

# A unified approach to the perception of motion, stereo, and static-flow patterns

THOMAS V. PAPATHOMAS

*Rutgers University, Piscataway, New Jersey*

ILONA KOVÁCS

*Rutgers University, Piscataway, New Jersey  
and Eotvos Lorand University, Budapest, Hungary*

ANDREI GOREA

*Université René Descartes  
and Centre National de la Recherche Scientifique, Paris, France*

and

BELA JULESZ

*Rutgers University, Piscataway, New Jersey*

This paper presents a unified approach for studying the perception of motion, stereopsis, and static-flow (Glass) patterns. The objective is to address the same issues across these “modalities.” To this end, a new class of stimuli and procedures were developed, the key feature of which is the incorporation of a forced-choice competition paradigm into the random-dot stimuli paradigm that has traditionally been used in these modalities. The two competing percepts that are pitted against each other are opposite directions for motion, near/far depth planes for stereopsis, and orthogonal global patterns for Glass patterns. The differences in qualitative predictions for competing hypotheses are generally well pronounced, and the results provide clear evidence for deciding between alternative hypotheses. This approach has been used to confirm that covariance is the preferred metric for eliciting global correlations in all the modalities and to investigate the nature of front-end processes in each modality. It has the potential for neurophysiological studies for both single-cell and neuronal ensemble recording.

The global percept in the “modalities” of motion, stereopsis, and spatial grouping is elicited by the correlation of visual attributes. This correlation is across frames in the spatiotemporal domain for motion, across the two monocular retinal images in stereopsis, or in two-dimensional space for Glass patterns (otherwise known as random-dot Moiré, or static-flow patterns). Thus, it is not surprising that there are striking analogies among motion, stereopsis, and Glass patterns (Anstis, 1970; Chubb & Sperling, 1988; Gorea & Papathomas, 1991a, 1991b; Kovács & Julesz, 1992). A unified approach would be desirable for investigating similarities and differences among the underlying

mechanisms for these three important modalities of early vision. The stimuli presented in this paper allow the use of similar spatiotemporal conditions and parameters across all modalities, enabling such comparative studies in psychophysics, as illustrated by two typical experiments in this paper. This new approach with well-parameterized stimuli is also suited for neurophysiological studies, because it enables the isolation of specific early vision mechanisms and pathways without the influence of higher level cognitive processes. The new unified approach combines the strengths of two important methodological breakthroughs in the psychophysics of early vision: the introduction of random-dot patterns, and the forced-choice competition technique. We will briefly consider the background behind each of these developments.

One of the greatest challenges in the study of early vision modalities, such as stereopsis, motion, spatial grouping, and so forth, has been how to remove all higher order “semantics” and familiarity cues, if possible, so as to isolate those mechanisms responsible for the percepts in each modality. Undoubtedly, the most successful method to meet this challenge has been the introduction of random-dot stimuli. Thus, in stereopsis, the introduction of random-dot stereograms (Julesz, 1960) succeeded in isolating binocular disparity from all other depth cues and demon-

---

The authors thank Akos Feher for implementing the computer programs, and Junqing Huang for developing software and administering experiments. We also acknowledge valuable suggestions by Charles Chubb, Zygmunt Pizlo, Peter Werkhoven, and an anonymous reviewer. This research was supported by Grant BNS-910938 from the National Science Foundation to T. V. P. and Grant 91-074 (DRET) to A. G. Preparation of this paper was supported by a grant from the National Science Foundation and OTKA (U.S.–Hungarian Science and Technology Joint Fund JF-360) to I.K., and a grant from the McDonnell-Pew Program in Cognitive Neuroscience to T.V.P., A.G., and B.J. Correspondence should be addressed to T. V. Papathomas, Laboratory of Vision Research, Psychology Building, Busch Campus, Rutgers, the State University of New Jersey, Piscataway, NJ 08854 (e-mail: papathom@gandalf.rutgers.edu).

strated that monocular form recognition is not a prerequisite to stereoscopic fusion. Random-dot stereograms were also used to provide behavioral (Bough, 1970) and physiological (e.g., Poggio, Gonzalez, & Krause, 1988; Poggio, Motter, Squatrito, & Trotter, 1985) evidence of stereoscopic vision in monkeys, and it allowed neurophysiologists to isolate disparity-tuned units in the visual cortex. Similarly, in motion, the random-dot cinematogram played a very important role in both psychophysical (e.g., Julesz & Payne, 1968; Anstis, 1970; Braddick, 1973, 1974) and neurophysiological (e.g., Movshon & Newsome, 1992; Orban, Gulyás, & Vogels, 1987; Schiller, Logothetis, & Charles, 1990) studies by isolating motion mechanisms in the absence of static (intraframe) cues, since the moving object is defined only by virtue of motion. Finally, in the area of spatial grouping, the introduction of random-dot Moiré patterns (Glass, 1969) started a new era in which quantitative approaches became possible (e.g., Prazdny, 1986a; Sagi & Kovács, 1993).

In the forced-choice competition methods Schmidt (1936) made the first attempt to utilize such stimuli for studying the role of form and color in apparent motion. A more recent set of methods is typified by that of Burt and Sperling (1981), who estimated the relative strength of visual attributes (such as luminance and orientation) in apparent motion by obtaining performance as a function of interframe displacement for each attribute. Barchilon (1992) took a similar approach for textural grouping. Prazdny (1984) was the first to introduce a forced-choice competition paradigm in the study of Glass patterns. Gorea and Papathomas developed a class of stimuli that allowed the relative strength of any two attributes to be compared directly for both motion and stereopsis (Papathomas & Gorea, 1988, 1989; Gorea & Papathomas, 1989). The key feature of their stimuli is that an attribute can be arranged in the spatiotemporal domain (or in the two monocular images) simultaneously with, but independently of, other attributes. In particular, two attributes can be arranged so as to elicit apparent motion (or depth) in opposite directions, enabling a direct comparison of the relative strength of any pair of attributes. Werkhoven and his colleagues (Werkhoven, Snippe, & Koenderink, 1990a, 1990b) and Nishida and Takeuchi (1990) developed a clever competitive-path paradigm, which is covered in some detail in the next section.

The importance of the forced-choice competition paradigm in all three modalities is related to the issue posed by the question: Given an element  $q$  and two competing candidate matches  $a$  and  $b$ , how does the visual system select the dominant match between the possibilities  $qa$  and  $qb$ ? To be specific, assume that elements  $a$ ,  $b$ , and  $q$  are of equal size, and have luminances  $L_a$ ,  $L_b$ , and  $L_q$ , respectively. One way of establishing which match dominates would be to determine a metric  $f(L_u, L_v)$  that satisfies the following criterion: Match  $qa$  dominates if and only if  $f(L_a, L_q) > f(L_b, L_q)$ ; match  $qb$  dominates if and only if  $f(L_b, L_q) > f(L_a, L_q)$ ; finally, an ambiguous percept is obtained if and only if  $f(L_a, L_q) = f(L_b, L_q)$ . In other words,  $f(L_u, L_v)$  represents the "motion strength" of the path defined by the luminances of elements  $u$  and  $v$ . Once the

form of  $f(L_u, L_v)$  is obtained, the next step would be to determine what type of physiologically plausible mechanisms could implement the function  $f$ . However, the very first need in pursuing a quest for  $f(L_u, L_v)$  is to design stimuli in which two competing matches  $qa$  and  $qb$  are presented to observers whose task is to report which of the two percepts dominates over the other. This paper presents a set of stimuli and techniques that address precisely this need for all three modalities by extending the two-path competitive paradigm of Werkhoven et al. (1990b) to random-dot stimuli in motion, as well as stereopsis and Glass patterns.

From the psychophysical and the computational perspectives, a central issue in these modalities is how the visual system detects these correlations in each modality. In particular, a question that has received a lot of attention is whether this detection is dictated by a similarity metric or a covariance metric. One obvious choice for a similarity metric, used by Werkhoven et al. (1990b), among others, is based on the physical similarity of targets in stimulus space. This similarity metric would predict that matching  $qa$  dominates over  $qb$  if  $q$  appears physically more similar<sup>1</sup> to  $a$  than to  $b$ . By contrast, what counts for a covariance metric is the combined "energy" of the pairs of targets. One of the most common forms used for this energy metric is the product of the individual responses elicited by each member of a target pair. If the covariance metric dictates grouping, then percept  $qa$  will win over  $qb$  if the energy of the  $qa$  pair is larger than that of  $qb$ . One way to test between the two alternatives is to set  $q = a$  in the competing paradigm above, and see whether the  $qa$  match dominates by virtue of the similarity of  $a$  and  $q$ , or whether there will be conditions for which the  $qb$  match wins when its covariance is larger than that of  $qa$ . This question has been answered in favor of the covariance, or energy, metric for first-order stimuli—that is, those that involve regions differing from the background in luminance (or color), for motion (Reichardt, 1961; Werkhoven et al., 1990b); for stereopsis (Gulick & Lawson, 1976; Julesz, 1961; Papathomas, Kovács, & Huang, 1993); and for Glass patterns<sup>2</sup> (Prazdny, 1984, 1986b). The methods proposed here address this issue very effectively, as illustrated in Experiment 1, presented in a subsequent section.

Another common property of the three modalities is that they all involve processes that can be thought of as taking place in two major stages. Thus, in motion, front-end filters process the input images and then feed their differentially delayed outputs to a higher order stage for extracting the motion signal (e.g., Reichardt, 1961). In stereopsis, the visual input to each eye is processed by monocular front-end filters, or cascade of filters, before being fed to binocular disparity-tuned units (Julesz, 1971; Julesz & Miller, 1975). Finally, in Glass patterns, there may well be two processes, an early one (front-end filter) to extract local orientation, followed by a second stage that integrates the outputs of the first one to derive the global pattern (e.g., Zucker, Stevens, & Sander, 1983). We will denote the response of the front-end stage to an element of luminance  $L$  by  $h(L)$ , where we assume that the size of the element is comparable to the filter's receptive field and that the element is

centered on the filter's receptive field. Obviously, there is a wide variety of such front-end units that are tuned to different spatial frequencies. The assumption of matched scales between elements and receptive fields is made only to obtain a simple approximation for near-optimal spatial tuning. We will use the notation  $h(L)$  for front-end filters in all the modalities (apparent motion, stereopsis, and Glass patterns). The properties of these front-end filters can be studied with the present stimuli. An example is described in Experiment 2 in a later section.

