of apparent motion displays that have features in common with both the short- and long-range processes. This class is defined by the global organization of an ensemble of individual apparent motion pairs. We discuss here the one that is closely allied to the short-range process, illusory occlusion.

Ramachadran, Inada, and Kiama (1986) described a display that consists of a number of apparent motion pairs (two spatially separated dots that blink in alternation) and a test dot this is paired with an occluder (a piece of paper taped onto the graphics screen will suffice). The perception of the test dot was ambiguous because it can either appear to blink without moving or appear to move behind the occluder. Within a window of spatial and temporal parameters, the unpaired blinking dot was entrained by the apparent motion of its neighbors and was observed to pass beneath the occluder. This phenomenon was termed illusory occlusion.

There is a strong family resemblance between illusory occlusion and random-dot kinematogram texture segregation. In the perception of segregating texture, individual dots do not emerge in isolation. Rather, what is observed is an entire texture in motion, and the impression is that not only are the stimulus elements themselves segregating but the unmarked space in between segregates as well. A similar percept is experienced in illusory occlusion. The entrainment of the unpaired test dot appears to be part of a field in which all of the apparent motion pairs are embedded. This field is quite similar to a random-dot kinematogram texture in that it does not exist in any particular frame; it arises out of the synchrony imposed by coordinating the onset and offset times of the apparent motion pairs. In this sense, both the illusory occlusion field and the random-dot texture are emergent properties of the temporal organization of an ensemble of elements. These similarities suggest that illusory occlusion may also show a sensitivity to the power spectrum of the individual stimulus elements comprising the apparent motion pairs.

There are, however, also important differences between illusory occlusion and texture segregation. The primary difference between a moving texture and a synchronized field of apparent motion pairs is that individual paths are perceived in the latter but not in the former. Path perception in the long-range process is highly labile; priming, experience, gender, and presumably imagination are all relevant to how long-range apparent motion is perceived (Proffitt, Gilden, Kaiser, & Whelan, 1988). None of these parameters are relevant to the short-range process. The individual differences that exist for the short-range process are trivial when compared with the interpretations that can arise for the long-range process.

It was conjectured by Ramachadran et al. (1986) that the global field percept that arises in illusory occlusion, and the consequent entrainment of the unpaired dot, may be mediated by low spatial frequencies. Although a formal algorithmic description of field formation has not been developed, this conjecture does satisfy some common-sense notions of how a field might be generated from isolated dot pairs. For example, it might be argued that large receptive fields that do not resolve individual dot pairs might constructively pool their outputs if the pair timing across the field was synchronized. In this way the impression of a global field might be generated. Cooperative models that are based on pooled outputs of spatially disparate detectors have also been suggested as the basis of texture segregation in random-dot kinematograms (Chang & Julesz, 1984, 1985).

An alternative model for illusory occlusion is that field generation is an example of figural organization, where the figure is specified by the simultaneous organization of discrete motion paths. In this model, the individual dots are regarded as objects and illusory occlusion arises out the visual system's tendency to enforce object permanence. Notions of object permanence do not arise from the dipole statistical properties of images, but rather are embedded in the higher order statistics that allow the visual system to segment the dot pairs. Consequently, if the figural properties of the dot paths are key to illusory occlusion, we would not expect transformations on the power spectrum to influence the percept. In particular, the existence of low spatial frequencies would not be crucial for field formation.

In this experiment, we assessed the independent contributions of spatial frequency content and temporal phase in the perception of illusory occlusion. Depending upon the amount of asynchrony, the instantaneous path directions between two pairs will agree for various fractions of the time consumed by the two-frame cycle. The degree of temporal phase synchronization between the paired field dots and the unpaired test dot is a potentially important variable for the strength of entrainment that occurs in illusory occlusion. As in our studies of random-dot kinematograms, local filtering allows two isomorphic displays to be constructed: one with unfiltered dots and one with balanced dots. This isomorphism permits a direct test of the influence of low spatial frequency upon the perception of illusory occlusion.

Method

Subjects. Three experienced psychophysical observers, naive as to the purpose of the experiment, served as subjects. These subjects had not participated in any of the previous studies.

