Perception, 1996, volume 25, pages 65–76

Psychophysical evidence of a sustained input to directionally selective motion mechanisms

Walter F Bischof, Vincent Di Lollo
Department of Psychology, University of Alberta, Edmonton, Alberta T6G 2E9, Canada
Received 25 April 1995, in revised form 20 October 1995

Abstract. Human psychophysical evidence congruent with neurophysiological findings of a sustained input to directionally selective motion sensors in cat visual cortex is reported. Apparent motion was produced by displaying a group of dots in two frames (F1 and F2), where F2 was a translated version of F1. All stimulus sequences included a period during which F1 and F2 were displayed concurrently (combined images) and a period during which only F1 or F2 was on display (single images). There were three stimulus sequences: a display beginning with combined and ending with single image, a display beginning with single and ending with combined image, and a display beginning with F1, continuing with combined image, and ending with F2. Six durations of single and of combined images (10, 20, 40, 80, 160, and 320 ms) were crossed factorially in each stimulus sequence. Directional motion was seen easily at long durations of the single image in all stimulus sequences, as would be expected on the basis of a sustained input to the directional-motion-sensing mechanisms. Perception of directional motion improved with the duration of single images, but declined as the duration of combined images was increased. Baker and Cynader's model could account for the effect of duration of single images, but not for the effect of duration of combined images. An elaborated version of the model provides a good qualitative match to all empirical findings.

1 Introduction

Apparent motion is seen if two brief stimuli are displayed in sequential frames, so that the onset of the leading frame (F1) precedes the onset of the trailing frame (F2) by a suitable interval (stimulus onset asynchrony, SOA). In a common paradigm, F1 contains a set of dots, and F2 is a translated version of F1. Under appropriate spatiotemporal conditions, the dots are seen as moving in the direction of the displacement. In traditional studies, the temporal relationship between F1 and F2 has been kept within relatively narrow limits. Notably, the SOAs have been relatively brief, and the stimuli have always been displayed in strict temporal succession. Theoretical models have reflected these restrictions. For example, models such as those proposed by Reichardt (1961), Adelson and Bergen (1985), van Santen and Sperling (1985), and Watson and Ahumada (1985) postulate that the production of motion signals depends critically on the interaction between the brief transient responses time-locked to the onsets of F1 and F2. For such an interaction to occur, it is necessary that at least some trace of the transient response triggered by F1 be still available when the corresponding F2 response is initiated. If the F1 response has abated before the initiation of the F2 response, no interaction can take place, and no motion signal can be generated. Such a decoupling can be achieved in either of two ways. One is to insert a sufficiently long temporal gap (interstimulus interval, ISI) between the offset of F1 and the onset of F2, during which the screen is blank and the transient F1 response is allowed to decay. The other is to use a long SOA and allow F1 to remain visible until (or beyond) the onset of F2; in this case, the exposure duration of F1 should be long enough to allow the onset-locked transient response to abate before F2 is displayed. Whether the temporal decoupling of the onsets is achieved through ISI or SOA, the theoretical expectation remains the same: as the temporal separation is increased, the strength of the motion signal should decrease.

This expectation has been confirmed in the case of long ISIs: perception of directional motion is progressively impaired as the duration of the ISI is increased (Baker and Braddick 1985). No such confirmation, however, has been reported with long SOAs. On the contrary, it has been shown that directional motion can still be seen at SOAs as long as 1.6 s, even if a brief ISI is inserted in the stimulus sequence (Strout et al 1994). Informal observations reveal that this is true even if F1 remains visible beyond the onset of F2. In this case, the transient response to the onset of F2 coincides not with the transient but with the ongoing sustained response to F1. Such observations lead us to question the role of transient responses as exclusive sources of motion signals. Instead, they suggest that motion signals can ensue from the interaction of transient and sustained responses, thus entailing the existence of a sustained input to the motion sensor.

Experimental confirmation of a sustained input to directional-motion-sensing units in cat visual cortex has been reported recently by Emerson et al (1992) and by Baker and Cynader (1994). Baker and Cynader employed a long-SOA paradigm that included some temporal overlap between F1 and F2. The results show that a motion signal can be triggered by the onset of F2 well after the F1 transient has abated, provided that F1 is still visible at the onset of F2. Baker and Cynader accounted for this result with a motion-sensing model described in detail below.

In the present work, we studied directional motion perception at long SOAs with human observers. The studies were designed to achieve two objectives; first, to explore the limits of directional-motion perception over a comprehensive range of SOAs, with stimuli in which the duration of temporal overlap between F1 and F2 was varied systematically. Second, to establish whether Baker and Cynader's model could be used—or adapted—to account for the psychophysical data.

2 Methods

2.1 *Observers*

The second author and a paid volunteer, both experienced observers, served in all experiments. Both had corrected-to-normal vision.

2.2 *Apparatus*

All stimuli were composed of dots displayed on a Tektronix 608 oscilloscope equipped with P15 phosphor. The *X*, *Y*, and *Z* (intensity) coordinates of each dot were stored in a fast plotting buffer that displayed them to the screen at the rate of 1 dot μs^{-1} (Finley 1985). Intensity of the dots was measured with a Minolta LS-110 luminance meter after the method described by Sperling (1971). Values of luminous directional energy were expressed in millicandela microseconds (mcd μs)—rather than the more common candela microseconds—in order to avoid very small numbers. Images were displayed within a square area in the centre of the screen. At the viewing distance of 57 cm, set by a headrest, the sides of square area subtended an angle of 2 deg.

