Pergamon

0042-6989(95)00096-8

Vision Res. Vol. 36, No. 2, pp. 207–222, 1996

0042-6989/96 $9.50 + 0.00

On the Relative Contributions of Motion Energy and Transparency to the Perception of Moving Plaids

DELWIN T. LINDSEY,*† JAMES T. TODD*

Received 14 April 1994; in revised form 30 August 1994; in final form 27 March 1995

Stoner, Albright and Ramachandran [(1990) *Nature*, *344*, 153–155] found that moving rectangular-wave plaid patterns that admitted a transparency interpretation appeared to segment and "slide" past one another as the plaids were translated, while the components of plaids that did not admit a transparency interpretation appeared to unify and move rigidly in the direction of translation of the plaid. In experiment I, we show that the magnitude of the effect reported by Stoner *et al.* is due largely to their repeated-trials experimental protocol, in which plaids moving in a particular direction, upward or downward, are repeatedly presented. This protocol leads to a direction-of-motion-specific adaptation that diminishes the effectiveness of processes that are presumably involved in the unification of the various sensory signals evoked by a moving plaid. In the second experiment, we measured frequencies of nonrigidity for a larger class of moving plaid-like patterns that moved either upwards or downwards on a pseudorandom schedule identical to that employed by Stoner *et al.* Some of the patterns admitted a transparency interpretation, while others did not. The overall pattern of results could not be accounted for within the context of Kim and Wilson's [(1993) *Vision Research*, *33*, 2479–2489] model of motion integration that considers only the oriented motion energy present in a moving plaid stimulus. The results indicate that additional factors, distinct from though perhaps related to the visual analysis of transparency, must also be incorporated into models of perceived plaid motion.

Motion Motion energy Transparency Moving plaids Rigid motion

INTRODUCTION

Many recent studies of motion perception in human observers have used moving plaids as stimuli. A plaid pattern is classically composed of two superposed gratings with different orientations. When the stimulus is repeatedly displaced over successive frames of an animation sequence, the observer may experience one of two types of percept, which depend upon the stimulus conditions: either the plaid pattern displaces rigidly as a whole (this is known as coherent motion), or else the gratings appear to move independently, sliding freely past each other (this is known as component motion). For plaids consisting of cosine gratings, coherent motion is generally seen unless the component gratings are markedly different from one another along the dimensions of spatial frequency, contrast, velocity (speed and/or direction), disparity, or color (e.g. Adelson & Movshon, 1982; Movshon, Adelson, Gizzi & Newsome, 1985; Heeger, 1987; Krauskopf & Farell, 1990; Kooi, DeValois, Switkes & Grosof, 1992; Stone, Watson & Mulligan, 1990).

Adelson and Movshon (1982, 1984; Movshon *et al.*, 1985) advanced a two-stage model of motion perception to account for the coherent motion most often seen in moving plaids consisting of physically-similar cosine grating components. The model was intended to demonstrate how diverse sensory information from an object in motion might be unified into coherent perceptual interpretations of that object. In the first stage, arrays of low-level spatial-frequency, orientation- and speed-tuned motion sensors extract the velocity of each cosine component in the direction normal to its orientation, which in turn defines a set of possible velocities for each 1-D component grating of the plaid. In the second stage, a global velocity is found which is simultaneously a member of each set of the possible velocities for each cosine grating. This is the so-called intersection-of-constraints solution (Fennema & Thompson, 1979). The model can account for the failure of some patterns to cohere by asserting that coherence will occur only when both grating components stimulate classes of motion sensors tuned to the same or similar spatial frequencies, speeds, disparities, colors, etc.

*Department of Psychology and The Center for Cognitive Science, The Ohio State University, 147 Townsend Hall, Columbus, OH 43210, U.S.A. [*Email* dlindsey@magnus.acs.ohio-state.edu].
†To whom all correspondence should be addressed.

While Adelson and Movshon and their collaborators have demonstrated how object velocity might be computed from the outputs of arrays of low-level motion sensors, a number of studies have suggested that factors unrelated to motion, *per se*, may also influence the perception of moving objects. Of particular interest in recent years has been the role played by surface segmentation cues associated with opaque surface occlusion (e.g. Shimojo, Silverman & Nakayama, 1989) and with transparency (e.g. Stoner *et al.*, 1990; Trueswell & Hayhoe, 1993).

Plaid stimuli such as those used in the present study are schematized in Fig. 1. Each panel is intended to represent one frame of an apparent motion sequence in which a plaid is translated behind a stationary circular aperture and along a heading ϕ. A plaid of the type depicted in Fig. 1 can be thought of as a discrete, periodic pattern created by repetitions of four contiguous, colored, parallelogram-shaped regions, labelled **a**, **b**, **c** and **d** in the figure. Following Grunbaum and Shephard (1989), we will call this regularly-repeated set of parallelograms the *motif* of a plaid. The borders of the regions that comprise the motif are all inclined at angles ± 22.5 deg from the vertical, and subjects generally perceive plaids of the sort depicted in Fig. 1 as consisting of a superposition of two rectangular-wave gratings (3:1 duty cycle in the figure) that differ in orientation θ by 45 deg. The plaids in Figs 1a and b differ only in the luminance that has been assigned to region **c** of each plaid's motif. From inspection of the plaids in Fig. 1, it can be seen that analysis of either plaid's upward trajectory will lead to two dominant local motion signals: one directed up and to the left, and the other directed upward and to the right.

Stoner *et al.* (1990) found that plaids of the sort depicted in Fig. 1a cohere under translation, as suggested by Adelson and Movshon's analysis, while plaids like that in Fig. 1b do not. They argued that this occurred because the latter plaids admit a transparency interpretation. Regions **b** and **d** can be thought of as belonging to strips of neutral density filter superimposed on a bright background. Region **c** is therefore a region where two strips overlap, and its luminance will depend upon the amount of light reflected from the strips, as well as upon the product of the transmittances of the neutral density strips. The luminances of regions **a**, **b** and **d**, therefore, impose upper and lower constraints on the luminance of region **c** that admits a transparency interpretation. The lower bound will correspond to the condition of zero reflectance and therefore purely multiplicative transparency and the upper bound will correspond to the condition of occlusion: $L_b^2/L_a \leqslant L_c < L_b$. This range of values for L_c was referred to as the "transparency zone" by Stoner *et al.* The plaid in Fig. 1b satisfies this constraint (i.e. falls within the transparency zone), while the plaid in 1a does not, since the luminance of region **c** in Fig. 1a is greater than L_b.

Classification of features

Stoner and Albright (1992, 1993) have taken this and other results as evidence that, while the processing of motion information in translating plaids may proceed initially in the manner suggested by Adelson and Movshon, the integration step is altered or "gated" by processes which contribute to the classification of image features into intrinsic and extrinsic (Shimojo *et al.*, 1989). A similar argument has been advanced by Trueswell and Hayhoe (1993). In the case of the moving plaids, the classification of features is governed by processes sensitive to patterns of excitation in the retinal image that are consistent with phenomenal transparency. Intrinsic features are those that may arise naturally from transitions in reflectance from one object to another. These transitions are a consequence of differences in the surface properties of the different objects. Extrinsic features may arise when one object overlays and partially obscures a portion of another object. A transition in luminance in this case is not due to the surface properties of a third object, but is instead a

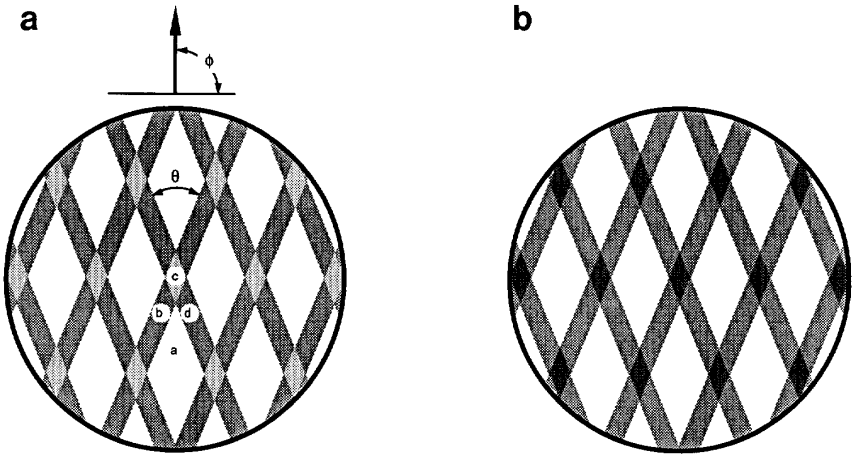

FIGURE 1. Examples of two plaid patterns, similar to those used in previous studies of plaid motion perception and used in Experiment I of the present study. See text for details.

consequence of the surface properties of the two objects and the spatial relationship of these objects to one another.

In the case of the plaid depicted in Fig. 1a, all of the boundaries are classified perceptually as intrinsic features because the patterns are inconsistent with a transparency interpretation, and thus arise from the juxtaposition of four different reflectances. Motion integration is not gated in this case, and a unitary percept of coherent motion is generated. In the case of the plaid depicted in Fig. 1b, the boundaries that define region **c** are classified as extrinsic; that is, as arising from the overlap of the gray bars of one orientation by partially-transmitting bars of another orientation. In this case, motion integration is gated, and two motion percepts are generated in parallel, each based on one of the two dominant signals present in the low-level representation of optical flow.