The rest of the paper is organized as follows: The next section discusses relevant approaches that motivated the development of our unified approach. Then, two sections present the new set of stimuli and illustrate their usage with representative experiments, respectively. Extensions of the basic approach and brief concluding remarks are given in the last section.

### Development of a Unified Approach

In the areas of motion and stereopsis, Gorea and Papathomas (1989; see also Papathomas & Gorea, 1988, 1989) used a forced-choice competition paradigm with a set of stimuli a typical member of which is shown in Figure 1 in  $x$ - $t$  space.

Each element is defined by the conjunction of two attributes, luminance (L) and color (C), where each attribute is allowed two values; one possibility is to have  $L_0$  and  $L_1$  represent luminances above and below the uniform luminance of the yellow background, respectively, and  $C_0$  and  $C_1$  to be red and green colors. Notice that the luminance

and color values are distributed over  $x$ - $t$  space to elicit motion to the right and to the left, respectively. Stimuli of this type were used by Papathomas, Gorea, and Julesz (1991) to show that color-based motion dominated luminance-based motion under certain conditions, indicating that chromatic signals do play a role in apparent motion. These stimuli can also be adapted for circular motion (Papathomas & Gorea, 1989), forming an extension of Green's stimuli (Green, 1989; Green & Odom, 1986) for multi-attribute motion perception. The same type of stimuli was modified and extended for textural grouping studies (Gorea & Papathomas, 1991a, 1991b; Werkhoven, Sperling, & Chubb, 1992).

Working independently, Werkhoven and his colleagues (Werkhoven et al., 1990a, 1990b) and Nishida and Takeuchi (1990) developed a set of stimuli the generic form of which is shown schematically in Figure 2.

Two competing motion paths are present: path  $p_a$ , based on matching  $aq$ , and path  $p_b$ , based on matching  $bq$ . These stimuli can be used to determine whether motion is based on matching of elements that appear similar in stimulus space (Werkhoven termed this the *similarity metric*; see note 1) or is dictated by some covariance metric. This test is not possible with the stimuli of Figure 1, as Werkhoven et al. (1990b) pointed out. To test for the two possibilities with Figure 2, element  $q$  is made identical to  $a$ , creating a homogeneous path  $p_a$  along identical elements  $aa$ , and a heterogeneous path  $p_b$  along dissimilar elements  $ba$ . The similarity metric predicts that the homogeneous path will always dominate, whereas the covariance metric predicts

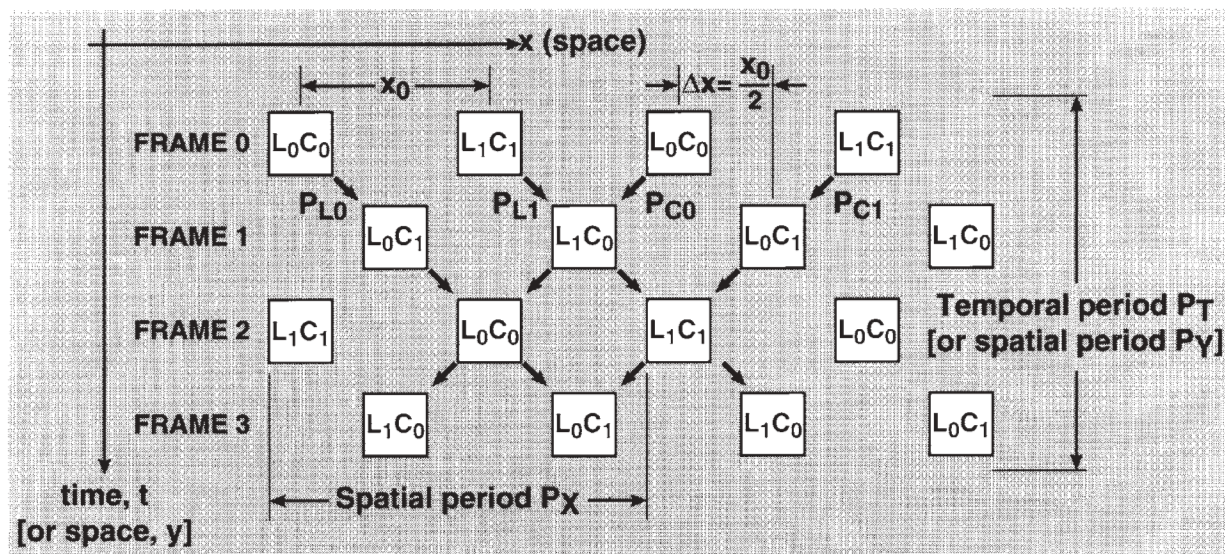


Figure 1. A typical member of the set of stimuli developed for studying the interaction of visual attributes in motion perception (Papathomas & Gorea, 1988). This schematic representation in the  $x$ - $t$  domain shows a set of frames in temporal sequence. Each element is defined by the conjunction of two attributes L and C, and each attribute can be assigned two values. For example, if L stands for luminance and C for "color," then  $L_0$  and  $L_1$  could be "bright" and "dim," respectively, and  $C_0$  and  $C_1$  could be "red" and "green," respectively. Thus, element  $L_0C_0$  stands for a bright red target,  $L_0C_1$  denotes a bright green target, and so forth. There are two types of competing paths: One type is favored by luminance correlations eliciting rightward motion (paths  $P_{L0}$  and  $P_{L1}$ ); the other type is color based and elicits motion to the left (paths  $P_{C0}$  and  $P_{C1}$ ). The same stimuli can be used for studies in textural grouping if displayed as an  $x$ - $y$  pattern (see note 4).



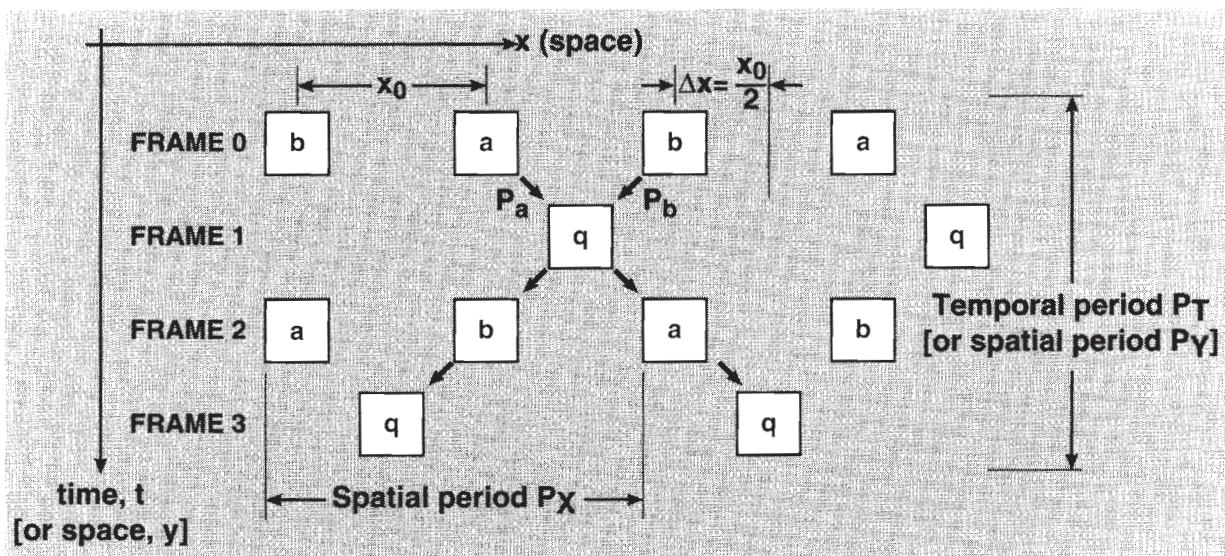


Figure 2. Schematic diagram of the stimuli developed by Werkhoven, Snippe, and Koenderink (1990b) and Nishida and Takeuchi (1990). Elements *a*, *b*, and *q* can be defined by the conjunction of visual attributes. Two competing motion paths are present: path  $p_a$ , based on matching *aq*, and path  $p_b$ , based on matching *bq*. The presence of only two motion paths, compared with four in Figure 1, is an advantage of this stimulus, because it simplifies the analysis. Another advantage is that these stimuli can be used to determine whether motion is based on matching of elements that appear similar in stimulus space or that have a higher correlation (see text for details). As with Figure 1, these stimuli can be used for studies in textural grouping (see note 4).

that the heterogeneous path will dominate when the “strength” of element *b* is larger than that of *a*. Werkhoven et al. (1990a, 1990b) and Nishida and Takeuchi (1990), working with luminance-defined elements, obtained strong evidence in favor of the covariance metric, one form of which could be the product of the “energy” of the elements. Namely, what matters in establishing a correspondence for matching two target elements with luminances  $L_u$  and  $L_v$  seems to be some form of a product  $h(L_u) \times h(L_v)$ , where  $h(\ )$  is the space-averaged response of the front-end filter to a well-aligned target (Werkhoven et al., 1990b; Werkhoven, Sperling, & Chubb, 1993; Papathomas et al., 1993). Thus, for Fourier<sup>3</sup> motion, one possible form of the function  $f(\ )$  that we introduced earlier is

$$f(L_u, L_v) = h(L_u) \times h(L_v). \quad (1)$$

Of course, the response  $h(L)$  depends on the luminance of the target, as well as on that of the background, although for simplicity the latter dependence is not made explicit in Equation 1.