Stimuli. Ten dot pairs were randomly positioned in a 3.75° × 3.75° area of the screen. Each dot pair was separated by 0.5°. The filtered and balanced dots were created as in the previous experiments. The balanced dots subtended 7 arc min, and the unfiltered dots subtended 3.5 arc min. One additional dot was placed near the center of the screen. A gray texture of dimensions 0.5° × 0.5° was placed on the screen centered on the position where its partner would have been. An example of this configuration is shown in Figure 4.

For a given dot pair, the ISI was zero, and each dot had on and off times of 330 ms. The dots that had partners we refer to as field dots. The unpaired dot we refer to as the test dot. All the field dots were synchronized so that their durations, onset times, and offset times were matched to produce identically directed motion paths.

The test dot was treated separately. Its partner, the occluding screen, was always on. The test dot, being to the left of the occluding screen, is nominally associated with the dots in the field that were to the left of their partners. The timing of the test dot was linked to the onset times of leftward field dots. Onset time of this dot lagged behind the onset times of the leftward field dots by either 0, 99, 198, or 297 ms. If we regard the time sequence of the dot illumination to be a square wave, then these lags correspond to phase angles of 0°, 54°, 108°, and 162°.

A given display consisted of 10 complete cycles of the field dot apparent motion sequence (6.6 s). All displays were composed of either unfiltered or balanced dots. Balanced dots were constructed by the procedures described in Experiment 1.

Procedure. Subjects dark-adapted for about 10 min in scotopic room conditions. Illusory occlusion was explained to the subjects using a diagram. Subjects were instructed to rate the quality of the illusory occlusion on a 7-point rating scale. A rating of 1 would mean that the unpaired dot was observed to just blink on and off. A rating of 7 would mean that the unpaired dot appeared to move behind the occluding screen, and that its path was no less salient than the paths associated with the dots that were paired. For the subjects to become familiar with the range of stimuli, and with the rating scale, they rated four blocks of trials as practice. A block consisted of eight illusory occlusion displays, four of which were composed of balanced dots and four of which were composed of unfiltered dots. Immediately following the practice, subjects rated an additional five blocks of stimuli. Viewing distance was 105 cm.

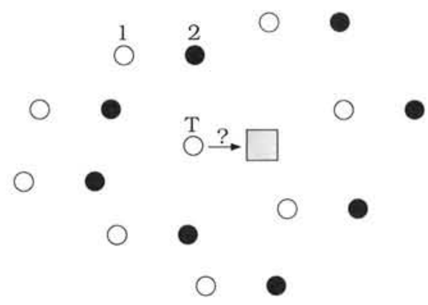

Figure 4. The geometry of illusory occlusion. The open circles represent stimuli that appear in the first frame, and the solid circles represent stimuli that appear in the second. The test dot marked T has no partner and is associated with an occluder. It appears either together with dots in Group 1 or after some time delay. The test dot may appear to either blink or move in apparent motion behind the occluder.

Results and Discussion

Kolmogorov–Smirnov tests on the rating distributions were performed to determine whether temporal asynchrony and high-pass spatial filtering were effective in degrading the percept of illusory occlusion. Overall ratings for balanced dots were not reliably distinguishable from those of unfiltered dots (maximum difference between the two cumulative distributions was .17, yielding a probability $p < .36$). There was, however, a detectable difference at the phase angle = 0°, where subjects unanimously gave the illusory occlusion the maximum rating of 7 for the unfiltered displays. The median ratings for temporally phase-shifted displays were a monotonically decreasing function of the shift magnitude. The rating distributions were generally distinguishable when considered pairwise; the probability for obtaining a difference in the cumulative distributions as large as was observed was, respectively, $p < .04$ for comparing lag = 0° with lag = 54°, $p < .02$ comparing lag = 54° with lag = 108°, but $p < .61$ for comparing lag = 108° with lag = 162°. The data are plotted in Figure 5, where these relationships are manifest. Error bars signify standard error of the median.