Elaborated Adelson and Movshon models

Recently, Kim and Wilson (1993) have proposed an alternative to the feature-classification scheme proposed by Stoner and Albright (1992, 1993) that is essentially an elaborated version of the scheme proposed by Movshon and Adelson. Kim and Wilson (see also Wilson, Ferrara & Yo, 1992) have proposed an initial pooling stage in which oriented motion energy from low-level sensors is pooled across space, followed by cooperative pooling across sensors with similar direction tuning and competitive pooling across sensors that are not tuned to similar directions of motion. If the pooling of motion responses across direction of motion were indiscriminate, only a single global motion solution (coherent motion) would be obtained, so long as the spatial and temporal frequency characteristics of the plaid components were similar. However, multiple simultaneous global motion solutions (component motion) are a direct consequence of Wilson's cooperative/competitive scheme. The number of solutions will depend on the distributions of oriented-motion responses. In general, unimodal or nearly unimodal response distributions will lead to a coherent global motion percept, while bimodal response distributions will yield a component motion percept.

An important prediction of this model is that the sliding effect should only occur when the components of a plaid are at relatively small angles to one another in the direction of coherent motion. In order to test this prediction, Kim and Wilson (1993) examined plaids with possible transparency interpretations whose components could be at two different relative orientations—43 or 136 deg. The results confirmed that small angle plaids usually appeared to slide, whereas the large angle plaids usually appeared coherent. On the basis of these findings, and on the basis of related results of Stoner and Albright (1992, 1993) that perceived sliding only occurs over a limited range of duty cycle of the components, it seems reasonable to question the extent to which this phenomenon is representative of the analysis of motion under more general environmental conditions.

Our interest in the robustness of this sliding effect was initially motivated by a series of pilot demonstrations performed in our laboratory using plaids composed of orthogonal gratings with the same four luminances depicted in Fig. 1b that were translated in a different direction each time the plaid was presented. The orientation of the plaid as a whole was also varied across trials so that the component gratings always had orientations of ± 45 deg with respect to the direction of translation. We discovered to our surprise that almost all of these displays appeared perfectly coherent. When we reoriented the component gratings to values of ± 22.5 deg relative to the direction of translation, as did Stoner *et al.* (1990), we were able to see some sliding, but even then, the effect did not begin to occur until after we had undergone a prolonged period of viewing many such displays in succession. Experiment I was designed therefore to examine more systematically how the appearance of sliding is influenced by an observer's viewing history.

GENERAL METHODS

Subjects

A single group of six subjects was run in each of the experiments described below. All were either emmetropic, or had normal visual acuity with appropriate corrective lenses.

Apparatus

The stimuli were generated on a SGI VGXT color graphics workstation (Silicon Graphics Inc.) and displayed on a 20 in. high-resolution RGB monitor (1280 h × 1024 v pixels; 8 bits/pixel; 60 Hz noninterlaced). At the viewing distance of 57.3 cm, each pixel subtended 1.54 min arc. The animation sequences consisted of presentations of a successively translated plaid pattern. Since the SGI workstation employs a double-buffering scheme, the transition from one presentation to the next occurred during the display's vertical-blanking interval.

A linear relationship between pixel value (0–255) and pixel luminance was established for our RGB monitor by creating a correction lookup table. Relative luminance values were measured with a PIN10 (United Detector Technologies) photo detector as a function of pixel value. The measurements were then digitized and averaged. This calibration was then inverted to produce the desired correction vector for the video display controller. Throughout this paper, luminance is specified in units equivalent to the linearized representation of an image pixel in the frame buffer (0–255). A specification of 100 units corresponds to a luminance of 76.5 cd/m^2.

Stimuli

The stimuli consisted of moving plaids with spatial configurations schematized in Fig. 1. The fundamental spatial frequencies of the gratings formed by the pattern motif were 0.8 c/deg. The orientation difference between gratings, θ, was 45 deg. The stimuli were presented in a

circular aperture 11.4 deg in diameter. Motion was produced by successive 0.1 deg translations of the plaid along the direction specified by the heading, ϕ. The animation rate was 30 times/sec, which corresponded to a velocity of 7.65 deg/sec.

General procedures

Each subject was tested individually. The order of conditions was the same across all subjects. The subject sat in front of the computer monitor in a room illuminated only by the light from the monitor. The viewing distance was 57.3 cm. The viewing distance was maintained and head movements by the subjects were minimized with a conventional chin/head restraint system. The subject fixated a small stationary red spot in the center of the stimulus display region of the RGB monitor throughout a trial. The SGI's 3-button mouse was used by subjects to initiate each 1.5 sec trial and then to record their response: "coherent plaid motion" or "component motion". Subjects were asked to base their responses on the category of perceived motion that tended to predominate during the trial. Subjects were given several practice runs of 280 trials each in which to establish confidence in their abilities to make this judgment reliably. All data shown below are averages based on 20 trials per condition per subject. All dependent parameters were varied in block-random fashion during an experimental session. Unless otherwise specified, each subject was tested on a single experimental condition per session, and at least 18 hr elapsed between experimental sessions.

EXPERIMENT I

In this experiment, we examined the effects of viewing history on the probability that component motion will be perceived.

Methods

Regions **b** and **d** of the plaid motif were assigned a luminance of 100 units, region **a** was assigned a luminance of 255 units, and the luminance of region **c** was varied from trial-to-trial between 10 and 140 units, inclusive.

On separate days, subjects viewed moving plaids presented according to one of three schedules. In the first schedule, referred to as the *UP–DOWN* schedule, the plaids could translate either along an upward- or along a downward-oriented heading. The actual heading varied pseudo randomly from trial to trial during the experimental session. This schedule of presentations was identical to that employed by Stoner *et al.* (1990). In the second schedule, the *UP-ONLY* schedule, the plaids were always presented moving on an upward trajectory. In the third schedule, each of four plaid headings—upward, downward, leftward or rightward—occurred with equal mean frequency, although the actual heading varied pseudo randomly from trial to trial. This was the *FOUR-WAY* schedule.

Both the total exposure time to stimuli moving in a particular direction and the average time between exposures to a particular direction of translation varied across the three schedules described above. As the total exposure time increased with schedule, the average time between exposures to a particular direction decreased correspondingly. Therefore, if the perception of coherent plaid motion were determined mainly or even partly by adaptation of a direction-of-motion-specific process, both factors would synergistically influence coherence perception in a way that varies systematically with heading schedule.

Results and discussion

The average results of Experiment I for the six subjects are shown in Figs 2 and 3. In Fig. 2, the fraction of trials on which component motion was seen is plotted as a function of the intensity of region **c** of the plaid. Filled triangles, squares, and circles plot the averages obtained when the heading schedule was, respectively, Up-Only, Up–Down, and Four-Way.

The solid line plotted in Fig. 2 is the average results obtained by Stoner *et al.* (1990). These data are in close agreement with those obtained from an identical heading schedule (Up–Down) in Experiment I, both with regard to the magnitude of the effect of intersection luminance, **c**, on the frequency of component motion, as well as with regard to the luminances of region **c** which maximize the probability that component motion will be seen. It can also be seen in Fig. 1 that the likelihood of component motion increases as the number of distinct directions of plaid motion in the experimental session decreases. The effect of heading schedule on perceived component motion in Experiment I was large: a given plaid viewed in the Up-Only session was, on average, three times more likely to be perceived as nonrigid than the same plaid viewed in the Four-Way session. This result is suggestive of some sort of adaptation process, although the results do not permit us to draw any firm conclusions about the nature of the mechanism or mechanisms that may be adapted, other than to point out that it must be sensitive

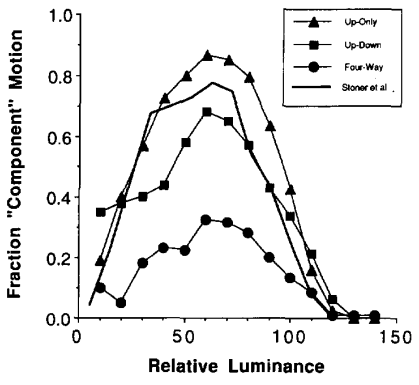

FIGURE 2. Results of Experiment I. Fraction of trials on which the dominant perception of a moving plaid was of two gratings "sliding" past one another, plotted as a function of the luminance of the **c** regions of the moving plaid. Each data point represents the average results from six subjects. Data are shown for three schedules of plaid heading: Up-Only, ▲; Up–Down, ■; Four-Way, ●. The solid line function is that derived from the study by Stoner *et al.* (1990).

to the *direction* of motion. It is not sensitive to repeated exposure to motion, *per se*, nor is it sensitive to repeated exposure to a particular spatial pattern. This follows from the fact that subjects were exposed to equivalent numbers of trials of moving stimuli and to equivalent plaids of the same spatial configuration in each of the three sessions in Experiment I, and the sessions differed only with respect to the frequency and recency of occurrence of plaid motion in any particular direction of motion during the experimental session (see Discussion below).

Further evidence for direction-of-motion-specific adaptation process is provided when the data from Experiment I are displayed in time series format, as in Fig. 3. The abscissa values correspond to the number of the block in which a particular trial occurred, and the ordinate values are the average numbers of trials (out of 14) in each of the 20 blocks on which component motion was seen, collapsed across all values of region c luminance.