2.3 *Visual displays*

All displays comprised two sets of dots, F1 and F2. Set F1 contained 160 dots distributed randomly on each trial over the 2 deg × 2 deg viewing area. Set F2 also contained 160 dots: 80 were distributed randomly on each trial over the 2 deg × 2 deg viewing area; the other 80 were identical to 80 randomly chosen dots in F1 except for a uniform horizontal displacement of 10 min arc; that is, only 50% of the dots were displaced coherently. This was done to lower the level of performance and thus prevent excessive ceiling effects. Direction of displacement (left or right) was determined randomly on each trial. The displacement occurred entirely within the spatial confines of the 2 deg × 2 deg viewing area, with conventional wraparound for dots displaced outside the viewing area. The size of each dot was approximately 0.25 min arc.

Every motion sequence had two components, the single image and the combined image. The single image consisted of a single set of dots, either F1 or F2, according to the display sequence. The combined image contained both F1 and F2, a total of 320 dots. There were three basic sequences, as illustrated in figure 1. In the 0-SOA sequence, the display began with the combined image and ended with a single image; in the 0-STA (zero stimulus termination asynchrony) sequence, the display began with a single image and ended with the combined image; in the asynchronous (ASYN) sequence, the display began with a single image (F1), continued with the combined image, and ended with a single image (F2).

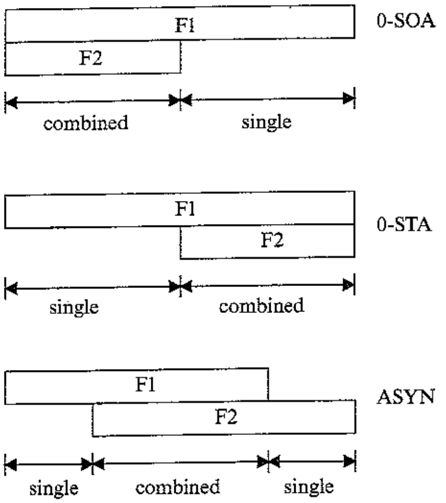

Figure 1. Schematic illustration of the temporal sequence of single and combined displays in the 0-SOA, 0-STA, and ASYN stimulus sequences.

There were six exposure durations of the single images and of the combined images: 10, 20, 40, 80, 160, and 320 ms. The six durations of the single images were combined factorially with the six durations of the combined images to yield a total of 36 conditions for each of the 3 basic sequences illustrated in figure 1, for a total of 208 temporal sequences.

2.3.1 *Intensity of the dots.* For all conditions in which the durations of F1 and F2 differed, an important consideration was time-intensity reciprocity known as Bloch's law. That is, up to a critical duration, the brightness (ie the perceived intensity) of a stimulus is determined jointly by its luminance and its duration. Such time-intensity reciprocity holds for intensities near threshold (Bloch's law) as well as above threshold (Butler 1975). This is a potential source of confounding and, as such, it deserves a detailed examination.

Suppose that F1 were longer than F2, and that both were displayed at the same luminance, time-intensity reciprocity would cause F1 to appear progressively brighter than F2 as its duration was increased. The resultant brightness mismatch between the two portions of the display would interfere with accuracy of performance (eg Eriksen and Collins 1967). Moreover, the effect of duration would be inextricably confounded with that of brightness. In earlier investigations, this confounding was resolved by displaying briefer stimuli at higher luminance values, so that all stimuli appeared of the same brightness regardless of duration (eg Bowen et al 1974; Servière et al 1977). This procedure yields equal-energy stimuli up to the critical duration. Its implementation is consistent with neurophysiological evidence showing that—within the critical duration—the magnitude of the visual response is determined not by stimulus luminance but by stimulus energy (eg Hartline 1934; Levick and Zacks 1970). Corroborating psychophysical evidence has been reported by Onley and Boynton (1962) who found that the magnitude of the ON response is related not to the luminance of the inducing stimulus but to its brightness. What is more, the latency of the visual response, which is known to differ for fixed-duration stimuli of different luminance, has been shown to be constant for brightness-matched stimuli of different durations. For example, Servière et al (1977) recorded the latency of cortical evoked-potential responses to brightness-matched stimuli that varied in duration from 0.5 to 320 ms. They found that the latency of the visual response was determined not by the luminance but by the brightness of the stimulus. That is, response latencies to brightness-matched stimuli were invariant with stimulus duration. Similar results, based on single-unit recordings, have been reported by Hartline (1934).

In accordance with earlier investigations, the present work employed stimuli of equal brightness. The process of brightness equalization was carried out in a separate psychophysical procedure similar to those described by Servière et al (1977) and by Di Lollo and Finley (1986). In this procedure, two identical random-dot stimuli are displayed briefly side by side for different durations (eg 20 and 30 ms, respectively). When the two are displayed at the same level of luminance, the longer stimulus appears brighter because of Bloch's law. The task of the observer is to adjust (reduce) the luminance of the longer stimulus until its brightness matches that of the shorter stimulus. The shortest stimuli (10 ms) were displayed at a luminous directional energy of 3.17 mcd μs; longer stimuli were displayed at corresponding lower levels of intensity, as shown in figure 2. The luminance of the background was below what could be measured reliably with our photometer; it was less than 0.1 cd m^{-2}.