There are three conclusions that we draw from this experiment. The first is that phase shifting the test dot relative to the field dots degrades perceived illusory occlusion. The phase shift and median ratings were related by a monotonically decreasing function for both dot types. The second conclusion is that the phase angle variable induces a larger interval of the rating scale to be used than the variable of dot type. We interpret this to mean that phase shift is a more important variable than the presence of low spatial frequencies. This comparison is meaningful because both variables were pre-

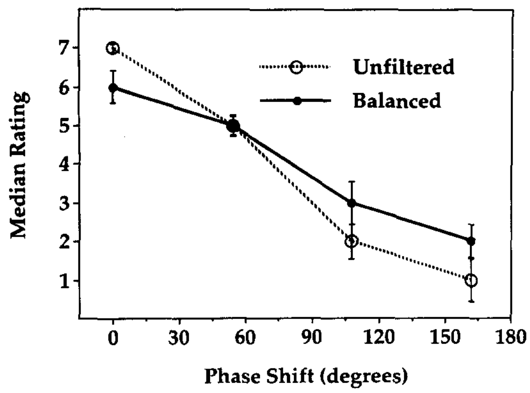

Figure 5. Results from the illusory occlusion study. The median ratings of the percept of illusory occlusion are shown as a function of dot type and the phase difference between the onset times of the field dots and the occluding dot. Error bars show the standard deviation of the median.

sented in conditions covering the extent of their possible ranges. Finally, when the entire display is synchronized, the illusion is somewhat stronger when low spatial frequencies are present. These data do not allow a calibration of how much stronger or weaker illusory occlusion is in one condition relative to another because differences may not be numerically meaningful.

Because of the measurement ambiguities inherent in rating reports, we debriefed our subjects on what they saw when they looked at the illusory occlusion displays. These subjects had not previously seen balanced dots, and because they do not occur in nature, they were an unusual stimulus. The lower ratings given to the balanced dot displays at small phase shifts (especially at zero phase shift) may be due to the strangeness of the display as opposed to a weaker motion signal. Similarly, we do not place much emphasis on the crossed interaction because at large phase shift (phase shifts exceeding 90°) the percept of illusory occlusion was reported to be weak in any case. The debrief reports thus provided convergent evidence that a large phase shift weakened the illusion much more than a change in the type of dot in the display. These naive reports were consistent with the authors' experience in working with these displays over several months.

A replication of this experiment using 5 naive subjects showed the same pattern of results. In the replication only the extreme phase shifts of 0° and 162° were used. Our primary result, that phase shift of the test dot induces ratings that cover the entire scale whereas balancing the dots induces only slightly different ratings, was replicated.

Experiment 4: Temporal Synchronization in Random-Dot Kinematograms

The use of temporal asynchrony as an independent variable permits a critical test of the primary thesis of this article: that the two-process distinction in apparent motion is in part a distinction in depth of computation through the *n*-point statistical hierarchy. It was shown in Experiment 3 that the introduction of temporal asynchrony caused a radical deterioration in the global field effect of illusory occlusion that was independent of the spatial frequency content of the display. Our interpretation of this result is that illusory occlusion, as an example of the long-range process, is computed primarily in terms of the 3-point and higher statistics and is therefore insensitive to manipulations in the point and dipole statistics. The natural question that arises here is what dependency on spatial frequency would be found when temporal asynchrony is introduced into random-dot kinematograms. The observations of Julesz and collaborators (Julesz, 1975, 1981; Julesz et al., 1973) that random-dot patterns are computed primarily in terms of their point and dipole statistics suggest that temporal asynchrony will have differential effects on unfiltered and high-passed random-dot kinematograms. In the following experiment this conjecture was tested.

Method

Subjects. Two of the observers in Experiment 2 served as subjects.

Stimuli. Random-dot kinematograms were created as in Experiment 2. There were two conditions that differed with respect to the phase of onset time for the coherent and incoherent regions of the random-dot kinematogram. In the asynchronous condition, the onset times were staggered by one half of the frame duration, or 33 ms. This corresponded to a phase angle of 180°. These displays require four frames to be specified because the random region is changed at a different time than the coherent region is shifted. In the synchronous condition, the onset times were simultaneous (i.e., only two frames were required to specify this stimulus).

In this experiment, the viewing distance was shortened to 83 cm. Note that at this viewing distance the balanced dot core subtends 4.43 arc min instead of 3.5 arc min as in the previous experiments. The short-range process was again assessed by performance in a form identification task where the displacement of the coherent region was the independent variable. Displacements of 2, 4, 6, and 8 core widths were used, corresponding (at 83 cm) to 8.86, 17.7, 26.6, and 35.4 arc min. In all other respects, the stimulus presentation and procedure were as in Experiment 2.

