Vision Res. Vol. 33, No. 17, pp. 2515–2534, 1993
Printed in Great Britain. All rights reserved

0042-6989/93 $6.00 + 0.00

Motion Perception with Spatiotemporally Matched Chromatic and Achromatic Information Reveals a "Slow" and a 'Fast" Motion System

ANDREI GOREA,* THOMAS V. PAPATHOMAS,† ILONA KOVACS†

Received 29 June 1992; in revised form 16 December 1992

Recent reports dealing with apparent motion challenged the standard view according to which motion processing should be impossible if the visual attributes matched across space and time are processed in independent channels (the *similarity principle*). The present work examines this possibility insofar as it relates to the spatiotemporal combination of pure chromatic and pure luminance information. The data indicate that the "similarity principle" is indeed infringed at low (⩽2.5 Hz, i.e. velocities of 2.5 deg/sec for spatial modulations of 1 c/deg, in this study) but not at high (⩾7.5 Hz) temporal frequencies. The fact that colour and luminance may or may not combine to yield motion perception depending on their temporal modulation reconciliates contradictory results in the literature and supports the idea of two motion systems, a "fast"/specific one, integrating information only from similar subunits, and a "slow"/unspecific one, integrating information across dissimilar subunits (in the present case, across the chromatic and achromatic "domains"). This dichotomy is also supported by the finding that chromatic reverse-phi (i.e. with equiluminant, red and green stimuli) can be observed at medium temporal frequencies but is replaced by direct motion at low temporal frequencies, presumably within the range of the "slow"/unspecific system. Using a modified "minimum motion" technique (referred to as the Reverse-Phi equiluminance method) we present data allowing to assess the relative weights of the two systems as a function of temporal frequency.

Motion processing Colour Luminance Covariance Similarity

INTRODUCTION

The two basic assumptions used in most of the recent models of motion perception are the *covariance* and the *similarity* principles. The former, originally proposed by Reichardt (1961) is but another spelling of the fact that perceived motion can be accounted for in terms of the oriented spatiotemporal energy of the physical stimulus (Adelson & Bergen, 1985; Van Santen & Sperling, 1985). The latter relates to the specificity of the sensory information used to compute spatiotemporal covariance (or energy). Until recently, the general consensus was that the sensory units (or filters) whose outputs provide the critical information to a "generalized Reichardt-detector" (Van Santen & Sperling, 1985) are necessarily of the same type, i.e. they share the same spatiotemporal filtering properties (e.g. same size, same spatial frequency, same chromatic or achromatic characteristics, etc.). Experimental arguments favouring this "similarity principle" are essentially based on apparent motion techniques (Green, 1986; Green & Odom, 1986; Watson, 1986; Gorea & Papathomas, 1987, 1989).

There are also experimental arguments against the similarity principle some of which appear to be based on a misunderstanding. The misunderstanding relates to the domain of application of the similarity metric itself, namely physical or sensory. The use of the first alternative led a few authors (Ullman, 1980; Shechter, Hochstein & Hillman, 1989; Werkhoven, Snippe & Koenderink, 1990a, b) to conclude against the similarity metric (without rejecting the covariance principle). Indeed, the experimental finding that the spatiotemporal correspondence of, for example, two very dissimilar orientations may still yield strong motion perception (Ullman, 1980) is a strong argument against the similarity metric insofar as one refers to a *physical* similarity. However, within the sensory space, we know that the dissimilar orientations may well be processed by lower-level spatially isotropic filters whose outputs can be used to compute spatiotemporal energy (or correspondence). Accordingly, such an experimental result should not be taken as evidence against a similarity metric defined in

---

*Laboratoire de Psychologie Expérimentale, Université René Descartes and CNRS, 75006 Paris, France.
†Vision Research Laboratory, Rutgers University, NJ 0893, U.S.A.

the *sensory space* (see also Nishida & Takeuchi, 1990). On the other hand, one may consider that the similarity metric (in the sensory space) is infringed if the relevant isotropic filters represent an even higher processing stage upon which converge all the oriented subunits at a given spatial scale. Mathematically, the two schemes are equivalent and the decision concerning the infringement of the similarity principle becomes a matter of convention. Nonetheless, our present knowledge of the hierarchical structure of the visual system render the second wiring scheme quite unlikely. This is not necessarily true for visual attributes other than orientation (see below).

More serious arguments against a similarity metric in the sensory domain were offered from two different perspectives. Using an oscillatory, two-stroke apparent motion paradigm and an adjustment procedure, Cavanagh, Arguin and von Grünau (1989) showed that motion perception can be elicited by a multitude of combinations of distinct stimuli (in a sensory sense) such as colour- and luminance-defined ones. Quite recently, using an objective procedure, very carefully calibrated stimuli and a new animation technique [originally proposed by Nishida and Takeuchi (1990) and extensively used by Werkhoven *et al.* (1990a) and Werkhoven, Sperling and Chubb (1993); see below], we have reported the absence of any motion perception yielded by this same colour-luminance combination and suggested potential stimulation artifacts to account for this discrepancy (Gorea, Kovacs & Papathomas, 1992).

The second line of argument against the similarity metric is based on a criticism of the traditional stimuli used in apparent motion studies (such as those of Green, 1986; Green & Odom, 1986; Watson, 1986; Gorea & Papathomas, 1987, 1989). Werkhoven *et al.* (1990a, 1993) and Werkhoven, Sperling and Chubb (1992a, b) drew attention to the fact that a "classical" periodic apparent motion stimulus whereby two arbitrary "tokens" A and B are systematically matched across space and time in one direction but alternate in the opposite direction (see the apparent motion studies cited above) contains by necessity more energy along the homogeneous (i.e. A–A and B–B) than along the heterogeneous (i.e. A–B and B–A) path.* The typical finding obtained with such stimuli, i.e. that the homogeneous path is always preferred to the heterogeneous one, might therefore be explained by a trivial mathematical inequality. By using a stimulus configuration which avoids this inconvenience, Werkhoven *et al.* (1992a, b, 1993) could show that, contrary to previous reports by Green (1986) and Watson (1986), there is a motion system which can combine spatial frequencies more than one octave apart. Whether this finding is to be taken as evidence against the similarity principle (in the sensory domain) depends, as for the orientation attribute, on the postulated hierarchy of the underlying processing stages.

A plausible scheme is that spatially band-pass filters converge on higher-order units which will consequently behave as low-pass filters. Since under this hypothesis spatial frequency information is first processed and subsequently lost at the motion processing stage, Werkhoven *et al.*'s result may be taken as evidence against the similarity principle (in the sense defined above). The alternative hypothesis is that low-pass filtering is present at the first processing stage and that the bifocal motion detector uses directly this type of information. Werkhoven *et al.*'s result will then be compatible with the similarity metric. The two schemes are difficult, if not impossible to discriminate since they both account for motion perception across spatial frequency in terms of low-pass filtering. However, low-pass filtering cannot be invoked as a sufficient requirement to account for any type of heterogeneous motion perception. It is widely accepted that, at the first processing stages, chromatic and achromatic information are processed in parallel streams (e.g. Livingstone & Hubel, 1984). Consequently, one should conclude to an infringement of the similarity principle (as defined) any time that motion perception can be elicited by a spatiotemporally heterogeneous stimulus combining colour and luminance information.

In the last few years, two of us developed and used a family of spatiotemporally discrete, apparent motion stimuli permitting the evaluation of the relative contributions to motion perception of chromatic and achromatic information (Gorea & Papathomas, 1987, 1989; Papathomas & Gorea, 1988; Papathomas, Gorea & Julesz, 1991). The interpretation of some of these results is critically dependent on the standard assumption of an underlying similarity metric for motion perception. In the particular case of luminance vs chromatic information, the similarity assumption states that the two types of information *cannot* combine to produce motion perception. The present study was meant to test this hypothesis. The data presented in a partial report of the present study (Gorea *et al.*, 1992) and obtained at relatively high temporal rates strongly supported the similarity assumption. Additional data obtained at much lower temporal rates (close to those used in Cavanagh *et al.* and in Werkhoven *et al.*'s studies) reject, however, the similarity assumption. We now present the whole set of experiments and argue in favour of two distinct motion systems, a "fast"/specific and a "slow"/unspecific one (which may or may not match the classical short- and long-range dichotomy; see Cavanagh & Mather, 1989). Based on the characteristic behaviour of the "fast" system and on the luminance reverse-phi phenomenon (Anstis, 1980) we propose a new method for assessing the relative strength of the two systems. Finally, we present a few observations on the reverse-phi phenomenon with both chromatic and achromatic stimuli which support the "fast"/"slow" dichotomy in motion processing.

---

*If A and B are intensities (or contrasts), the overall energies along the homogenous and heterogeneous paths are $A^2 + B^2$ and 2AB, respectively. It is always true that $A^2 + B^2 \geqslant 2AB$.

GENERAL METHODS

Stimuli

The stimuli were vertical square-wave luminance and/or chromatic gratings displayed on a Sony Trinitron monitor (GDM 1601/1950) driven by an Adage PG-90/10 graphic card under the control of a LEANORD-386 AT computer. The Adage card provides 8 bits modulation (256 levels) per gun. The individual bars composing a stimulus could be red, green and/or yellow with CIE x and y coordinates (0.611, 0.353) for red, (0.285, 0.597) for green and (0.448, 0.475) for yellow as measured with a Minolta Chroma Meter CS-100. They were always displayed on a yellow background set at 20 cd/m^2. Their width subtended 0.25 deg of visual angle at 114 cm from the observer. The red and green bars could be equiluminant between themselves and relative to the yellow background or not. The luminance of the yellow bars was always above or below the luminance of the background. The cone contrast of the equiluminant red/green modulation (see Cole & Hine, 1992) depended on the specific equiluminant settings of each observer and was close to an average of 28 or 22% depending on whether the contribution of the S cones was or was not considered, respectively.

The spatial configuration of the "gratings" was unconventional so as to permit a variety of spatiotemporal modulations, as described below. With one exception (see below), the spatial (and temporal) sequence of red, green and/or yellow bars in the stimuli defined a four-stroke period. Thus the maximum drift rate that can be obtained for such stimuli with a 60 Hz video raster (i.e. 16.6 msec/frame) is 15 Hz (i.e. with zero ISI). Directional performances were measured for drift rates of 2.5, 3.75, 7.5 and 15 Hz. The first three modulations were obtained by increasing the number of "stimulus (ON) frames" and/or of "background (OFF) frames". Most of the data presented in the remainder were obtained with balanced ON/OFF duty-cycle ratios (1/1, 2/2 and 3/3 for modulations of 7.5, 3.75 and 2.5 Hz, respectively; by necessity, the 15 Hz modulation requires an 1/0 ON/OFF ratio). Other duty-cycle ratios have also been used yielding, with one exception, practically identical results.

The fundamental spatial frequency of the square gratings was 1 c/deg. The stimulus used to assess equiluminance by means of heterochromatic flicker photometry (HFP) had a spatial and temporal period of two. It was displayed at a temporal rate of 30 Hz. The grating stimuli subtended a 6.5 × 6.5 deg area at 114 cm from the observer with the yellow background extending over 17 × 13.5 deg.

General procedure

With the exception of the HFP experiments where the luminance of one colour was adjusted to minimize flicker perception, observer's task in all the experiments described below was to specify the perceived direction of the stimulus (i.e. leftward vs rightward). Each stimulus presentation consisted of eight "stimulus frames" (i.e. two temporal periods). Datum points were obtained by means of a two-alternative forced-choice (2AFC) procedure with constant stimuli and were expressed as percentages correct, if there was a "correct" answer, or as "preference" percentages for the direction of a given "motion-carrier". With the exception of Expt 3, percentages "correct" were obtained as a function of the luminance contrast of either the chromatic (red and/or green) or "achromatic" (i.e. yellow) bars relative to the luminance of the yellow background. Luminance contrasts used in the main experiments were always estimated in preliminary experiments by the method of adjustment so as to bracket the directional threshold or the point of subjective equality. Luminance contrasts of yellow lights were measured as the ratio between the luminance increments (or decrements) relative to the yellow background luminance. For red and green lights, luminance contrasts were measured as above but with reference to the red and green lights equiluminant with the yellow background (see below). In Expt 3 chromatic and luminance contrasts were fixed.

In each experimental session, percentages "correct" were always computed out of 50 trials per datum point. Each session was repeated at least three times so that final performances were computed out of at least 150 trials per condition. More typically, sessions were repeated four times (200 trials/condition) and in some cases up to 16 times (800 trials/condition).

The first author, an emmetrope, served as an observer in all experiments. Two naive but well trained graduate students (females) served as observers in the most critical conditions. Their vision was corrected to normal. Vision was binocular with natural pupils.

Initial equiluminance settings

Since the demonstration that chromatic contrast may or may not combine with luminance contrast is critically dependent on the absence of all luminance cues in the chromatic stimulus, particular care was devoted to the HFP procedure. The stimuli were vertical gratings composed of red–green, red–yellow or green–yellow bars of the same spatial extent as those used in the main experiments. The bars were modulated in counterphase at a temporal frequency of 30 Hz. Observer's task was to adjust the luminance of one of the two types of bar used in the stimulus in order to minimize flicker perception. The initial luminance level of the adjustable field was randomly set above or below the minimum flicker range (MFR) at an arbitrary level.

Given that we were interested in assessing red and green equiluminant points relative to a yellow light and that the latter was to be obtained by summing equiluminant red and green lights, their relative efficiency, $E\,(E = L_G/L_R)$, was estimated by means of a partially recurrent HFP procedure. The progressive estimation of E was terminated when it did not vary from one assessment to another. For each new assessment of E, the observer produced five adjustments of both the upper and the lower bounds of the MFR. The equiluminance point was taken as their arithmetic mean. The step-by-step estimations of E proceeded as follows.