The stimulus of Figure 2 can be regarded as a major extension of that shown in Figure 1. In fact, its advantages render the stimulus of Figure 1 obsolete. Indeed, it is possible to test “similarity versus covariance” only with the stimulus of Figure 2, but not with that of Figure 1. This is because Figure 1 predicts the same outcome for both similarity and covariance. Namely, if attribute *C* in Figure 1 is the same as that of the background, then the targets are discriminated from the background only by attribute *L*. The added strength of the two rightward paths is  $h^2(L_0) + h^2(L_1)$ , which is never less than the strength of the two leftward paths  $2h(L_0)h(L_1)$ . Thus covariance predicts mo-

tion to the right, just as similarity does. This argument was developed by Werkhoven et al. (1990b). Another reason why the stimulus of Figure 2 can be regarded as an extension of the stimuli of Figure 1 is that the competing paths of Figure 1 can be implemented in Figure 2 by letting  $a = L_0C_0$ ,  $b = L_1C_1$ , and  $q = L_0C_1$ ; elements of type  $L_1C_0$  of Figure 1 are no longer needed in Figure 2. An additional advantage of the stimulus of Figure 2 is that it contains only one type of path to the left,  $p_b$ , and only one path to the right,  $p_a$ . This makes the mathematical analysis much simpler to deal with.

Studies that examined the role of attributes in textural grouping and Glass patterns for determining the dominant percept between two opposing alternatives followed similar studies in apparent motion much later. Thus, Zucker et al. (1983) investigated the relation between proximity and brightness similarity in textural grouping and Glass patterns, presenting evidence that percepts in both modalities are the result of the same kind of (or even the very same) underlying mechanisms and processes. Prazdny (1984, 1986a, 1986b) introduced a forced-choice competition paradigm in Glass patterns with stimuli in which there are two orthogonal groupings. The main feature of the stimuli is that there are two orthogonal transformations that give rise to two dots, *a* and *b*, for every random dot *q* in the original pattern. Both of these are equally likely matches. Thus, there are three sets of dots in such a stimulus: The set *Q* of dots in the original pattern, and the sets *A* and *B* of points defined by the two orthogonal transformations. The luminance of all the dots in a given set is the same throughout the image. By selecting appropriate luminances for *a*, *b*, and *q*, Prazdny (1984, 1986a, 1986b) obtained ev-

idence that the “energy” of the pairing determines which pair (qa or qb) dominates. His definition of energy is some function of the size and the contrast magnitude of each dot (Prazdny, 1986a). This is true as long as the dots in the pair have the same luminance polarity (i.e., sign of contrast). Glass and Perez (1973) reported that no global percept is perceived for dot pairs of opposite polarity. Nevertheless, when the dot density is very high (Anstis, 1970; Kovács & Julesz, 1992) or when the contrast is extremely low (Papathomas, Ramanujan, & Gorea, 1994), reverse-polarity Glass patterns and textural groupings result in a global pattern that is orthogonal to the veridical grouping that would be expected from the physical transformation of the dots.

In the modality of stereopsis, there is evidence that correspondence matching is also governed by a covariance principle for luminance-defined stimuli (Cormack, Stevenson, & Schor, 1991; Gulick & Lawson, 1976; Julesz, 1960; Papathomas et al., 1993). In fact, this covariance metric could be modeled as a product of front-end responses, just as in the case of motion (see Equation 1), provided that both targets have the same luminance polarity (Cormack et al., 1991; Julesz, 1963). The stimuli proposed in this paper make it possible to test this hypothesis directly with modified random-dot stereograms.

### Stimuli for a Unified Approach

The new techniques introduced in this paper incorporate a forced-choice competition paradigm into random-dot stimuli that have traditionally been used in these modalities. This competition paradigm has been extended to random-dot cinematograms in the case of apparent motion, to random-dot stereograms for stereopsis, and to modified random-dot Glass patterns for static-flow patterns. These are discussed separately in the following subsections.<sup>4</sup>

**Stimuli for apparent motion.** Figure 3a shows schematically the two-frame sequence of the new random-dot cinematogram (RDC) stimuli for motion studies. As in the stimuli of Figure 2, there are three types of elements, denoted by the letters a, b, and q. The RDC is composed of element triplets {q, a, b} of which q is shown in one frame (0 or 1), while elements a and b are shown in the other frame (1 or 0).

Let us first describe frame 0: Elements q occur isolated, whereas elements a and b are always displayed side by side, separated by a center-to-center distance  $2\Delta x$ . The average distance between elements q and ab pairs is much larger than  $\Delta x$ . Frame 1 is generated from frame 0 as follows: Each ab pair of frame 0 is replaced in frame 1 by an element q, placed in the location indicated by the solid dot; this solid dot, exactly in the middle of a and b, is not displayed in the frame, but it is used in the illustration only for notational convenience. Also, the location of each q in frame 0 (denoted by a solid dot in frame 1) is spanned on either side by elements a and b in frame 1, each placed a distance  $\Delta x$  from the small solid dot. Notice that a is to the left of b in frame 0, whereas the converse is true in frame 1. There are two competing motion paths in this two-frame stimulus. One is defined by aq matches and another by bq matches, generating motion to the right and to the left, re-

spectively. One advantage of the stimulus of Figure 3a is that it eliminates the secondary rightward (bb) and leftward (aa) paths that are formed by targets in frames 0 and 2 of Figure 2. These secondary paths contribute undesirable motion signals outside the temporal range of interest.

This RDC stimulus enables one to try experiments that are completely analogous to those conducted with the periodic stimulus of Figure 2, since the same three types of elements are used in both, defining the same competing pair-wise matches. In fact, Werkhoven et al. (1992, 1993) applied such stimuli to pattern orientation (1992) and Fourier/non-Fourier motion (1993). Such experiments, as well as extensions of previous ones, are presented in the next section.

**Stimuli for stereopsis.** The stimuli here are generated by ambiguous random-element stereograms (RESs), the generic schematic of which is shown in Figure 3b. The RES is composed of element triplets {q, a, b} of which q is shown to one eye, while a and b are shown to the other eye. The four possible types of triplets are shown in Figure 3b for illustration purposes. The actual stimuli are composed of a large number of such triplets. The distribution of qs and (a,b) pairs is random within each eye's image, in approximately equal proportions, to maintain spatial uniformity. Elements a and b are always displayed side by side, separated by a center-to-center distance  $2\Delta x$ . As with the motion stimuli above, the solid dot shown exactly in the middle of each ab pair is not displayed in the actual image, but it is used in the illustration only for notational convenience: It marks the location occupied by the q element of the triplet {q, a, b} in the other eye's image. The average distance between elements q and ab pairs is much larger than  $\Delta x$ . There are two competing matches for q, that is, aq and bq, that have opposite disparities of equal magnitude  $\Delta x$ . Matchings aq have disparities that cause the top half of Figure 3b to appear in front of the bottom half; of course, bq matches have the converse disparities. Braddick showed that stereograms like the ones on the top (or bottom) half of Figure 3b elicit the percept of two fronto-parallel depth planes if a and b are comparable to q (cited in Marr, 1980, p. 141). The new element here is to have both halves adjacent to each other and to vary the “strengths” of a and b in order to produce a global horizontal depth edge. We have thus created stereo stimuli that are completely analogous to those used in motion, where the two competing matches had opposite velocities. This allows us to design experiments that have direct correspondence with those in motion.

**Stimuli for Glass patterns.** The stimuli have the general form of those in Figure 3c. Instead of element pairs, as in classical Glass patterns (Glass, 1969), triplets {q, a, b} are used, which form two competing patterns: a horizontal one, due to matching aq, and a vertical one, due to bq (radial and tangential patterns can also be used). Prazdny (1984) used similar competing Glass patterns. The extension here is that all four possible arrangements of elements a, b, and q (shown in Figure 3c) are scattered randomly throughout the image in equal proportions. The arrangements are formed by the  $2 \times 2$  combinations of placing