In this context, two things should be noted; first, contrast levels were far above threshold for all stimuli. Second, it is known that the motion system saturates rapidly at low levels of contrast (Campbell and Maffei 1981). Thus, all stimuli were equally suitable as potential motion stimuli. For this reason, no provision for differential intensities was made in the theoretical model reported below.

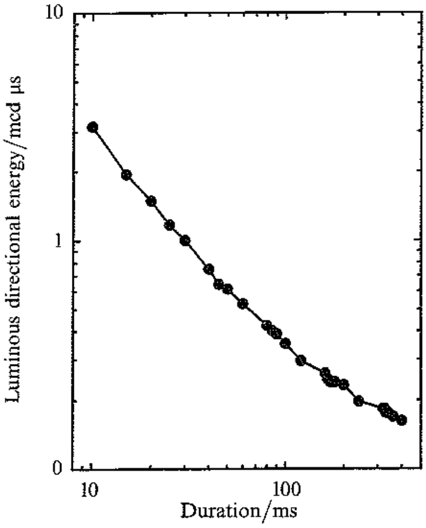

Figure 2. Luminous directional energy to which the observers set a comparison stimulus displayed for the duration shown on the abscissa, so that it appeared of the same brightness as a standard stimulus displayed for 10 ms at a luminous directional energy of 3.17 mcd μs. Each point defines a combination of intensity and duration that yields the same brightness as any other point.

2.4 *Procedure*

Observers sat in a dimly-lit observation chamber. On each trial, the sequence of events was as follows: a dim fixation cross (0.5 deg × 0.5 deg) was displayed in the centre of the screen. When ready, the observer pressed a push button on a hand-held box to initiate a trial. Upon a button press, the fixation cross disappeared and, after a 200 ms delay, the appropriate sequence of stimuli was displayed. Next, the observer reported (or guessed) the direction of motion (left or right) by pressing the appropriate button on the box. Finally, the fixation cross reappeared to indicate readiness for the next trial. An experimental session lasted about one hour. Each observer made 100 observations in each of the 108 experimental conditions.

3 Results

Individual results are shown in figure 3. Performance reached high levels in all three motion sequences (0-SOA, 0-STA, ASYN), thus confirming that directional motion