Step 1. The luminance of the green bars, L_G, was set at 20 cd/m^2 while the luminance of the red bars, L_R, was adjusted for minimum flicker ($L_{G1} = 20$, $L_{R1} =$ adj.). E_1 was defined as the ratio between L_{G1} and the midpoint of the adjusted MFR, L_{R1}.

Step 2. Using the above notation: $L_{R2} = 20$, $L_{G2} =$ adj.; $E_2 = L_{G2}/L_{R2}$.

Steps 3–4. Yellow bars were generated so that $L_Y = L_{R3} + L_{G3} = 20$ cd/m^2 and $L_{G3}/L_{R3} = (E_1 + E_2)/2 = E_3$. L_{R4} and L_{G4} were separately adjusted to match L_Y; $E_4 = L_{G4}/L_{R4}$.

Step 5. A new yellow was generated satisfying the conditions $L_{Y'} = L_{R5} + L_{G5} = 20$ and $L_{G5}/L_{R5} = E_4$ ($= E_5$). Again, L_{R6} and L_{G6} were separately adjusted to match $L_{Y'}$. At this point $E_4 = E_6 = 1.156$ (averaged across the three observers) and the procedure was completed. The mean luminances of the red and green phosphors yielding the measured equiluminance with a 20 cd/m^2 yellow were 18.55 and 21.45 cd/m^2, respectively. Since these values are the midpoints between the lower and higher bounds of the MFR, they do not correspond necessarily to actual values obtainable on the 8 bits/phosphor Adage card used in these experiments.

MFR (averaged across the successive steps of the procedure) expressed as a percentage of its mean was 34.3% (5.5 cd/m^2) for red and 24.5% (5.9 cd/m^2) for green. Despite the rapid convergence of the E estimations, the chances that the equiluminance condition might have not been achieved by just taking the midpoint of these very large MFRs are to be considered. Moreover, *a priori* negligible differences (at most 1.1%) between the computed midpoints and the actual luminance values obtainable with the 8 bits/phosphor Adage card may also introduce luminance artifacts. To avoid them, the critical experiments were also run under conditions of spatiotemporal *luminance noise*.

Notice that red and green elements were never simultaneously present in one stimulus frame and that, with the exception of Expt 3, they were never present together in one stimulus sequence either. Thus, chromatic aberration related artifacts were unlikely with most of the present stimuli. In any event, manipulation of the luminance noise presumably counteracts them.

EXPERIMENT 1. HOMOGENEOUS (CHROMATIC-CHROMATIC) VS HETEROGENEOUS (CHROMATIC-ACHROMATIC) MOTION PATHS

Equivalent luminance contrast

One of the convenient things about a covariance metric is that the perceptual "strength" of a given moving component (in a complex stimulus) can be directly predicted by multiplying the activities elicited by the moving "tokens" along the motion path (see Adelson & Bergen, 1985; Werkhoven *et al*., 1992a, b, 1993). To the extent that these activities are linearly related to the contrast of these tokens, the energy along (and thus the perceptual strength of) the given path is readily quantifiable. For a system which ignores the similarity metric, directional perception based on the coherent spatiotemporal matching of chromatic and luminance gradients should override the directional perception based on only chromatic (or only luminance) gradients any time that the covariance product in the former case is larger than in the latter. In order to check this hypothesis, one should first be able to express the chromatic contrast in terms of *equivalent luminance contrast* (EqLC). This was achieved by means of a motion cancellation technique (Anstis, Cavanagh, Maurer, Lewis, MacLeod & Mather, 1986) slightly modified to allow separate estimation of the EqLC for red/yellow and green/yellow chromatic contrasts (Agonie & Gorea, 1993). The stimulus configuration used in this preliminary experiment and schematically illustrated in Fig. 1 has the same spatial and temporal period (i.e. 4) as all the stimuli used in the main experiments.

In Fig. 1 the "motion cancellation" stimulus is displayed in a space–time plane with each row showing one stimulus frame. Each stimulus (or ON-) frame was presented for a variable number of raster frames (i.e. $N \times 16.6$ msec). The interstimulus interval was null (i.e. 0 OFF-frames). The durations of the ON-frames were set to yield 2.5, 7.5 and 15 Hz modulations (i.e. 6, 2 and 1 raster frames, respectively).

In Fig. 1, the chromatic contrast (CC), is systematically matched to the right (i.e. rightward motion) while the luminance contrast (LC) is matched to the left (leftward motion). In the actual experiments, the directions of the spatiotemporal matchings were randomly swapped. The 20 cd/m^2 yellow background is shown as an homogeneous grey. The "CC" patches (vertical bars in the actual experiment) could be red or green and were set at the equiluminance point (with respect to the yellow background). The "LC" patches were yellow with a variable, positive LC with respect to the background.

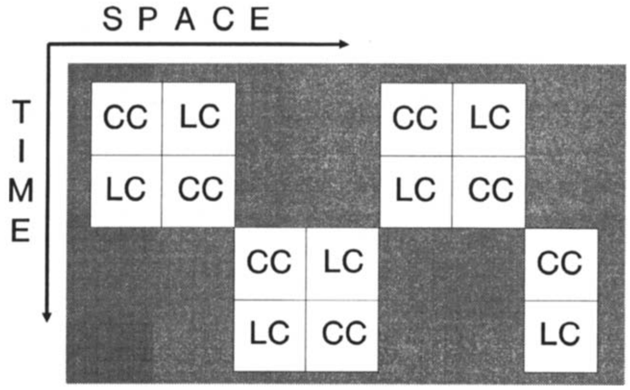

FIGURE 1. Schematic spatiotemporal representation of the stimulus used to assess the *equivalent luminance contrast* (EqLC) of a given chromatic contrast, CC. Each row represents one stimulus frame. In the experiments, one stimulus frame consisted of CC (red or green) and LC (luminance contrast; dark or bright) bars on a yellow background (shown here in grey) equiluminant with the CC elements. Both the spatial and temporal periods of this stimulus are equal to four. In the illustration, the chromatic spatiotemporal path (CC-CC) is oriented to the right while the luminance spatiotemporal path (LC-LC) is oriented to the left. CC-CC and LC-LC paths were randomized in the actual experiments.

(Control experiments demonstrated that the measured EqLC is independent of the sign of the "LC" elements.) "Chromatic motion preferences" (percentages) were measured as a function of LC for red/yellow and green/yellow CCs in independent, randomized sessions. Each session was repeated four times so that each experimental point shown in Fig. 2 is based on 200 trials. They were fitted with psychometric functions of the form $p = 1 - 2^{[(C/\alpha)^{\beta}]}$ (with C, the contrast, α, the threshold and β, the slope; Quick, 1974) by means of a procedure

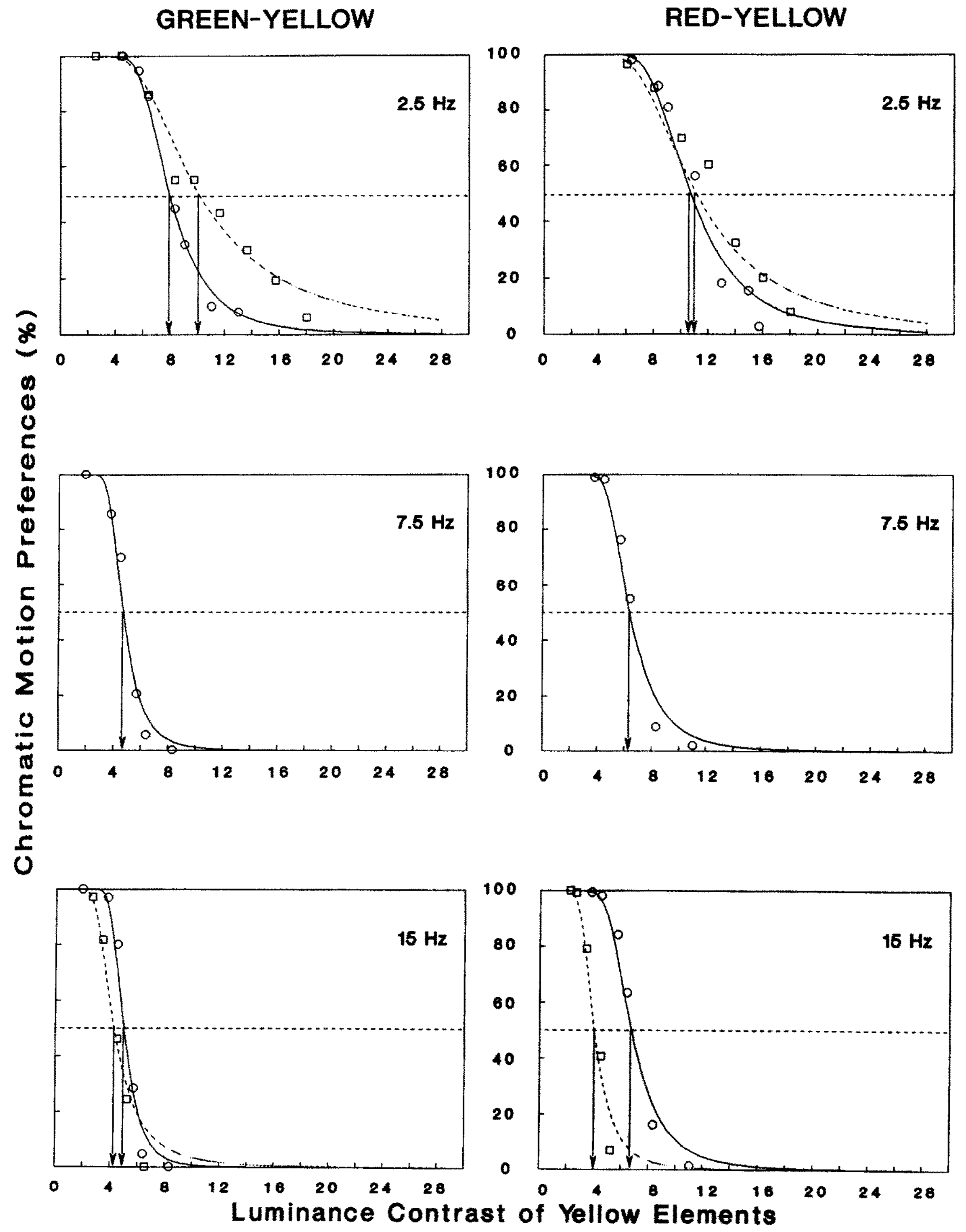

FIGURE 2. Percentages of chromatic motion preferences (CC-CC path) measured with the stimulus of Fig. 1 as a function of the positive luminance contrast (LC+) of the "LC" elements at three temporal frequencies (increasing from top to bottom). Circles and squares are for observers AG and CA. Data obtained with green–yellow and red–yellow CCs are shown in the left and right panels, respectively. The continuous and dashed sigmoids are psychometric functions fitted for each observer. (CA was not run at 7.5 Hz). Arrows show the EqLCs, namely the luminance contrasts yielding 50% CC-CC (or LC-LC) directional preferences.

adapted from (Watson, 1981) with both α and β as free parameters. The EqLC was defined as the LC producing 50% chromatic motion preferences.

The data and the fitted psychometric functions are shown in Fig. 2 for two observers (AG and CA, circles and continuous curves and squares and dashed curves, respectively; CA did not run the 7.5 Hz condition). EqLCs measured at 2.5 Hz are significantly higher (by a factor of about 1.65 and 2.5 for AG and CA, respectively) than EqLCs measured at 7.5 and 15 Hz (see also Cavanagh & Anstis, 1991). For the two highest frequencies, the overall red/green EqLC (i.e. the sum of red/yellow and green/yellow) is about 11.3 and 9.3% for AG and CA, respectively, well within the range of previously measured EqLCs (Anstis *et al.*, 1986; Agonie & Gorea, 1993). At 2.5 Hz, the overall red/green contrast is about 18.5 and 20% for AG and CA, i.e. much closer to the cone contrast of our particular red/green modulation (i.e. 22% under the assumption that the S cones do not contribute to the luminance signal). This suggests that, at low temporal frequencies, colour- and luminance-based motion information are about equally efficient (see Agonie & Gorea, 1993). AG displays a noticeable asymmetry between the red/yellow (2.5 Hz: 10.6%; 7.5 and 15 Hz: 6.5%) and green/yellow (2.5 Hz: 7.9%; 7.5 and 15 Hz: 4.8%) EqLCs. This suggests that the adapting background was slightly off the unique yellow for this observer. In contrast, CA's settings are quite symmetrical about yellow.

The "Werkhoven" experiment

Having assessed the EqLC for both red and green modulations on an equiluminant yellow background, we are now in position of testing the main claim of the "similarity principle". Figure 3 illustrates the "Werkhoven" stimulus (Werkhoven *et al.*, 1990a, 1992a, b, 1993) in the space–time plane. As above (see Fig. 1), CC and LC stand for chromatic (red or green) and luminance (negative or positive) contrast modulations, while the grey area represents the 20 cd/m^2 yellow background. Also as above, stimulus drift rates could be 2.5, 7.5 or 15 Hz obtained with a balanced duty–cycle ratio. (Practically identical results were obtained with unbalanced duty–cycle ratios.)

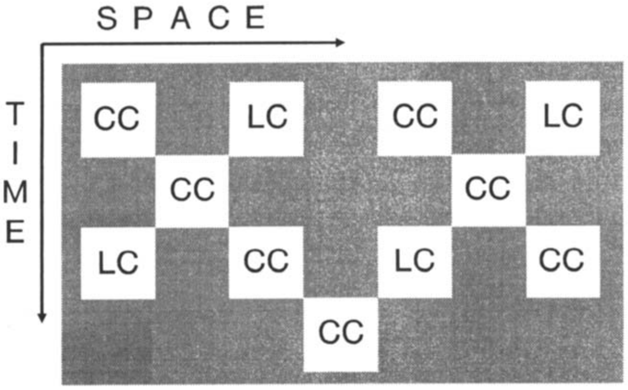

FIGURE 3. Schematic spatiotemporal representation of the "Werkhoven-stimulus". All details are as for Fig. 1. Notice the "conflicting" motion paths: to the right, the CC-CC, homogeneous path and to the left, the LC-CC, heterogeneous path. According to the "similarity hypothesis" (see text) the LC-CC path should never override the CC-CC path. If the similarity metric is irrelevant, the LC-CC path should take over the CC-CC path any time that the LC × CC energy product is larger than the CC × CC product (in equivalent luminance contrast units—a covariance metric).