Figure 3 shows clearly that performance on the Up–Down schedule used in Experiment I of the present study and used by Stoner *et al.* (1990) is not uniform across time. At the beginning of an experimental session, subjects are more likely to see coherent as opposed to component motion of the plaids. However, as the experiment progresses, subjects begin to see component motion more and more often. This trend is also seen clearly in the data obtained for the Up-Only and Four-Way schedules. The results of a 3 Heading Schedules × 20 Blocks × 6 Subjects ANOVA indicated that both heading schedule and block number ($F_{\text{heading}} = 42.0$, $P < 0.001$; $F_{\text{block}} = 8.139$, $P < 0.001$) contribute significantly to the likelihood of component motion in our displays. Significant interactions between block number and heading schedule were not revealed by the ANOVA ($F_{\text{heading} \times \text{block}} = 1.332$, $P > 0.1$). However, there is compelling evidence from the data shown in Fig. 3 that heading schedule affects not only asymptotic performance, but also affects the rate at which asymptotic performance is achieved. In particular, note that in the first block of trials that followed the Four-Way schedule of heading (circles), the average number of trials on which component motion was seen by our six subjects was only 0.5. Only two of these subjects reported any component motion in this first block, and only three reported any component motion in the first 3 blocks of trials. On the other hand, in the Up-Only sessions (triangles), all six subjects reported component motion for some of the plaids in the first as well as subsequent blocks of trials. The solid line in each panel is based on a simple exponential model incorporating both asymptotic performance as well as a rate constant as parameters for each heading schedule. The values of the parameters for the fit to each of the three data sets are indicated in the figure caption. The solid lines are meant to be suggestive only and are not offered as a rigorous quantitative account of the time-dependent processes underlying the perception of component motion in our displays. The experimental sessions in which the data were collected were self-paced by the subjects, and therefore the block number parameter only crudely estimates a constant interval in time across all subjects and conditions.

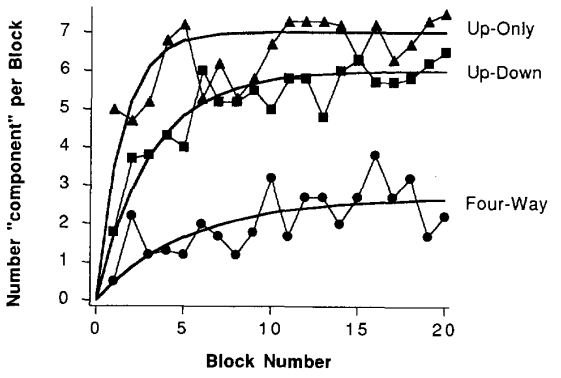

FIGURE 3. Time-series plots of the results from Experiment I. The abscissa values correspond to the block of trials (1–20) in which a particular plaid stimulus was presented; the ordinate values correspond to the total number of trials (out of a total of 14) within the block on which subjects reported component motion of the moving plaid. Separate plots are shown for data collected using each of three presentation schedules: Up-Only, ▲; Up–Down, ■; Four-Way, ●. The functions (thick lines) drawn through each plot are derived from the following simple exponential model: $y = Q(1 - e^{-kt})$. The Q parameter for each of the three functions is the average performance on the last 5 blocks of trials on each schedule. The k parameters were determined by least squares fits of the model to the data. The values of Q and k are, respectively: Up-Only: 7.0, 0.683; Up–Down: 6.0, 0.323; Four-Way: 2.7, 0.186.

These results are reminiscent of those reported by Movshon *et al.* (1985), who investigated the perception of moving plaids under a wide variety of conditions. They postulated that motion is analyzed using a 2-stage process. According to this view, the first stage is performed by an array of independent mechanisms, each of which is selectively tuned to a preferred contour orientation and a preferred direction of motion perpendicular to that orientation. When stimulated by a moving plaid, this stage would generate separate responses for each of the component gratings, which would then be combined at the second stage to determine a unique global velocity for the pattern as a whole. Based on their empirical observations, Movshon *et al.* (1985) suggested that this second stage process may require that the individual component gratings of a plaid have similar contrasts, spatial frequencies, and speeds. Whenever this condition is satisfied, they observed, a plaid will appear as a coherent pattern moving rigidly in a single direction. If, on the other hand, their contrasts, spatial frequencies, or speeds are sufficiently different, then the plaid will appear as two overlapping gratings moving in different directions.

In an effort to test this hypothesis, Movshon *et al.* (1985) developed an ingenious adaptation paradigm. Suppose, for example, that an observer undergoes a prolonged period of adaptation to a vertically oriented grating moving in a rightward direction, and is then tested with a plaid composed of diagonal gratings

moving in a rightward direction. Movshon *et al.* reasoned that the response of the first stage to the moving plaid would be primarily mediated by mechanisms tuned to diagonal orientations, which would be relatively unaffected by the vertical adapting stimulus. However, if the rightward sensitive mechanisms in the second stage are orientation invariant, then they would be significantly fatigued by the adapting stimulus, and the perception of coherent motion might therefore become less tolerant to differences in contrast or spatial frequency between the plaid's component gratings. Movshon *et al.* performed several experiments that confirmed this prediction.

Although the effects of viewing history in the present experiment were most likely due to some form of adaptation, there were several differences that deserve to be highlighted between our procedure and the one used by Movshon *et al.* (1985). First, the relative orientations of our component gratings were always the same for any given direction of motion, so that both stages of processing should have, or at least in principle could have, been adapted equally over successive presentations. Second, the component gratings of our plaids all had identical contrasts, spatial frequencies, and speeds, but the displays used by Movshon *et al.* only appeared to slide when the overlapping sine waves differed along one or more of these dimensions. Finally, the component gratings used in our displays all had a fundamental frequency of (0.7 cpd), whereas Movshon *et al.* obtained only minimal adaptation effects with cosine gratings less than 2.0 cpd.

Even though we cannot identify the precise mechanism by which prolonged viewing of motion in one direction alters observers' perceptions in this paradigm, there is one important conclusion that seems perfectly clear: In the absence of adaptation, coherent motion is the preferred perceptual interpretation for all of these displays, regardless of the luminance in the region of intersection. To the extent that image segmentation cues specifying surface transparency can lead to the perception of component motion, our results suggest that their effects are limited to those situations where the preferred interpretation of coherent motion has somehow been fatigued.

EXPERIMENT II

When considered as a whole, the results of Experiment I together with those of Kim and Wilson (1993) and Stoner and Albright (1993) provide a growing body of evidence that the transparency based sliding effects first reported by Stoner *et al.* (1990) appear to be limited to a relatively narrow range of display parameters. Nevertheless, when plaids are constructed with rectangular-wave gratings of appropriate duty cycles and relative orientations, and are viewed for a sufficiently long period of time, the effect can be quite compelling. That is to say, over a certain range of luminances in the region of intersection, the component gratings will appear to slide past one another in opposite directions, whereas for intersection luminances outside this range the gratings will appear to be moving together in the same direction.

Given that all the other constraints needed to perceive sliding are satisfied, we were curious to discover if the effect is due to an underlying analysis of transparency as suggested by Stoner *et al.* (1990) or whether it could be explained adequately using a low-level analysis of motion energy as hypothesized by Kim and Wilson (1993). In order to address this question, we decided to examine a much broader range of plaid motifs than have been used in previous investigations, and we specifically manipulated the overall patterns of luminance among the different regions of each plaid to influence their appearance of static transparency. A number of previous empirical studies have shown that transparency will be perceived when certain figural and intensity constraints are satisfied at the boundaries that define the opaque and transparent portions of the stimulus. Figural constraints refer to those global pictorial properties of a stimulus that evoke a "double presence" (Kanizsa, 1979); i.e. that support the interpretation of a region such as **c** in Fig. 1 as belonging simultaneously to two different surfaces. Intensity constraints are those constraints on the relationships of order and magnitude among the four luminances that define x-junctions in a stimulus that may support a transparent interpretation.

The design of the plaids used in Experiment II was guided by two analyses of achromatic transparency in stationary stimuli that have been proposed by Metelli and by Beck and their respective collaborators (Metelli, 1974, 1985; Beck, Pradzny & Ivry, 1984; Beck & Ivry, 1988). Both analyses are based on extensive empirical studies, and yield quantitatively similar results, although each was designed to deal with perceptual transparency that arises from different physical conditions. Metelli's model is specifically designed to account for diaphanous transparency, also referred to as screen-door transparency in the computer graphics literature. In diaphanous transparency, the transparent material is perforated, much like a screen door, with holes too small to be individually resolved by the eye. The light reaching the eye from this material is sum of the light passing through the perforations, and the light reflected from the material itself. In Beck's analysis, the transparent material obeys the Beer–Lambert relations, and the light reaching the eye is the sum of reflected light and the fraction of light from the opaque underlayer that is transmitted by the transparent material. The test stimuli in Experiment II could be analyzed within the framework of either description of the physical causes of transparency.