Results and Discussion

The data were analyzed using a repeated measures ANOVA, with dot type (2), displacement size (4), and timing (2) as the independent variables. There were significant main effects for dot type, $F(1, 1) = 235.1$, $p < .042$, displacement size, $F(3,3) = 73.7$, $p < .005$, and timing, $F(1, 1) = 210.3$, $p < .044$. The two-way interactions were also significant; Dot Type × Timing, $F(1, 1) = 60.5$, $p < .029$, Dot Type × Displacement Size, $F(3, 3) = 13.1$, $p < .041$, and Timing × Displacement Size, $F(3, 3) = 21.4$, $p < .014$. The three-way interaction was marginal, $F(3, 3) = 5.20$, $p < .11$.

The results of this experiment are summarized in Figure 6. In the synchronous condition, the results are quite similar to those of Experiment 2, except that at the smallest displacement (8.85 arc min) the observers are better than chance in discriminating forms in balanced random-dot kinematograms. Recall that in Experiment 2 observers were roughly at chance in all balanced-dot displacement conditions (however, segregation was observed at a displacement of 3.5 arc min but with possible confounding form cues). We believe that the

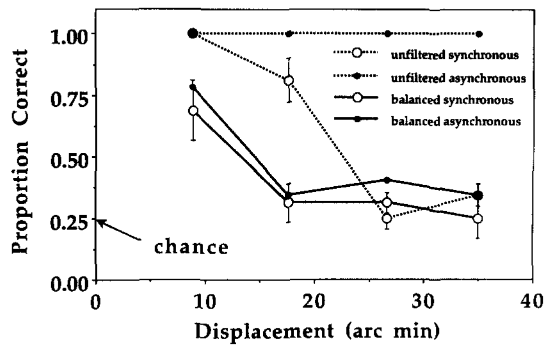

Figure 6. Results from Experiment 4. Plotted are the proportion of correct responses in the identification of a missing quadrant in a random-dot kinematogram. Note that the introduction of temporal asynchrony has different effects with respect to dot type. Balanced-dot identifications are not significantly improved by asynchronization, whereas unfiltered-dot identifications are improved to ceiling.

differences found in the balanced-dot conditions are primarily due to the 20% reduction in viewing distance (83 cm here and 105 cm in Experiment 2). Unlike unfiltered dots, the low spatial frequency power for balanced dots is quite sensitive to viewing distance. Balanced-dot power increases with decreasing viewing distance to the fourth power at low spatial frequencies. Unfiltered dot power, however, is independent of viewing distance to lowest order. A simple calculation using the analytic expressions for balanced-dot power (see Appendix) shows that a 20% reduction in viewing distance boosts the power in the low frequency range (less than 2 cycles/degree) by more than a factor of 3. Increased power in the low spatial frequencies is expected to make balanced-dot textures more visible in random-dot kinematograms; this is in fact the implied result of Experiment 2. Relatively small changes in viewing distance are not expected to change the discrimination function for unfiltered dot textures, and in fact no change was detected.

The main result of this experiment is the divergent behavior of balanced and unfiltered textures in the asynchronous condition. The unfiltered textures were always discriminated perfectly; there were no errors at any measured displacement. However, there was no detectable difference in the discriminability of the balanced dot textures. The difference between the synchronous and asynchronous conditions as a function of dot type is captured by the significant Dot Type × Timing interaction and is clearly visible in Figure 6.

We contrast the large Dot Type × Timing interaction that is manifest in the context of random-dot kinematograms with the negligible interaction in illusory occlusion displays. In the random-dot kinematogram, not only does balancing the displays make it difficult to achieve texture segregation, but the introduction of temporal asynchrony affects the balanced and unfiltered displays in markedly different ways. This result supports our claim that the depth of processing is different in illusory occlusion and random-dot kinematograms. In particular, it gives evidence that processing of random-dot kinematograms terminates near the dipole level, whereas the processing of illusory occlusion commences at the 3-point statistics.

