Vision Res. Vol. 17, pp. 893 to 902. Pergamon Press 1977. Printed in Great Britain.

PATTERNS OF TEMPORAL INTERACTION IN THE DETECTION OF GRATINGS[1]

ANDREW B. WATSON and JACOB NACHMIAS
Department of Psychology, University of Pennsylvania, Philadelphia, PA 19174, U.S.A.

(*Received* 4 *June* 1976; *in revised form* 19 *October* 1976)

Abstract—Threshold contrasts were determined for pairs of brief, temporally separated gratings of the same frequency, same or opposite phase, and various contrast ratios. For low spatial frequencies, a range of separations is found that results in facilitation between opposite-phased pairs and inhibition between same-phased pairs. For higher spatial frequencies, this range is absent or nearly absent. The relation between the contrasts of the two gratings in a threshold pair of fixed separation suggests either a detector operating on the squared, integrated output of a linear temporal filter, or probability summation over time of the noise-perturbed output of such a filter.

INTRODUCTION

Over the past decade much work has been directed towards a comprehensive theoretical description of the sensitivity of the eye to spatial distributions of luminance. For the most part, candidate theories have embodied one or more spatial filters. Similar efforts have been directed towards an understanding of the temporal sensitivity of the eye, and here models have invoked one or more temporal filters. In each case, attempts have been made to specify the form of the filter or filters primarily by measuring the visibility of spatial and temporal sinusoids.

However, the form of the spatial modulation sensitivity function is dependent upon the temporal distribution of the stimulus, or equivalently, the form of the temporal modulation sensitivity function depends upon the spatial distribution of the stimulus (Kelly, 1959; Schober and Hilz, 1965; Robson, 1966; Nachmias, 1967). The nature of this interaction between spatial and temporal sensitivities was first clearly shown by Robson, who measured the detectability of patterns whose luminance varied sinusoidally in both space and time. As the temporal frequency was reduced, a decline was observed in the sensitivity to low spatial frequencies. Likewise, as the spatial frequency was reduced, a decline was seen in the sensitivity to low temporal frequencies.

In this paper, further evidence is provided for a distinction among the temporal responses of the visual system to gratings of different spatial frequencies. In contrast to the simple determination of the visibility of spatial and temporal sinusoids, the method employed here taps more directly the principle of superposition in the temporal domain, and allows a more detailed analysis of those stages of the detection process which follow the initial spatio-temporal filtering of the stimulus. The procedure is adapted from the work of Rashbass (1970) in which the sensitivity to a pair of brief, temporally separated changes in the luminance of a large uniform field was used to infer the nature of the temporal detection process. In essence, the present experiments extend this technique to the case in which the stimulus contains variations in both space and time.

METHODS

Observers

Three observers (RP, OZ, and AW) participated in this study. The first two are emetropic; AW, the first author, is a well-corrected myope.

Stimuli

Sinusoidal gratings were generated on the screen of a cathode ray tube (P31 Phosphor) by means of Z-axis modulation of a high frequency raster. The screen subtended 2° by 1.5° and was surrounded by a 6° surface of approximately the same color and average luminance (approx 15 cd/m^2). It was viewed binocularly from a distance of 285 cm with a chin rest and natural pupils. A small spot at the center of the screen was used for fixation.

For a vertical sinusoidal grating, the equation

$$L(x) = L_0(1 + m \sin 2\pi fx) \quad (1)$$

describes the luminance at a horizontal distance x from some arbitrary spatial origin, where L_0 is the space average luminance of the screen, f is the spatial frequency, and m is the contrast of the grating. All stimuli in the present experiments can be represented as temporal variations in the value of m. Each stimulus consisted of a pair of brief presentations of a grating, both of the same spatial frequency, and each of duration δ, with their onsets separated by an interval τ. The temporal waveform of the stimulus is then a pair of rectangular pulses, as shown in Fig. 1. The size of each pulse represents the contrast of the grating, while positive and negative signs indicate spatial phases 180° apart. For fixed values of f, τ, and δ, a stimulus may then be specified by a point in a two-dimensional coordinate space; the first coordinate, a, represents the contrast and phase of the first grating, the second coordinate, b, represents the contrast and phase of the second grating.

Ratios of a to b of 0.5, −0.5, 1, −1, 2, or −2 were used. In addition, either a or b might be set to zero. When the sign of a or b is considered, these conditions yield a total of 16 stimulus types. When viewed in the coordinate

[1]Some of this work was reported in April 1976 at the annual meeting of the Association for Research in Vision and Ophthalmology.

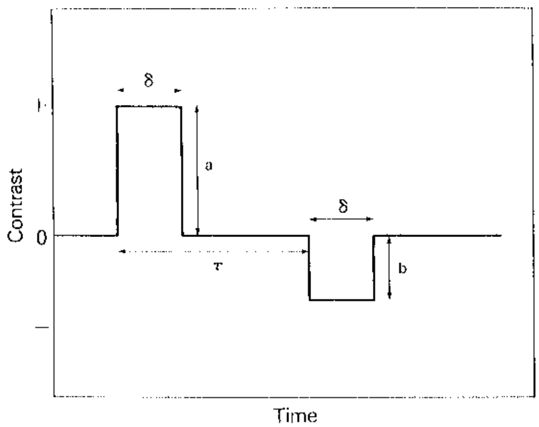


Fig. 1. Schematic representation of the stimuli used in these experiments. The height of the heavy line shows the amount of contrast present in a grating at each point in time. Positive and negative contrasts indicate spatial phases 180° apart.

space, each type constitutes a radius extending outward from the origin.

During the experiment the absolute position (phase origin) of the grating was altered by a random amount between each stimulus presentation, in order to prevent local adaptation and phase dependent effects due to fixation. The distribution of phase origins was approximately uniform over 0.5°.

Spatial frequencies (f) of 1.75, 3.5, 7.0, and 10.5 c/deg, and pulse onset asynchronies (τ) between 10 and 200 msec were used. The method of producing these gratings (intensity modulation of individual frames of a raster) requires that durations of stimuli be measured in integral multiples of the frame period. For RP, each pulse was one frame in duration, with a frame period of 5.26 msec. For OZ and AW the pulse duration was two frames, with a frame period of 2.86 msec, for a nominal duration of 5.72 msec.