Studies by Beck and by Metelli, in which subjects judged whether or not a particular simulation evoked the distinct perception of transparency, indicate that transparency will be perceived when the following two intensity constraints are satisfied:

(i) The polarity of luminance change across the border separating two homogeneous opaque surfaces should not be reversed when it is partly overlaid by a homogeneous transparent surface. Following

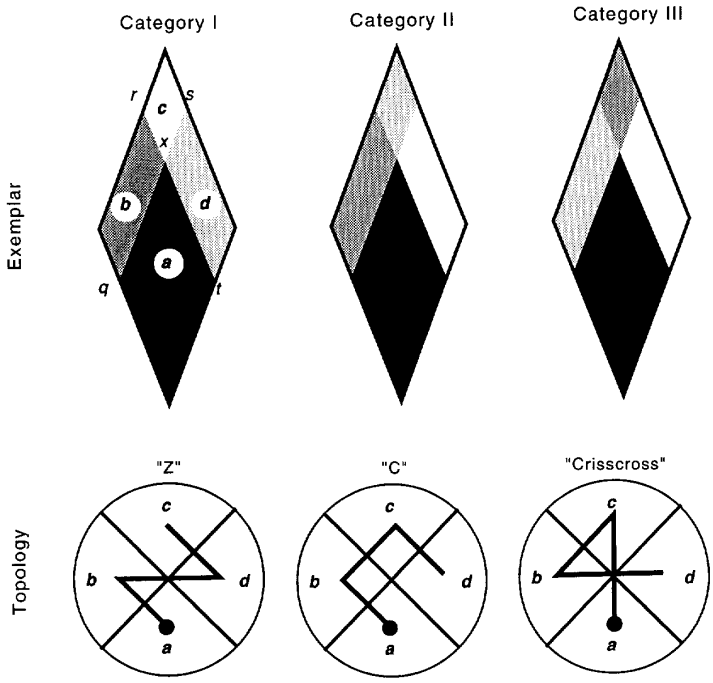

FIGURE 4. Overview of rationale for design of stimuli used in Experiment II. The top row depicts the motif of a plaid from each of three stimulus categories. Shown (left to right) are motifs from plaids **abdc**, **abcd** and **acbd**. The bottom row of the figure depicts the order of luminances around the x-junction in each motif shown in the top row. We refer to this arrangement of luminances as the topology of the motif. All plaids in Category I had the Z topology, those in Category II all had the C topology, while plaids in Category III all had the "crisscross" topology.

Beck *et al.* (1984), we refer to this constraint as the *order* constraint.

(ii) The change in luminance or lightness across the border created by juxtaposing two opaque surfaces must decrease when the border is overlaid by a homogeneous transparent surface. We refer to this constraint as the *magnitude* constraint.

Beck and Metelli have not agreed on precisely how the magnitude constraint should be implemented. Metelli has argued that the visual system specifically measures and analyzes the luminance differences across opaque and transparent borders, while Beck has argued that a nonlinear transformation akin to that which underlies the perception of lightness precedes the analysis of differences across the opaque and transparent borders. In the spirit of Beck's ideas concerning lightness differences, more recent work on the magnitude constraint that has been specifically related to the factors that govern perceptual coherence in moving plaids (Stoner *et al.*, 1990; Trueswell & Hayhoe, 1993), suggests that the local *contrasts* across opaque and transparent borders are the magnitudes that should be evaluated. We therefore evaluated the patterns of results from Experiment II with regard to both luminance and contrast differences. For most of the test plaids, either method of implementing the magnitude constraint led to the same prediction. For some of the plaids that we studied, however, consideration of luminance and contrast differences led to different predictions of transparency.

Methods

Twenty-four plaid stimuli, equally divided among three stimulus categories, were used in Experiment II. Plaids in Category I had two valid transparency interpretations, as specified by the order and magnitude constraints described above. Plaids in Category II had one valid transparency interpretation, while plaids in Category III had no transparency interpretation.

Plaids were generated from two palettes, each containing four luminances. The luminances in the first palette were 0, 128, 192 and 255 units; those in the second were 0, 64, 128, and 255 units. Each palette was used to construct four plaids in each of the three categories by assigning the four luminances in different ways to the four regions of the plaid motif (see Fig. 4). Thus, the 12 plaids derived from each palette, differed from one another only in the arrangement of luminances in the motif. All plaids were geometrically equivalent to those used in Experiment I. Thus, two of the lines of translational symmetry of the plaid patterns were ± 22.5 deg from the vertical. The plaids were presented on an Up–Down heading schedule only, and the frequencies of component motion were measured for all 24 stimuli in the same experimental session.

The motifs of all of the Category I plaids were based on the topology of the exemplar depicted in the upper-left panel of Fig. 4. Note the x-junction, labelled x in the figure, formed by the intersection of regions **a**, **b**, **c** and **d**. This x-junction identifies a locus of possible

transparent overlap in the plaid, by either of the two surfaces that share edge *qs* or those surfaces that share edge *tr*. By applying the order and magnitude constraints described above to the motif, it can be shown that the two valid interpretations are as follows: (1) **b** and **c** belong to the transparent surface and **a** and **d** are opaque surfaces and (2) **d** and **c** belong to the transparent surface and **a** and **b** are opaque surfaces. The bottom-left panel of Fig. 4 depicts the topological feature that all plaids in Category I shared: when the regions of the motif are considered in ascending order of luminance, **abdc** in the case of the motif shown in the top-left panel of Fig. 4, the path followed in traversing the *x*-junction is *Z*-shaped (see Beck & Ivry, 1988, for further discussion of *x*-junction topology).

Three other plaids, all with the same *Z* topology were constructed from each of the two palettes: **bcad**, **cdba** and **dacb**. Portions of these six plaids, along with the two plaids derived from the prototype motif, are shown in Fig. 5. All had two valid transparency interpretations, but differed with regard to the regions that, according to the order and magnitude constraints, could be considered opaque or transparent. The valid interpretations for each of the 8 plaids in Category I are indicated in Table 1. Note that the table includes entries based on each of two methods of applying the magnitude constraint. In the first method, comparison was made of the luminance differences across the appropriate boundaries; i.e. for a border formed by juxtaposition of two regions, 1 and 2, $L_{region1} - L_{region2}$ was calculated. In the second method, the local contrast of the border separating regions 1 and 2 was calculated: $(L_{region1} - L_{region2})/(L_{region1} + L_{region2})$. Note that these two methods lead to different predictions of the transparent and opaque regions for the Category I plaids derived from palette 2.

The plaids in Category II all had motifs topologically equivalent to the motif shown in the upper-middle panel of Fig. 4. This topology supports only one valid transparency interpretation, which in the case of the motif shown in Fig. 5 is: **b** and **c** belong to a transparent surface and **a** and **d** belong to opaque surfaces. Note that all plaids in Category II had a *C*-shaped ordering of luminances around the *x*-junctions, as depicted in the lower-middle panel of Fig. 4. Portions of the eight plaids in Category II, are shown in Fig. 6. The valid transparency interpretations of these plaids are listed in Table 1.

Finally, Category III plaids all had motifs that were topologically equivalent to the one depicted in the upper-right panel of Fig. 4. They all had the "crisscross" ordering of luminances depicted in the lower-right panel of Fig. 4, that Beck and Ivry (1988) identified as diagnostic of *x*-junctions for which no transparency interpretation is possible. This topology does not admit a transparency interpretation, since all pairwise comparisons of luminance around the *x*-junction violate the order constraint. Portions of the eight plaids in Category III, are shown in Fig. 7.

It can be seen in Figs 5–7 that manipulation of the luminances assigned to the motif of a plaid, in the ways described above, leads to a number of different perceptual organizations of the static plaids that in most cases are consistent with the order and magnitude constraints. In some of the plaids in our study—for example, **abcd**—the narrow bars, either light or dark gray, appeared transparent. In other plaids—for example, **cdab** and **dabc**—the wide white or gray bars appeared transparent. In still other plaids, in particular those in Category III, none of the regions were phenomenally transparent.

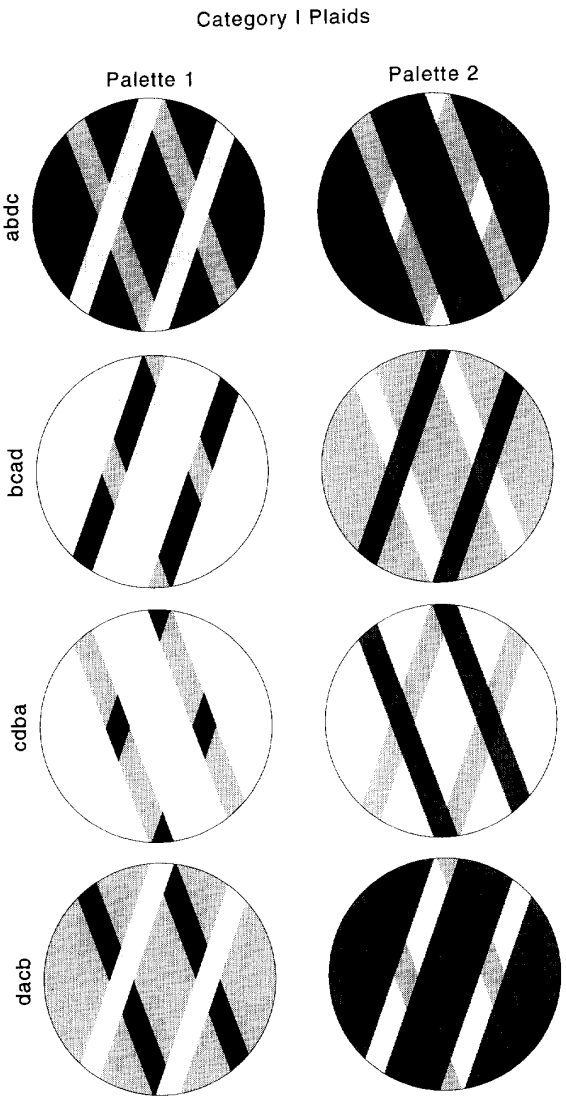

FIGURE 5. Halftone samples of each of the Category I plaids. Each of these plaids had the *Z* arrangement of luminances in their motif. Plaids in the left column were generated on a CRT using the palette containing luminances of 0, 128, 192 and 255 units; for those plaids shown in the righthand column, the palette contained luminances of 0, 64, 128 and 255 units. The reflectances of the original halftone images closely matched the relative luminances of the two palettes; some loss in fidelity may have occurred in reproducing the original figure. Only a portion of each of the experimental plaids is represented in the figure; the actual stimuli contained more repetitions of the motif (see General Methods).