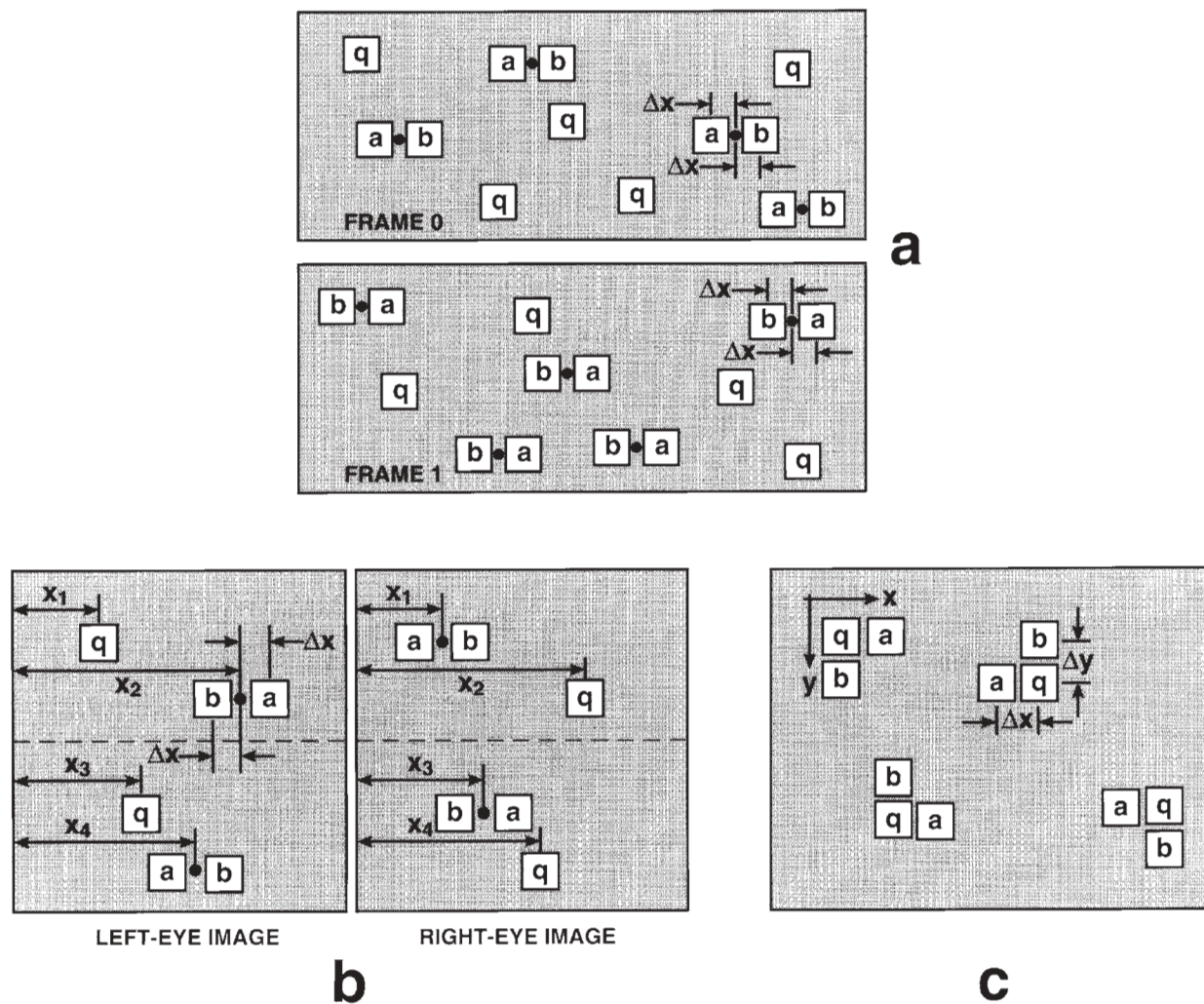


Figure 3. Schematic diagrams for the random-dot stimuli that were adapted for a forced-choice competition paradigm. The notation  $(u, v)$  is used for a pair of elements  $u$  and  $v$  that are aligned horizontally and are  $2\Delta x$  units apart along the horizontal direction, with  $u$  displayed to the left of  $v$ . (a) A two-frame sequence of a random-dot cinematogram. Each element  $q$  in frame 0 is replaced, in register, by a  $(b, a)$  pair in frame 1. Conversely, each  $(a, b)$  pair in frame 0 is replaced, in register, by an element  $q$  in frame 1. Matchings  $aq$  and  $bq$  elicit motion to the right and to the left, respectively, if frame 1 follows frame 0 in temporal sequence. (b) Typical triplets  $\{q, a, b\}$  for the construction of a random-element stereogram. Element  $q$  is shown to one eye, while  $a$  and  $b$  are shown to the other eye. The actual stimuli are composed of a large number of such triplets. The disparities of  $aq$  matches cause the top half to appear in front of the bottom half in depth; of course,  $bq$  matches have the opposite disparities. (c) The four possible arrangements of  $\{q, a, b\}$  triplets that generate modified Glass patterns with competing global percepts: horizontal for matches  $aq$ , and vertical for matches  $bq$ . Such triplets are arranged randomly in space to produce the global pattern.

element  $a$  to the left or to the right of  $q$ , and  $b$  above or below  $q$ . The presence of four different patterns rather than a single one (as used by Prazdny, 1984), such as having element  $a$  to the left of  $q$  and  $b$  above  $q$ , should limit the probability that the observer reports the local patterns rather than the desired global pattern.

#### Using the Stimuli

In this section we report on two related experiments, each conducted for the three modalities of motion, stereopsis, and Glass patterns, which illustrate typical uses of the new set of stimuli. In both experiments, one of two opposite arrangements was shown in each trial, selected ran-

domly across trials. One arrangement is shown in Figures 3a, 3b, and 3c, for motion, stereopsis, and Glass patterns, respectively. This arrangement elicits the following percepts, if the  $aq$  matching dominates  $bq$ : motion to the right for Figure 3a, a depth edge with the top half in front of the bottom half for the RES of Figure 3b, and a horizontal global pattern for Figure 3c; naturally, this same arrangement produces the competing percept if  $bq$  dominates over  $aq$ . The opposite arrangement for each modality is obtained when the positions of elements  $a$  and  $b$  in each  $\{q, a, b\}$  triplet are interchanged. These opposite arrangements produce opposite percepts in the three modalities. Thus, if match  $aq$  is stronger, the percepts are as follows: motion

to the left for Figure 3a, a depth edge with the top half behind the bottom half for the RES of Figure 3b, and a vertical global pattern for Figure 3c. Stimuli were displayed tachistoscopically to limit the role of eye movements, and the observer reported which of the two competing percepts was obtained in a two-alternative forced-choice paradigm. The observer responded about the direction of motion, the depth polarity of the global stereo edge, or the global Glass pattern. We recorded the percentage of responses that favored the aq matches. Each point on the graphs of Figures 4 and 5 represents the average of at least 100 trials.

Before presenting details on the two separate experiments, we discuss their common points. Stimuli were generated on Silicon Graphics IRIS raster display systems. Accurate timing is ensured within 16.667 msec because the IRIS provides access to the vertical blanking signal at a video rate of 60 Hz. Five observers participated, two of whom were naive as to the purpose of the experiments. Each

had normal, or corrected to normal, vision. The results of both experiments were very similar across observers in all three modalities; thus we will present typical data from individual observers. One observer (J.H.) does not possess normal stereopsis, so we have no such results with this observer. The background was uniform gray with a luminance  $L_G$ , and it subtended a square field with a side of  $13.37^\circ$  from a viewing distance of 1.00 m. Elements a, b, and q were also gray of uniform luminance, the value of which was varied, depending on the experiment. They were square in shape, with their side subtending  $5.45'$  of arc. The stimulus duration was 166.7 msec in all the experiments. The interframe displacement  $\Delta x$  of Figure 3a was  $7.26'$  min of arc in all the motion experiments, as was the disparity  $\Delta x$  of Figure 3b in all the stereopsis experiments. The displacements  $\Delta x$  and  $\Delta y$  of Figure 3c were equal to each other, and they also subtended  $7.26'$  in all the Glass pattern experiments. This is an example of using the

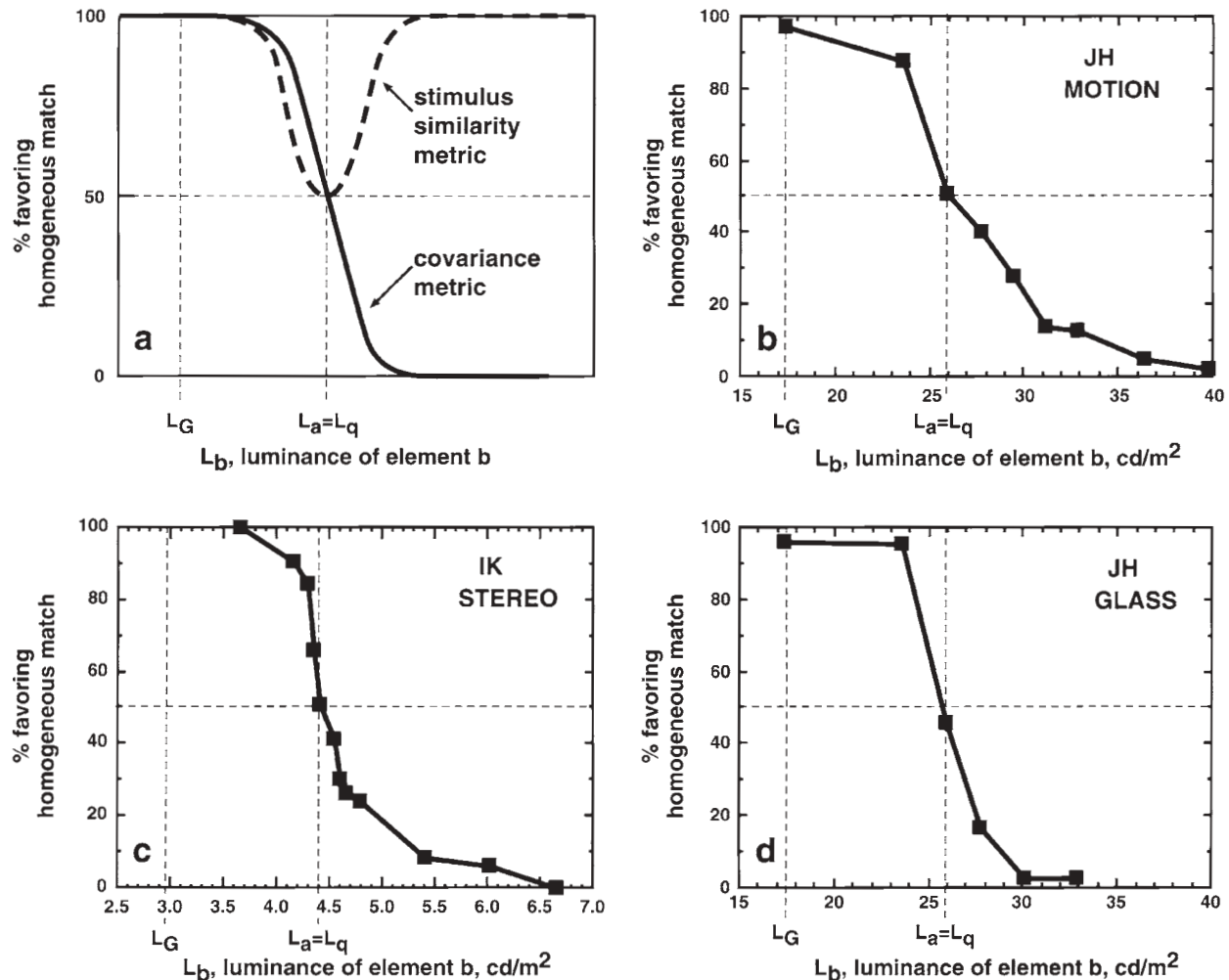


Figure 4. Predictions and results for Experiment 1 with the stimuli of Figures 3a, 3b, and 3c.  $L_G$ ,  $L_a$ ,  $L_b$ , and  $L_q$  are the luminances of the background, and elements a, b, and q, respectively. Vertical dashed lines indicate the values of  $L_G$ ,  $L_a$ , and  $L_q$  on the horizontal axis for reference. Percentages favoring the aq match are plotted as a function of  $L_b$ . (a) Qualitative predictions for the similarity and covariance metrics. (b) Results with the motion stimuli of Figure 3a for Observer J.H. (c) Results with the stereo stimuli of Figure 3b for Observer I.K. (d) Results with the modified Glass patterns of Figure 3c for Observer J.H.