Procedure

Thresholds were determined by the method of adjustment. The adjustment was made during repetitive presentation of one stimulus type. Each presentation was accompanied by a tone marker, and was separated from the next presentation by an interval of 1 sec. During an experimental session, the values of f, τ, and δ were fixed. A session was composed of three sets of 16 adjustments. Within each set, threshold was determined once for each of the stimulus types, in a pseudo-random order.

The use of the random phase shift theoretically obviates the need for eight of the 16 stimulus configurations. For this reason certain sessions were composed of five sets of eight adjustments. The results did not appear different in any way.

In different sessions, observers were asked to use one of two threshold criteria. In the "detection criterion" condition, the observer was instructed to adjust the stimulus contrast until the screen was seen as either spatially or temporally inhomogeneous, that is, until it was seen as differing from a screen whose brightness did not vary in either space or time. In the "pattern criterion" condition, the adjustment was made to the point at which the vertical striations of the grating became just evident.

RESULTS

In Fig. 2a we have plotted the mean threshold adjustments of observer RP for the set of 16 stimulus types obtained with a pulse onset asynchrony (τ) of 10.5 msec and a spatial frequency of 3.5 c/deg. The pattern criterion was used. The horizontal and vertical axes represent the contrasts and phases of the first and second grating pulses, respectively, scaled by a factor whose derivation will be given below.

Examine first the points lying on the axes. Each represents the mean threshold contrast for a single grating pulse; the gratings to which they refer differ only in spatial phase and temporal position. It is not surprising that their thresholds are very similar. The points lying within the first and third quadrants represent the thresholds for pairs of gratings of the same phase. It is evident that for an onset asynchrony of 10.5 msec, nearly linear summation is obtained between same-phased grating pairs. The second and fourth quadrants contain thresholds for opposite-phased grating pairs. At this asynchrony we observe a severe antagonism between the elements of the pair.

Figures 2b–d indicate the result obtained under the same conditions but with increasing pulse onset asynchrony. At 37 msec (Fig. 2b) the points are arrayed in a nearly circular configuration about the origin. Note that in this case a same-phased grating pair is about as detectable as an opposite-phased pair of equivalent contrast. At 58 msec (Fig. 2c), it can be seen that *less* contrast is required to detect an *opposite*-phased grating pair than a same-phased pair. In Fig. 2d, with a separation of 180 msec, the threshold points array themselves in a "rounded square" fashion about the origin. At this point it will suffice to note that results of this sort are consistent with the notion that the two pulses are detected independently, interacting only through probability summation.

Recently, Rashbass (1970) proposed a model for the detection of brief-luminance changes in a spatially homogeneous visual field. The model consists of a linear temporal filter with impulse response $h(t)$, a squaring operation, followed by integration over a fixed interval T. According to this theory, 2 brief luminance changes of duration δ and onset asynchrony τ are at threshold when their amplitudes, a and b, satisfy the following equation:

$$\int_0^T [a\delta h(t) + b\delta h(t-\tau)]^2 \, dt = 1. \quad (2)$$

Where the expression in parentheses is zero outside the epoch of integration, i.e. where

$$h(t-\tau) = 0 \qquad \text{for} \qquad t > T \quad (3)$$

the combination of amplitudes a and b at the threshold for the pair will describe an ellipse:

$$a^2\lambda^{-2} + b^2\lambda^{-2} + 2ab\lambda^{-2}\kappa(\tau) = 1. \quad (4)$$

The parameter λ is easily shown to be the threshold for a single pulse, while $\kappa(\tau)$ is equal to $(\lambda\delta)^2$ times the autocorrelation of the impulse response. This model provided a reasonable account of Rashbass' data, that is, threshold contours appeared well fit by ellipses.

In Figs. 2a–c the continuous curves are ellipses, fitted to the points under the following constraints:

(1) the sum of the squared departures of the points from the curve, along the appropriate radii, was minimized;

(2) the center of the ellipse was at the origin; and

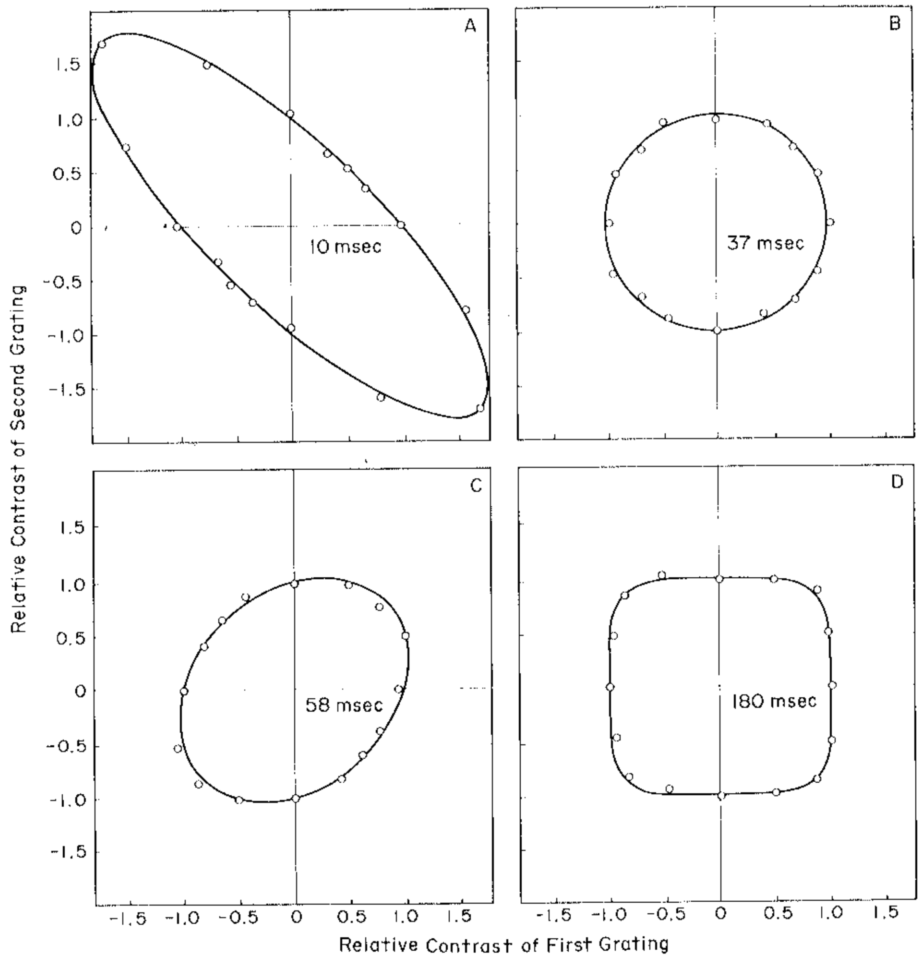


Fig. 2a–d. Each figure is a plot of the contrasts of first and second gratings in threshold pairs of fixed onset asynchrony. The onset asynchrony is indicated near the center of each figure. Each point is the mean of three adjustments. The data are from observer RP, at a spatial frequency of 3.5 c/deg, and using the pattern criterion. The smooth threshold contour curves are described in the text.