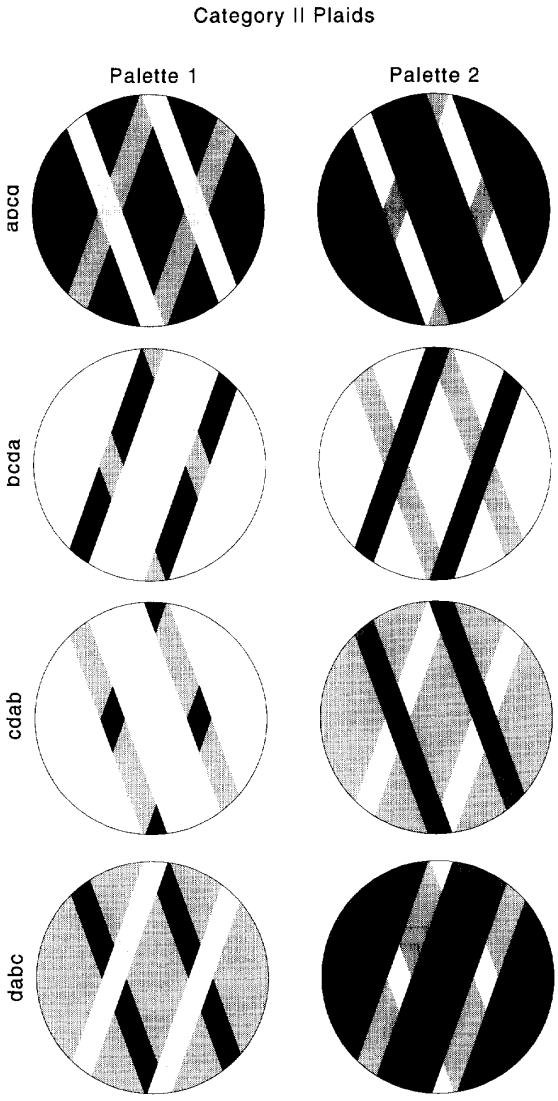

FIGURE 6. Halftone samples of each of the Category II plaids. Each of these plaids had the C arrangement of luminances in their motif. Consult caption of Fig. 5 for details.

To summarize, 24 moving plaid stimuli were employed in Experiment II, 8 in each of three categories. Four of the plaids in each category were derived from each of two palettes of luminances. Category I plaids had two valid transparency interpretations, Category II plaids had one transparency interpretations, and Category III plaids had no transparency interpretations. The null hypothesis in Experiment II was that the rates of component motion perception should be uniformly high for those plaids in categories I and II, and uniformly low for those plaids in category III.

Results and analysis

The results of Experiment II are shown in Fig. 8. The graphs plot average fractions of trials (out of 20) on which component motion was seen when the specified stimulus was presented. Data obtained from Category I, II, and III plaids are shown, respectively, in the top, middle, and bottom rows of graphs. The left and right columns of graphs plot data from plaids derived from the two palettes of luminances. The significance of the open and closed symbols and the arrows will be discussed later.

As expected, component motion was rarely seen by any of the subjects when Category III stimuli were presented. However, plaid category did not otherwise meaningfully partition those plaids that did frequently "slide" from those that did not slide. If subjects' performances were determined solely by the existence of valid transparency interpretations, the fractions of trials on which component motion was elicited by Category I and II plaids should have been uniformly high. This

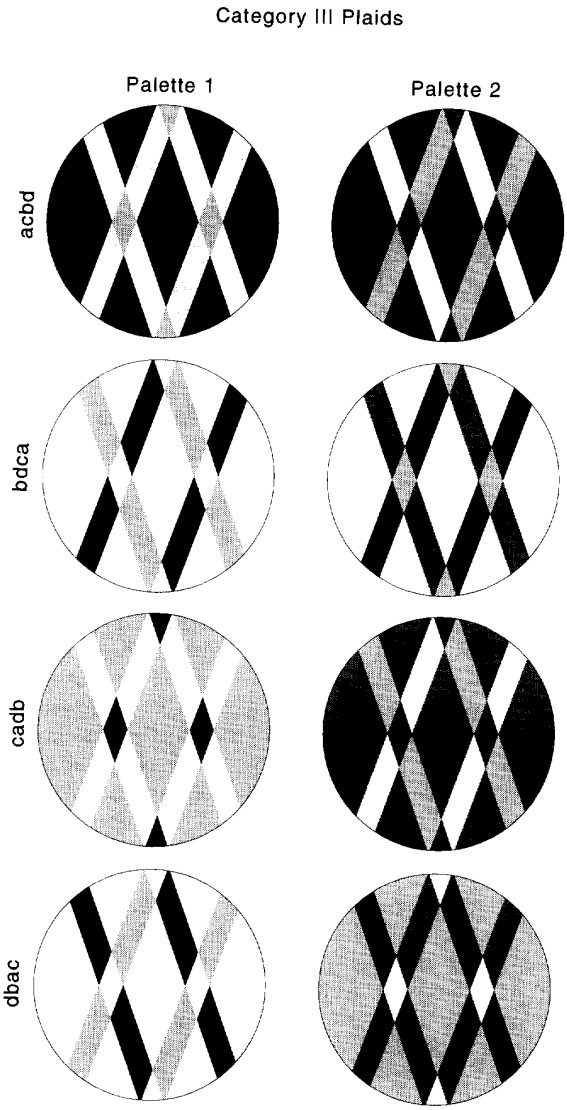

FIGURE 7. Halftone samples of each of the Category III plaids. Each of these plaids had the "crisscross" arrangement of luminances in their motif. Consult caption of Fig. 5 for details.

TABLE 1. Valid transparent interpretations of the plaid stimuli in Experiment II

		Luminance difference	Local contrast
Category I			
Palette 1	**abdc**	N+ N−	N+ N−
	bcad	N+ W−	N+ W−
	cdba	W+ W−	W+ W−
	dacb	W+ N−	W+ N−
Palette 2	**abdc**	W+ W−	N+ N−
	bcad	W+ N−	N+ W−
	cdba	N+ N−	W+ W−
	dacb	N+ W−	W+ N−
Category II			
Palette 1	**abcd**	N−	N−
	bcda	N+	N+
	cdab	W−	W−
	dabc	W+	W+
Palette 2	**abcd**	N−	N−
	bcda	N+	N+
	cdab	W−	W−
	dabc	W+	W+
Category III		No valid transparent interpretations	

N+ = narrow bars +22.5 deg from vertical; N− = narrow bars −22.5 deg from vertical.
W+ = wide bars +22.5 deg from vertical; W− = wide bars −22.5 deg from vertical.

expectation is obviously not borne out by the data. Nor are any clear-cut differences in performance seen between Category I and II plaids that might be specifically attributed to differences in topology of the Category I and II motifs.

Analysis of foreground/background relationships

Of particular interest were the results obtained with **cdba**, **cdab** and **dabc** plaids. Why was component motion rarely seen in these moving plaids, while the rate of component motion was generally high when the other plaids in Categories I and II were presented? A modified version of the simple transparency hypothesis was therefore examined. Stoner and Albright (1992, 1993) have argued that figural cues for foreground/background assignment, when they conflict with those for transparency, tend to abolish the perception of component motion in moving plaids. According to these authors, subjects tend to perceive region **a** in the plaid motif as part of the background, regardless of the brightness of that region. This tendency leads to a conflict in the assignment of order in depth when transparency cues signal that **a** is a region of transparent overlap between two surfaces, since the latter interpretation would require that **a** lie in the foreground. This additional constraint on valid transparency interpretations in plaids was not consistent with our own experiences when viewing stationary versions of some of the plaids employed in Experiment II. Nevertheless, we examined the possibility that this additional figural constraint might account for our results.

The filled circles in Fig. 8 correspond to plaids that had at least one valid transparent-narrow-bars interpretation, when the magnitude constraint was evaluated in terms of local contrast. The open circles in Fig. 8 correspond to plaids that did not have a valid transparent-narrow-bars interpretation. A very good correspondence can be seen between this partitioning of the plaids, and the frequencies of occurrence of coherent and component motion in the plaids. An obvious conclusion is that plaids that do not have a transparent-narrow-bars interpretation tend to cohere, while plaids that do admit this interpretation tend to evoke component motion frequently.