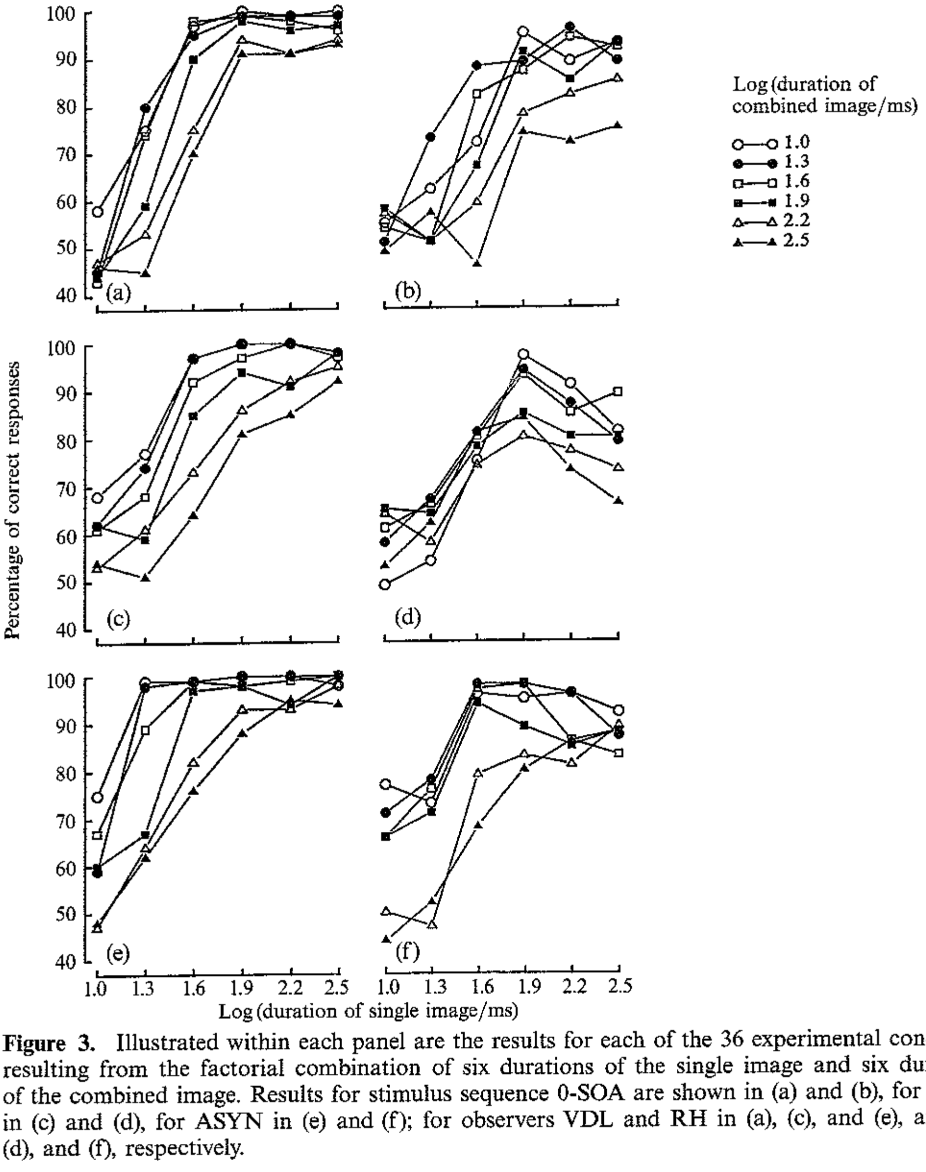

Figure 3. Illustrated within each panel are the results for each of the 36 experimental conditions resulting from the factorial combination of six durations of the single image and six durations of the combined image. Results for stimulus sequence 0-SOA are shown in (a) and (b), for 0-STA in (c) and (d), for ASYN in (e) and (f); for observers VDL and RH in (a), (c), and (e), and (b), (d), and (f), respectively.

can be seen at long SOAs or STAs. At long durations of the single images, clear directional motion was seen with each of the three stimulus sequences illustrated in figure 1. Motion was seen in the direction F2 – F1 with 0-SOA sequences, and in the direction F1 – F2 with 0-STA and ASYN sequences. In the remainder of this paper, these are referred to as the 'correct' directions of motion.

Perception of directional motion with a 0-SOA sequence had been reported earlier (Di Lollo et al 1993); the outcome of the 0-STA sequence is—to our knowledge—a novel result. Corresponding results with 0-STA sequences have been obtained in temporal-order judgments (Happ and Pantle 1987). The phenomenological appearance of the displays is worth noting; at very brief durations of the single component, observers saw a field of pairs of dots that appeared to remain relatively static. At longer durations of the single component (40 – 80 ms), the dots were seen in smooth, coherent motion whose quality was indistinguishable from that seen with strictly sequential displays. At very long durations of the single component, motion was seen reliably, but its direction was harder to detect.

The performance curves in figure 3 reflect the phenomenological appearance of the displays. In all conditions, performance improved as the duration of the single component was increased to about 40 ms. At longer durations, performance became asymptotic except in the 0-STA sequence for observer RH, where it declined slightly. In contrast to the progressive improvement as a function of the duration of the single component, performance decreased monotonically as the duration of the combined image was increased. The monotonic decrement can be seen in figure 3 by noting the order of the curves within each panel, particularly in condition 0-STA for observer VDL.

4 Discussion

4.1 *Baker and Cynader's model*

Baker and Cynader's model of directional motion selectivity [Baker and Cynader 1994; figures 4 and 5 (section 4.2)] can be represented as in figure 4, which is a schematic representation of a directional motion sensor.

In figure 4, two-sequential stimuli, F1 and F2, are assumed to activate the motion sensor through spatially separate inputs I_L and I_R, respectively. Each stimulus is assumed to lead to optimal stimulation of a pair of temporal filters, S and T, at each spatial position. Filters T_L and T_R are assumed to generate transient responses, whereas S_L and S_R are assumed to generate sustained responses. More precisely, the transient filter T, taken from Adelson and Bergen (1985), is defined by the filter equation

$$T(t) = \left(\frac{t}{\tau}\right)^n \exp\left(-\frac{t}{\tau}\right)\left[\frac{1}{n!} - \frac{t^2}{\tau^2(n+2)!}\right], \tag{1}$$

with $n = 3$ and $\tau = 10$ ms. The sustained filter S is given by the integral of equation (1),

$$S(t) = \int T(t)\mathrm{d}t = \left(\frac{t^5}{5!\,\tau^4} + \frac{t^4}{4!\,\tau^3}\right)\exp\left(-\frac{t}{\tau}\right), \tag{2}$$

again with $n = 3$ and $\tau = 10$ ms. The output of each filter is half-rectified before being combined in C_L and C_R, with C_L responding optimally to motion in the direction $I_\mathrm{R} - I_\mathrm{L}$ and C_R responding optimally to motion in the direction $I_\mathrm{L} - I_\mathrm{R}$. Following Baker and Cynader (1994), the components C_L and C_R are modelled by means of heterosynaptic facilitation, with $C = T(1 + S)$. Directional selectivity is given by the difference $C_\mathrm{R} - C_\mathrm{L}$ in the output unit Σ. Finally, the response of Σ, integrated over time, is a predictor of the directional selectivity of a given motion-sensing neuron. It is obvious from figure 5 that Baker and Cynader's model reduces to that of Reichardt

(1961), given removal of the half-wave-rectification stage, and redefinition of S as $T(t - \Delta)$, where Δ corresponds to the temporal constant of a delay unit.