(3) the axes of the ellipse had slopes of +1 and −1.

Such an ellipse has the form of equation (4). The curve fitting procedure described above can be shown to provide best estimates of the parameters κ and λ, which reflect respectively the eccentricity and size of the ellipse. In this and all following estimation procedures a fast, general minimization program was used (Chandler, 1965).

Visual inspection of the data from three observers suggest that the ellipse provides a good fit at the four spatial frequencies tested, provided that the pulse onset asynchrony is less than 100 msec. The quality of fit may be quantified by means of a likelihood ratio statistic whose asymptotic distribution is χ^2. The essence of the statistic is a comparison of the hypotheses that the true threshold for each stimulus type is given by the mean of the three adjustments or by the value predicted by the best fitting ellipse. Sixty-eight sessions were obtained in which 16 stimulus types were used. For 39 of the 48 sessions in which τ was less than or equal to 100 msec, the fit cannot be rejected at the 0.05 level of confidence. Of the 20 remaining sessions only 7 meet this criterion. It is evident that the fit declines at long separations. [A modified version of this statistic has been applied to the data of sessions in which five adjustments of eight stimulus types were made, and no appreciable differences were found.]

Quite apart from any theoretical interpretation it may be given, the relatively good fit of the elliptical threshold contour justifies the use of the parameter κ as an index of the degree of summation at a particular onset asynchrony. The theoretical limits of κ are +1, indicating complete linear addition, and −1, indicating complete linear subtraction. In Figs. 2a–c the obtained values of κ are +0.828, +0.021, and −0.261, respectively.

Recall also that λ provides an estimate of the threshold contrast for a single pulse, and should thus be independent of the pulse onset asynchrony. This expectation was generally confirmed, and so we may use λ^{-1} as an index of sensitivity at each spatial frequency. In Figs. 2a–d the axes are expressed in units of λ.

The remaining data for observer RP at 3.5 c/deg using the pattern criterion are summarized by the filled circles in Fig. 3b, where the estimated values

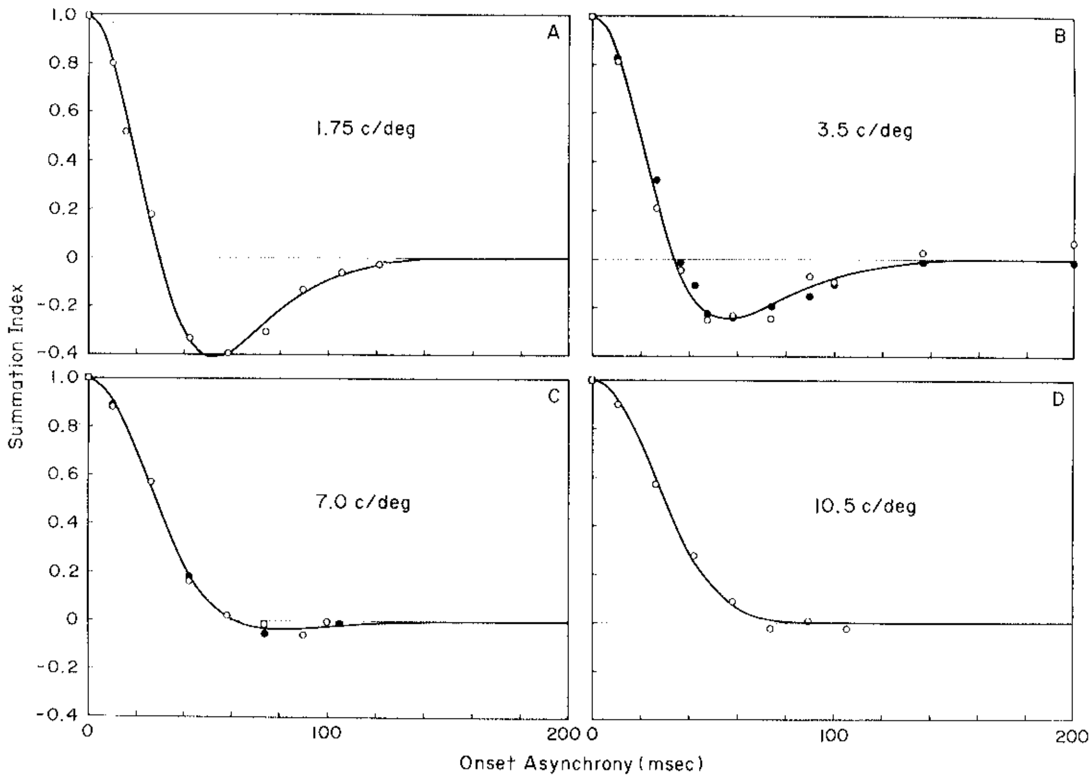


Fig. 3. Summation index as a function of onset asynchrony for four spatial frequencies for observer RP. Each value of the index is obtained from least squares fit of an ellipse to the data from one session. Each point is the mean of one or two values. Open circles: detection criterion; filled circles: pattern criterion. Open squares at 0 msec are theoretical. Each continuous curve is the difference between 2 gaussians, as explained in the text.

of the parameter κ are plotted as a function of pulse onset asynchrony (τ). Each point is the mean of one or two sessions. In this and succeeding figures, the plotted point at 0 msec asynchrony was not experimentally determined, but rather represents the value of κ that should *theoretically* be obtained when the two pulses are superimposed in time. This summation function is seen to proceed from a large positive value to zero at about 35 msec, to a negative minimum at about 60 msec and to return to zero at something over 100 msec. The continuous curves that appear in this and succeeding figures are the best fitting members of the family of curves described by the equation

$$\kappa = c_1 \exp(-c_2^2\tau^2) + (1 - c_1)\exp(-c_3^2\tau^2), \quad (5)$$

where c_1, c_2, and c_3 are constants. This is simply the sum of two arbitrary Gaussians, constrained only in that κ must equal 1 when $\tau = 0$. Gaussians were used only for convenience.