The logic of this stimulus is as follows. CC modulations are set at the equiluminance point with the background, while LC is variable. If directionality is determined by the "largest energy" product across frames *independently of the nature of the signals* being processed (i.e. chromatic or luminance), the CC-CC (homogeneous) path should dominate any time that the absolute LC is below the EqLC of a given CC modulation. Any time $|LC| > |EqLC|$, the LC-CC (heterogeneous) path should dominate. If performances "correct" are expressed in terms of CC-CC path preferences, they should be above 50% for $|LC| < |EqLC|$ and below 50% for $|LC| > |EqLC|$.

As mentioned in the General Methods section, the equiluminance points obtained by means of the HFP procedure do not necessarily correspond to the precise values that can be obtained with our 8 bits/phosphor Adage card. For the red phosphor and AG, the measured equiluminant point (with respect to the 20 cd/m^2 yellow background) was 17.89 cd/m^2. The closest obtainable value was 18.08 cd/m^2 value, i.e. 1.06% error. For the green phosphor, the measured equiluminant point was 20.82 cd/m^2 and the closest obtainable value was 20.76 cd/m^2 value, i.e. 0.3% error. To counteract these and other potential luminance artifacts, *luminance noise* was introduced as a parameter. It was defined as the spatial and temporal random variation of the luminance of all (i.e. "LC" and "CC") stimulus elements within a predefined range around their nominal LC value. Given the technical limitations of our graphic card, only three luminance noise levels, within a given noise range, could be presented per frame. The noise range could be 0% (i.e. no noise), ±5% or ±15% variation. These three noise ranges were used only with AG at 15 Hz modulation. For the remaining two drift rates, this observer was run under 5% noise conditions. CA was run under 0% noise conditions at 15 Hz and under 5% noise conditions at 2.5 Hz.

The data shown in Figs 4 (AG) and 5 (CA) were collected as above, i.e. by means of a constant stimuli, 2AFC procedure with each session being run with a set of five (out of 15 or 13, for AG and CA, respectively), randomly presented LCs, with a fixed CC and a fixed luminance noise range. The set of five LCs in each session, CC (red or green) and luminance noise (0, 5 or 15%) were randomly varied across sessions.

Figures 4 and 5, display "chromatic motion preferences" as a function of LC. In the two figures, the vertical lines on both sides of 0 LC represent ±EqLCs for green (left panels) and red (right panels) CCs. A covariance metric infringing the similarity principle predicts that colour-carried motion preferences (i.e. the CC-CC path) should be above 50% for $|LC| < |EqLC|$

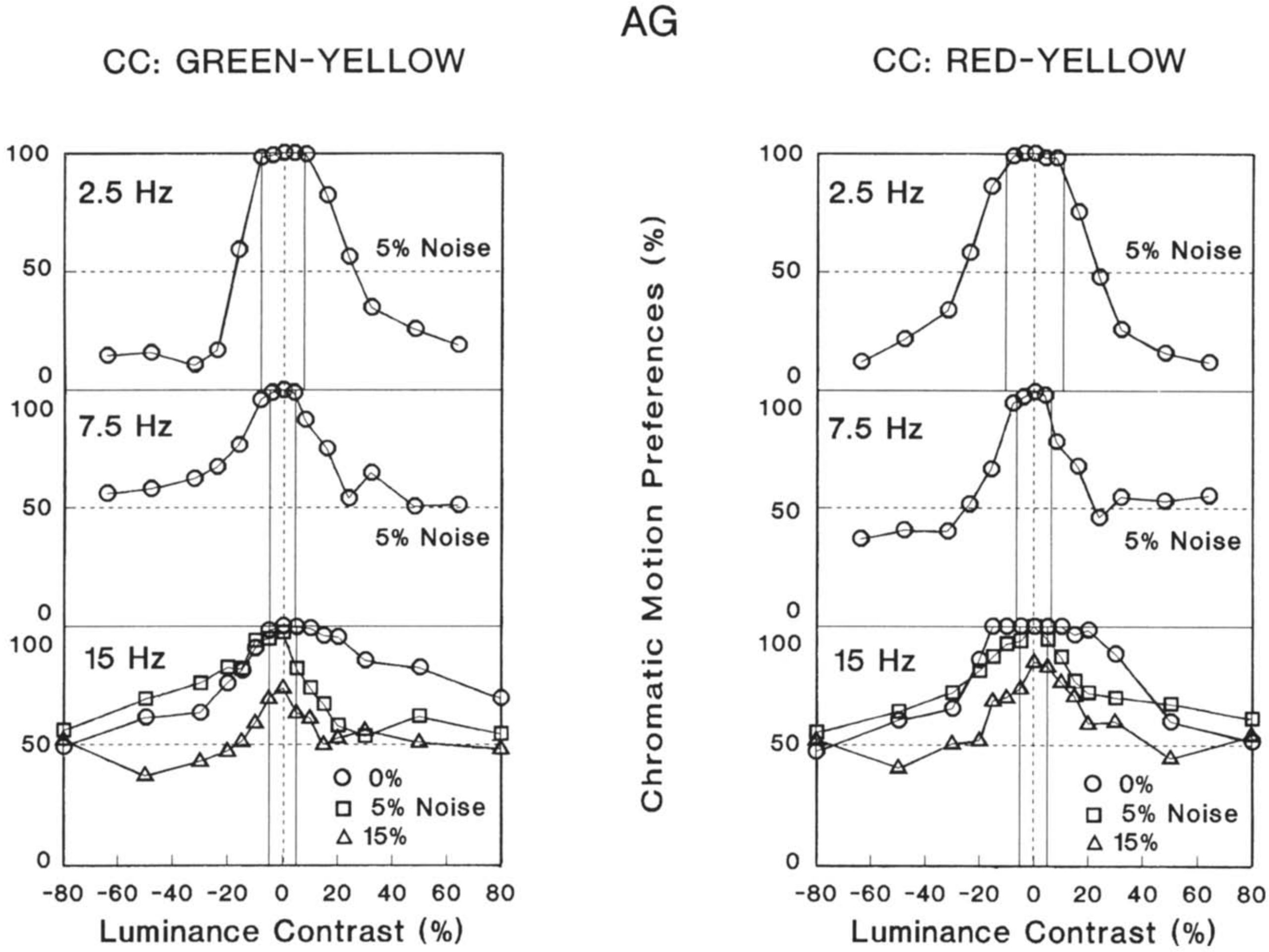

FIGURE 4. Percentages of chromatic motion preferences (CC-CC path) measured with the stimulus of Fig. 3 as a function of the luminance contrast (LC) of the "LC" elements at three temporal frequencies (increasing from top to bottom). The observer is AG. Data obtained with green–yellow and red–yellow CCs are shown in the left and right panels, respectively. "CC" elements were equiluminant with the yellow background. Circles, squares and triangles show performances obtained under 0%, 5 and 15% luminance-noise conditions. Vertical lines symmetrical about 0% LC show the EqLCs from Fig. 2. The horizontal dashed lines show the 50% preferences.

and below 50% for |LC| > |EqLC|. Clearly, this prediction is not verified at any of the three studied temporal drift rates. At 2.5 Hz and for LCs about twice larger than the measured EqLCs, the heterogeneous path is indeed preferred to the homogeneous one. This observation provides strong evidence *against* the similarity principle while suggesting, as a first interpretation, that the efficiency of the "heterogeneous motion system" is less than expected on the basis of a straightforward covariance metric. At higher temporal rates, however, the similarity principle appears to be respected since, within the range of the studied LCs (up to ±80%, i.e. for LCs more than 12–20 times the estimated EqLCs) the hetero-path never overrides the homo-path (i.e within experimental error, performances are never below 50%). Thus, colour and luminance information *do* combine to yield motion perception at low temporal rates but appear to be processed independently at medium to high temporal rates. Increasing the luminance noise range from 0 to 15% (Fig. 4, AG, 15 Hz) decreases the LC range over which CC-CC motion is perceived but does not favour the relative strength of the LC-CC path.

Taken together, the data suggest the existence of two motion systems, a "slow" system which operates at low drift rates and infringes the similarity principle and a "fast" system optimally activated at higher temporal rates and respecting the similarity principle. The latter is necessarily a *chromatic* system since it responds to the equiluminant CC-CC path. The fact that, at low temporal rates, the heterogeneous motion takes over the homogeneous motion at LCs higher than expected on the basis of a pure covariance metric does not necessarily mean that the "slow" system infringes such a covariance metric. While both "slow" and "fast" systems may be active along the homo-path, only the former should be activated by the hetero-path. Thus, the hetero-path should take over the homo-path when the covariance product within the "slow" system responsive to one direction out-weighs the *sum* of the covariance products within both "slow" and "fast" systems responsive to the opposite direction. Along this line of argument, the additional LC required to reverse the perceived direction of motion may be taken as an estimate of the relative strength of the "fast" system at low temporal rates. This point will be expanded in the next section.

The data displayed in Figs 4 and 5 were obtained after the careful assessment of equiluminance described in the Methods. However, a very different picture was initially obtained subsequent to a simplified (i.e. not recurrent) equiluminance assessment procedure having yielded slightly different equiluminance values. To illustrate the type of artifactual data one may obtain with the "Werkhoven" stimulus if equiluminance is not properly assessed, colour-carried motion preferences of AG were measured once more at 15 Hz with 0% noise for green and red modulations slightly off their equiluminance point relative to the yellow background (±1 bin out of 256 bins, for the green gun—yielding ±1.6% luminance contrast variation, and ±3 bins for the red gun—yielding ±3.3% luminance contrast variation; remember that, like many of the current graphic systems used in the related literature, the Adage graphic card used here allows 8 bits modulation per gun). These "artifactual" data are shown in Fig. 6.

As in Figs 4 and 5, vertical lines on either side of 0% luminance contrast show ±EqLC points. Solid and open symbols show directional preferences with "CC" elements set at luminances below and above the actual equiluminance point, respectively. The data are strikingly different from those presented in Fig. 4. In particular, colour-carried motion preferences may reach 0% for either negative or positive LCs depending on whether the equiluminant point was under- or overestimated, respectively. From a methodological point of view, the main observation here is that, for green/yellow CCs, a deviation from the equiluminance point of only 1 bin (out of 256) is sufficient to completely distort the aspect of the data and to lead to conclusions opposite to those drawn from Fig. 4. In particular, the data shown in Fig. 6 appear to reject the similarity metric since, as expected on this account, the hetero-path appears to be preferred for a wide range of LCs. However, the drop of performances below 50% can be easily accounted for in terms of the LC-LC matching which overrides the CC-CC matching: when the LC of the yellow bars has the *same sign* as the artifactual LC added to the CC bars, luminance-carried motion will take over the colour-carried motion (i.e. performances will drop below 50%) any time the LC × LC product is larger than the CC × CC product expressed in EqLC units. This prediction is more or less verified when one considers the estimated EqLCs for both green/yellow and red/yellow chromatic modulations and the LC artifacts added to each of them.

The 100% branches of the colour-carried motion preference functions shown in Fig. 6 are due to a combination of two factors. The first is directly related to the CC × CC product which is the basis of the colour-carried motion *per se*. The second is very probably related to a *reverse-phi* phenomenon (Anstis, 1980). Indeed, note that these branches occur when the LC added to the chromatic bars and the LC of the yellow bars are of opposite signs, i.e. the required condition for reverse-phi (see Expts 2 and 3). The data show a more or less pronounced tendency to approach 50% performances for very high ±LCs of the "LC" elements. Under these conditions, the ratio between the luminance

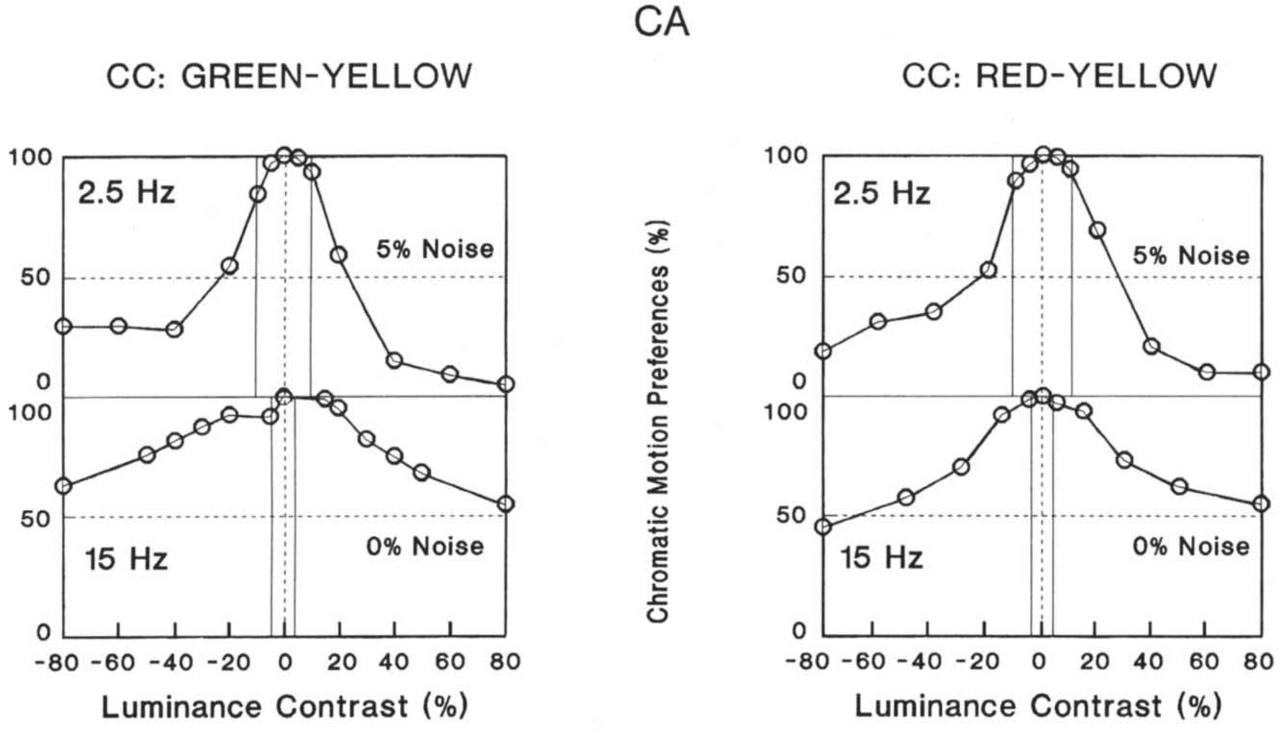

FIGURE 5. Same as Fig. 4 but for observer CA. Only two temporal frequencies were tested.