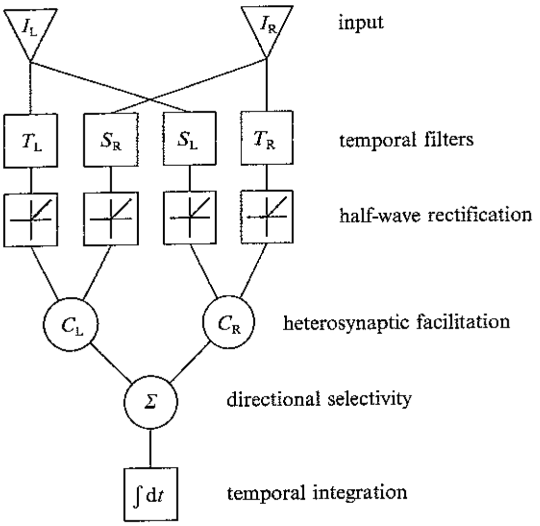

Figure 4. Schematic illustration of a modified Reichardt motion sensor. Each input to the sensor passes through a spatial filter (I_L or I_R) and a pair of temporal filters, one transient (T_L or T_R) the other sustained (S_L or S_R), followed by half-wave rectification. Two mirror-symmetrical correlators (C_L and C_R) multiply one transient signal by the other sustained signal and provide input to Σ, where the two products are subtracted one from the other ($C_R - C_L$). Finally, the difference is integrated over time to produce a motion signal. Positive, negative, or zero sums indicate motion to the right, to the left, or no motion, respectively.

4.2 *Applying Baker and Cynader's model to psychophysical data*

Application of the model to human psychophysical data is predicated on the assumption that the output of the motion sensor, integrated over time as in figure 4, can be taken as an index of the strength of the motion signal, and hence as a predictor of performance in the motion-detection task. As did Baker and Cynader, we aimed at providing a qualitative account of the experimental outcome, as distinct from a quantitative fit of the data.

As a first approximation, we considered only average levels of performance in each of the experimental conditions illustrated in figure 3. Average performance curves for observer VDL are shown in figures 5, 6, and 7 for conditions 0-SOA, 0-STA, and ASYN, respectively. Performance averaged over all durations of the combined image is shown in figures 5a, 6a, and 7a. Performance averaged over all durations of the single image is shown in figures 5d, 6d, and 7d. The curves for observer RH were very similar to those for VDL, and have not been included in the figures. Corresponding averages of simulated scores from Baker and Cynader's model are shown in panels (b) and (e) of figures 5, 6, and 7. A different arbitrary scaling factor was used in each of the figures, with the constraint that, within each figure, the same factor was used in panels (b) and (e).

Bearing in mind that it had been developed as an account of single-unit responses in the cat visual system, the model meets with credible—though mixed—success in accounting for human psychophysical performance. For each of the three stimulus sequences—0-SOA, 0-STA, and ASYN—the model succeeds in predicting performance as a function of the duration of the single image. This is seen by comparing panels (a) and (b) in figures 5, 6, and 7; in every case, the theoretical curves provide good qualitative matches to the empirical curves. The predictions of the model fail, however, if the data are plotted as a function of the duration of the combined image. This is seen by comparing panel (d) with panel (e), separately in figures 5, 6, and 7. In every case, the empirical curves show a steady decline as the duration of the combined image is increased. By contrast, the predicted curves show an inverted-U function for the 0-SOA sequence (figure 5e) and a steady increment for the 0-STA and ASYN sequences (figures 6e and 7e). With reference to the performance curves in figure 3, it can be said that the model predicts accurately the progressive improvement in performance as a function of the duration of the single image, but the order of the curves within each panel (ie the effect of duration of the combined image) is mostly reversed. Perhaps the most orderly example of such reversal can be seen in the results of VDL with the 0-STA sequence in figure 3. At the single-image duration of 40 ms, performance was determined largely by—and was related inversely to—the duration of the combined image.

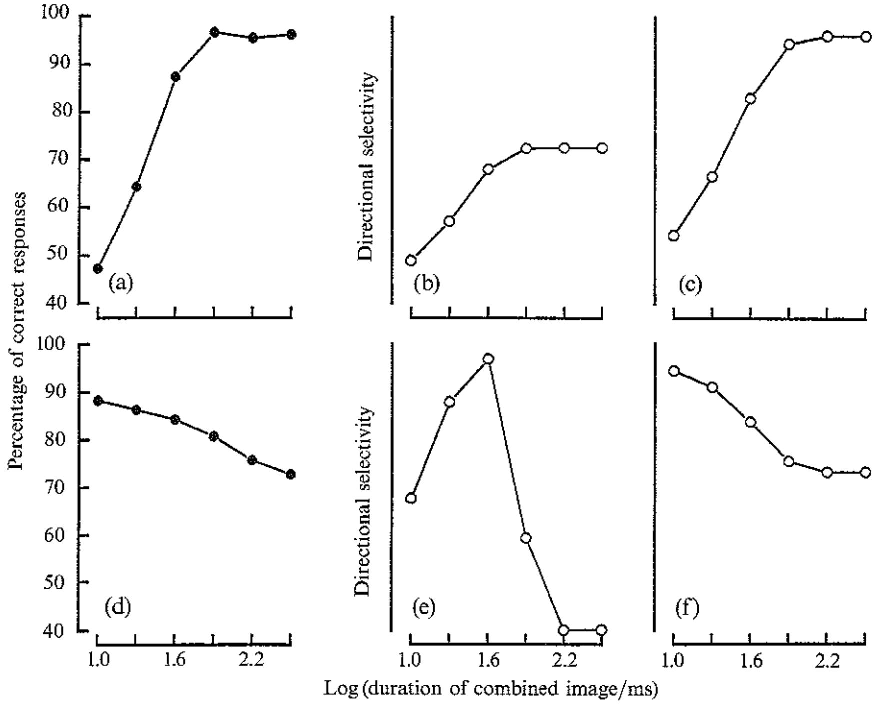

Figure 5. Obtained (for observer VDL) and simulated results with the 0-SOA stimulus sequence. (a) Percentage of correct responses at each duration of the single image, averaged over all durations of the combined image. (d) Percentage of correct responses at each duration of the combined image, averaged over all durations of the single image. (b) and (e) correspond to (a) and (d), respectively, and show simulated results obtained by using Baker and Cynader's (1994) original model. The ordinates are scaled in arbitrary units; the same scaling factor was used in both panels. (c) and (f) correspond to (a) and (d), respectively, and show simulated results obtained by using the elaborated model described in the text. The ordinates are scaled in arbitrary units; the same scaling factor was used in both panels.