Figures 3a–d summarize all of the results for observer RP. Open circles indicate use of the detection criterion, filled circles, the pattern criterion. A spatial frequency of 1.75 c/deg (Fig. 3a) generates a summation function that is biphasic, with a large negative lobe. At 3.5 c/deg (Fig. 3b) the negative lobe remains but is less prominent. At still higher spatial frequencies (Figs. 3c and d) the negative phase is absent or nearly absent. It should be noted that though the points at asynchronies of greater than 100 msec are obtained from poorly fitting ellipses, they continue to reflect the relative sensitivity to same and opposite phased grating pairs.

It is evident from the data of Figs. 3b and c that the change from detection to pattern criterion produced little or no change in the summation function at these two frequencies, though average sensitivity (λ^{-1}) was reduced by 2.9 and 1.3 dB, respectively.

Data from the remaining two observers are shown in Figs. 4a–c. With the exception of a less pronounced negative lobe at 3.5 c/deg, these results are very similar to those of observer RP.

DISCUSSION

Conceptual framework

It will be useful in the discussion of these results to make explicit a model of spatio-temporal visual analysis. A block diagram of the model is shown in Fig. 5. We suppose a number of independent channels, which differ in their spatial sensitivities. Support for this notion is present in the considerable body of work suggesting a range of spatially selective channels (for example, Graham and Nachmias, 1971; Thomas, 1970). Where the transfer function of a spatio-temporal filter may be expressed as the product of spatial and temporal transfer functions, the filter may be decomposed into an equivalent sequence of a purely spatial and a purely temporal filter. Where this is possible the order of the component filters is immaterial. For theoretical convenience, a separability of variables will be assumed. We thus decompose

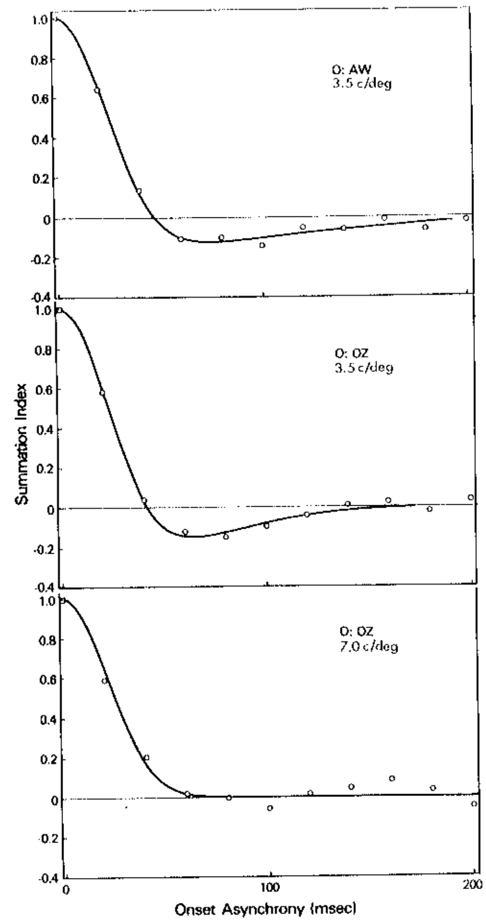


Fig. 4. Summation index as a function of onset asynchrony for observers AW and OZ. Top figure: 3.5 c/deg; middle: 3.5 c/deg; bottom: 7.0 c/deg. Details as in Fig. 3.

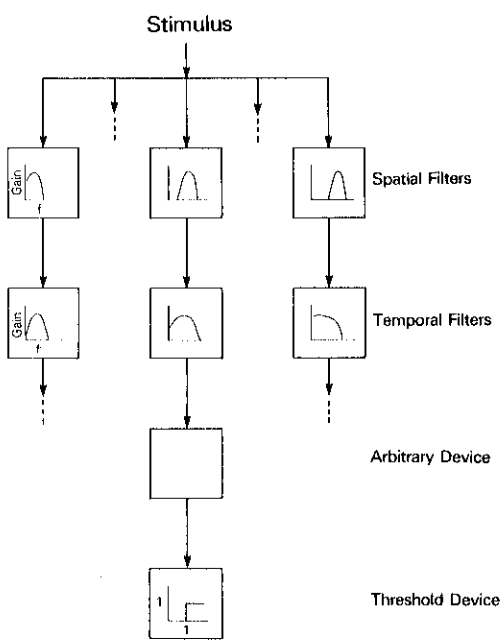


Fig. 5. Block diagram of a model of detection of spatio-temporal luminance distributions. See text for details.

each channel into a spatial and a temporal filter; arbitrarily placing the spatial filter first in the sequence. Each channel is then provided with an arbitrary device, which may, for example, impose a non-linear operation upon its input. The particulars of this element will be investigated below. We then terminate each channel with a threshold device.

This assumption of separability may well be false. If we suppose the response of a channel to be some combination of the responses of a variety of sub-elements (e.g. receptors, cells), then a change in the spatial input to a channel would likely change the population of sub-elements contributing to the response. Unless all sub-elements have a common temporal response, this is likely to alter the temporal form of the combined response. It is clear that for a filter in which spatial and temporal components are not separable, no unique temporal response exists that will completely characterize the temporal behavior of a channel.

Relation between spatial and temporal filters

In seeking to characterize the temporal behavior of certain visual cells, physiologists have adopted the terms "transient", indicating a system whose step response vanishes outside of some interval, and "sustained", indicating a response which does not so vanish. The behavior of a linear temporal filter may be entirely specified by its response to a unit impulse, its impulse response, $h(t)$ or by the Fourier transform of this function, the system function, $H(\omega)$. In these terms, transient behavior is defined by an impulse response whose integral is zero, or, equivalently, by an amplitude response, $|H(\omega)|$, of zero at zero frequency. Such conditions will be met for example, by an impulse response which is biphasic, and whose positive and negative lobes are of equal area. Sustained behavior is defined by an impulse response whose integral is non-zero, or by a non-zero amplitude response at zero frequency.