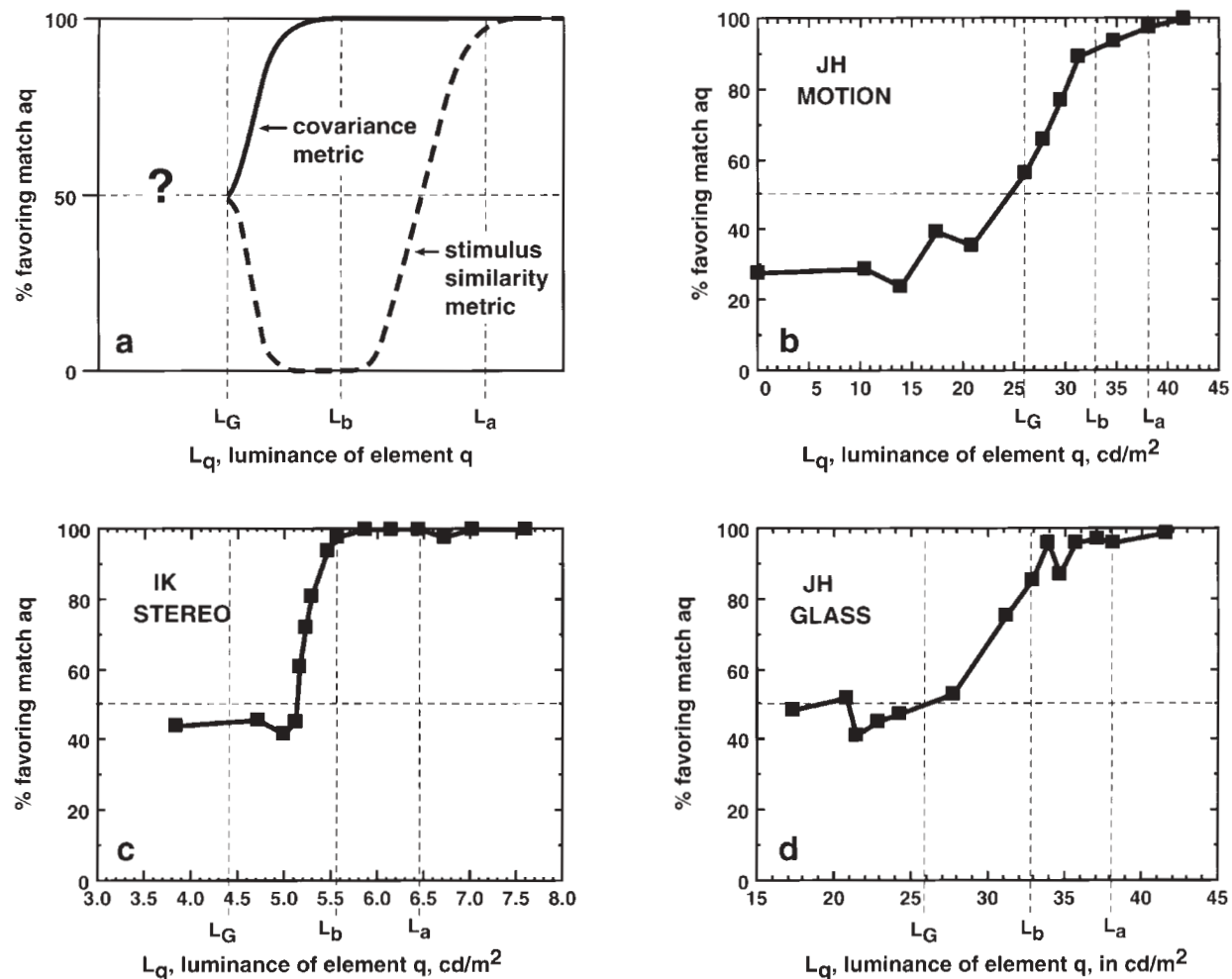


Figure 5. Predictions and results for Experiment 2 with the stimuli of Figures 3a, 3b, and 3c, with the same notation as in Figure 4. Vertical dashed lines indicate the values of  $L_G$ ,  $L_a$ , and  $L_b$  on the horizontal axis for reference. Percentages favoring the aq match are plotted as a function of  $L_q$ . (a) Qualitative predictions for the similarity and covariance metrics. (b) Results with the motion stimuli of Figure 3a for Observer J.H. (c) Results with the stereo stimuli of Figure 3b for Observer I.K. (d) Results with the modified Glass patterns of Figure 3c for Observer J.H.

same spatiotemporal parameters for all the modalities. The average distance between elements q and ab pairs (within a motion frame in apparent motion, or within one eye's image in stereopsis) was 34.5' of arc. The average distance between {q, a, b} triplets for Glass patterns was also 34.5'.

### EXPERIMENT 1

In this first group of experiments we utilized the random-dot stimuli of Figure 3a to repeat Werkhoven et al.'s (1990b) motion study, which was carried out with the periodic stimuli of Figure 2. This can be seen as a step for "calibrating" the stimuli, just as was done with the Glass patterns stimuli of Figure 3c. These results were expanded in Experiment 2. In addition, we applied the RES stimulus of Figure 3b to address the same issue that was studied in apparent motion and Glass patterns. The issue is whether

correspondence matching is dictated by covariance or similarity in stimulus space (see note 1). To investigate this, we made elements a and q identical by fixing their luminance at 25.95 cd/m<sup>2</sup>, and we varied the luminance of b, keeping the background luminance at  $L_G = 17.3$  cd/m<sup>2</sup>. The similarity metric predicts that the homogeneous path aq = aa will always dominate over bq (except when  $b = a = q$ ), whereas the covariance metric predicts that the heterogeneous path will dominate when the "strength" of element b is larger than that of a. As a result, this stimulus gives two distinctly different predictions. These qualitative predictions are shown in Figure 4a in the form of percentages of responses favoring the aq path as a function of the luminance  $L_b$  of the b element. The values of  $L_G$ ,  $L_a$ , and  $L_q$  are marked off on the horizontal axis, and vertical dashed lines are added for reference. The similarity and covariance metrics' predictions are shown in bold dashed and solid lines, respectively, in Figure 4a. The results, presented in



the subsections below, give strong evidence that a covariance metric determines the dominant percept in all the modalities.

### Apparent Motion

Two frames of the type shown in Figure 3a were displayed in rapid temporal sequence. Each frame was displayed for 166.7 msec, without an interstimulus interval (ISI). The results of the first experiment with the stimulus of Figure 3a are shown in Figure 4b, and are clearly in favor of a covariance metric, in agreement with Werkhoven et al. (1990b).

### Stereopsis

A pair of RES images of the type shown in Figure 3b was displayed for 166.7 msec. The IRIS workstations are equipped with Stereographics peripherals, which alternate left- and right-eye images at a rate of 120 Hz; these are separated for the two eyes by liquid-crystal display (LCD) glasses, which are synchronized to the display via an infrared signal. These glasses attenuate the luminance by a factor of 5.89, as seen in the abscissa values of Figure 4. The results of the first experiment with the stimulus of Figure 3b are shown in Figure 4c, indicating that a covariance metric dictates the percept.

### Glass Patterns

A frame composed of a large number of the four types of triplets shown in Figure 3c was displayed for 166.7 msec. As expected from similar studies (see, e.g., Prazdny, 1984), the results of this experiment, shown in Figure 4d, also offer clear evidence in favor of a covariance metric.

## EXPERIMENT 2

The scope of Experiment 1 is extended in this experiment, in which we remove the restriction  $a = q$  and consider what happens when we attempt to match elements of opposite luminance polarity. To this end, we fixed the luminance of elements  $a$  and  $b$  in each triplet  $\{q, a, b\}$  of Figure 3a, 3b, and 3c, so that  $L_a = 38.06 \text{ cd/m}^2 > L_b = 32.87 \text{ cd/m}^2 > L_G = 25.95 \text{ cd/m}^2$ . This time, however, it was the luminance  $L_q$  of the  $q$  element that was the independent variable, assuming values above as well as below the background luminance  $L_G$ . All the other conditions were exactly the same as those in Experiment 1. Again, we plotted the observer's preference for the  $aq$  path as a function of  $L_q$ . Qualitative predictions are shown in Figure 5a in the form of percentages of responses favoring the  $aq$  path as a function of the luminance  $L_q$  of the  $q$  element. Vertical dashed lines are drawn at the values of  $L_G$ ,  $L_a$ , and  $L_b$  on the horizontal axis for reference. In view of the rejection of the notion of physical similarity governing the percept in Experiment 1, we should not even consider its predictions, but we do so for completeness. The similarity metric predicts that the path  $aq$  will dominate over  $bq$  when element  $q$  looks "more similar" to element  $a$  than to  $b$ . However, note that the similarity is now much more difficult to define than it was in Experiment 1. One may say

that  $q$  will be more similar to  $a$  than to  $b$  when  $L_q \geq L_a$ , since  $L_a > L_b$ ; also,  $q$  may appear more similar to  $a$  when  $L_q$  is between  $L_a$  and  $L_b$ , but "closer to"  $L_a$ , even though we do not know how to measure "closeness." Along the same lines, path  $aq$  will be dominated by  $bq$  when  $L_G < L_q \leq L_b$ , and also when  $L_q$  is between  $L_a$  and  $L_b$ , but "closer to"  $L_b$ . The similarity metric's predictions are shown in dashed lines in Figure 5a, but such predictions for the case where  $L_q < L_G$  are even more difficult to speculate on: How does one judge physical "similarity" between elements of opposite luminance polarity? This is the reason for the question mark in Figure 5a.