4.3 *An elaboration of Baker and Cynader's model*

In a nutshell, the model responds appropriately to changes in duration of the single—but not of the combined—images. This partial failure is not surprising, given the sources of the two sets of data; one represents the output of a single unit, the other that of a total system. Just as important, Baker and Cynader employed only ASYN-type sequences in their experiments, and thus did not obtain data with 0-SOA and

0-STA sequences that demand a more comprehensive account. Baker and Cynader's original model can be elaborated so as to account for the psychophysical results, as is shown below.

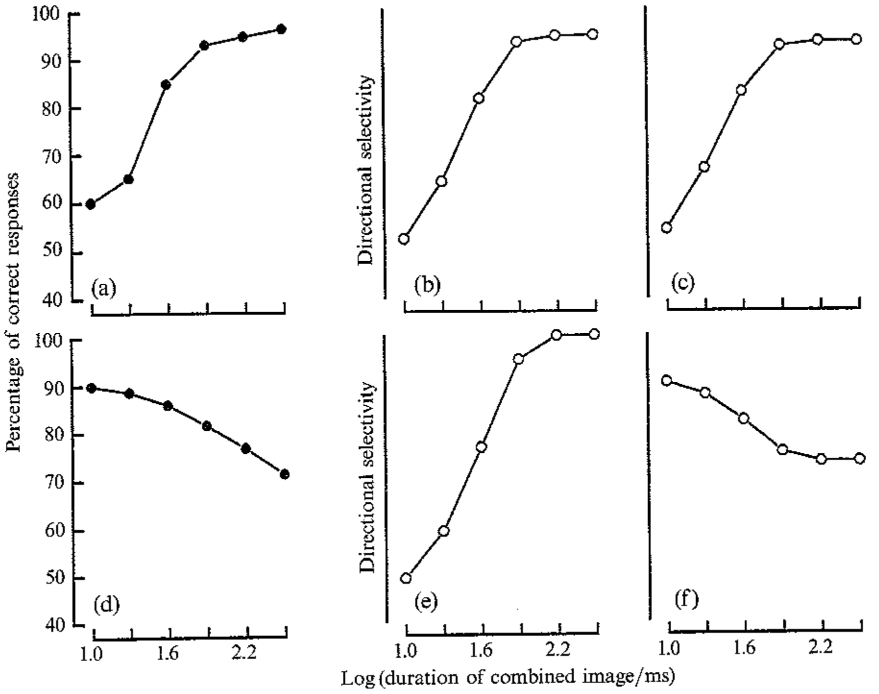

Figure 6. As figure 5, for the 0-STA stimulus sequence.

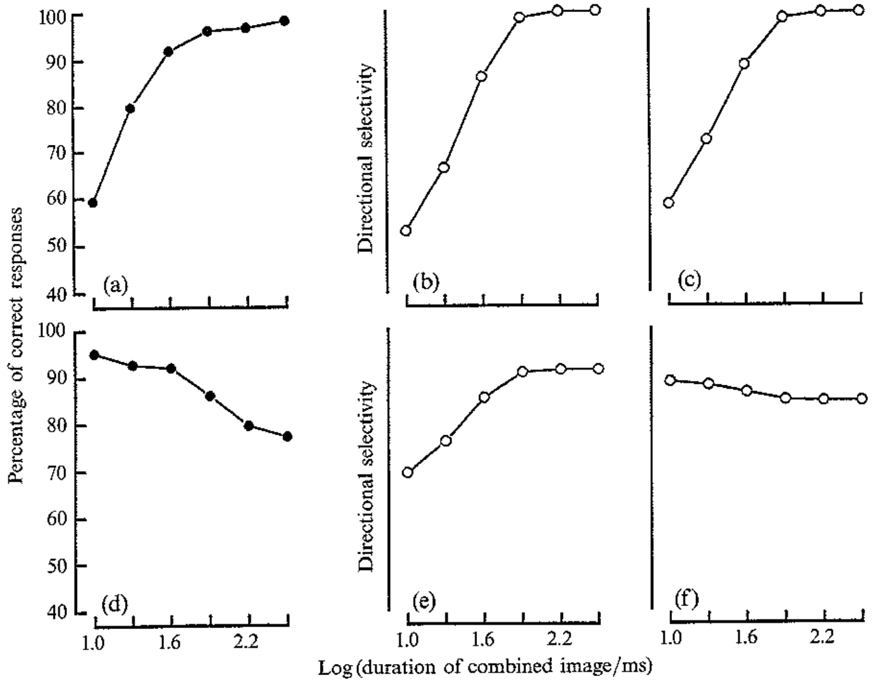

Figure 7. As figure 5, for the ASYN stimulus sequence.