Since by these definitions a transient system will be quite rare, while sustained systems will be almost always encountered, it will be more descriptive to speak of a system as being *purely* transient, *purely* sustained, or primarily transient or sustained. By purely transient we again mean that case in which the integrated impulse response is zero. By purely sustained we indicate that case in which the absolute value of the integral of the impulse response is equal to the integral of the absolute value of the impulse response, in other words, in which the impulse response is all of one sign. We may note that the step response in a purely transient system approaches zero in the limit, while that of the purely sustained approaches a constant, which is its maximum. In this light, we may now regard a particular system as relatively transient or sustained, in the degree to which it approaches one or the other of these extremes.

Although a detailed relation between the summation function and the impulse response of the temporal filter must await a specification of the arbitrary device of Fig. 5, it can be shown that for the conditions to be considered in the sequel, a purely sustained system will yield an all positive summation function; that is, only facilitation will be observed between inputs of the same sign.

Thus of the summation functions obtained, only that of observer RP at a spatial frequency of 10.5 c/deg (Fig. 3d) indicates the presence of a purely sustained temporal filter. In contrast, the summation functions shown in Figs. 3a and b and 4a and b show that at lower spatial frequencies, certain temporal asynchronies give rise to inhibition between gratings of the same phase and facilitation between gratings of opposite phase. It is clear that as the spatial frequency is decreased, the form of the summation function departs farther and farther from that appropriate to the purely sustained case.

This result suggests a correlation between the spatial sensitivities of channels and their temporal responses. The channels predominantly sensitive to high spatial frequencies respond in a primarily sustained fashion, while those sensitive to lower spatial frequencies respond in a primarily transient fashion.

This conclusion is in agreement with those of a number of previous investigators. By modulating the contrast of a grating sinusoidally in time, Robson (1966) has demonstrated the presence of a decline in sensitivity at low temporal frequencies at 0.5 c/deg—consistent with transient behavior—and the lack of one at 4 c/deg and above—consistent with purely sustained behavior. Kelly (1971), too, has shown that the pronounced low frequency decline obtained with flickering uniform fields is effectively abolished when a flickering 3 c/deg square wave grating is used.

A similar distinction between the temporal responses to high and low spatial frequencies has been observed by Tolhurst (1975a, b). The detectability of a brief temporal increment introduced at various times during an extended sub-threshold grating presentation was found to be influenced only in the vicinity of the onset and offset of the sub-threshold grating when a low spatial frequency was used. When a higher spatial frequency was used, detectability was enhanced for the entire duration of the sub-threshold grating. Further, an examination of reaction times to gratings indicates a bimodal distribution to low spatial frequencies, and unimodal to high. This finding is compatible, respectively, with biphasic and monophasic responses of the underlying mechanisms.

It should be noted, however, that at each spatial frequency a different summation function was obtained. Thus while the response to a particular spatial frequency may be described as primarily transient or sustained, the change from one sort of temporal response to the other appears to be gradual, rather than abrupt. This, too, has its counterpart in the results of Robson (1966) in which the sensitivity to sinusoidally modulated gratings revealed a graded appearance of a low temporal frequency decline as the spatial frequency of the grating was decreased.

In terms of the model, this would suggest that each channel is distinguished in both spatial and temporal sensitivities. It must be allowed, however, that both sustained and transient channels may be involved in the detection of a single spatial frequency. Kulikowski and Tolhurst (1973) have demonstrated an apparent overlap in the spatial sensitivities of the two sorts of channels. If this were the case, then a single spatial frequency might be detected by either sort of channel, depending upon its temporal distribution. In particular, at an appropriate spatial frequency and onset asynchrony, we might suppose that a pair of gratings of equal contrast and the same phase would be detected by a sustained mechanism, a pair of equal contrast but opposite phase by a transient mechanism. If detected by different mechanisms, we might suppose they would be discriminable at, or near, threshold.

In order to investigate this possibility, an observer (RP) was presented with the two types of stimuli, intermixed and at several contrasts about their respective thresholds, in a two-interval, temporal forced-choice procedure. On each trial, the observer was asked to select the interval in which the stimulus had occurred, and to identify the stimulus. For both stimuli, preliminary results appear to show that the psychometric functions describing correct selection of interval and correct identification of stimulus are separated by a relatively constant 4–6 dB. In other words, these stimuli are identified only when at about twice the contrast at which they are detected. This did not appear to vary with spatial frequency or with onset asynchronies of less than 100 msec. For comparison, we note that Tolhurst and Dealy (1975), in an analogous procedure, found that for edges of opposite polarity the two psychometric functions are separated by less than 1 dB. Thus identification of stimuli at threshold apparently *is* possible under some circumstances. Evidently, in the present case, the same mechanism detects both same and opposite phased grating pairs.

Detection criteria

Many investigators have noted that temporal modulation of spatial stimuli may give rise to two separate visual appearances: one, a spatial inhomogeneity in the image, or *pattern*; and one a temporal inhomogeneity, or *flicker*. The effects appear to have different thresholds; as contrast is increased from zero, one or the other appearance may occur first, depending upon the spatial and temporal distribution of the stimulus. By instructing the observer to adhere to one criterion, it was possible to trace out the sensitivity of the mechanisms responsible for each effect as a function of spatial and temporal frequency (Kulikowski and Tolhurst, 1973). The general result may be stated as follows: flicker detection has a temporal sensitivity that declines at both low and high frequencies, and a spatial sensitivity that declines only at high frequencies; pattern detection has a temporal sensitivity that declines only at high frequencies, and a spatial sensitivity that declines at both low and high frequencies. The congruence of these sensitivities with those assumed for transient and sustained channels, has led to the proposal that the perception of flicker is exclusively provided by the operation of transient channels, and pattern by sustained channels. If this were so, the use of a pattern criterion should have provided a summation function characteristic of sustained behavior; that is, with little or no negative lobe. It is clear from Fig. 3b that it did not. Although thresholds were on average 2.9 dB higher for the pattern criterion than for the detection criterion, the summation index (κ) was the same at each asynchrony.

Several explanations are possible. It may be that the pattern criterion was not, in fact, employed by the observer. This would be peculiar in view of the observers' stated awareness of the two criteria. On the other hand, it may be that the observers' thresholds are "locked" to the more sensitive criterion. When pattern is asked for, the observer adjusts all of the stimuli of that session (for example, those of Fig. 2a) so that pattern is evident, *and* so that all provide *equal* flicker.

However, this explanation is far-fetched. An explanation of this peculiarity may have to wait upon the development of some more effective means of enforcing the use of one or the other criterion.