We also incorporated the additional constraint into a second analysis of transparency in which luminance differences, rather than contrast, were considered. The circles without arrows indicate those plaids that admitted a transparent-narrow-bars interpretation and the arrows above circles correspond to those plaids that did not admit a transparent-narrow-bars interpretation. We found that the predictions from this analysis did not correspond to the average performance of our subjects on two plaids derived from palette 2: **abdc** and **cdba**. The analysis of luminance differences predicted that **abdc** should have evoked mostly coherent motion, while **cdba** should have evoked component motion. Precisely the opposite pattern of response was seen in the data for

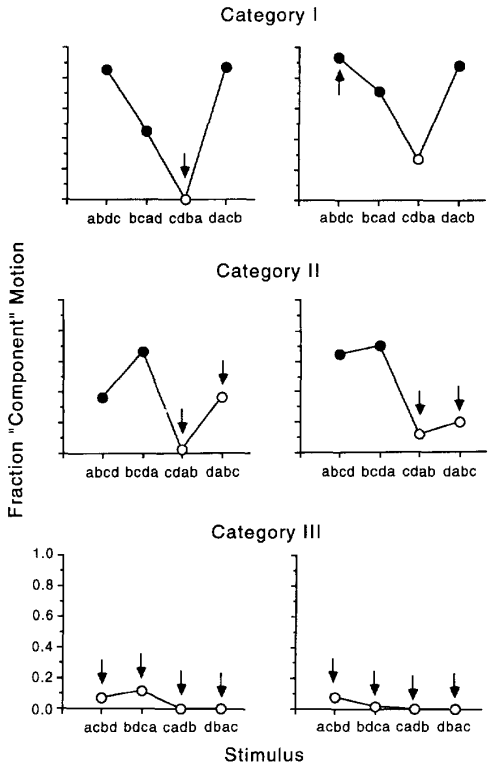

FIGURE 8. Fractions of trials on which component motion was seen for each of the 24 stimuli presented in Experiment II. Data points are averages of results from 6 subjects, 20 trials per subject. Stimuli are identified by luminance configuration, as described in the text. Left column: stimuli derived from palette containing luminances of 0, 128, 192 and 255 units (palette 1). Right column: palette contained luminances of 0, 64, 128 and 255 units (palette 2). See text for meaning of the filled and unfilled symbols and of the arrows.

these two plaids. We conclude from this analysis that, to the extent that perceived motion in plaids is determined by processes that embody figural and intensity constraints on transparency, the substrate for these processes appears to respond to local contrast and not to local differences in luminance. However, we emphasize that while the figure–ground analysis successfully accounts for perceived sliding, it does not account for phenomenal transparency of static plaids. In particular, the wide bars of the plaids **cdab** and **dabc** of Category II distinctly stand out as transparent, and yet these two plaids evoked generally low frequencies of component motion from our subjects.

Analysis of motion energy distributions

We next examined whether an approach to motion integration based on distributions of motion energy in the moving plaids, as described by Wilson *et al.* (1992) and Kim and Wilson (1993), might also account for the results of Experiment II. In this model, motion energy is measured by arrays of directionally-selective motion sensors, each tuned to a different spatial frequency band and a different direction of motion energy. There are two parallel global motion energy pathways, one that analyzes first-order or Fourier optical flow, and another that analyzes second-order of non-Fourier optical flow. The outputs of these two pathways are fed forward into a neural network that resolves its inputs into a much reduced number of resultant velocity signals which correspond to predictions of global motion perception. In the case of plaid stimuli, the neural network converges on one or two resultant signals, corresponding, respectively, to coherent or component motion perception (see Kim & Wilson, 1993).

Our analysis incorporated two major simplifications of the one proposed by Kim and Wilson. First, we only evaluated the visual responses to first-order motion. Our justification for doing so is based on Kim and Wilson's demonstration that the results of Stoner *et al.* (1990) could be qualitatively accounted for by considering only first-order motion. Second, we did not implement the neural network stage of Kim and Wilson's model. Instead, we concluded from their study that the distributions of motion responses were the key determinant of component or coherent motion perception of moving plaids. In brief, plaids that yield bimodal motion response distributions should yield significantly higher rates of component motion perception than plaids that do not.

In our analysis, motion energy was sampled across each of 4 spatial-frequency ranges and across 13 directions that spanned the stimulus space from 0 (rightward) to 180 deg (leftward) in 15 deg increments for a plaid moving in an upward (90 deg) trajectory. The spatial and temporal properties assigned to each of these 52 mechanisms were similar to those used by Kim and Wilson (see also: Wilson *et al.*, 1992; Wilson & Gelb, 1984; Phillips & Wilson, 1983). Briefly, the line spread function (LSF) of each orientation and spatial-frequency tuned mechanism was based on difference-of-Gaussian functions published by Wilson and Gelb (1984) and Phillips and Wilson (1983). Thus, the simulated motion sensors had approximately ± one octave spatial frequency passbands centered at 0.8, 1.7, 2.8, or 4.0 cpd. The orientation bandwidths of the 52 mechanisms were all approximately 45 deg (full-width, half-amplitude). The peak sensitivities that we assigned to these mechanisms, were, from low to high center frequency, 50, 100, 150 and 200.

All computations were performed in the Fourier domain, following transformation of 256 × 256 pixel images (spatial resolution: 0.0266 deg/pixel) of the stimuli and the mechanism LSFs into their Fourier-domain counterparts. The computation of motion energy was based on a simple bilocal cross-correlation scheme of the type first proposed by Reichardt (1961) and incorporated into the model of motion perception proposed by Wilson *et al.* (1992). The critical step in the motion energy computations we performed is given by:

$$M_{ij} = \sum_{k=0}^{255} \sum_{l=0}^{255} (\mathrm{P}_{kl}\mathrm{LSF}_{ijkl})^2 \times \sin(2\pi\omega_{kl} v \Delta t_i)\sin[2\pi\omega_{ij}\Delta x_i \cos(\varphi_{jkl})] \quad (1)$$

where i and j are indices of spatial-frequency (0 . . . 3) and orientation tuning (0 . . . 12) of the mechanisms, and k and l are indices of summation in the Fourier domain. M_{ij} is the motion energy sampled by the ij^{th} spatial frequency and orientation-tuned mechanism, and P and LSF are the spatial Fast Fourier transforms, respectively, of the stimulus and each of the 52 mechanisms. Δx and Δt are the span and time delay parameters of the bilocal detector, and v is the velocity of the plaid pattern in the direction of translation: 7.65 deg/sec. ω_{kl} is the 2-D spatial frequency specified by k and l, and ϕ is the difference in orientation between ω_{kl} and the orientation of peak response of the ij^{th} mechanism. Only those terms in the summation for which both sine terms were positive were included in the computation.

In the final computational step, a compressive nonlinearity similar to that published by Wilson *et al.* (1992) was applied to each M_{ij}, so that at contrast levels above about 20 times threshold contrast, modest increases or decreases in contrast produce little variation in mechanism response.

$$R_{ij} = \frac{M_{ij} R_{\max}}{R_{\max} - 1 + M_{ij}} \quad (2)$$

where R_{ij} is the response of the ij^{th} spatial frequency- and direction-selective mechanism, and $R_{\max}$ is the maximum mechanism response.

The distributions of motion responses, R_{ij}, derived from each of the Category I through III stimuli are shown, respectively, in Figs 9–11. For the purposes of simulation, all plaids were assumed to be moving on an upward (i.e. 90 deg) trajectory at a speed of 7.65 deg/sec. Separate motion response distributions for mechanisms with passbands centered at 0.8, 1.7, 2.8 and 4.0 cpd are indicated by separate symbols in the graph for each stimulus. The abscissa of each graph identifies the directions of optimal response of each of the 13

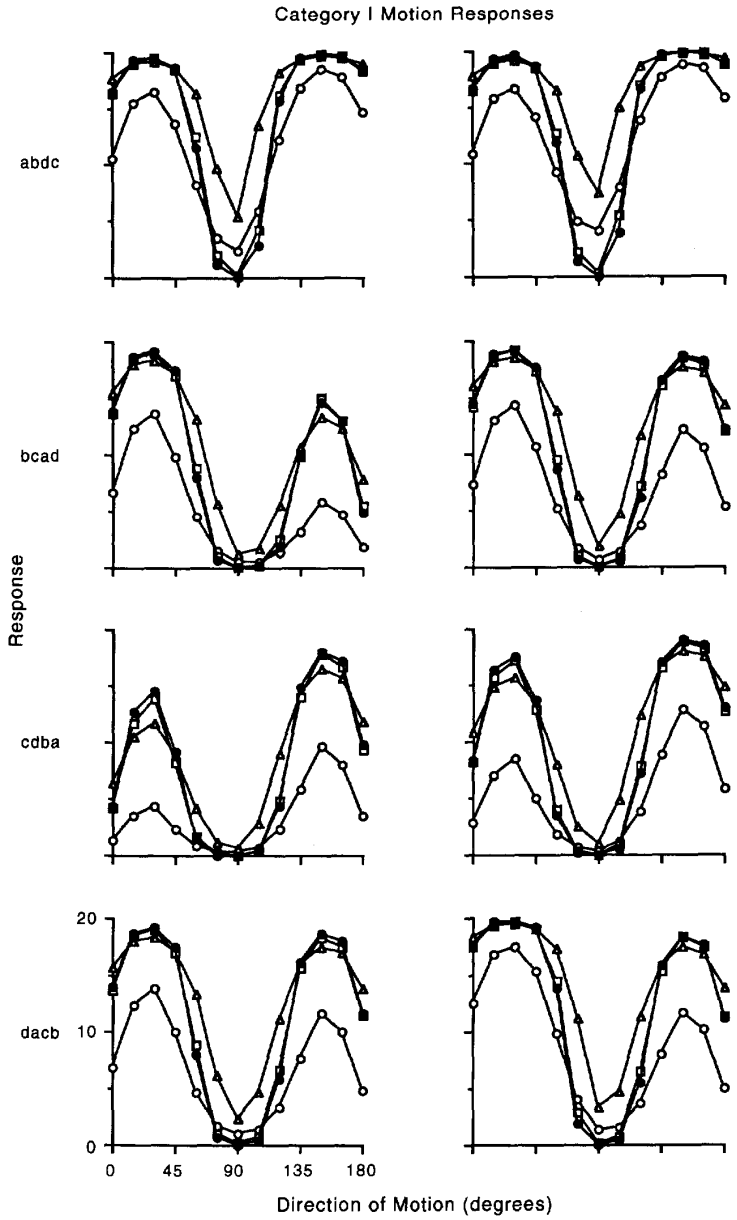

FIGURE 9. Motion response distributions obtained from Category I stimuli: **abdc**, **bcad**, **cdba** and **dacb**. Left column: stimuli derived from palette containing luminances of 0, 128, 192 and 255 units (palette 1). Right column: palette contained luminances of 0, 64, 128 and 255 units (palette 2). The four functions in the graph for each stimulus are response distributions for different spatial frequency tuned mechanisms: 0.8 cpd, ○; 1.7 cpd, △; 2.8 cpd, □; 4.0 cpd, ●.