Two changes were made to the original model. First, the rectification stage was omitted, and, second, the magnitude of the motion signal was normalized as follows: instead of defining the output of the motion sensor as $\Sigma(t) = C_{\mathrm{R}}(t) - C_{\mathrm{L}}(t)$, we define it as

$$\Sigma(t) = \frac{C_{\mathrm{R}}(t) - C_{\mathrm{L}}(t)}{|C_{\mathrm{R}}(t)| + |C_{\mathrm{L}}(t)|}, \tag{3}$$

where C_{R} and C_{L} are the same as in figure 4.

Simulated outputs from the elaborated model are shown in panels (c) and (f) of figures 5, 6, and 7. In each case, the simulated curve provides a good qualitative match to the corresponding empirical curve. In particular, the progressive impairment in performance with increasing duration of the combined image is predicted correctly. This improvement in the model was achieved by omitting rectification and by introducing normalization. We now consider the rationale for the two changes and their significance for models of motion perception.

4.3.1 *Removal of half-wave rectification.* Half-wave rectification is not an essential requirement of the original model. Baker and Cynader (1994, page 1089) note that similar results are obtained by omitting the rectification stage, as might be the case in a simulation of opposite-signed LGN cells operating in 'push–pull'. We have found that this holds true for the ASYN sequence (the only sequence used by Baker and Cynader) and for the 0-STA sequence. In the 0-SOA sequence, however, any form of rectification leads to incorrect predictions. This is so because, at long durations of the combined image in the 0-SOA sequence, the only transient event that is available to trigger the motion signal is an OFF response (see figures 1 and 4). Thus, given half-wave rectification, no motion signal is generated because the OFF response cannot pass the rectification stage. This effect can be seen in figure 5e where, at the longest durations of the combined image, the motion signal has value of zero. The problems with half-wave rectification cannot be remedied by implementing full-wave rectification. In the 0-SOA sequence, full-wave rectification yields reverse-direction motion signals at all durations of the combined image because the polarity of the OFF response is reversed by rectification.

In brief, motion signals can be produced in 0-SOA sequences, provided that rectification is omitted, thus enabling the OFF responses. Absence of rectification does not prevent the triggering of motion signals in other stimulus sequences, as has also been noted by Baker and Cynader (1994).

4.3.2 *Normalization.* In Baker and Cynader's original model, the strength of the motion signal is determined solely by the difference $C_{\mathrm{R}} - C_{\mathrm{L}}$ (see figure 4). By this rule, the model predicts that strength of the motion signal increases both with the duration of the single image (see figures 5b, 6b, and 7b) and, allowing for one minor exception (figure 5e), with the duration of the combined image (see figures 5e, 6e, and 7e). As is obvious from the simulated results, the rule is successful with single images but not with combined images; the empirical data show that performance actually declines as duration of the combined image is increased (figures 5d, 6d, and 7d).

To account for the effect of the combined image, we first consider a thresholding mechanism. It is possible that the simultaneous presence of both images acts as a suppressor in the production of the motion signal. More specifically, it could be surmised that, in order to trigger a motion signal, the magnitude of the imbalance between the inputs must exceed a given threshold level. On the assumption that prolonged concurrent input from the combined image acts to increase the threshold level, progressive hindering of the motion signal would follow. Although appealingly simple, this option cannot be true. As seen in figures 5d, 6d, and 7d, directional-motion perception—and, by inference, strength of the motion signal—is impaired whether the period of concurrent activation precedes (0-SOA) or follows (0-STA) the $C_R - C_L$ imbalance. In the 0-STA sequence, the motion signal is triggered at the onset of the combined image and, therefore, *before* the threshold could possibly have been increased by the concurrent activation of the inputs.

The threshold-increment rule is denied by the empirical evidence. At the same time, the evidence suggests that factors antagonistic to the strength of the motion signal must be at work. Further, the evidence suggests that the prominence of those factors increases with the period for which C_R and C_L are active simultaneously (ie as the duration of the combined image is increased). Normalization—as in equation (3)—captures the interaction of the opposing factors: the agonistic effect of the $C_R - C_L$ imbalance expressed in the numerator is modulated by the antagonistic effect of simultaneous activation expressed in the denominator. The strength of the ensuing motion signal increases with the $C_R - C_L$ imbalance, and decreases with the period of concurrent activation of the two inputs.

In practice, normalization permits a good qualitative account of the effects of duration both of the single and of the combined images [see panels (c) and (f) in figures 5, 6, and 7]. Our approach to normalization [equation (3)] is closely related to Heeger's (1992) divisive normalization and to Geisler and Albrecht's (1992) multiplicative contrast-gain control mechanisms. Our equation is simpler than either Heeger's or Geisler and Albrecht's because we did not take into account contrast-response and saturation effects. We could omit these factors because our stimuli were far above threshold, and it is known that the motion system saturates rapidly at low contrast levels.

4.4 *Related developments*

In the present work, we have studied motion perception with nonconventional temporal sequences, 0-SOA, 0-STA, and ASYN. Each had been used separately in earlier studies: 0-SOA sequences were used by Di Lollo et al (1993) in psychophysical studies of motion perception and metacontrast masking; 0-STA sequences were used by Happ and Pantle (1987) in temporal-order judgments; ASYN sequences were used by Baker and Cynader (1994). In the present experiments, all three sequences were investigated systematically so as to permit a comprehensive assessment of the temporal factors at work. The results lead to two main conclusions: first, that there is a sustained input to directional-motion mechanisms and, second, that a normalization component must be part of any satisfactory theoretical account.