Eye movements

It is clear that eye movements occurring between the first and second elements of a grating pair may alter the phase relationship between their retinal images, particularly at higher spatial frequencies. We may ask whether such an alteration might be responsible for some of the results obtained here; in other words, might a single form of temporal response, in conjunction with the distribution of eye movements characteristic of binocular fixation, result in the different patterns of temporal interaction obtained with high and low spatial frequencies. Consider the results of observer RP at 3.5 and 7.0 c/deg and a separation of 60 msec (Figs. 3b and c). A monophasic temporal response might give the obtained results for the lower spatial frequency if the typical displacement of the retinal image between the two grating pulses was such as to shift the second by 180° in phase, i.e. by about 9′. This is well beyond the range of eye movements to be expected within a 60 msec interval (Riggs, Armington and Ratliff, 1954). Alternatively, a biphasic response might give the results obtained for the higher frequency if the typical shift was 4.3′. This also is too large to be accounted for by eye movements occurring during 60 msec of fixation.

It is, however, quite probable that eye movements do influence the measured summation function. As the spatial frequency is increased, the size of the average spatial phase shift between the elements of a grating pair will increase. If the effect of such phase shifts is to reduce the magnitude of κ, then we may underestimate the spatial frequency at which the temporal response changes from transient to sustained. A suggestion that this may be so is given by Kulikowski (1971) who showed that image stabilization may shift to a higher spatial frequency the transition between apparently transient and sustained responses to flickered gratings. This, however, is in contrast to other results which show little effect of stabilization upon such responses (Keesey, 1972).

The model of Rashbass

We have noted that the elliptical threshold contours obtained with onset asynchronies of less than 100 msec are compatible with the assumption that the output of a temporal filter is squared and integrated, as proposed by Rashbass (1970). In terms of the present model, as schematized in Fig. 5, the squaring and integration would enter as the third stage of each channel.

In this context, the progression of $\kappa(\tau)$ towards zero at asynchronies of about 100–200 msec indicates an impulse response of a similar duration, excluding any constant latency. The fit to an ellipse deteriorates at asynchronies greater than 100 msec. If we suppose that the declining quality of fit is due to a progressive violation of the condition imposed by equation (3), that is, that the responses to the two pulses no longer fall entirely within the epoch of integration, then an epoch of some 200–300 msec is indicated.

By Fourier transforming the summation function for his data, Rashbass was able to show the amplitude response, $|H(\omega)|$, of the assumed temporal filter. We may perform a similar operation, transforming for convenience the Gaussians that have been fitted to the individual summation functions, to obtain the amplitude response functions displayed in Fig. 6.

In theory these functions should allow one to predict the sensitivity to sinusoidally modulated gratings of the appropriate spatial frequency. Specifically,

$$\alpha^{-1} = |H(\omega)|\sqrt{\frac{T}{2} + \frac{1}{4\omega}\sin\omega T}, \qquad (6)$$

where α is the threshold amplitude of temporal modulation, ω the radian temporal frequency of modulation, T the epoch of integration, and $|H(\omega)|$ the amplitude response as indicated in Fig. 6.

In practice, preliminary results suggest that for at least one spatial frequency, equation (6) predicts a temporal contrast sensitivity function that is uniformly higher than the measured values. This can be remedied only by allowing T to assume a value that is less than one-half the estimate made earlier of 200–220 msec. The issue deserves further investigation.

Probability summation

Consider the manner in which the effects of two gratings interact when separated by a long interval of time, say 180 msec, as in Fig. 2d. It is clear that at separations of this order, the contrast required in the first grating is independent of the contrast of the second grating, and vice versa. This is true, however, only where the contrast of one or the other pulse predominates. When their contrasts are equal, the

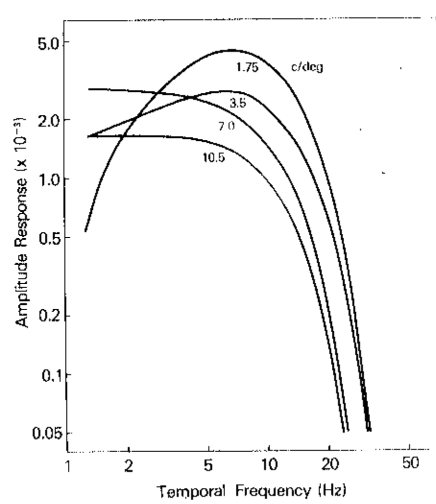


Fig. 6. Theoretical amplitude response functions, $|H(\omega)|$, derived from the data of Fig. 3.

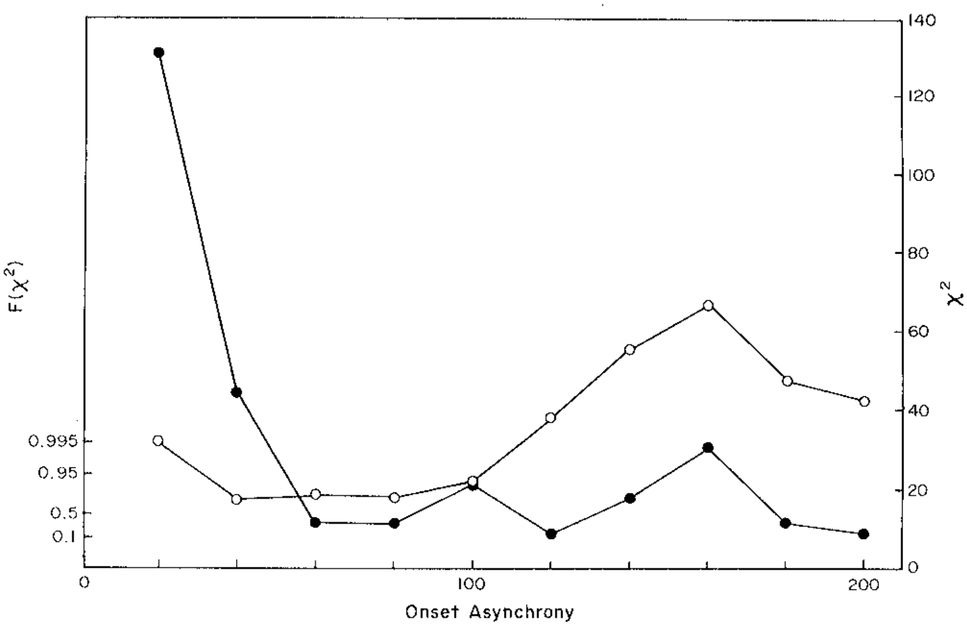


Fig. 7. Quality of fit of Rashbass' model (open circles) and of the probability summation model assuming no overlap of responses (filled circles) as a function of onset asynchrony, for observer OZ at a spatial frequency of 7.0 c/deg. Points give the value of the χ^2 statistic which describes the fit of each model. The right ordinate gives the value of the statistic, the left give the probability that the value on the right ordinate will not be exceeded.

pair is distinctly more detectable than either grating alone. This results in the "rounding" of the corners of the square, which can be attributed to probability summation. The response to each pulse, as perturbed by noise, has some probability of exceeding threshold. The probability of either or both responses exceeding threshold is then greater than the probability for a single response. The character of this effect will be the same whether or not squaring and integration precede the threshold device.