General Discussion

In this article, we presented critical tests for the dependence upon low spatial frequencies in the perception of apparent motion; for the short-range process using random-dot kinematograms and for global organization of the long-range process using illusory occlusion. These tests required the creation of high-pass filtered displays that preserved pointwise correspondence with unfiltered displays that had low spatial frequencies. The balanced dot allows this criterion to be met. Experiment 1 demonstrated that a psychophysical determination of the balancing condition creates dots with the expected properties. Experiment 2 showed that low spatial frequencies are critical in determining the maximum displacement possible for the perception of texture segregation in random-dot kinematogram, and that this sensitivity is not related to the grain size of the display or to the dot number density. Experiment 3 showed that temporal asynchrony degrades the perception of illusory occlusion independently of spatial frequency. Experiment 4 showed that temporal asynchrony effects d_{max} differentially depending upon the spatial frequency content of the display.

A key distinction that is addressed by these experiments is whether or not the stimulus that exists in a momentary time slice permits perceptual organization as a figure. What constitutes a figure or shape is quite subtle, but it is clear that perceptual organization must always involve some sort of summarizing or "chunking" of the information that describes the light distribution in detail. In the sense of coding theory (Restle, 1982), perceptual organization always involves an interpretation of a stimulus that contains fewer primitives than a point-by-point enumeration of the light intensity. The frames of a random-dot kinematogram are unique in this regard because they do not allow a description that is more succinct than a complete listing of the dot positions. The information content of a random-dot pattern is much larger than that contained in the same amount of light composing a figure.

The visual system recognizes the informational differences between random-dot patterns and figural contour in terms of the statistical moments that are computed in brief displays. The information associated with contour is contained in the 3-point and higher correlation functions that form the phase spectrum. The perception of contour is not adversely affected by radical transformations of the power spectrum. Random-dot textures, conversely, are processed by the visual system in terms of the point and dipole statistics that form the power spectrum. Consequently, transformations on the power spectrum have striking effects on texture segregation in random-dot kinematograms.

An important distinction that arises in consideration of random-dot kinematograms is the descriptive level at which spatial frequency analysis occurs. The contours that arise in random-dot kinematograms are subsequent to a motion analysis by a perceiver; these contours do not exist in the light distribution at any one time. The output of this analysis is a texture that can in turn be analyzed within the context of the Fourier theory, but it is implicit that this stage of analysis is not mediated by detectors responding to light on individual time slices. Fourier decomposition of a segregating texture must be performed on an abstract representation of the texture. Thus, our work as well as that of Chang and Julesz (1983) show the importance of the power spectrum in the light distribution as an input to a motion analysis of random pattern. Similarly, Prazdny's (1986) experiments show that path preference in the long-range process is determined by processes that are subsequent to spatial frequency analysis on the light distribution. Neither of these sets of studies addresses the role of the power spectrum of a segregating texture as an input to other processes, in particular the long-range motion process. However, it must be recognized that the motivation for introducing spatial frequency analysis into the theory of the long-range process was the desire to describe the motion process in terms of the light distribution. The physiology of spatial frequency channels is fundamentally related to the processing of light energy.

The distinction between spatial analyses on individual time slices and spatial analyses on abstract representations that require integration (or covariation) over time is particularly important in the consideration of global field effects. Global field effects share both time-integrated organizational properties with texture segregation in random-dot kinematograms and time-sliced organizational properties of static contours. In the time domain, random-dot kinematograms and field effects are produced by the synchronization of individual elements. This organization leads to the perception of a coherent moving texture in one case and to a motion field in the other. However, an illusory occlusion display has, in addition, figural information at many levels. First, the patch associated with the test dot is not simply an array of pixels, it is an occluder. In fact, Ramachadran et al. (1986) reported that a piece of paper pasted to the screen enhances the percept of occlusion. Second, individual dots in the apparent motion field have identities: In the long-range process, the impression is always of something moving; in the random-dot texture segregation, specific dots are not individually identified apart from the texture. It is this link with static configural properties that makes illusory occlusion immune to transformations on the power spectrum.