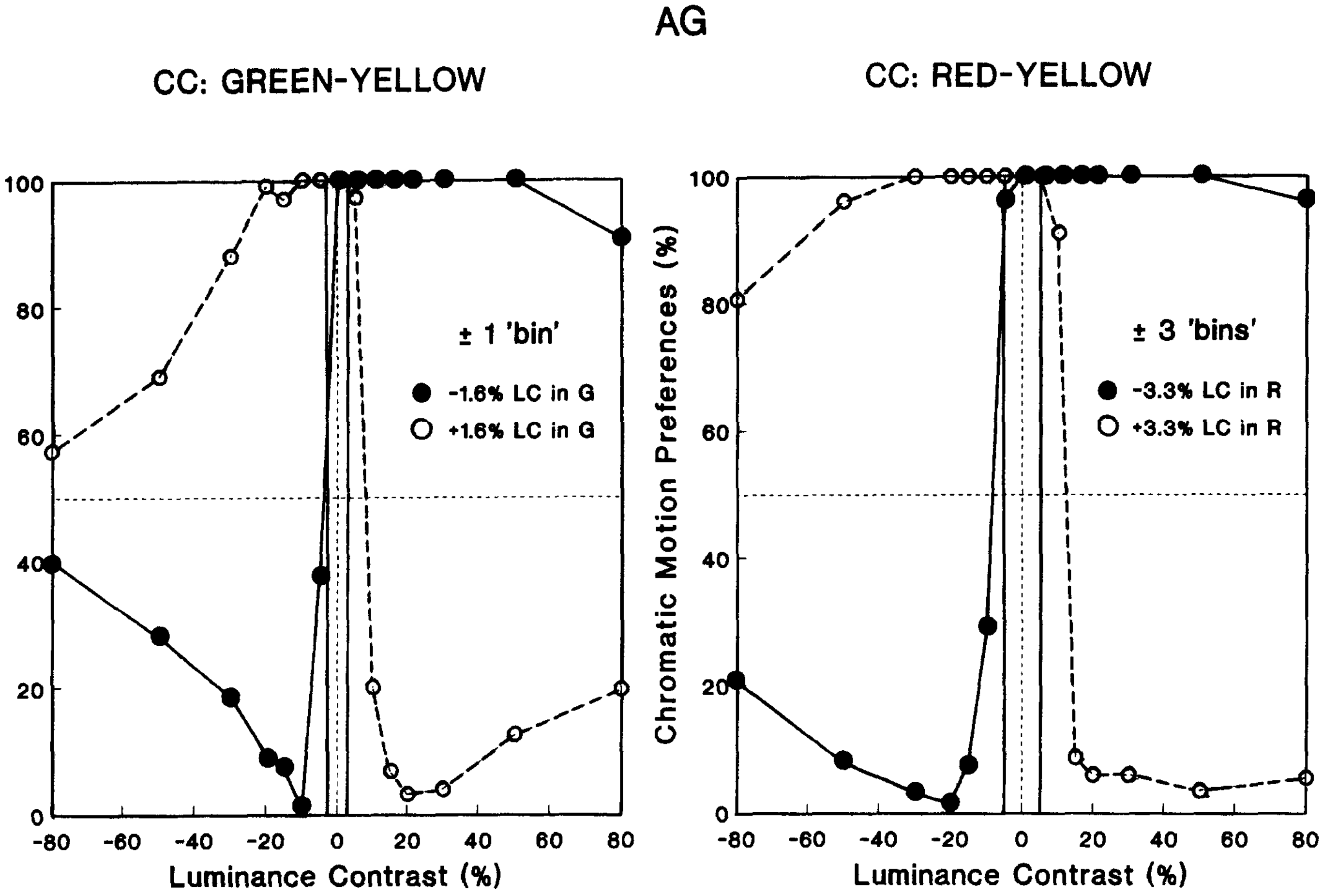

FIGURE 6. Same as Fig. 4 (AG) but with "CC" elements slightly off their equiluminant point relative to the yellow background. Data were obtained at 15 Hz. Solid and open symbols show chromatic motion preferences when the "CC" elements are set at negative and positive luminance contrasts relative to the background, respectively. The negative and positive luminance contrasts were obtained by adding or subtracting 3 or 1 bins (out of 256) to the equiluminant bit-values of the green (+1.6%) or red (±3.3%) guns.

contrasts in the "LC" and "CC" elements is very high and the overall spatiotemporal oriented energies along the LC-LC leftward and rightward paths (see Fig. 3) are quite comparable. It is likely that, as a consequence, ambiguous luminance-carried motion (i.e. counterphase modulation) will prevail over directional perception.

To conclude, artifactual luminance contrasts as small as 1.6% (1/256 error in the green gun) or 3.3% (3/256 in the red gun) may lead to very different conclusions related to the similarity principle. Errors of this kind may easily occur due both to inaccuracies in the standard HFP procedure (minimum flicker ranges of about 30%—see General Methods) and to the limitations of 8-bits graphic cards.

EXPERIMENT 2. ADDITIONAL EVIDENCE FOR TWO MOTION SYSTEMS AND THEIR RELATIVE WEIGHTS ASSESSED WITH THE REVERSE-PHI PROCEDURE

Rather than using the relatively elaborated "Werkhoven" stimulus configuration illustrated in Fig. 3 to test the similarity principle, one might have taken advantage of the simpler configuration illustrated in Fig. 7. Under the "similarity hypothesis", if the "CC" elements are set at the "real" equiluminance point (with respect to the yellow background), the human observer should never perceive a coherent motion whatever the luminance of the "LC" elements. In a left/right response task, observers' performances should be at 50%. If this happens to be the case, i.e. implying that luminance and chromatic contrasts do not combine to yield motion perception, the stimulus configuration shown in Fig. 7 could be used to

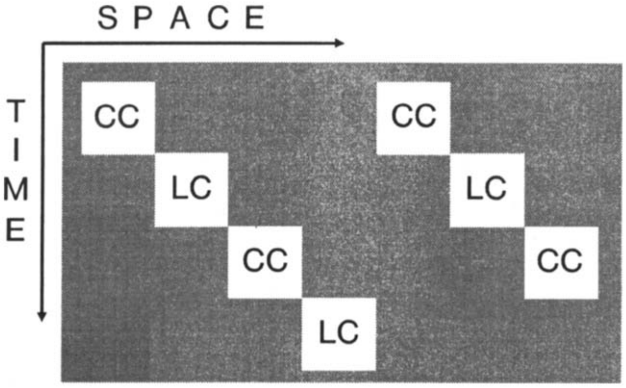

FIGURE 7. Schematic spatiotemporal representation of the stimulus used in Expt 2 (the reverse phi method). All details are as for Fig. 1.

assess equiluminance itself. Indeed, any luminance information added to the "CC" elements should bias observers' performances toward 0 or 100% depending on the *signs* of the LCs of the "CC" and "LC" elements. If the two LCs are of the *same sign*, luminance-carried motion should be perceived along the direction of the physical displacement of the stimulus (i.e. rightward in Fig. 7) and performances should be biased toward 100% "correct". If the two LCs are of *opposite signs*, luminance-carried motion should be perceived along the direction opposite to the physical displacement of the stimulus given the reverse-phi phenomenon (i.e. leftward in Fig. 7) and performances should be biased toward 0% "correct". The LC of the "CC" elements yielding exactly 50% "correct" responses should then be taken as the equiluminant point. The similarity principle will be verified if this point coincides with the equiluminant point assessed by means of HFP. Hereafter, this new method of assessing equiluminance will be referred to as the reverse-phi (RP) method.

Three observers, those from the previous experiments and a third naïve observer (SJ), were used in this experiment. The LC of the "LC" elements was fixed at +5% or −5%. The luminance of the "CC" elements was variable and spanned a range crossing the equiluminance point as measured by means of HFP. Close to this point, the luminance step was the smallest possible on a 8 bits graphic card (i.e. about 1.6 and 1.1% for the green and red phosphors, respectively). Larger steps were used for luminances far away from the equiluminant point. Left/right responses were measured by means of a 2AFC procedure whereby up to five predefined luminances of the "CC" elements (out of as many as 10) were randomly presented within each session. Directional performances were assessed at four temporal frequencies (2.5, 3.75, 7.5 and 15 Hz) with different duty–cycle ratios. Data are presented only for balanced duty–cycles (excepting the 15 Hz modulation where the ON/OFF duty-cycle ratio was 1/0). With one exception (see below), performances did not depend on the duty–cycle ratio. Chromatic contrasts (green– and red–yellow), temporal frequencies and the specific luminances of the "CC" elements within one session were randomly varied across sessions. Each session was repeated at least three times such that each datum point presented in Figs 8, 9 and 10 (for AG, SJ and CA, respectively) is based on at least 150 trials. In these figures percentages "correct" indicate directional preferences coinciding with the physical displacement of the stimuli, i.e. the CC-LC hetero-path (rightward in Fig. 7). Data obtained with "LC" elements of positive and negative LCs ($\pm\mathrm{LC}_{\text{"LC"}}$) are shown as open and closed symbols, respectively. Sigmoids are psychometric functions fitted by Watson's (1981) algorithm with both α (the "threshold" or "point of subjective equality", PSE, at 50% "correct") and β (the slope) as free parameters. Performances obtained with green–yellow and red–yellow CCs are shown in the left and right panels of each figure, respectively. Temporal frequency increases from top to bottom. The data display the following main characteristics.

(1) As a general observation, we note the trend of the psychometric functions obtained with positive and negative "LC" elements to intersect precisely at (or very close to) the 50% "correct" point for high temporal frequencies (7.5 and 15 Hz) and to progressively spread apart at low temporal frequencies (2.5 and 3.75 Hz). This is true for the three observers.

(2) Not only are the PSE obtained at high temporal rates (7.5 and 15 Hz) with $\pm\mathrm{LC}_{\text{"LC"}}$ practically identical, but they are also very close to the equiluminance points obtained with the HFP method: differences averaged over red–yellow and green–yellow equiluminant points are less than 0.6, 2.8 and 3.1% for AG, SJ and CA, respectively. At medium to high temporal rates, the RP method can thus be used to assess equiluminance. Moreover, it provides significantly smaller uncertainty luminance ranges than the HFP procedure: the psychometric functions shown in Figs 8–10 cover the 0–100% "correct"-responses range for luminance variations of <5% while the minimum flicker uncertainty range with the HFP procedure was larger than 24% (see Methods).

(3) With the exception of the 15 Hz modulation, Figs 8–10 display only data obtained with balanced temporal duty–cycles (1/1, 2/2 and 3/3 for 7.5, 3.75 and 2.5 Hz modulations). With one exception (see below), other duty–cycles provided practically identical results.

(4) Within the studied luminance range, the RP method cannot provide PSEs for the 2.5 Hz and, in a few cases (see Figs 8 and 9), for the 3.75 Hz modulations: the luminance reverse-phi direction cannot override the hetero-path even for luminance contrasts as high as 40% (i.e. the fixed 5% $\mathrm{LC}_{\text{"LC"}}$ added to about 35% $\mathrm{LC}_{\text{"CC"}}$, computed in reference to the equiluminant point measured at 15 Hz). The consistent choice (frequently close to 100%) of the hetero-path at 2.5 Hz is direct evidence that chromatic and luminance information *do* combine at low temporal rates corroborating the results obtained with the "Werkhoven" stimulus (Expt 1). For this slow modulation rate, the same pattern of results was obtained for the three observers when tested with an ON/OFF duty–cycle of 1/5. The situation was different, however, for a 6/0 duty–cycle. For AG and CA (but not for SJ) the measured psychometric functions did cross the 50% point allowing for performances close to 0% (i.e. reverse-phi wins), although the crossing point was significantly displaced away from the equiluminant point assessed at 15 Hz (see below). This duty–cycle effect must be related to the spectral composition of the stimuli in the time domain differentially activating first- and second-order mechanisms (e.g. Georgeson & Harris, 1990; Nishida & Sato, 1992b).