The effects of probability summation need not be limited to equal responses at large separations. We may ask whether the patterns of interaction observed between pulses of brief separation may be similarly accounted for, without recourse to the model proposed by Rashbass. The logic to be applied here is similar to that used by Tolhurst (1975a).

Suppose that the temporal waveform of the stimulus passes through a linear filter whose output is perturbed by noise which is independent from moment to moment. The stimulus is detected if the magnitude of the noise-perturbed output waveform exceeds some fixed threshold criterion level at least once during the observation interval. Therefore, the probability of detecting the stimulus is the joint probability that threshold has been exceeded at least once during this interval. In other words, the probability that the stimulus is detected (P) is given by

$$P = 1 - \Pi(1 - p(t_i)) \tag{7}$$

where $p(t_i)$ is the probability that threshold is exceeded at time t_i. Let us assume that an equation recently proposed by Quick (1974) as descriptive of psychometric functions obtained in a wide variety of situations also describes the dependence of $p(t_i)$ upon the output of the temporal filter at time t_i, $[f(t_i)]$. Then,

$$p(t_i) = 1 - 2^{-|f(t_i)|^\beta} \tag{8}$$

where β is a constant.[2] Combining equations (7) and (8) we obtain

$$P = 1 - 2^{-\Sigma|f(t_i)|^\beta}. \tag{9}$$

If the observer adjusts the contrast until the stimulus is seen as often as not, $P = 0.5$, and we have

$$1 = \Sigma|f(t_i)|^\beta, \tag{10}$$

or passing to the integral,

$$1 = \int |f(t)|^\beta \, dt. \tag{11}$$

We shall consider the behavior of this model in three cases: where the input is a single pulse; where the input is two well-separated pulses; and where the separation between two input pulses is brief. In the first case, making use of equation (11):

$$1 = \int_{-\infty}^{\infty} |\lambda\delta h(t)|^\beta \, dt \tag{12}$$

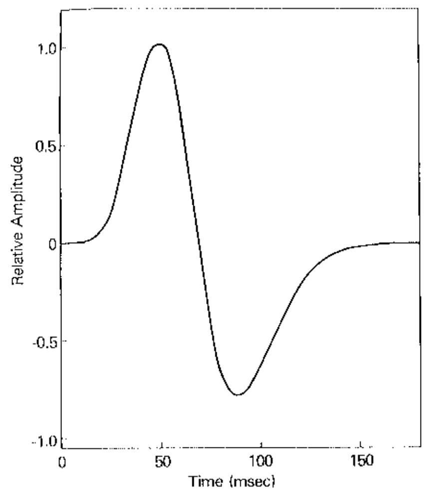
or

$$[\lambda\delta]^{-\beta} = \int_{-\infty}^{\infty} |h(t)|^\beta \, dt. \tag{13}$$

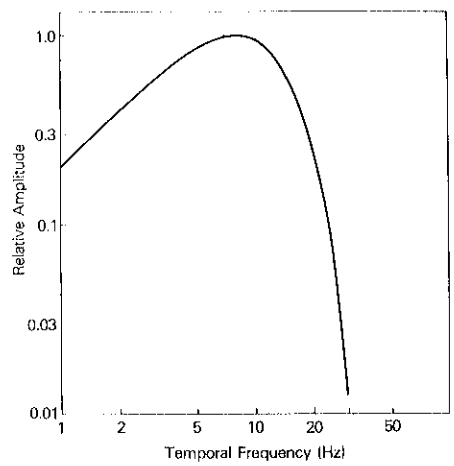
In the second case, by expression (8), we have

$$1 = \int_{-\infty}^{\infty} |a\delta h(t) + b\delta h(t - \tau)|^\beta \, dt. \tag{14}$$

[2] The 2 in this expression may be replaced with any positive number, without altering the substance of this argument.



(a)



(b)

Fig. 8(a). Hypothetical impulse response used to generate threshold contours with a probability summation model. (b) Amplitude response of a filter whose impulse response is given in (a). This is the modulus of the Fourier transform of the curve in (a).

Where the two input pulses are sufficiently well separated (such that $h(t)h(t-\tau) = 0$, for all t), we may write equation (14) as

$$1 = [|a|\delta]^{\beta} \int_{-\infty}^{\infty} |h(t)|^{\beta}\, dt + [|b|\delta]^{\beta} \int_{-\infty}^{\infty} |h(t-\tau)|^{\beta}\, dt. \quad (15)$$

Making use of equation (13) we arrive at the following expression for the threshold contrasts a and b of a pair of well-separated pulses, where λ is the contrast of a single pulse

$$1 = \lambda^{-\beta}[|a|^{\beta} + |b|^{\beta}]. \quad (16)$$

The curve fitting procedure described in the results section is easily adapted to provide fits of this expression to the data from individual sessions. The continuous curve in Fig. 2d is an example, the estimate of β in this case being about 4.5. The exponent β may also be estimated by fitting equation (8) to data which empirically describe the rise in the probability of detection with increases in signal strength, that is, to psychometric functions. Using this procedure on data obtained with comparable stimuli and experimental conditions, we find estimates of β that are typically between 3 and 6. A comparison of the fits yielded by Rashbass's model and the probability summation model (assuming no overlap of responses) is provided for one subject in Fig. 7. Within the regions where each is appropriate, the fits are of about equal quality. The data from the two other subjects support the same conclusion. Note that the poor fit of the probability summation model at brief asynchronies does not argue against this model, but only suggests that the responses to the two pulses overlap.