Our evidence for the existence of a sustained input to directional motion mechanisms was not totally unanticipated. Aside from the work of Baker and Cynader (1994), Emerson et al (1992) have proposed a sustained component in a model of directionally selective complex cells. Congruent psychophysical evidence has been provided by Wilson (1985) in a model of directional selectivity in threshold motion perception, and by Strout et al (1994) in an account of interframe effects in apparent motion.

In the course of developing a sustained–transient (ST) model, we considered and evaluated other approaches involving either transient–transient (TT) or sustained–sustained (SS) interactions. Among the TT models, we tested that of Adelson and Bergen (1985) in which both channels are described by equation (1). We also considered the variation proposed by Emerson et al (1992), and varied the transiency parameter β over the entire range, thus covering the spectrum between TT ($\beta = 1$) and SS ($\beta = 0$). Among the SS models, we tested Strout et al's (1992) 'phase II' model, which is simply Emerson et al's model with $\beta = 0.1$. In every case, we analyzed model predictions with and without normalization. We found that, at best, all models provided only partial accounts of the data. With normalization, the most common failure was in respect to the duration of the single images. Also, reverse motion was predicted in several instances.

Without normalization, all models failed to account for the duration of the combined images. Amongst all the models that we considered, only the ST model reported above provided a complete and satisfactory account of the empirical results.

Acknowledgements. This work was supported by Grants OGP38521 and OGP0006592 from the Natural Sciences and Engineering Research Council of Canada to W F Bischof and to V Di Lollo, respectively.

References

AdelsonE H, Bergen J R, 1985 "Spatiotemporal energy models for the perception of motion" *Journal of the Optical Society of America A* **2** 284–299

Baker C L Jr, Braddick O, 1985 "Temporal properties of the short-range process in apparent motion" *Perception* **14** 181–192

Baker C L Jr, Cynader M S, 1994 "A sustained input to the direction-selective mechanism in cat striate cortex neurons" *Visual Neuroscience* **11** 1083–1092

Bowen R W, Pola J, Matin L, 1974 "Visual persistence: Effects of flash luminance duration and energy" *Vision Research* **14** 295–303

Butler T W, 1975 "Luminance-duration relationships in the photopic ERG and the apparent brightness of flashes" *Vision Research* **15** 693–698

Campbell F W, Maffei L, 1981 "The influence of spatial frequency and contrast on the perception of moving patterns" *Vision Research* **21** 713–721

Di Lollo V, Bischof W F, Dixon P, 1993 "Stimulus-onset asynchrony is not necessary for motion perception or metacontrast masking" *Psychological Science* **4** 260–263

Di Lollo V, Finley G, 1986 "Equating the brightness of brief visual stimuli of unequal durations" *Behavior Research Methods, Instruments, and Computers* **18** 582–586

Emerson R C, Bergen J R, Adelson E H, 1992 "Directionally selective complex cells and the computation of motion energy in cat visual cortex" *Vision Research* **32** 203–218

Eriksen C W, Collins J F, 1967 "Some temporal characteristics of visual pattern perception" *Journal of Experimental Psychology* **74** 476–484

Finley G, 1985 "A high-speed point plotter for vision research" *Vision Research* **25** 1993–1997

Geisler W S, Albrecht D G, 1992 "Cortical neurons: isolation of contrast gain control" *Vision Research* **32** 1409–1410

Happ A, Pantle A, 1987 "Visual motion processing and display design", in *Ergonomics and Human Factors: Recent Research* Eds L S Mark, J S Warm, R L Huston (New York: Springer) pp 85–94

Hartline H K, 1934 "Intensity and duration in the excitation of single photoreceptor units" *Journal of Cellular and Comparative Physiology* **5** 229–247

Heeger D J, 1992 "Normalization of cell responses in cat striate cortex" *Visual Neuroscience* **9** 181–197

Levick W R, Zacks J L, 1970 "Responses of cat retinal ganglion cells to brief flashes of light" *Journal of Physiology (London)* **206** 677–700

Onley J W, Boynton R M, 1962 "Visual responses to equally bright stimuli of unequal luminance" *Journal of the Optical Society of America* **52** 934–940

Reichardt W, 1961 "Autocorrelation a principle for the evaluation of sensory information", in *Sensory Communication* Ed. W A Rosenblith (Cambridge, MA: MIT Press) pp 303–317

Santen J P H van, Sperling G, 1985 "Elaborated Reichardt detectors" *Journal of the Optical Society of America* **24** 300–320

Servière J, Miceli D, Galifret J, 1977 "A psychophysical study of the visual perception of 'instantaneous' and 'durable'" *Vision Research* **17** 57–63

Sperling G, 1971 "The description and luminous calibration of cathode ray oscilloscope visual displays" *Behavior Research Methods and Instrumentation* **3** 148–151

Strout J J, Pantle A, Mills S L, 1994 "An energy model of interframe interval effects in single-step apparent motion" *Vision Research* **34** 3223–3240

Watson A B, Ahumada A J, 1985 "Model of human motion-vision sensing" *Journal of the Optical Society of America* **24** 322–341

Wilson H R, 1985 "A model for direction selectivity in threshold motion perception" *Biological Cybernetics* **51** 213–222

© 1996 a Pion publication printed in Great Britain