Observations (2) and (4) have two strong implications. First, they indicate that the hypothetical mechanism which integrates chromatic and achromatic information along the heterogenous path does not display a (luminance) reverse-phi behaviour, or if it does, this type of response must be very weak (see Expt 3). Presumably, at this processing stage, positive and negative luminance contrasts are rectified before being integrated. A mechanism of this kind has already been suggested by a number of authors and was referred to as a "contrast (or second-order) motion mechanism" (Chubb & Sperling, 1988, 1989; Derrington & Badcock, 1985; Derrington, Badcock & Holroyd, 1992; Nishida & Sato, 1992a, b; Werkhoven *et al.*, 1990a, 1993). Second, the question still remains whether the classical "luminance, first-order motion mechanism" (i.e. a mechanism with a linear behaviour before the multiplication, or energy extraction stage, and displaying therefore a reverse-phi behaviour) is active at all within this temporal frequency

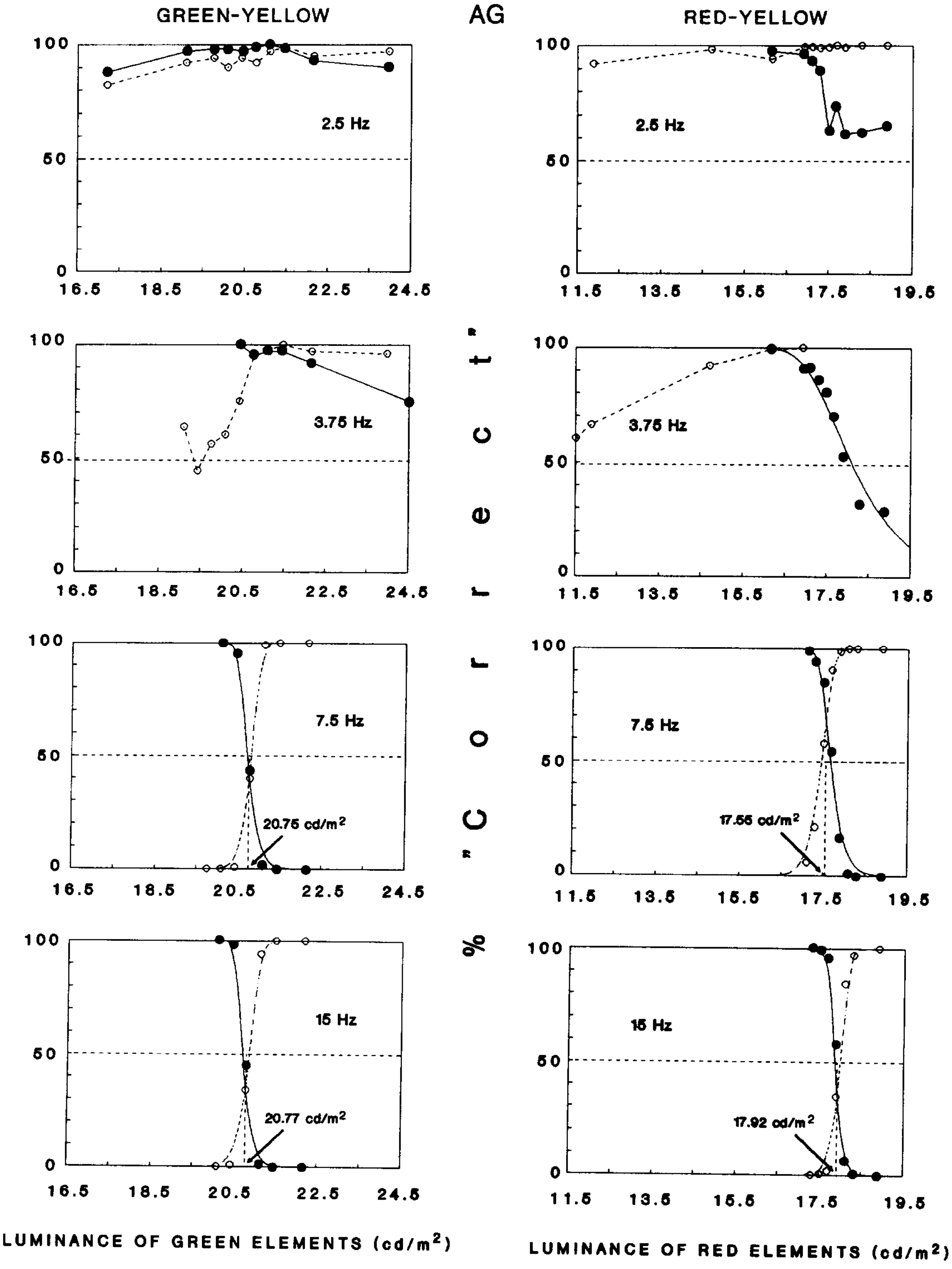

FIGURE 8. Percentages "correct" obtained with the stimulus illustrated in Fig. 7 as a function of the luminance contrast of the "CC" elements and with temporal frequency (increasing from top to bottom) as a parameter. Left and right panels show measurements with green and red "CC" elements, respectively. Open (dashed lines) and solid (continuous lines) symbols refer to data obtained with positive (+5%) and negative (−5%) luminance contrasts of the "LC" (yellow) elements. Percentages higher than 50% indicate preferences for the direction coinciding with the actual displacement of the stimulus. Percentages lower than 50% indicate preferences for the direction opposite to the actual displacement of the stimulus. The latter are accounted for in terms of luminance-based reverse-phi. Sigmoids are psychometric functions fitted to the data (only at 7.5 and 15 Hz; at 3.75 Hz only for red "CC" and negative "LC" elements). Arrows show the equiluminant point. Observer is AG.

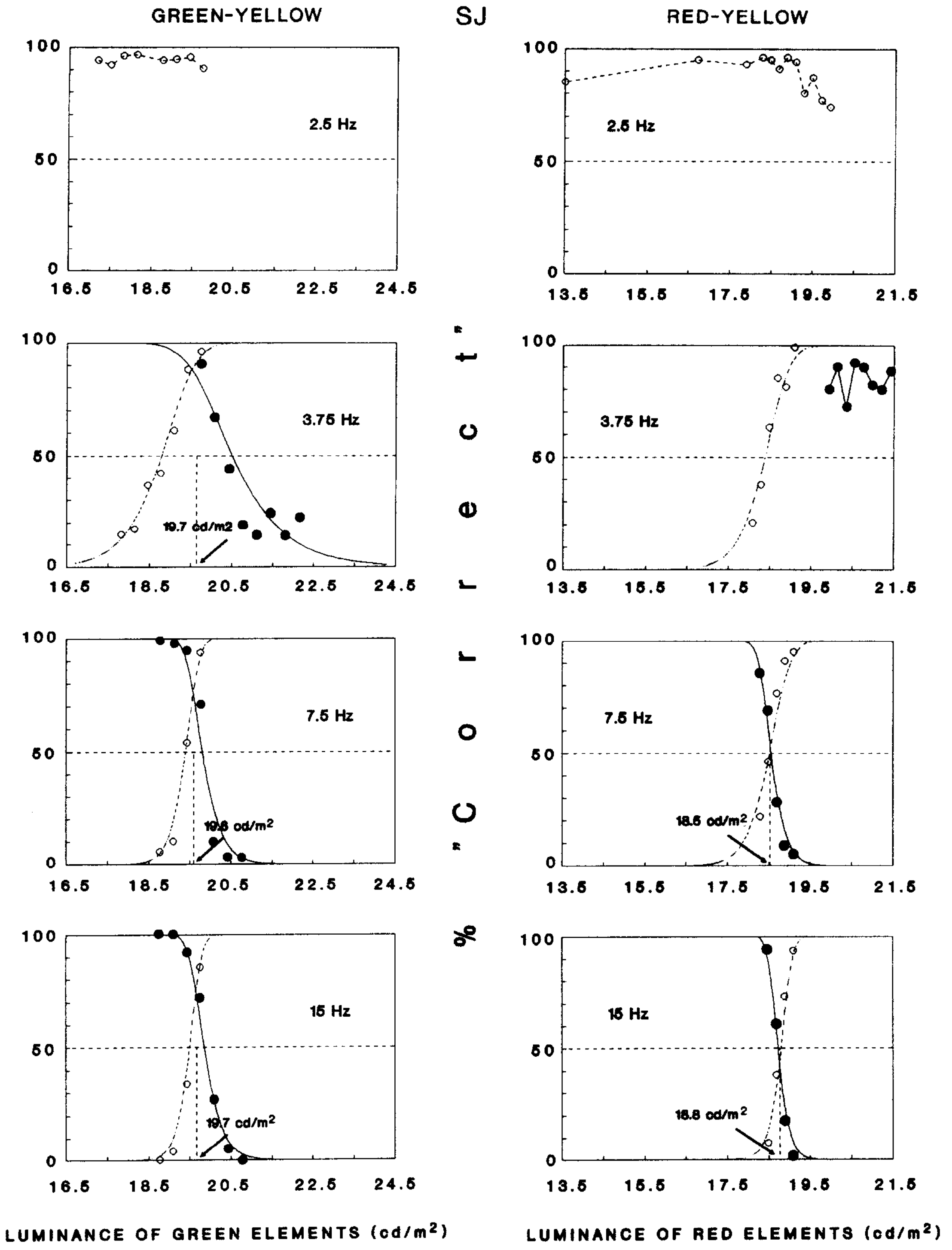

FIGURE 9. As for Fig. 8 but for observer SJ. Psychometric functions could be fitted for the highest three temporal frequencies with the exception of the red–yellow stimulus at 3.75 Hz and negative "LC" elements. This observer was not run with negative "LC" elements at 2.5 Hz.

range. If it is, how come that even for very large, opposite-sign luminance contrasts its output cannot override the response of the "contrast motion mechanism"? While the literature provides a clearly positive answer to the first question, the answer to the second is still a matter of debate. The present results together with previous studies (see the Discussion in Derrington *et al.*, 1992) suggest that, within the low temporal frequency range, the output of the "fast" first-order *luminance* mechanism is substantially weaker than the output of the "slow" (contrast or second-order) mechanism. Along this line of reasoning, it is interesting to note that at these low temporal rates the three observers reported the presence of two motions of opposite directions which they perceived in transparency. While they chose the perceptually dominant direction which happened to coincide in most of the cases with the hetero-path, the existence of an opposite motion suggests the activation of a second mechanism presumably displaying the reverse-phi phenomenon.

We interpret the general trend of the data displayed in Figs 8–10 [observation (1)] as follows. Given observation (2) and the rationale of the RP method, the data obtained at 7.5 and 15 Hz imply that colour and luminance information do not combine (i.e. are not integrated) to yield motion perception. However, these informations appear to be combined at the lowest (2.5 and 3.75 Hz) modulation rates. As the energy along the heterogeneous path increases, the energy along the reverse-phi path should increase correspondingly in order to counterbalance it. Within a covariance metric, the reverse-phi related energy is proportional to the absolute product of the fixed $LC_{\text{“LC”}}$ (i.e. + or −5%) and the variable, opposite sign $LC_{\text{“CC”}}$. When $LC_{\text{“LC”}} > 0$, an absolute increase of this product requires a *decrease* of the luminance of the “CC” elements, $L_{\text{“CC”}}$, i.e. an increase of $|-LC_{\text{“CC”}}|$. It follows that the displacement $(-\Delta)$ of the PSE point obtained with positive “LC” elements (open symbols in Figs 8–10) toward lower luminances ($L_{\text{“CC”}}$) as temporal frequency decreases may be taken as direct evidence of the progressive increase in the “strength” of the hetero-motion. The same argument accounts for the progressive *increase* $(+\Delta)$ of the luminance threshold obtained with *negative* “LC” elements. Thus, PSEs obtained at low temporal rates with positive and negative “LC” should be symmetrical about the “real” equiluminance point (EqL) and the *luminance contrast* between the PSE and the “real” equiluminant point ($\pm\Delta/\text{EqL}$) should measure the strength of the hetero-motion given the specific stimulus used.

Figure 11 displays these $|\Delta/\text{EqL}|$ ratios (circles) averaged across observers, chromatic contrasts (i.e. red–yellow and green–yellow), duty-cycles and $\pm LC_{\text{“LC”}}$ as a function of the temporal modulation rate. Conditions where the Δ/EqL ratio could not be measured, i.e. where performances did not drop below 50%, were

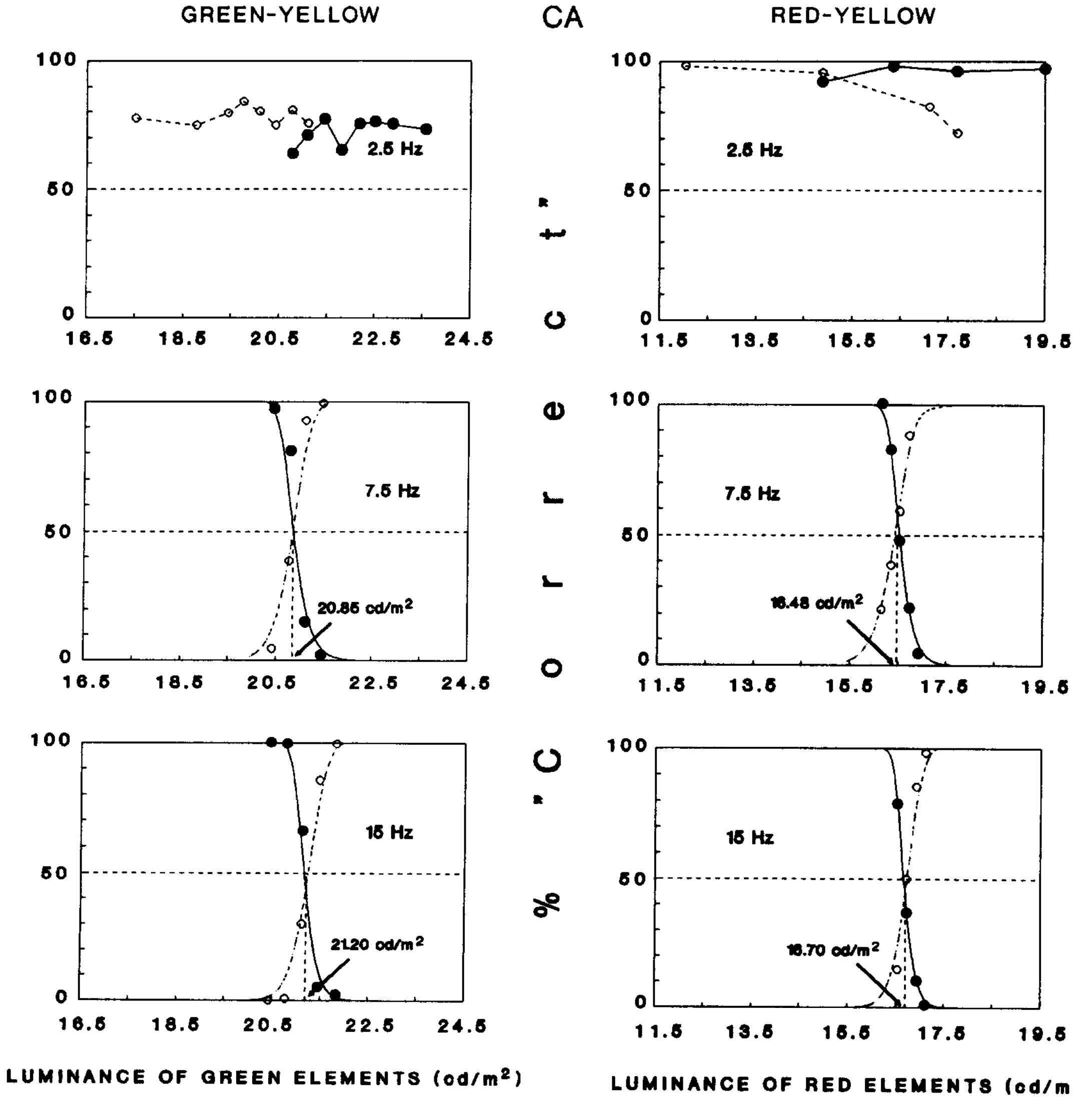

FIGURE 10. As for Fig. 8 but for observer CA. This observer was not run at 3.75 Hz.