A fundamental question in perceptual organization is the role that time plays in the formation of groups. We distinguish here the problem of grouping in time from problems in the detection of motion for which there are formal models (Adelson & Bergen, 1985; van Santen & Sperling, 1985; Watson & Ahumada, 1985) based essentially on notions of temporal covariance. Grouping, on the other hand, has not received formal treatment, being a much more difficult problem involving issues of symmetry, similarity, and form. Note that our demonstration that temporal asynchrony decouples the test dot from the field in illusory occlusion does not address whether field formation is a cooperative process operating in early vision, or whether it arises from an interpretation of the display in terms of object permanence. Both models would suggest that asynchrony would disrupt grouping. However, even when the test dot and field are synchronized, reports of occlusion are not universal, and it is possible to prime subjects to see the test dot blinking on and off without occlusion. Cooperative algorithms seem to lead to obligatory perceptions that are not subject to large individual differences, such as the perception of texture segregation in random-dot stereograms.

A related global field effect has been discussed by Ramachadran and Anstis (1985) in the context of bistable quartets. Bistable quartets are defined by four dots positioned at the vertexes of a square, where the two dots connected by a diagonal are shown on alternate frames. Such quartets permit both horizontal and vertical motions (also compatible but not often reported are clockwise and counterclockwise motions). When a group of bistable quartets are shown together in synchrony, the directional ambiguity is resolved for the entire field in common; all the quartets move vertically or they all move horizontally. We refer to this phenomenon as orientation capture. This percept does not appear to be labile. In pilot studies we have found, in agreement with Ramachadran and Anstis (1985), that observers cannot split the field up so that horizontal and vertical motion coexist. In addition, we found that orientation capture also occurs in fields of balanced-dot bistable quartets.

In further pilot studies, motivated by our work on illusory occlusion, we found that the introduction of asynchrony between different groups of quartets can in some cases lead to the simultaneous presence of both vertical and horizontal motion interpretations within the quartet field. For these studies, we created two groups of bistable quartets. The quartets within a group were synchronized, but there was a time lag between the two groups. The possibility of orientation decoupling between the two groups depended critically on the manner in which the two groups were distributed throughout the field. If the groups were isolated and distributed along separate sections of a figure, such as the two line elements of the letter *T*, orientation decoupling was possible. However, if the groups were intermixed at random, then decoupling was not possible. These studies will be reported in a future publication.

In conclusion, these studies were designed to clarify the relation between perceptual organization in apparent motion and the components of the Fourier transform. Our experiments suggest that temporal organization can create percepts that exist as a result of motion perception per se, but it does so using the information available in individual time slices. The sorts of computations performed depends critically on the spatial structure at each moment of time. Vision apparently does not use the phase information in random-dot patterns and computes up to the autocorrelation. Brief presentations of textures, whether static or in apparent motion, are perceptually bound to the information in the power spectrum. Scenes in which contour recognition and object segmentation take place require information that exists in the phase spectrum. The perception of a moving contour, even in apparent motion, is subsequent to a calculation of the phase spectrum. The higher level reached in the n-gon statistical hierarchy for computations of spatial organization leads to invariance under transformations on the power spectrum.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America, 2*, 284–299.

Anstis, S. M. (1980). The perception of apparent movement. *Philosophical Transactions of the Royal Society of London, 290*, 153–168.

Badcock, D. R. (1989). Phase- or energy-based face discrimination: Some problems. Unpublished manuscript.

Baker, C. L., & Braddick, O. J. (1982). The basis of area and dot number effects in random dot motion perception. *Vision Research, 26*, 1253–1259.

Becker, M. F., & Knopp, J. (1978). Processing of visual illusions in the frequency and spatial domains. *Perception & Psychophysics, 23*, 521–526.

Braddick, O. (1974). A short-range process in apparent motion. *Vision Research, 14*, 519–527.

Braddick, O. J. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society of London B, 290*, 137–151.

Carlson, C. R., Moeller, J. R., & Anderson, C. H. (1984). Visual illusions without low spatial frequencies. *Vision Research, 24*, 1407–1413.

Carrasco, M., Figueroa, J. G., & Willen, J. D. (1986). A test of the spatial-frequency explanation of the Muller–Lyer illusion. *Perception, 15*, 553–562.

Chang, J. J., & Julesz, B. (1983). Displacement limits for spatial frequency filtered random-dot cinematograms in apparent motion. *Vision Research, 23*, 1379–1385.

Chang, J. J., & Julesz, B. (1984). Cooperative phenomena in apparent movement perception of random-dot cinematograms. *Vision Research, 24*, 1781–1788.