directionally-selective motion sensors assigned to each of the four spatial frequency passbands.

The distributions of responses to Category I stimuli (Fig. 9), taken as a group, are strikingly homogeneous. The four response functions for each stimulus are all bimodal. The minimum of each function occurs in the direction of translation of the plaid (90 deg) and the maxima (30 and 150 deg) occur for mechanisms tuned to motion in directions nearly orthogonal to the orientations of the gratings formed by the plaid motifs. These bimodal patterns of responses, if processed by the neural network proposed by Kim and Wilson (1993) would all be expected to yield component motion solutions. We cannot see any differences among the eight sets of motion response functions shown in Fig. 9 that might explain why plaids **abdc** and **dacb** gave uniformly high component motion rates, while plaids **cdba** gave uniformly low rates of component motion. It seems unlikely that the differences in subjects' performances on these plaids is due to differences in their second-order motion energy. Plaids **abcd** and **cdba**, for example, have very similar spatial Fourier amplitude spectra, with the

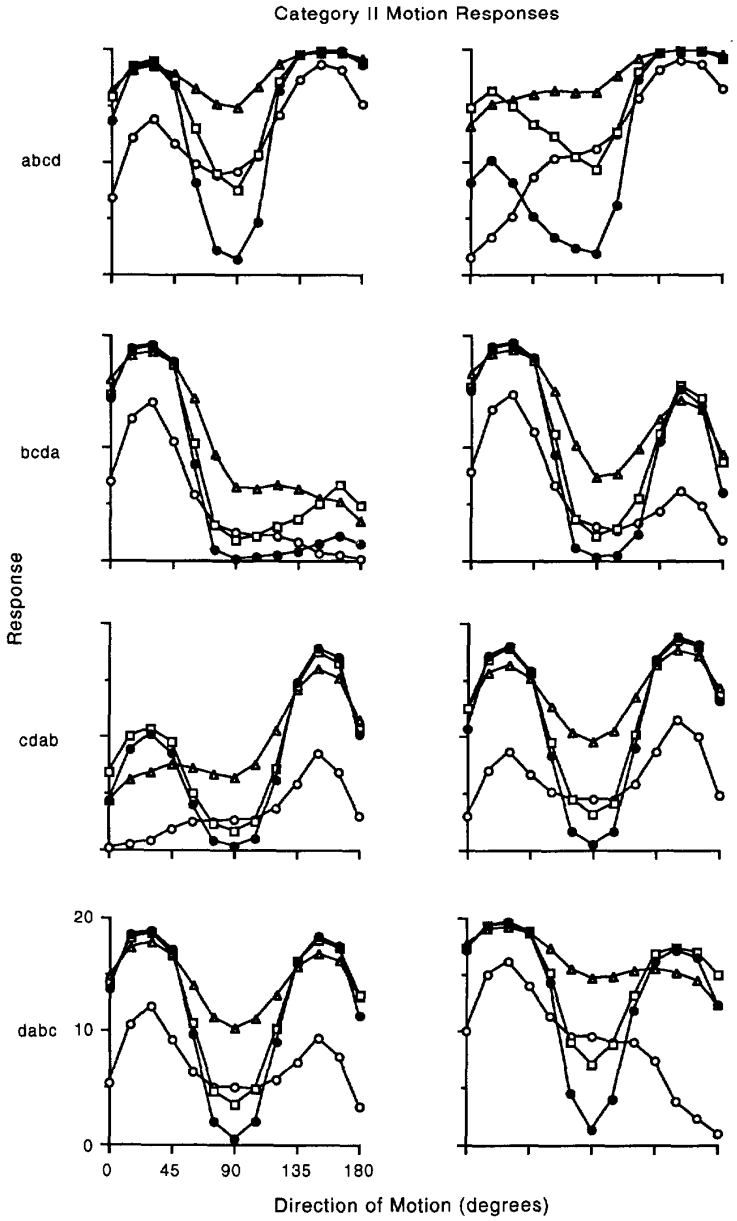

FIGURE 10. Same as Fig. 9, except motion response distributions were derived from Category II stimuli: **abcd**, **bcda**, **cdab** and **dabc**.

notable exception that the spatial contrast of **cdba** plaids is uniformly lower than that of the **abdc** plaids. Since the second order motion energy for a given plaid is related monotonically to its amplitude spectrum, we would expect the effects of second-order motion to be less for **cdba** than for **abdc** plaids. While second-order motion is likely to increase the frequency of perceived coherent motion in plaids, the expected differences in second-order motion energy between these two plaids is precisely opposite of what is needed to explain our data.

In contrast to the response distributions for Category I stimuli, those computed for Category II stimuli (Fig. 10) are, as a group, rather complex and heterogeneous. A feature common to many of the Category II distributions is the dissociation in the shape of the response distributions across spatial frequency. The response distributions for mechanisms tuned to the two highest spatial frequency bands are generally bimodal, while those for mechanisms tuned to the lower spatial frequency bands are, as a class, more variable in shape. Category II plaids tend to have more Fourier energy at 90 deg than do Category I plaids, which explains the low spatial frequency response functions. We attribute the higher spatial frequency bimodal functions to the velocity tuning of these mechanisms. They are much less sensitive to the relatively high velocities of Fourier

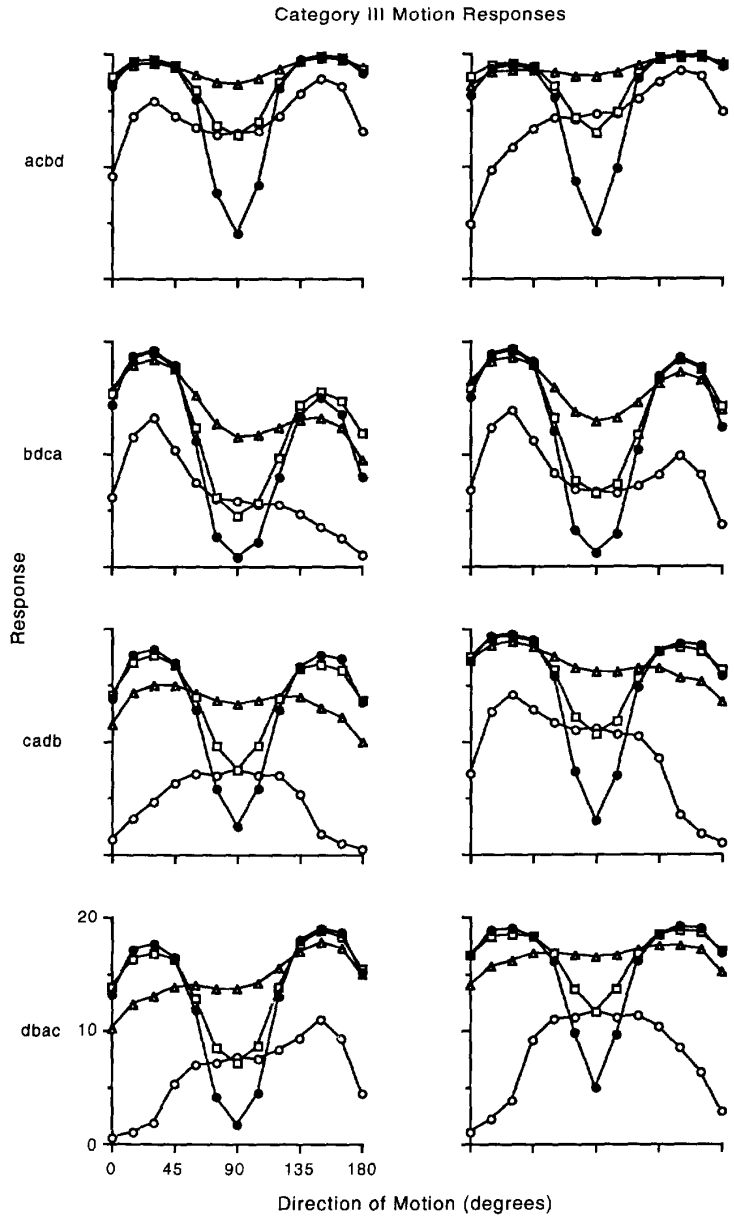

FIGURE 11. Same as Fig. 9, except motion response distributions were derived from Category III stimuli: **acbd**, **bdca**, **cadb** and **dbac**.

components at 90 deg produced by vertical translation of the plaid than they are to the relatively low velocities of Fourier components near 0 and 180 deg. This heterogeneity in response across spatial frequency is qualitatively consistent with our finding generally lower rates of component motion perception with Category II than with Category I plaids. However, this heterogeneity also makes it difficult to explain why the frequency of perceived component motion was high for some of the plaids in Category II and low for others.