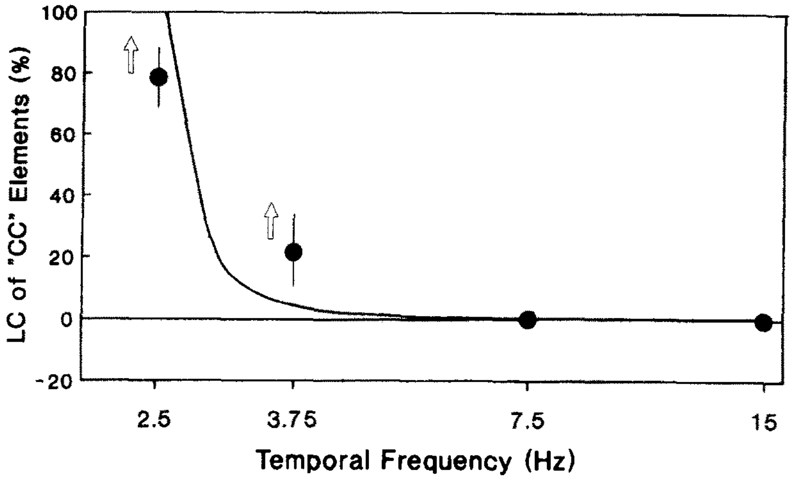

FIGURE 11. Estimated (circles) and predicted (continuous curve) $\mathrm{LC}_{\text{“CC”}}$ required by the "fast" mechanism to yield a reversal of the perceived direction as a function of temporal frequency. Estimated $\mathrm{LC}_{\text{“CC”}}$ is computed as the absolute difference between the luminances yielding 50% performances (with positive and negative "LC"s) in Figs 8–10 and the "real" equiluminance points normalized by the latter (i.e. Δ/EqL ratios). Each estimate (circle) is the average across observers, chromatic contrasts (green–yellow and red–yellow), positive and negative "LC"s and duty-cycles. Vertical bars are standard errors of these means. Given that Δ/EqL ratios could not always be measured for the two lower temporal frequencies, the corresponding estimates might have been underestimated. The empty arrows indicate this possibility. Predictions (continuous curve) are from equations (3) and (3′). See text for details.

arbitrarily given a value of 100% contrast.* Obviously, this is a cavalier procedure which might have been avoided by extending the studied $\mathrm{LC}_{\text{“CC”}}$ range to $\pm 100\%$. Figure 11 also displays the predicted Δ/EqL ratios (continuous curve) based on a simple model described below.

The relative strength of the hetero-motion as a function of temporal frequency can be easily accounted for if one postulates the existence of a "slow" and a "fast" motion mechanism with overlapping temporal frequency (TF) bandwidths as illustrated in Fig. 12. The "slow" and "fast" mechanisms have sensitivities $S_S = f(\mathrm{TF})$ and $S_F = g(\mathrm{TF})$. Posing that both mechanisms respect a covariance metric, with the former being sign insensitive (i.e. full-wave rectification) and indifferently summing chromatic and achromatic information, their activities, R_S and R_F, will be given by:

$$R_S = S_S(|\mathrm{CC}| + |\,\mathrm{LC}_{\text{“CC”}}|) \times |\mathrm{LC}_{\text{“LC”}}| \tag{1}$$

$$R_F = S_F \times \mathrm{LC}_{\text{“CC”}} \times \mathrm{LC}_{\text{“LC”}} \tag{2}$$

where CC should be measured in *equivalent luminance contrast* units (see Expt 1)†. The PSE points in Figs 8–10 are obtained when R_S and R_F are equal and of opposite signs. Since R_S is by necessity positive (i.e. rightward motion in Fig. 7), R_F must be negative (i.e. reverse-phi). This requires that $\mathrm{LC}_{\text{“CC”}}$ and $\mathrm{LC}_{\text{“LC”}}$ be of opposite signs. It then follows from equations (1) and (2) that, at 50% "correct"

$$\mathrm{LC}_{\text{“CC”}} = +\Delta/\mathrm{EqL} = S_S \times \mathrm{CC}/(S_F - S_S)$$
$$\text{with } S_F > S_S \text{ (for } \mathrm{LC}_{\text{“LC”}} < 0) \tag{3}$$

$$\mathrm{LC}_{\text{“CC”}} = -\Delta/\mathrm{EqL} = -S_S \times \mathrm{CC}/(S_F - S_S)$$
$$\text{with } S_F > S_S \text{ (for } \mathrm{LC}_{\text{“LC”}} > 0). \tag{3′}$$

The continuous, heavy curves in Figs 11 and 12 are $\mathrm{LC}_{\text{“CC”}}$ values computed by means of equations (3) and (3′) with the arbitrary sensitivities profiles shown in Fig. 12. The only constraints needed to provide *qualitative* agreement with the data are that the $S_F(\mathrm{TF})$ and $S_S(\mathrm{TF})$ functions [Gaussians on a log(FT) scale] intersect at temporal frequencies where PSEs could not be measured (i.e. $\pm\Delta/\mathrm{EqL} = \infty$; around 2.5 Hz) and that S_S be virtually zero at temporal frequencies where $\pm\Delta/\mathrm{EqL} = 0$ (i.e. around 7.5 Hz). Obviously, better agreement with the data (Fig. 11) could be obtained by adjusting the σ of the Gaussians in Fig. 12. Given the uncertainty in the measured Δ/EqL ratios themselves

---

[*] In fact, the Δ/EqL ratios under these conditions should have been assigned an infinite value. Since Δ/EqL ratios measured for a given temporal frequency could vary across observers, chromatic contrasts, duty-cycles and $\pm\mathrm{LC}_{\text{“LC”}}$, such a procedure would have prevented their averaging across these factors.

[†] If one applies the same logic to the stimulus used in Expt 1 to assess EqLC (see Fig. 1), the responses of the "slow" and "fast" mechanisms at 50% chromatic motion preferences are given by

$$R_S = S_S(|\mathrm{EqLC}_S|^2 - |\mathrm{LC}_{\text{“CC”}}|^2) \tag{i}$$

$$R_F = S_F(\mathrm{EqLC}_F^2 - \mathrm{LC}^2_{\text{“CC”}}) \tag{ii}$$

with EqLC_S and EqLC_F standing for the equivalent luminance contrasts of the "slow" and "fast" mechanisms which, given the results of Expt 1, are probably unequal (see also the Discussion). It then follows from equations (i) and (ii) that the measured EqLC is given by

$$\mathrm{EqLC} = [(S_S \mathrm{EqLC}_S^2 + S_F \mathrm{EqLC}_F^2)/(S_S + S_F)]^{0.5}. \tag{iii}$$

EqLC_S or EqLC_F can be derived from equation (2) only if they are equal or if S_S or $S_F = 0$. In equations (1–3′), only the "slow" mechanism uses the chromatic information. Moreover, at 2.5 Hz S_F is virtually zero. Thus at this temporal rate, $\mathrm{EqLC} \approx \mathrm{EqLC}_S$, which is the value used in our computations. At any rate, an incorrect choice of EqLC in equations (3) and (3′) will entail only a vertical shift of the $\mathrm{LC}_{\text{“CC”}}$ function of temporal frequency.

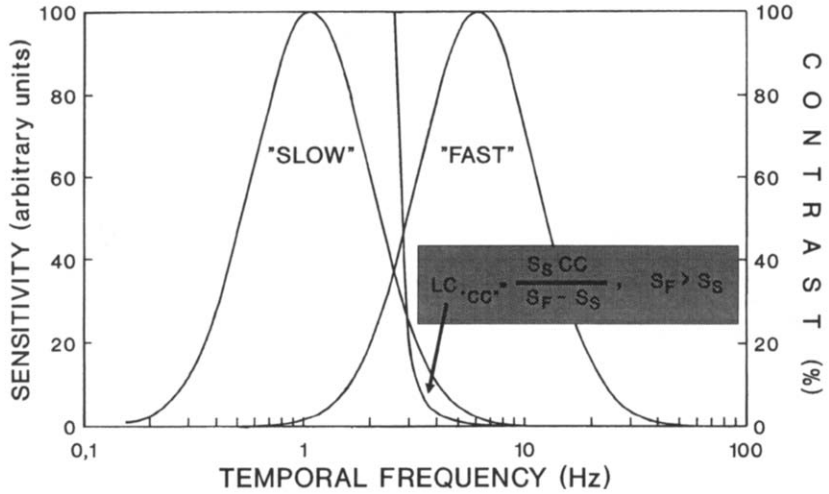

FIGURE 12. Arbitrary, Gaussian (on a log abscissa) sensitivity profiles of the "slow" and "fast" mechanisms as a function of temporal frequency (left ordinate) and the predicted $LC_{"CC"}$ (Δ/EqL ratios) required by the "fast" mechanism to yield a reversal of the perceived direction (curve pointed by the arrow, computed as shown in the insert; right ordinate). See text for details.

(see above), we did not give priority to this enterprise. In fact parabolic (also on a log(FT) scale) S_F(TF) and S_S(TF) functions provide equally good qualitative agreement with the data. Further research should provide more accurate characterizations of these functions.

The representation of the "slow" and "fast" motion mechanisms displayed in Fig. 12 is clearly reminiscent of the one offered a while ago by Thompson (1982) and by Murray, MacCana and Kulikowski (1983) in quite different experimental contexts. It is not clear, however, whether the two motion systems revealed by the present and previous data necessarily coincide. In fact, the idea of a dichotomy within the motion system has been entertained for many years under different labellings such as transient-sustained, fast-slow, short-long-range, magno-parvo, etc. The common denominator of these descriptions is the assumption that the two motion systems are parallel and that their respective outputs originate at processing stages of the same hierarchical order (e.g. first-order mechanisms). The data presented here argue however in favour of two hierarchically ordered stages of motion processing of the type already proposed by Cavanagh *et al.* within similar experimental contexts (Cavanagh *et al.*, 1989; Cavanagh & Mather, 1989) and by many others within different experimental contexts (Chubb & Sperling, 1988, 1989; Derrington & Badcock, 1985; Derrington *et al.*, 1992; Nishida & Sato, 1992a, b; Ogata & Sato, 1992; Werkhoven *et al.*, 1990a, 1992a, b, 1993; Wilson, 1992, etc.). Future work will decide whether this dichotomy is or is not best described in terms of the specific metrics (similarity, covariance) applicable to each mechanism (see related information in the last section of this study). Taken together, Expts 1 and 2 point to the existence of a unique "slow" system and of two, a chromatic and an achromatic, "fast" systems.