Chang, J. J., & Julesz, B. (1985). Cooperative and non-cooperative processes in apparent movement of random-dot cinematograms. *Spatial Vision, 1*, 39–45.

Chen, L. (1985). Topological structure in the perception of apparent motion. *Perception, 14*, 197–208.

De Valois, R. L., & De Valois, K. K. (1980). Spatial vision. *Annual Review of Psychology, 31*, 309–341.

Ginsburg, A. P. (1975). Is the illusory triangle physical or imaginary? *Nature, 257*, 219–220.

Ginsburg, A. P. (1986). Spatial filtering and visual form perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (pp. 34-1–34-41). New York: Wiley.

Ginsburg, A. P., & Evans, P. W. (1979). Predicting visual illusions from filtered images upon biological data. *Journal of the Optical Society of America, 69*, 1443.

Graham, N. (1981). Psychophysics of spatial-frequency channels. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 1–26). Hillsdale, NJ: Erlbaum.

Green, M. (1986a). Correspondence in apparent motion: Defining the heuristics. *Proceedings of Vision Interface, '86*, 337–342.

Green, M. (1986b). What determines the correspondence strength in apparent motion? *Vision Research, 26*, 599–607.

Hood, D. C., & Finkelstein, M. A. (1986). Sensitivity to light. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (pp. 5-1–5-66). New York: Wiley.

Julesz, B. (1975). Experiments in the visual perception of texture. *Scientific American, 232*, 34–43.

Julesz, B. (1981). Figure and grouped perception in briefly presented isodipole textures. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 27–54). Hillsdale, NJ: Erlbaum.

Julesz, B., Gilbert, E. N., Shepp, L. A., & Frisch, H. L. (1973). Inability of humans to discriminate between visual textures that agree in second-order statistics-revisited. *Perception, 2*, 391–405.

Kleiner, K. A., & Banks, M. S. (1987). Stimulus energy does not account for 2-month-old's face preferences. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 594–600.

Lappin, J. B., & Bell, H. H. (1976). The detection of coherence in moving random-dot patterns. *Vision Research, 16*, 161–168.

Marr, D. (1982). *Vision.* San Francisco: W. H. Freeman.

Palmer, S. E., Kube, P., & Kurschke, J. K. (1987). Testing the low spatial frequency conjecture for a gestalt effect. *Bulletin of the Psychonomic Society, 25*, 333.

Peebles, P. J. E. (1980). *The large-scale structure of the universe.* Princeton, NJ: Princeton University Press.

Petersik, J. T. (1989). The two-process distinction in apparent motion. *Psychological Bulletin, 106*, 107–127.

Piotrowski, L. N., & Campbell, F.W. (1982). A demonstration of the visual importance and flexibility of spatial-frequency amplitude and phase. *Perception, 11*, 337–346.

Prazdny, K. (1986). What variables control (long-range) apparent motion? *Perception, 15*, 37–40.

Proffitt, D. R., Gilden, D. L., Kaiser, M. K., & Whelan, S. M. (1988). The effect of configural orientation on perceived trajectory in apparent motion. *Perception & Psychophysics, 43*, 465–474.

Ramachadran, V. S., & Anstis, S. M. (1985). Perceptual organization in multistable apparent motion. *Perception, 14*, 135–143.

Ramachadran, V. S., Ginsburg, A. P., & Anstis, S. M. (1983). Low spatial frequencies dominate apparent motion. *Perception, 12*, 457–461.

Ramachadran, V. S., Inada, V., & Kiama, G. (1986). Perception of illusory occlusion in apparent motion. *Vision Research, 26*, 1741–1749.

Restle, F. (1982). Coding theory as an integration of gestalt psychology and information processing theory. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 31–56). Hillsdale, NJ: Erlbaum.

Rubin, J. M., & Kanwisher, N. (1985). Topological perception: Holes in an experiment. *Perception & Psychophysics, 37*, 179–180.

Shapley, R., & Lennie, P. (1985). Spatial frequency analysis in the visual system. *Annual Review of Neuroscience, 8*, 547–583.

Tyler, C. W. (1977). Is the illusory triangle physical or imaginary? *Perception, 6*, 603–604.

van Santen, J. P. H., & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America, 2*, 300–320.