On the other hand, the covariance metric predicts that the  $aq$  path will dominate as long as  $L_q > L_G$ , because the "energy" of pair  $aq$  is larger than that of  $bq$ . This prediction is shown with a solid line in Figure 5a. Let us assume, for the sake of simplicity, that the outputs of the front-end filters are multiplied together by a higher order correlator. What would be the prediction of the covariance metric for the condition  $0 < L_q < L_G$ ? Given that now both the  $aq$  and the  $bq$  pairs are composed of elements of opposite polarity, there can be at least three possibilities: (1) The front-end filters that feed their outputs to the higher order covariance detectors contain both *on* and *off* pathways that remain separate; that is, there are two classes of higher order detectors, those that are fed exclusively by *on* units, and those fed exclusively by *off* units. In this case, neither pair ( $aq$  or  $bq$ ) will elicit a percept and we will get an ambiguous performance at 50%. (2) The front-end filters are linear, in which case they will produce opposite-signed responses  $h(L_a)$  and  $h(L_q)$  for the opposite-polarity pair  $aq$ ; ditto for  $bq$ . As a result, the products  $h(L_a) \times h(L_q)$  and  $h(L_b) \times h(L_q)$  will both be negative, but  $h(L_b) \times h(L_q)$  will dominate, because it is algebraically larger than  $h(L_a) \times h(L_q)$ , since  $L_b < L_a$ . Thus, performances will favor the  $bq$  match, and we expect the curve to drop below the 50% level. (3) The front-end filters rectify the signal; that is, they give the same response for targets of opposite polarity, provided they have the same magnitude of contrast. In this case,  $h(L_a) \times h(L_q)$  and  $h(L_b) \times h(L_q)$  will both be positive, and  $h(L_a) \times h(L_q)$  will dominate because  $L_a > L_b$ . This will cause the curve to stay above the 50% ambiguous level. The results, presented in the subsections below, provide evidence for the form of the front-end filter in the various modalities.

### Apparent Motion

The results of Experiment 2 with the stimulus of Figure 3a are shown in Figure 5b and agree with those predicted by a covariance metric for  $L_q \geq L_G$ . What is of most interest is the performance for  $L_q < L_G$ , which drops below 50%, indicating that the  $bq$  path dominates  $aq$ . In turn, this is strong evidence for the presence of linear front-end filters, as discussed in Item 2 above.<sup>5</sup> These linear mechanisms are also thought to be responsible for producing the well-known reverse-phi phenomenon in motion, in which a target that moves as it alternates its contrast polarity in every frame appears to move in the reverse direction (Anstis, 1970; Adelson & Bergen, 1985; Sperling, 1989).<sup>6</sup>

### Stereopsis

As expected, the results of Experiment 2, shown in Figure 5c, offer additional evidence that a covariance metric dictates the percept for  $L_q \geq L_G$ . As in the case of apparent motion above, it is the range  $0 < L_q < L_G$  that was targeted to provide clues about the properties of the front-end monocular filters. The fact that performance stays at the 50% ambiguous level is evidence that we have the first of the three possibilities above—that is, separate, noninteracting *on* and *off* pathways for stereopsis. There is additional recent psychophysical (see, e.g., Harris & Parker, 1993) and physiological (Ohzawa, De Angelis, & Freeman, 1990) evidence supporting this hypothesis. This is a restatement of the reason why reverse-polarity random-dot stereograms cannot be fused (Julesz, 1963) (a reverse-polarity random-dot stereogram is a conventional bilevel random-dot stereogram, but with the dot contrasts in one eye's image complemented).

### Glass Patterns

Since the covariance metric governs the percept in this modality, the results of Experiment 2, shown in Figure 5d, also follow the pattern of those in motion and stereopsis, above. In the particular range of interest for this experiment, that is,  $0 < L_q < L_G$ , performance stays at 50%, indicating separate *on* and *off* channels, just as in stereopsis.

## DISCUSSION/EXTENSIONS

A few words are necessary for explaining our rationale for selecting the competing percepts for the three modalities. Basically, they were chosen on the basis of well-documented psychophysical evidence for opponent mechanisms in each modality. Thus, for motion, the paths are in opposite directions because motion mechanisms seem to operate in opponency along  $180^\circ$  paths, as evidenced by motion aftereffects (the well-known waterfall illusion) as well as by the reverse-phi phenomenon (Anstis, 1970). In stereo, there is some evidence for near-far opponency, hence our choice of near-far competing percepts. For Glass patterns, we selected orthogonal competing groupings because reverse-polarity stimuli elicit a percept that is orthogonal to that produced by same-polarity stimuli (Kovács & Julesz, 1992).

The fact that the covariance metric is more suitable than the similarity metric is predicted by a wide class of motion models that respect the motion-from-Fourier-components principle (Adelson & Bergen, 1985; Reichardt, 1961; van Santen & Sperling, 1985; Watson & Ahumada, 1985). In particular, the multiplicative form of the covariance metric is explicit in Reichardt's covariance model, which extracts Fourier motion—that is, motion that contains energy in preferred directions in the spatiotemporal frequency domain. It achieves this by multiplying the delayed response of a spatial filter at one location with the undelayed response of a similar filter at a nearby location.

In the motion experiments reported here, the duration of each frame (166.7 msec) is rather long to favor Fourier motion. Non-Fourier mechanisms are certainly activated at

the low temporal frequency corresponding to this duration. Thus the final percept is the result of both Fourier and non-Fourier mechanisms. Appropriate choices of spatiotemporal conditions can examine the relative strengths of these two types of units (Papathomas, Gorea, & Chubb, 1994).

A question closely related to the issue of covariance versus similarity is whether correspondence is detected by early mechanisms or by an attention-based system that tracks similar features (Cavanagh, 1992). Again, the evidence is overwhelmingly in favor of early mechanisms for first-order stimuli. Whether this is also the case for second-order stimuli is currently the subject of very active psychophysical research, especially in motion (see, e.g., Cavanagh, 1994). Second-order stimuli involve regions that do not differ from the background in first-order statistics regarding luminance or color, that is, they are characterized by the same frequency of distribution of luminance (or color) values as the background; they differ from it in properties that are derived from these primary attributes, such as texture, contrast, flicker rate, and so on (Chubb & Sperling, 1989). Werkhoven et al. (1993) obtained evidence for a covariance metric in contrast-driven second-order motion.

So far, we have limited the scope of the stimuli and the techniques to experiments in which the elements are identical among themselves in all respects except for their luminance. Naturally, the stimuli of Figures 3a, 3b, and 3c can be extended to experiments with second-order attributes, as well as to studies in which the goal is to examine the interaction of visual attributes (luminance, contrast polarity, flicker rate, color, orientation, etc.) in the perception of motion, stereopsis, and Glass patterns. This is achieved by defining each element of the triplet  $\{q, a, b\}$  by the conjunction of these attributes, in a manner similar to our earlier class of stimuli (Gorea & Papathomas, 1991a; Papathomas & Gorea, 1988). We have already applied similar techniques using modified versions of the periodic stimuli of Figure 2 to examine the structure of the luminance and the chromatic motion pathways (Gorea, Papathomas, & Kovács, 1993a, 1993b). These techniques can easily be extended and applied to the random-element stimuli presented in this paper to study the mechanisms and pathways underlying the perception of depth and static-flow patterns.

We conclude with two specific extensions of the stimuli in the domains of motion and stereopsis.

### Apparent Motion

One disadvantage of the stimulus of Figure 3a, as compared to that of Figure 2, is that motion is elicited by two frames only, which is known to exhibit significant differences from multiframe stimuli in perceptual properties (Sperling, 1976). Both types of stimuli, that of Figure 3a and that of Figure 2 in its multiframe version, share another disadvantage. Motion to the right is based on two types of spatiotemporal matching:  $a \rightarrow q$  and  $q \rightarrow a$ , where  $u \rightarrow v$  denotes that element  $u$  in frame  $k$  is displaced to the right and replaced by element  $v$  in frame  $k + 1$ ; similarly,  $u \leftarrow v$  denotes motion to the left. It is quite possible that the motion strength of matching  $a \rightarrow q$  is different from that of