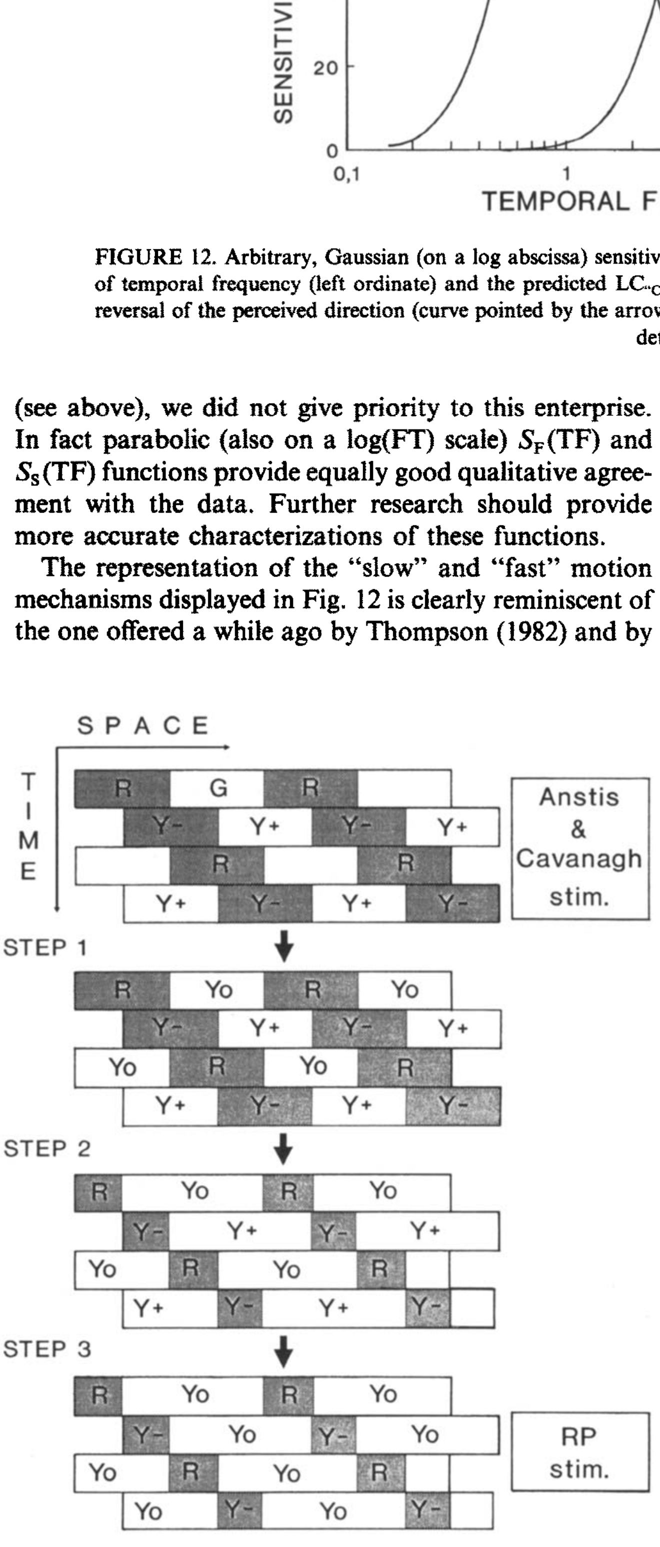

FIGURE 13. A three step transformation of the "minimum motion" stimulus of Anstis and Cavanagh (1983) and the stimulus shown in Fig. 7.

Before concluding this section, however, we believe it worthwhile to compare the RP method presented here with the "minimum motion" (MM) technique of Anstis and Cavanagh (1983). Figure 13 displays the original MM stimulus (top) and the transformations required to obtain the RP stimulus used in the present study (bottom). These transformations are as follows. *Step 1*: Change Green (or Red) elements (with an arbitrary Luminance = L_0) to Yellow (with the same arbitrary $L = L_0$). Such a stimulus configuration has actually been used in the original Anstis and Cavanagh study). *Step 2*: Change spatial duty–cycle from 1/1 to 1/3, while keeping spatial frequency constant. *Step 3*: Set the "new" Yellow (ex-Green) and "Y+" at the same luminance, L_0. It is obvious that the two stimuli are intimately related. They will both yield minimum motion whenever $L_R = L_G$, for the MM stimulus, and when $L_R = L_Y$, for the RP stimulus. As far as it stands, the critical difference between the two stimuli relates to the *a priori* impossibility for the MM technique to set the "G" and "Y+" elements at equiluminance, such as to isolate a pure reverse-phi component. The MM stimulus will provide a unique EqL point for any condition where $L_G < > L_{Y+}$, which can be taken as an advantage, but it will not be able to isolate the contribution of the "slow" mechanism unless $L_G = L_Y$. As a final remark, one may note that the psychometric functions presented in the original "MM paper" (1983) have slopes about half of those obtained here with the RP technique. While this difference may be related to the different spatial duty-cycles used in the two cases, it is more likely that it is related to the lower mean luminances used in the original MM study (a factor of two lower than here).

EXPERIMENT 3. CHROMATIC AND LUMINANCE REVERSE-PHI MOTION

By the nature of the stimuli, the "fast" mechanism described in the previous experiment was a *luminance* mechanism. It is a *sign* sensitive mechanism behaving in accord with the similarity and covariance principles. Combining (i.e. multiplying) opposite polarity luminance information accounts for the reverse-phi phenomenon (Anstis, 1980; Adelson & Bergen, 1985; Chubb & Sperling, 1988). Our question here bears on the existence of such a mechanism in the *chromatic* domain and on its temporal frequency characterization relative to the luminance mechanism. By analogy with a luminance Reichardt-detector, a sign sensitive, chromatic mechanism should process inputs from *chromatic double-opponent units* (see for a review Lennie & D'Zmura, 1988). While the existence of such units has been recently questioned (Lennie, Krauskopf & Sclar, 1990), they appear to be quite useful in modelling texture discrimination with chromatic stimuli (Gorea & Papathomas, 1993). Moreover, reverse-phi with equiluminant chromatic stimuli has already been observed (Sato, 1988).

Figure 14 illustrates the stimuli used here to compare chromatic- and luminance-based reverse-phi phenomena. The chromatic stimulus consists of red (CC$_R$) bars on a yellow background which, when shifted from one frame to the next by one-quarter of a spatial cycle, become green (CC$_G$; i.e. of reversed chromatic polarity with respect to the yellow background), and so forth. Red and green bars are equiluminant. In the luminance stimulus, the red and green bars are replaced by dark (LC−) and bright (LC+) yellow bars. The luminance contrast between the dark and bright bars was set at 10% which is close to the *equivalent luminance contrast* of the red–green modulation as measured at 15 Hz (see Expt 1).

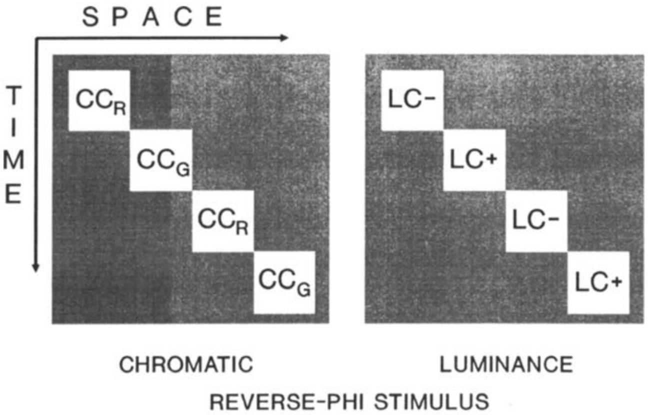

FIGURE 14. Schematic spatiotemporal representation of a chromatic (left panel) and luminance (right panel) "reverse-phi stimulus". Other details are as for Fig. 1.

Experiments were run at three drift rates (2.5, 7.5 and 15 Hz) and at three luminance-noise levels (0, 5 and 10%—see Expt 1). The equiluminant points were set according to the measurements of Expt 2 at 15 Hz. At 2.5 and 7.5 Hz, data are presented only for balanced duty–cycle ratios (i.e. 3/3 and 1/1, respectively) but other duty–cycles provided similar results. As in the previous experiments left/right performances were measured with temporal frequencies, noise-levels and stimulus type (colour and luminance) randomized across sessions. Each datum point is based on at least 300, but more frequently on 700 trials. Only the first author served as an observer.

Percentages "correct" (i.e. coinciding with the physical displacement of the stimulus) are shown in Fig. 15. Performances below 50% indicate reverse-phi perception. The data display the following characteristics.

Chromatic reverse-phi (left panel in Fig. 15)

With the pure chromatic stimuli, reverse-phi is observed only at 7.5 Hz. Reverse-phi is replaced by 100% direct-phi at 2.5 Hz and by chance responses (i.e. close to 50%) at 15 Hz. The luminance noise has practically no effect in the direct-phi range and only a moderate effect at 15 Hz, while it strongly perturbs the reverse-phi responses obtained at 7.5 Hz. Taken together, these data suggest the existence of two distinct chromatic motion mechanisms, a "slow" one, insensitive to the sign of the chromatic contrast (i.e. green–yellow vs red–yellow) and little perturbed by luminance noise and a "fast" one, operating optimally at medium temporal frequencies, sign sensitive (i.e. displaying reverse-phi) and strongly perturbed by luminance noise. At the highest temporal rate studied (15 Hz), the "fast" chromatic-motion mechanism appears to be still active although its output is considerably attenuated (36% "correct") and virtually null (i.e. 50% "correct") under 10% luminance noise conditions. The data cannot reveal whether the chromatic-luminance interactions are genuinely different in the two mechanisms or artifactually due to a ceiling effect at the lowest modulation rate where performances are practically at 100%. As a consequence, the data cannot reveal either whether the "slow" chromatic mechanism is a purely chromatic or whether it also responds to luminance information. If this "slow" mechanism is a subset of the "slow" mechanism revealed in the previous experiment, then the latter alternative should be true.

Luminance reverse-phi (right panel in Fig. 15)

With the pure luminance stimuli, reverse-phi is observed in all experimental conditions but one, i.e. at 2.5 Hz and 10% luminance noise where direct-phi is observed. The data indicate that a "first-order", sign sensitive, luminance-motion mechanism is active over the whole temporal frequency range studied and that, as expected, its response is highly sensitive to luminance noise. The data also suggest that the output of this mechanism is significantly weaker at 2.5 Hz where an increase in luminance noise from 0 to 10% switches motion perception from a reverse-phi to a direct-phi mode. This latter observation nicely corroborates the data obtained in Expt 2 since it accounts for the fact that the luminance reverse-phi path could not take over the LC-CC hetero-path at these low temporal rates (see top, panels in Figs 8–10). It also indicates the existence of a sign insensitive (second-order), luminance-motion mechanism operating at low temporal rates, only moderately (or not at all) perturbed by luminance noise and whose response appears to be masked by the "first-order" mechanism under 0% noise conditions and unmasked when the response of this "first-order" mechanism is perturbed under 10% noise conditions. Along the same argument as for the chromatic reverse-phi data, it is quite likely that this "slow", sign-insensitive motion mechanism is not luminance specific but that it responds to any type of modulation, whatsoever, i.e. that it does not respect the similarity principle.

Taken together, the chromatic and luminance data corroborate the distinction between "slow" and "fast" mechanisms made in the first two experiments of this study. They clearly suggest that the "fast" mechanism consists of two distinct subsets, namely a pure chromatic and a pure luminance, sign sensitive (first-order) mechanism whose behaviours are in accord with both the similarity and the covariance principles. The two "fast" mechanisms appear to differ in their temporal tuning characteristics, with the luminance mechanism responding to higher temporal rates than the chromatic mechanism. These reverse-phi data cannot be used by themselves to specify whether or not the "slow" mechanism also reunites a pure chromatic and a pure luminance mechanism, but Expts 1 and 2 speak in favour of a *unique* "slow" mechanism which infringes the similarity principle since it can indifferently combine luminance and chromatic information.

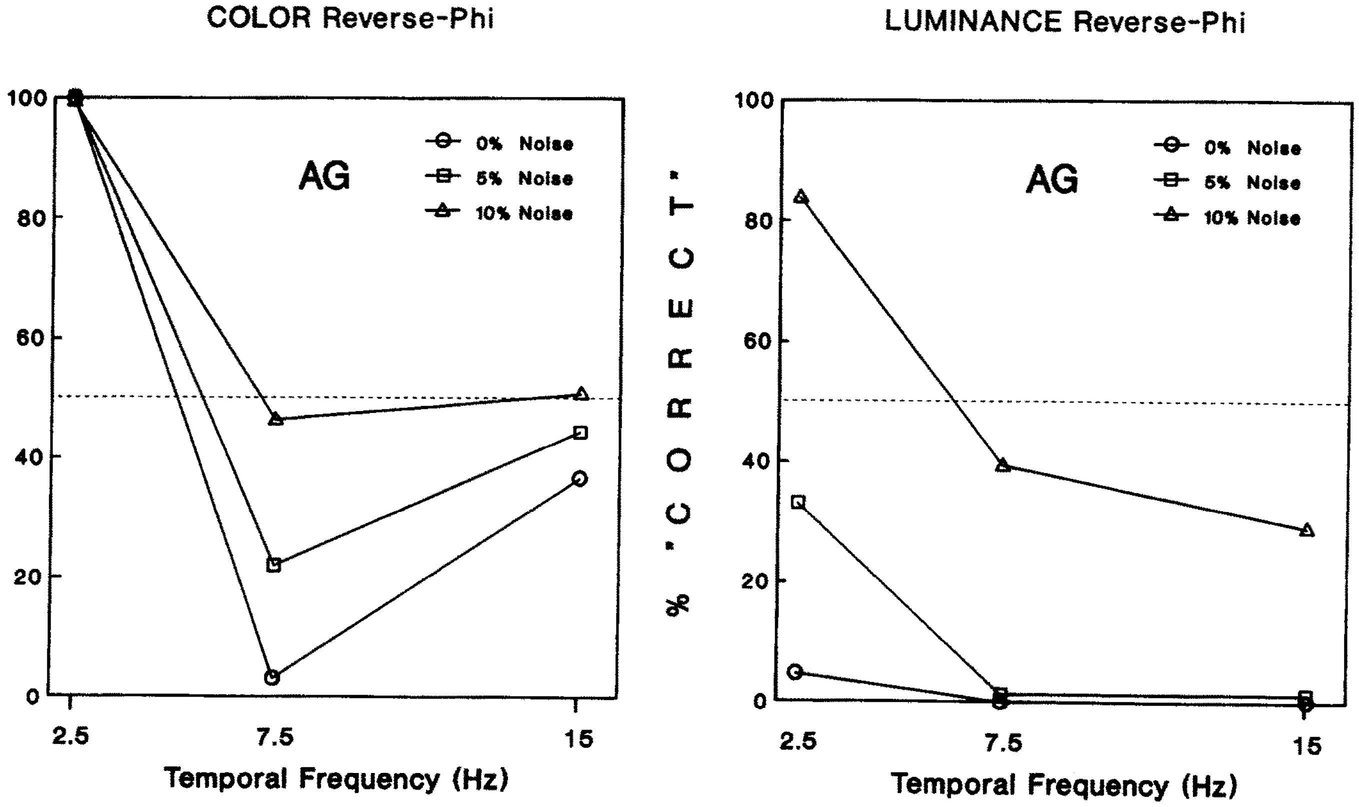

FIGURE 15. Percentages "correct" obtained with the stimuli illustrated in Fig. 14 as a function of temporal frequency. Data obtained with pure colour and pure luminance stimuli are shown in the left and right panels, respectively. Conditions with 0, 5 and 10% luminance-noise are shown as circles, squares and triangles, respectively. Percentages lower than 50% indicate preferences for the direction opposite to the actual displacement of the stimulus, i.e. reverse-phi. Observer is AG.