Watson, B. A., & Ahumada, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America, 2*, 322–341.

Wilson, H. R., & Bergen, J. R. (1979). A four mechanism model for threshold spatial vision. *Vision Research, 19*, 19–32.

(Appendix follows on next page)

Appendix

Balanced dots are constructed so as to have zero power at zero spatial frequency. We follow the method of Carlson et al. (1984) by forming balanced dots through dot subtraction. The dots created for these studies were necessarily square because of the pixel geometry.

The Fourier transform of a uniformly illuminated square with luminance L_1 and length $2a$, centered on the origin, is the two-dimensional sinc function

$$F_1(k_x, k_y) = 4L_1/(k_xk_y)\,\sin(k_x\,a)\,\sin(k_y\,a),$$

where k_x and k_y are the components of the spatial frequency vector in the x and y directions. The power at zero spatial frequency is the contrast of the dot, defined as the integral over the intensity distribution:

$$F_1(0, 0) = 4L_1a^2.$$

A balanced dot is constructed by subtracting from the first dot a second dot with luminance L_2 and length $2b$, also centered on the origin. Dot subtraction will yield a balanced dot when the following condition is satisfied:

$$L_1a^2 = L_2b^2.$$

Phenomenally, the appearance of a balanced dot consists of a bright center with a dark surround. Practically, the balancing condition requires that the area weighted averages of the luminance of the surround and center equal the ambient luminance of the screen.

The balancing condition eliminates the *DC* point of the balanced dot. Centering the dots at a common origin eliminates terms linear in spatial frequency, and so the first nonzero term in the power spectrum is quadratic:

$$F_{bal} = 2/3\ k_r^2(L_1a^4 - L_2b^4) = 2/3\ k_r^2a^2(a^2 - b^2),$$

where the k_r is the radial wave number, defined by $k_r^2 = k_x^2 + k_y^2$. For balanced dots constructed so that $a = 2b$, the spectrum peaks at a frequency roughly at

$$f_{bal}(\max) = k_r(\max)/2 = 0.1/a \text{ (cycles/degree).}$$

It is necessary to choose the dot size, a, sufficiently small (or the viewing distance sufficiently large) so that the peak occurs at a frequency of at least 5 to 10 cycles/degree.

The power spectrum of a balanced dot, together with the spectrum of the inner bright center alone (the unfiltered dot), is shown in Panel A of Figure A1. Both of the spectra are plotted on the same scale on the interval (-20 cycles/degree $< k_x, k_y <$ 20 cycles/degree). The numerical relations are clarified in panel B of Figure A1 (*B*) where the spectra are replotted on a slice defined by $k_x = k_y$, on the interval ($0 < k_r <$ 20 cycles/degree). This figure is computed assuming the core subtends 3.5 arc min (as in Experiments 1–3). The analytic estimate given previously predicts maximum power at 6.9 cycles/degree, in good agreement with the computed maximum at 6.4 cycles/degree.

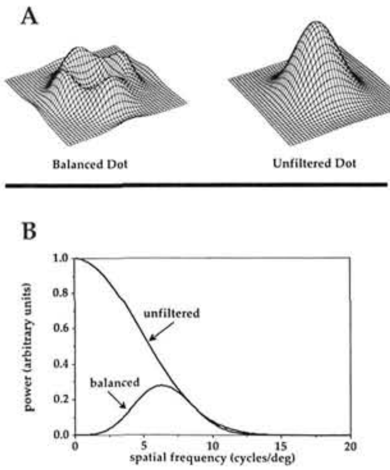

Figure A1. Panel A shows the two-dimensional power spectra of balanced and unfiltered dots. Both figures are displayed on the same scale in the domain (-20 cycles/degree $< k_x, k_y <$ 20 cycles/degree). The symmetry evident in the balanced-dot spectrum is due to their square shape, being constructed from square pixels. This symmetry exists in the unfiltered-dot spectrum, although it is not as apparent. Panel B shows a diagonal slice along $k_x = k_y$ in the interval ($0 < k_y <$ 20 cycles/degree). The figures are computed assuming a core that subtends 3.5 arc min.

Received January 31, 1989
Revision received February 5, 1990
Accepted February 5, 1990 ■