Finally, the response distributions for Category III plaids are shown in Fig. 11. As a class, they are, like the response distributions for Category II plaids, complex and heterogeneous. The distributions also reflect the greater amounts of motion contrast at 90 deg observed in the Fourier spectra of Category III plaids, as compared to the spectra of either Category I or II plaids. This finding is qualitatively consistent with the result that a component motion percept was rarely elicited by Category III plaids.

GENERAL DISCUSSION

It has long been known that plaids composed of gratings that differ along one of several stimulus dimensions frequently appear nonrigid when they are

uniformly translated (e.g. Adelson & Movshon, 1982, 1984; Movshon *et al.*, 1985; Heeger, 1987; Krauskopf & Farell, 1990; Kooi *et al.*, 1992; Stone *et al.*, 1990). This phenomenon, in addition to the phenomenon of rigid motion when the components are physically similar to one another, was easily explained by a number of early models of motion integration based on ideas proposed by Movshon and his colleagues (e.g. Movshon *et al.*, 1985; Heeger, 1987). These models, however, could not account for the Stoner *et al.* (1990) finding that nonrigid motion could be reliably elicited in some plaids composed of perceptually identical gratings. Stoner *et al.* proposed that the motion integration system was capable of using cues for phenomenal transparency, in addition to those signals arising from low-level motion sensors, during the motion integration process.

Kim and Wilson (1993) have argued, in contrast, that a consideration of surface transparency or other types of image segmentation cues is unnecessary to account for the sliding effect. They have proposed a specific computational model for the low-level analysis of Fourier and nonFourier motion energy, which they have applied to plaid patterns similar to those used by Stoner *et al.* (1990). An important prediction based on this model is that the sliding effect should only occur when the component gratings of a moving plaid are at relatively small angles to one another in the direction of motion. To confirm this prediction, they demonstrated empirically that the appearance of sliding is essentially abolished when the orientation difference between the individual components is increased to 136 deg. We have obtained similar results in our pilot experiments using orientation differences as small as 90 deg.

Although it is attractive as a parsimonious explanation of the sliding effect, the Kim and Wilson model in its present form does not provide a satisfactory account for the variations in perceived coherence among the different displays employed in experiment II of the present series. In examining the various distributions of motion energy shown in Figs 9–11, there are no obvious criteria for distinguishing those patterns that are frequently perceived to slide from those that do not. If this is a valid model for the analysis of low-level motion in human vision, then there must clearly be some higher-level process at work in the determination of motion coherence.

Could this process involve an analysis of surface transparency as suggested by Stoner *et al.* (1990)? One aspect of the present results that is supportive of this hypothesis, is that all of the displays that had no physically possible transparency interpretations (i.e. those in Category III) appeared perfectly coherent. It is important to keep in mind, however, that although a physically possible transparency interpretation may be a necessary condition for the perception of component motion, it is certainly not a sufficient condition. Several of the displays in Categories I and II appeared consistently coherent as well, yet all of these displays had at least one possible transparency interpretation. Thus, if perceived sliding is due to an analysis of surface transparency, then it must also be influenced by additional constraints.

A similar conclusion has been arrived at previously by Stoner and Albright (1993). They used plaids similar to the one shown in Fig. 1b with three different luminances that were arranged in two possible configurations—one that could be interpreted as narrow, transparent bars that overlapped in region **c**, and another that could be interpreted as wide, transparent bars that overlapped in region **a**. Component motion or sliding was only observed when there was a possible narrow bar interpretation of transparency. Stoner and Albright interpreted this result as being due to an additional figure-ground constraint. Owing to the larger size of **a** in comparison to **c**, they argued, observers would tend to assign **a** to the background and **c** to the foreground regardless of the luminances of these regions. Since a transparent surface must by definition appear in the foreground, this tendency would restrict the perceptual analysis to a narrow bar interpretation.

In one respect, the results of the present experiments are strongly supportive of this hypothesis. Of all the displays in Categories I and II, the ones that were most resistant to sliding were the ones that did not allow a narrow bar interpretation of transparency. There is another aspect of the results, however, which leads us to be suspicious of a figure–ground constraint as a valid explanation of why some plaids appear to slide while others do not. If one examines the various plaids in Figs 5 and 6, there are several in which the wide bars appear to stand out as a transparent surface (e.g. see **cdab** and **dabc** of Category II). If a wide bar interpretation is acceptable for static transparency, then why must it be excluded from the analysis of motion?

To summarize, out results show clearly that perceived sliding of moving plaids cannot be simply explained by the motion energy model proposed by Kim and Wilson (1993), and that the conditions for transparency may also play a role in this phenomenon. The best conditions for sliding are those that obey the Metelli/Beck rules, and which further obey the figural constraint of having narrow bars in front. However, transparency conditions are by no means sufficient to guarantee that sliding will occur. Other important factors include the orientations and velocities of the component gratings, and the prior exposure of the observer to moving patterns. Unless these conditions are in the right range, no sliding will be seen, even though the plaid may appear transparent when viewed statically.

REFERENCES

Adelson, E. H. & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *300*, 523–525.

Beck, J. & Ivry, R. (1988). On the role of figural organization in perceptual transparency. *Perception & Psychophysics*, *44*, 585–594.

Beck, J., Pradzny, K. & Ivry, R. (1984). The perception of transparency with achromatic colors. *Perception & Psychophysics*, *35*, 407–422.

Fennema, C. L. & Thompson, W. B. (1979). Velocity determination in scenes containing several moving objects. *Computer Graphics and Image Processing*, *9*, 301–315.

Grunbaum, B. & Shephard, G. C. (1989). *Tilings and patterns: An introduction*. New York: Freeman.

Heeger D. J. (1987). Model for the extraction of image flow. *Journal of the Optical Society of America A, 4*, 1455–1471.

Kanizsa, G. (1979). *Organization in vision: Essays on gestalt perception*. New York: Praeger.

Kim, J. & Wilson, H. (1993). Dependence of plaid motion coherence on component grating directions. *Vision Research, 33*, 2479–2489.

Kooi, F., DeValois, K. K., Switkes, E. & Grosof (1992). Higher order factors influencing the perception of sliding and coherence of a plaid. *Perception, 21*, 583–598.

Krauskopf, J. & Farell, B. (1990). Influence of color on the perception of coherent motion. *Nature, 348*, 328–331.

Marr, D. & Ullman, S. (1981). Bandpass channels, zero-crossings, and early visual information processing. *Journal of the Optical Society of America, 69*, 914–916.

Metelli, F. (1974). The perception of transparency. *Scientific American, 230*, 90–98.

Movshon, J. A., Adelson, E. H., Gizzi, M. S. & Newsome, W. T. (1985). The analysis of moving visual patterns. In Chagas, C., Gattass, R. & Gross, C. (Eds) *Pattern recognition mechanisms*. Rome: Vatican Press (*Pont. Acad. Sci Scr. Varia, 54*, 117–151.)

Phillips, G. C. & Wilson, H. R. (1983). Orientation bandwidths of spatial mechanisms measured by masking. *Journal of the Optical Society of America A, 1*, 226–232.

Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In Rosenblith, W. A. (Ed.) *Sensory communication*. New York: Wiley.

Shimojo, S., Silverman, G. H. & Nakayama, K. (1989). Occlusion and the solution to the aperture problem for motion. *Vision Research, 29*, 619–626.

Stone, L. S., Watson, A. B. & Mulligan, J. B. (1990). Effect of contrast on the perceived direction of a moving plaid. *Vision Research, 30*, 1049–1067.

Stoner, G. R. & Albright, T. D. (1992). The influence of foreground/background assignment on transparency and motion coherency in plaid patterns. *Investigative Ophthalmology and Visual Science (Suppl.)*, 1050.

Stoner, G. R. & Albright, T. D. (1993). Image segmentation cues in motion processing: implications for modularity in vision. *Journal of Cognitive Neuroscience, 5*, 129–149.

Stoner, G. R., Albright, T. D. & Ramachandran, V. S. (1990). Transparency and coherence in human motion perception. *Nature, 344*, 153–155.

Trueswell, J. C. & Hayhoe, M. M. (1993). Surface segmentation mechanisms and motion perception. *Vision Research, 33*, 313–328.

Wilson, H. R. & Gelb, D. J. (1984). Modified line element theory for spatial frequency and width discrimination. *Journal of the Optical Society of America A, 1*, 124–131.

Wilson, H. R., Ferrera, V. P. & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience, 9*, 79–97.

Acknowledgements—This research was supported in part by an AFOSR grant (No. F49620-93-1-0116) to James Todd. The authors wish to thank Angela M. Brown for her critical reading of earlier versions of the manuscript. The authors also wish to thank the anonymous referees, particularly Reviewer B, and the Section Editor of *Vision Research* for their helpful comments on the manuscript.