DISCUSSION

The first two experiments of this study presented convincing evidence that the human motion system operates in two distinct modes, one of which infringes the similarity principle (as defined in the Introduction) and is optimally active at low temporal rates (the "slow" system) and another one which behaves in accord with the similarity principle and prefers medium to high temporal rates (the "fast" system). Experiment 3 provided additional evidence supporting this dichotomy and demonstrated that the "slow" (unlike the "fast") system does not display a reverse-phi behaviour which is an indication that its input signal is fully rectified before direction computation proper. The data of Expt 3 also support the idea that the "fast" system is in fact the union of two distinct, chromatic and achromatic "fast" systems with the latter being significantly "faster" than the former. In contrast, the present experiments taken together suggest that the "slow" system is unique and fairly low-level. The distinction between the two systems is also supported by the observation that the *equivalent luminance contrast* of our red-green stimuli strongly depends on their temporal modulation rates (see Equivalent Luminance Contrast subsection of Expt 1 and Anstis & Cavanagh, 1983). At low modulation rates, EqLC is practically identical with the estimated cone contrast of the stimuli, the implication of which is that chromatic and luminance information yield equal efficiencies for the "slow" system. On the other hand, the "fast" chromatic system is substantially less efficient (a factor of two or more) than the "fast" luminance system insofaras directionality is concerned (see, however, Agonie & Gorea, 1993).

The present results support the rationale of Gorea *et al.*'s studies (see the Introduction) based on the similarity principle as well as their distinction between a chromatic and an achromatic motion system (see also Cavanagh & Anstis, 1991) since their experiments were performed with temporal modulations within the range of the "fast" system (6–7.5 Hz). This distinction between two "fast" systems does not preclude the possibility that, after the independent extraction of spatiotemporal oriented energy in the chromatic and luminance pathways, the two types of information combine at a later processing stage. In fact, this stand, originally proposed by Cavanagh and Favreau (1985) has received since then conclusive support (Gorea & Papathomas, 1989; Gorea, Lorenceau, Bagot & Papathomas, 1993; Papathomas *et al.*, 1991).

One may challenge the notion that a motion mechanism that combines colour and luminance information necessarily infringes the similarity principle. According to the definition given in the Introduction, this would be the case if the motion detector were fed from the very beginning by subunits responding to both chromatic and luminance modulations. Given the separate processing of the two types of information in early vision, this is quite unlikely an hypothesis. Moreover, the neurophysiology offers no evidence that units responding to both chromatic and achromatic modulations are more low-pass in the temporal (or spatial) domain than pure chromatic and pure achromatic units. Despite these circumstantial arguments, one may still find difficult to decide whether a system respects or infringes the similarity metric as defined. We therefore propose an alternative distinction in terms of *specific* ("fast") and *unspecific* ("slow") motion systems.

While the literature provides reasonable evidence that the "fast"/specific system(s) can be modelled in terms of a covariance metric (e.g. Adelson & Bergen, 1985; Van Santen & Sperling, 1985; Watson & Ahumada, 1985; Chubb & Sperling, 1988; Werkhoven *et al.*, 1993), much less is known in this respect about the behaviour of the "slow"/unspecific system. Recent data by Werkhoven *et al.* (1993) argue in favour of the covariance metric insofaras it concerns the combination of different spatial frequency patches along a hetero-motion path. The data of Expt 1 do not contradict the possibility that the covariance metric also applies across (colour and luminance) "domains". Under the hypothesis of a strict covariance metric, one might have expected that the heterogeneous motion (in the "Werkhoven" stimulus—see Fig. 3) would take over the homogeneous motion anytime that the covariance product along the hetero-path (chromatic and achromatic information confounded, i.e. within the "slow" mechanism) out-weighs the covariance product along the homo-path (i.e. the *sum* of the covariance products within both the "slow" and "fast" systems). At the lowest temporal rates used in this study (i.e. 2.5 Hz) the heterogeneous motion does indeed take over the homogeneous motion but for covariance products larger than expected, i.e. for luminance contrasts about twice the estimated equivalent luminance contrast of the chromatic modulation (see top panels in Figs 4 and 5). It is easy to argue that the additional contrast required to reverse the perceived direction in favour of the hetero-path is equivalent to the additional contribution of the "fast" mechanism which is still active in this temporal frequency range. While additional studies are required to firmly assess the validity of a covariance metric across "domains", the data of Expt 2 (Figs 8–10) and their qualitative agreement with the simple model we proposed [together with Werkhoven *et al.*'s (1993) study] provide a convincing clue in favour of this hypothesis. Indeed, the reverse phi method we proposed for assessing equiluminance in this experiment could also be used to estimate the relative strength of the "slow" and "fast" systems as a function of temporal frequency and to predict, at least qualitatively, the additional luminance contrast required to switch the dominant direction from the hetero- to the homo-path. This method, in conjunction with a more elaborated model should enable one to accurately characterize the temporal (and spatial) characteristics of the two motion systems. We note, as a conclusion, that the general behaviour of the "slow"/unspecific system is well suited to account for the spatiotemporal combination of potentially any type of visual information (see Cavanagh *et al.*, 1989) and deserves in that sense the generic name of a high-order motion system (Cavanagh & Mather, 1990).

As a final remark, we draw attention to the fact that the "slow"–"fast" terminology adopted here might be misleading since it makes direct reference to the notion of speed. While the present experiments did manipulate stimulus speed, they provide no information as to the behaviour of the two systems with regard to spatial frequency. Experiments in progress suggest that the "slow", as opposed to the "fast" system is optimally activated within the low temporal *and* spatial frequency range. If this finding were to be confirmed, the two systems could not be discriminated on the basis of their speed preferences.

CONCLUSION

(1) There are two visual motion systems, a "slow"/unspecific and a "fast"/specific one.

(2) The "slow"/unspecific system is unique: it combines all sources of information (within "domains" across channels as well as across "domains") and, in that sense, does not respect the similarity principle. The "slow"/unspecific system is sign-insensitive and, to a first approximation, behaves in accord with a covariance metric.

(3) There are at least two "fast"/specific systems, a chromatic and an achromatic one. (Within each of these two "domains" there are probably as many "fast"/specific systems as discriminable visual dimensions.) The achromatic system is "faster" than the chromatic system. Both behave in accord with the similarity and covariance principles.

REFERENCES

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

Agonie, C. & Gorea, A. (1993). Equivalent luminance contrast of red–green drifting stimuli: dependency on luminance–color interactions and on the psychophysical task. *Journal of the Optical Society of America A, 10,* 1341–1352.

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

Anstis, S. M. & Cavanagh, P. (1983). A minimum motion technique for judging equiluminance. In Mollon, J. D. & Sharpe, L. T. (Eds), *Colour vision: Physiology and psychophysics* (pp. 155–166). New York: Academic Press.

Anstis, S. M., Cavanagh, P., Maurer, D., Lewis, T., MacLeod, D. A. I. & Mather, G. (1986). Computer-generated screening test for colorblindness. *Color Research and Application (Suppl.), 11,* S63–S66.

Cavanagh, P. & Anstis, S. M. (1991). The contribution of color to motion in normal and color-deficient observers. *Vision Research, 31,* 2109–2148.

Cavanagh, P. & Favreau, O. E. (1985). Color and luminance share a common pathway. *Vision Research, 25,* 1595–1601.

Cavanagh, P. & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision, 4,* 103–129.

Cavanagh, P., Arguin, M. & von Grünau, M. (1989). Interattribute apparent motion. *Vision Research, 29,* 1197–1204.

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,* 1986–2007.

Chubb, C. & Sperling, G. (1989). Apparent motion derived from spatial texture. *Investigative Ophthalmology and Visual Science (Suppl.), 30,* 425.

Cole, G. R. & Hine, T. (1992). Computation of cone contrasts for color vision research. *Behavior Research Methods, Instruments and Computers, 24,* 22–27.

Derrington, A. M. & Badcock, D. R. (1985). Separate detectors for simple and complex grating patterns? *Vision Research, 25,* 1869–1878.

Derrington, A. M., Badcock, D. R. & Holroyd, S. A. (1992). Analysis of motion of 2-dimensional patterns: Evidence for a second-order process. *Vision Research, 32,* 699–707.

Gorea, A. & Papathomas, T. V. (1987). Form and surface attributes in motion perception studied with a new class of stimuli: A basic asymmetry. A T&T Bell Laboratories Technical Memorandum, 11233-870921-2TM, N.J.: Murray Hill.

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. (1993). "Double-opponency" as a generalized concept in texture segregation illustrated with color, luminance and orientation defined stimuli. *Journal of the Optical Society of America A.* In press.

Gorea, A., Kovacs, I. & Papathomas, T. V. (1992). Against a covariance metric for motion perception: Chromatic and luminance information do not combine. *Technical Digest, Optical Society of America, Annual Meeting,* 215.

Gorea, A., Lorenceau, J., Bagot, J. D. & Papathomas, T. V. (1993). Sensitivity to colour- and orientation-carried motion respectively improves and deteriorates under equiluminant background conditions. *Spatial Vision. 6,* 285–302.

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

Green, M. & Odom, J. V. (1986). Correspondence matching in apparent motion: Evidence for three-dimensional spatial representation. *Science, 233,* 1427–1429.

Lennie, P. & D'Zmura, M. (1988). Mechanisms of color vision. *CRC Critical Reviews in Neurobiology, 3,* 333–400.

Lennie, P., Krauskopf, J. & Sclar, G. (1990). Chromatic mechanisms in striate cortex of macaque. *Journal of Neuroscience, 10,* 649–669.

Livingstone, M. S. & Hubel, D. H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neurosciences, 4,* 309–354.

Murray, I., MacCana, F. & Kulikowski, J. J. (1983). Contribution of two movement detecting mechanisms to central and peripheral vision. *Vision Research, 23,* 151–159.

Nishida, S. & Sato, T. (1992a). Two mechanisms for motion perception in the human visual system. *IEEE International workshop on Robot & Human Communication,* 294–299.

Nishida, S. & Sato, T. (1992b). Positive motion after-effect induced by bandpass-filtered random-dot kinematograms. *Vision Research, 32,* 1635–1646.

Nishida, S. & Takeuchi, T (1990). The effects of luminance on affinity of apparent motion. *Vision Research, 30,* 709–721.

Ogata, M. & Sato, T. (1992). Motion-detection model with two stages: Spatiotemporal filtering and feature matching. *Journal of the Optical Society of America A, 9,* 377–387.

Quick Jr, R. F. (1974). A vector-magnitude model of contrast detection. *Kybernetic, 16,* 65–67.

Papathomas, T. V. & Gorea, A. (1988). Simultaneous motion perception along multiple attributes: A new class of stimuli. *Behavioral Research Methods, Instruments and Computers, 20,* 528–536.

Papathomas, T. V., Gorea, A. & Julesz, B. (1991). Two carriers for motion perception: Color and luminance. *Vision Research, 31,* 1883–1891.

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

Sato, T. (1988). Direction discrimination and pattern segregation with isoluminant chromatic random-dot cinematograms. *Investigative Ophthalmology and Visual Science, (Suppl.) 29,* 449.

Shechter, S., Hochstein, S. & Hillman, P. (1989). Size, flux and luminance effects in the apparent motion correspondence process. *Vision Research*, *29*, 579–591.

Thompson, P. (1982). Perceived rate of movement depends on contrast. *Vision Research*, *22*, 377–380.

Ullman, S. (1980). The effect of similarity between bar segments on the correspondence strength in apparent motion. *Perception*, *9*, 617–626.

Van Santen, J. P. H. & Sperling, G. (1985) Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, *2*, 300–321.

Watson, A. B. (1979). Probability summation over time. *Vision Research*, *19*, 515–522.

Watson, A. B. (1986). Apparent motion occurs only between similar spatial frequencies. *Vision Research*, *26*, 1727–1730.

Watson, A. B. & Ahumada, A. J. (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). Metrics for the strength of low-level motion perception. *Journal of Visual Communication and Image Representation*, *1*, 176–188.

Werkhoven, P., Snippe, HP. & Koenderink, J. J. (1990b). Effects of element orientation on apparent motion perception. *Perception & Psychophysics*, *47*, 509–525.

Werkhoven, P., Sperling, G. & Chubb, C. (1992a). The dimensionality of motion-from-texture. *Investigative Ophthalmology and Visual Science (Suppl.)*, *33*, 1049.

Werkhoven, P., Sperling, G. & Chubb, C. (1992b). Energy computations in motion and texture. *Technical Digest, Optical Society of America, Annual Meeting*, 216.

Werkhoven, P., Sperling, G. & Chubb, C. (1993). The dimensionality of texture-defined motion a single channel theory. *Vision Research*. In press.

Wilson, H. R. (1992). Interaction of first- and second-order processes in 2D motion perception. *Technical Digest, Optical Society of America, Annual Meeting*, 35.

Acknowledgements—This work was supported by grant 91-074 (DRET) for A. Gorea and by grants BNS-9109384 (NSF) and N00014-92-J-1312 (ONR) for T. V. Papathomas. Parts of the present material have been presented at the *Annual Meeting of the Optical Society of America*, 1992, 20–25 September, Albuquerque, N.M., U.S.A. We are grateful to an unknown referee who, among other things, drew our attention to the matter exposed in the footnote on p. 2528.