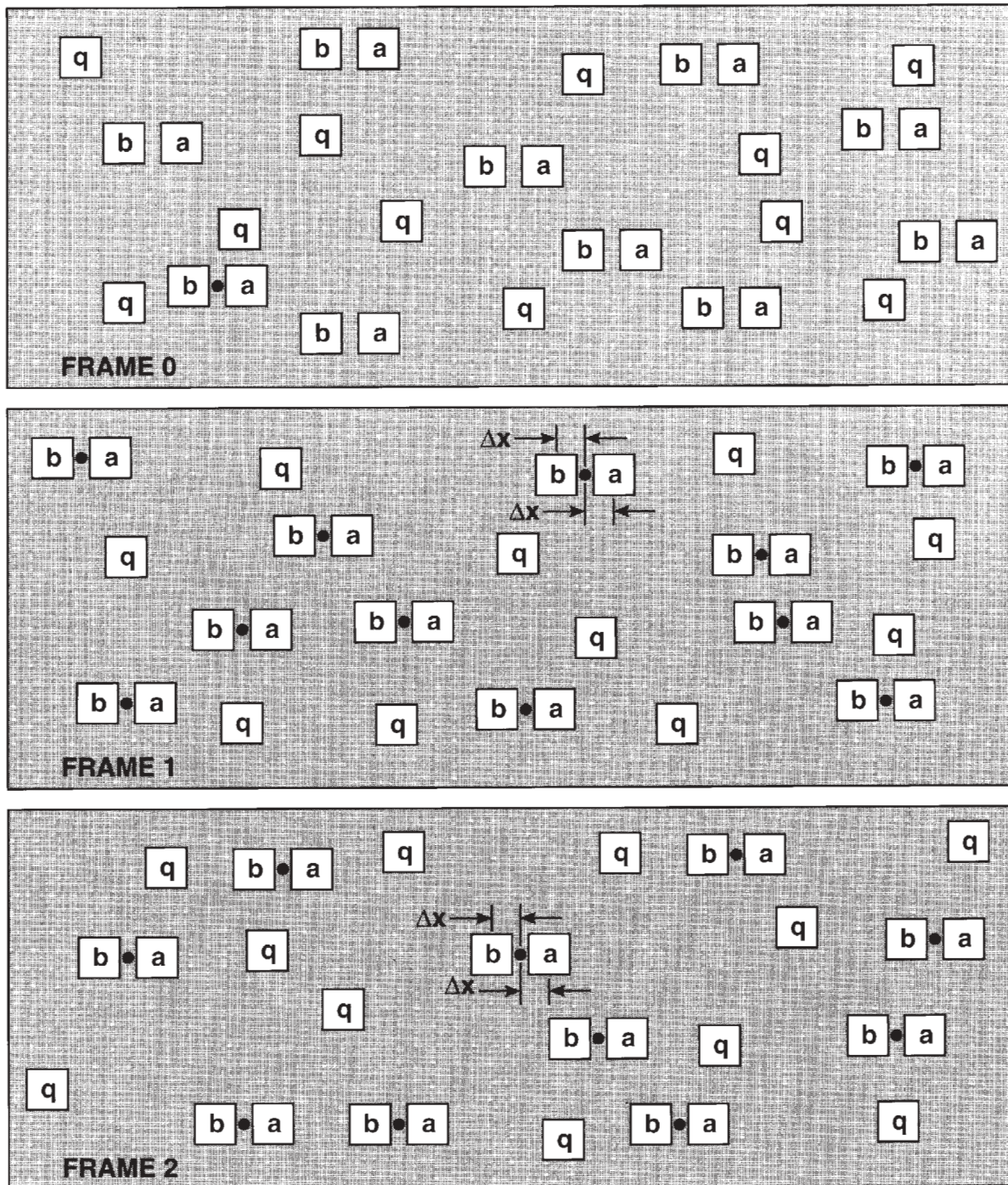


Figure 6. Schematic diagram showing three frames of a multiframe sequence of a random-dot cinematogram. Frame 0 is shown at  $t=0$ , and subsequent frames are shown at regular time intervals,  $\Delta t$  sec apart. Element  $q$  of frame  $k$  is always followed by a  $(b, a)$  pair in register in frame  $k+1$ . Pairs  $(b, a)$  of frame  $k$  die out in frame  $k+1$ , and elements  $q$  are born to replace them in random locations in frame  $k+1$ . The advantage over Figure 3a, in addition to the presence of multiple frames, is that the stimuli of this figure enable one to study the effect of the temporal sequence of the targets on the strength of the motion path (see text for details).



$q \rightarrow a$  (of course, in such a case the function  $f(L_u, L_v)$  that we referred to earlier is not well defined). The only remedy for the stimuli of Figure 2 to isolate paths  $q \rightarrow a$  and  $b \leftarrow q$  is to confine the apparent motion sequences to two frames. The advantage of the stimulus of Figure 3a is that it can be modified to isolate paths  $q \rightarrow a$  and  $b \leftarrow q$  as shown in Figure 6, which displays the first three frames of a sequence that can consist of an arbitrarily large number of frames. As in Figure 3a, the three types of elements are denoted by the letters a, b, and q, but elements a and b always occur in pairs, separated by a center-to-center distance  $2\Delta x$ . Frame 0 contains elements q and ab pairs in equal proportions, interspersed randomly throughout the image. Frame 1 is generated from frame 0 as follows: Each q in frame 0, denoted by a small solid dot in frame 1, is now replaced by an ab pair, where element a is always to the right of b, and each is placed a distance  $\Delta x$  from the small solid dot (the dot is not displayed in the frame, but it is used only for denoting the location of q in the previous frame). All ab pairs of frame 0 are erased in frame 1 and an equal number of elements q are now generated in random locations. Frame  $k+1$  is generated from frame  $k$  in exactly the same manner as frame 1 was obtained from frame 0. Thus, elements q of frame  $k$  are replaced by pairs ab in register in frame  $k+1$ ; pairs ab of frame  $k$  die out, and new elements q are born in frame  $k+1$ . Since these new elements are placed in random locations, they do not contribute to a consistent motion direction. The only consistent directions are generated by competing matchings,  $q \rightarrow a$  to the right and  $b \leftarrow q$  to the left, as desired. This is akin to stimuli for studying the kinetic-depth effect, in which dots have a limited lifetime. They appear in one frame and they are displayed in a small number of subsequent frames. When they die out, they are replaced by other dots in random locations (Doshier, Landy, & Sperling, 1989).

### Stereopsis

In studying the form of the covariance function  $f(L_u, L_v)$ , there is an analogy to the question of which of two target elements appeared temporally first or second in apparent motion. In stereopsis, it is the question of which of two targets was displayed to the dominant eye. In apparent motion, the stimulus of the previous paragraph allowed only the matches  $q \rightarrow a$  and  $b \leftarrow q$  and eliminated matches  $a \rightarrow q$  and  $q \leftarrow b$ . In a manner analogous to motion, there may be cases where one is interested in investigating what governs correspondence matching between an element q in the dominant eye and two competing matches a and b in the other eye. A slight modification is needed in Figure 3b for generating such a stimulus: Simply keep all the triplets in which q appears in the dominant eye, and eliminate all the rest.

### SUMMARY

This paper suggests a unified approach to study several perceptual phenomena in the modalities of motion, stereopsis, and Glass patterns. It is designed so that the results from applying this approach to each modality can be di-

rectly interpreted as supporting (or contradicting) a given theory. This approach can thus be regarded as "theory oriented" (Pizlo, personal communication, 1994). It should be viewed as complementary to, rather than as a substitute for, conventional "result oriented" approaches. These approaches are based on using several different methods to examine a single perceptual phenomenon, to exclude the possibility that this phenomenon is an artefact of a particular method.

The proposed approach is an extension of the methods and stimuli of Werkhoven et al. (1990b) and Nishida and Takeuchi (1990), which is achieved by the introduction of random-dot patterns in their stimuli. The extension results in two main advantages: First, in the motion domain, it enables the isolation of paths  $a \rightarrow q$  and  $q \rightarrow a$ , if we use the notation of the previous section. Second, it allows the application of their methods in the stereopsis domain, which, combined with the applicability to Glass patterns, makes possible a unified approach across all three modalities.

Two experiments, each applied in the three modalities, have been presented to illustrate typical uses of the proposed approach. The results from the first experiment verify a basic similarity among the three modalities, in that they are all governed by the same principle of covariance. The role of polarity was studied in the second experiment, and the results suggest that there are important differences in the front-end mechanisms and processes across the modalities.

### REFERENCES

- ADELSON, E. H., & BERGEN, J. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, **2**, 284-299.
- ANSTIS, S. M. (1970). Phi movement as a subtraction process. *Vision Research*, **10**, 1411-1430.
- BARCHILON, B. M. (1992). *The role of similarity, proximity and attention in visual grouping*. Unpublished doctoral dissertation, Weizmann Institute of Science, Rehovot, Israel.
- BOUGH, E. W. (1970). Stereoscopic vision in the macaque monkey: A behavioural demonstration. *Nature*, **225**, 42-43.
- BRADDICK, O. (1973). The masking of apparent motion in random-dot patterns. *Vision Research*, **13**, 355-359.
- BRADDICK, O. (1974). A short-range process in apparent motion. *Vision Research*, **14**, 103-129.
- BURT, P., & SPERLING, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review*, **88**, 171-195.
- CAVANAGH, P. (1992). Attention-based motion perception. *Science*, **257**, 1563-1565.
- CAVANAGH, P. (1995). Is there low-level motion processing for non-luminance-based stimuli? In T. V. Papathomas et al. (Ed.), *Early vision and beyond* (pp. 113-119). Cambridge, MA: MIT Press.
- CHUBB, C., & SPERLING, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A*, **5**, 1986-2006.
- CHUBB, C., & SPERLING, G. (1989). Second-order motion perception: Space-time separable mechanisms. In *Proceedings: Workshop on visual motion* (pp. 126-138). Washington, DC: IEEE Computer Society Press.
- CORMACK, L. K., STEVENSON, S. B., & SCHOR, C. M. (1991). Interocular correlation, luminance contrast and cyclopean processing. *Vision Research*, **31**, 2195-2207.
- DOSHER, B. A., LANDY, M. S., & SPERLING, G. (1989). Kinetic depth effect and optic flow: I. 3D shape from Fourier motion. *Vision Research*, **29**, 1789-1813.

- GLASS, L. (1969). Moiré effect from random dots. *Nature*, **223**, 578-580.
- GLASS, L., & PEREZ, R. (1973). Perception of random-dot interference patterns. *Nature*, **246**, 360-362.
- GOREA, A., & PAPATHOMAS, T. V. (1989). Motion processing by chromatic and achromatic pathways. *Journal of the Optical Society of America A*, **6**, 590-602.
- GOREA, A., & PAPATHOMAS, T. V. (1991a). Extending a class of motion stimuli to study multiattribute texture perception. *Behavioral Research Methods, Instruments, & Computers*, **23**, 5-8.
- GOREA, A., & PAPATHOMAS, T. V. (1991b). Texture segregation by chromatic and achromatic visual pathways: An analogy with motion processing. *Journal of the Optical Society of America A*, **8**, 386-393.
- GOREA, A., PAPATHOMAS, T. V., & KOVÁCS, I. (1993a). Motion perception with spatiotemporally matched chromatic and achromatic information reveals a 'slow' and a 'fast' motion system. *Vision Research*, **33**, 2515-2534.
- GOREA, A., PAPATHOMAS, T. V., & KOVÁCS, I. (1993b). Two motion systems with common and separate pathways for color and luminance. *Proceedings of the National Academy of Sciences*, **90**, 11197-11201.
- GREEN, M. (1989). Color correspondence in apparent motion. *Perception & Psychophysics*, **45**, 15-20.
- GREEN, M., & ODOM, J. V. (1986). Correspondence matching in apparent motion: Evidence for three-dimensional spatial representation. *Science*, **233**, 1427-1429.
- GULICK, W. L., & LAWSON, R. B. (1976). *Human stereopsis: A psychophysical analysis*. New York: Oxford University Press.
- HARRIS, J. M., & PARKER, A. J. (1993). Independent stereoscopic processing for signals of different contrast polarity. *Investigative Ophthalmology & Visual Science*, **34**, 1437.
- JULESZ, B. (1960). Binocular depth perception of computer generated patterns. *Bell System Technical Journal*, **39**, 1125-1162.
- JULESZ, B. (1961). Binocular depth perception and pattern recognition. In C. Cherry (Ed.), *Information theory* (pp. 212-224). London: Butterworth.
- JULESZ, B. (1963). Stereopsis and binocular rivalry of contours. *Journal of the Optical Society of America*, **53**, 994-999.
- JULESZ, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.
- JULESZ, B., & MILLER, J. E. (1975). Independent spatial frequency tuned channels in binocular fusion and rivalry. *Perception*, **4**, 125-143.
- JULESZ, B., & PAYNE, R. A. (1968). Differences between monocular and binocular stroboscopic movement perception. *Vision Research*, **8**, 433-444.
- KOVÁCS, I., & JULESZ, B. (1992). Depth, motion and static-flow perception at metaisoluminant color contrast. *Proceedings of the National Academy of Sciences*, **89**, 10390-10394.
- MARR, D. (1980). *Vision*. San Francisco: W. H. Freeman.
- MOVSHON, J. A., & NEWSOME, W. T. (1992). Neural foundations of visual motion perception. *Current Directions in Psychological Science*, **1**, 35-39.
- NISHIDA, S., & TAKEUCHI, T. (1990). The effects of luminance on affinity of apparent motion. *Vision Research*, **30**, 709-721.
- OHZAWA, I., DE ANGELIS, G. C., & FREEMAN, R. D. (1990). Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. *Science*, **249**, 1037-1041.
- ORBAN, G. A., GULYÁS, B., & VOGELS, R. (1987). Influence of moving textured background on direction selectivity of cat striate neurons. *Journal of Neurophysiology*, **57**, 1792-1812.
- PAPATHOMAS, T. V., & GOREA, A. (1988). Simultaneous motion perception along multiple attributes: A new class of stimuli. *Behavioral Research Methods, Instruments, & Computers*, **20**, 528-536.
- PAPATHOMAS, T. V., & GOREA, A. (1989). A new paradigm for testing human and machine motion perception. *Proceedings: SPIE-International Society for Optical Engineering*, **1077**, 285-291.
- PAPATHOMAS, T. V., GOREA, A., & CHUBB, C. (1994, September). *One or multiple motion systems?* Paper presented at the 17th annual European Conference on Visual Perception, Eindhoven, The Netherlands.
- PAPATHOMAS, T. V., GOREA, A., & JULESZ, B. (1991). Two carriers for motion perception: Color and luminance. *Vision Research*, **31**, 1883-1891.
- PAPATHOMAS, T. V., KOVÁCS, I., & HUANG, J. (1993). Similarity vs. covariance in motion, depth and static flow (Glass) patterns. *Investigative Ophthalmology & Visual Science*, **34**, 786.
- PAPATHOMAS, T. V., RAMANUJAN, K. S., & GOREA, A. (1994). Reverse grouping in texture: Fourier and non-Fourier mechanisms. *Investigative Ophthalmology & Visual Science*, **35**, 1667.
- POGGIO, G. F., GONZALEZ, F., & KRAUSE, F. (1988). Stereoscopic mechanisms in monkey visual cortex: Binocular correlation and disparity selectivity. *Journal of Neuroscience*, **8**, 4531-4550.
- POGGIO, G. F., MOTTER, B. C., SQUATRITO, S., & TROTTER, Y. (1985). Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. *Vision Research*, **25**, 397-406.
- PRAZDNY, K. (1984). On the perception of Glass patterns. *Perception*, **13**, 469-478.
- PRAZDNY, K. (1986a). Psychophysical and computational studies of random-dot Moiré patterns. *Spatial Vision*, **1**, 231-242.
- PRAZDNY, K. (1986b). Some new phenomena in the perception of Glass patterns. *Biological Cybernetics*, **53**, 153-158.
- REICHARDT, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W. A. Rosenblith (Ed.), *Sensory communication*. New York: Wiley.
- SAGI, D., & KOVÁCS, I. (1993). Long-range processes involved in the perception of Glass patterns. *Investigative Ophthalmology & Visual Science*, **34**, 1130.
- SCHILLER, P. H., LOGOTHETIS, N. K., & CHARLES, E. R. (1990). Role of the color-opponent and broad-band channels in vision. *Visual Neuroscience*, **5**, 321-346.
- SCHMIDT, B. (1936). Reflektorische Reaktionen auf Form und Farbe und ihre Typologische Bedeutung. *Zeitschrift für Psychologie*, **137**, 245-310.
- SERLING, G. (1976). Movement perception in computer-driven visual displays. *Behavioral Research Methods & Instrumentation*, **8**, 144-151.
- SERLING, G. (1989). Three stages and two systems of visual processing. *Spatial Vision*, **4**, 183-207.
- VAN SANTEN, J. P. H., & SERLING, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, **2**, 300-321.
- WATSON, A. B., & AHUMADA, A. J., JR. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, **2**, 322-342.
- WERKHOVEN, P., SNIPPE, H. P., & KOENDERINK, J. J. (1990a). Effects of element orientation on apparent motion perception. *Perception & Psychophysics*, **47**, 509-525.
- WERKHOVEN, P., SNIPPE, H. P., & KOENDERINK, J. J. (1990b). Metrics for the strength of low-level motion perception. *Journal of Visual Communication & Image Representation*, **1**, 176-188.
- WERKHOVEN, P., SERLING, G., & CHUBB, C. (1992, September 25). *Energy computations in motion and texture*. Paper presented at the annual meeting of the Optical Society of America, Albuquerque, NM.
- WERKHOVEN, P., SERLING, G., & CHUBB, C. (1993). The dimensionality of texture-defined motion: A single-channel theory. *Vision Research*, **33**, 463-485.
- ZUCKER, S. W., STEVENS, K. A., & SANDER, P. (1983). The relation between proximity and brightness similarity in dot patterns. *Perception & Psychophysics*, **34**, 513-522.

## NOTES

1. Although in this paper the term *similarity* denotes physical similarity between targets, we normally use the term *similarity principle* with respect to the sensory space of front-end processes that feed higher order mechanisms in each modality (Gorea, Papathomas, & Kovács, 1993a). Thus, in our terminology, the similarity principle would be obeyed in motion if the front-end filters that provide their input to a motion analyzer shared the same tuning properties (i.e., same receptive field size, spatial frequency, wavelength sensitivity, etc.); the similarity principle would be violated if the two front-end filters were quite different. Also, when we ask whether the similarity principle holds in stereopsis, we mean to ask whether the monocular front-end filters that feed their outputs to binocular disparity-tuned units share similar properties in sensory space. Our use of the term is thus quite different from common use denoting the physical similarity of stimuli (Werkhoven et al., 1990b).
2. The only case where color does not seem to play a role is in Glass patterns, where the evidence to date is that such patterns cannot be formed by purely chromatic input (Prazdny, 1986b; Kovács & Julesz, 1992).
3. The terms *Fourier* and *first-order* are used in the literature to characterize motion mechanisms in which the front-end stage involves only lin-

ear filters. We use these terms interchangeably in this paper. Similarly, the terms *non-Fourier* and *second-order* imply nonlinear front-end processes.

4. Because textural grouping is usually associated with nonrandom patterns, the new set of stimuli does not extend into this modality. Instead, the stimuli of Figures 1 or 2 can be used for textural grouping, with the vertical axis representing the spatial variable  $y$ .

5. Path bq never completely dominates over path aq in Figure 5b; that is, performance never drops to 0% for  $L_q < L_G$ . This result may well be due to the activation of full-wave rectifying second-order motion mechanisms (Chubb & Sperling, 1989), which would favor the aq path, in opposition to first-order mechanisms, which favor the bq path. Since performance is below 50%—that is, in favor of path bq—the first-order mechanisms seem to dominate under these experimental conditions.

6. Another instructive but not straightforward way to view why bq-defined motion dominates aq-defined motion in Figure 5b for  $L_q < L_G$  is to note that both matchings aq and bq produce reverse-phi for  $L_q < L_G$ . However, the reverse-phi due to aq dominates over that of bq because element a has more “energy” than b. Noting that reverse-phi signals motion in the opposite direction, the net result is that motion opposite to that intended by the aq matching is predicted—that is, performances below 50%, as indeed is the case.

(Manuscript received May 2, 1994;  
revision accepted for publication October 4, 1994.)