The third case, in which the responses to the two pulses overlap, is again described in equation (14). In this case, however, predictions cannot be made without knowledge of the form of the impulse response. We may nonetheless gain some idea of the model's behavior by applying to equation (14) an arbitrary but plausible version of $h(t)$. An example of such a function, and its amplitude response $|H(\omega)|$, are given in Fig. 8. The analytic expression for this function is

$$\begin{aligned} h(t) &= 0, & t \leq 0 \\ &= t^{n-1} e^{-t/\Omega}, & 0 < t \leq \sigma \\ &= t^{n-1} e^{-t/\Omega} - (t-\sigma)^{n-1} e^{-(t-\sigma)/\Omega}, & \sigma < t \end{aligned} \quad (17)$$

This is simply the impulse response of an n-stage low-pass filter added to a delayed and inverted replica of itself. In the present case, n was set to 10, and the remaining parameters were chosen in order to give $|H(\omega)|$ some resemblance to empirical temporal modulation sensitivity.

Examples of threshold contours computed using the impulse response of equation (17), and the probability summation model described by equation (14) with an exponent (β) of 4 are shown in Figs. 9a–d.

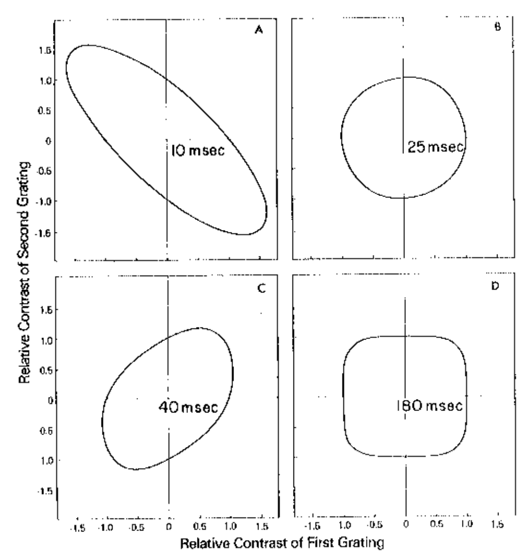


Fig. 9. Threshold contours predicted by the probability summation model and the impulse response of Fig. 8a. Note the similarity to the empirical threshold contours of Figs. 2 a–d.

Note first the close resemblance between these curves and the configurations of thresholds actually obtained at various onset asynchronies, as shown in Figs. 2a–d. Second, note that these computed contours are quite similar to ellipses, at least at small asynchronies. Additional computations confirm that for brief asynchronies, the form of the threshold contour is relatively insensitive to the value of the exponent. Examination of equations (2) and (14) will show that Rashbass's model and the probability summation model differ analytically only in the value of their exponents (provided that equation (3) is satisfied). Thus the contours that are generated with an exponent of 2, which are ellipses as prescribed by the model of Rashbass, are almost indistinguishable from those generated with an exponent of 4, as prescribed by the probability summation model and as shown in Figs. 9a–c.[3]

It should be noted that we have not attempted to actually fit the curves of Fig. 9 to the data. A comparison of the asynchronies at which the curves of Figs. 2 and 9 were generated, will show that such a fit would in fact be poor. Our intent here is only to show the general form of the threshold contours that result from a plausible impulse response in combination with the probability summation model. What must yet be demonstrated is that there exists for each spatial frequency an impulse response which will, in conjunction with equation (14), yield threshold contours that are reasonable fits to the data. This is presently being attempted.

The preceding argument supposes that the determination of threshold by method of adjustment is probabilistic. In the present experiments the stimulus was presented briefly at regular intervals, and it is not inconceivable that the observer adjusted the contrast until the stimulus was seen on some fixed proportion of trials. Such a strategy would indeed reveal the effects of probability summation. Nonetheless, other strategies of adjustment are imaginable, and it is thus important to note that the basic phenomena of these experiments have been replicated with a staircase procedure.

We may summarize these points in the following way. The model proposed by Rashbass accounts satisfactorily for the form of the empirical threshold contour at brief ($\tau < 100$ msec) separations. It fails to account for the threshold contour at long separations, and may yield contrary estimates for the length of its integration interval when both pulsed and sinusoidal stimuli are considered. The probability summation model accounts for the form of threshold contours at long separations, and may account for their form at brief separations. Finally, the probability summation model provides an estimate of the exponent β that is in keeping with that obtained in other experiments.

[3] Elliptical threshold contours are also predicted by a model proposed by Cohn and Lasley (1975). Indeed, it appears that their model predicts ellipses even at long separations, rather than the rounded square found in the present results. However, their model may be modified so as to yield the appropriate result at long separations.

Acknowledgement This work was supported by NSF grant No. BMS75-07658 to the second author.

REFERENCES

Chandler J. P. (1965) STEPIT. *Quantum Chemistry Program Exchange*. Department of Chemistry, Indiana University, Bloomington, IN, U.S.A.

Cohn T. H. and Lasley D. J. (1975) Spatial summation of foveal increments and decrements. *Vision Res.* **15**, 389–400.

Graham N. and Nachmias J. (1971) Detection of grating patterns containing two spatial frequencies: a comparison of single-channel and multiple-channel models. *Vision Res.* **11**, 251–259.

Keesey U. T. (1972) Flicker and pattern detection: a comparison of thresholds. *J. opt. Soc. Am.* **62**, 446–448.

Kelly D. H. (1959) Effects of sharp edges in a flickering field. *J. opt. Soc. Am.* **49**, 730–732.

Kelly D. H. (1971) Theory of flicker and transient responses—II. Counter-phase gratings. *J. opt. Soc. Am.* **61**, 632–640.

Kulikowski J. J. (1971) Effect of eye movements on the contrast sensitivity of spatio-temporal patterns. *Vision Res.* **11**, 261–273.

Kulikowski J. J. and Tolhurst D. J. (1973) Psychophysical evidence for sustained and transient channels in human vision. *J. Physiol., Lond.* **232**, 149–163.

Nachmias J. (1967) Effect of exposure duration on visual contrast sensitivity with square-wave gratings. *J. opt. Soc. Am.* **57**, 421–427.

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

Rashbass C. (1970) The visibility of transient changes of luminance. *J. Physiol., Lond.* **210**, 165–186.

Riggs L. A., Armington J. C. and Ratliff F. (1954) Motions of the retinal image during fixation. *J. opt. Soc. Am.* **44**, 315–321.

Robson J. G. (1966) Spatial and temporal contrast-sensitivity functions of the visual system. *J. opt. Soc. Am.* **56**, 1141–1142.

Schober H. A. W. and Hilz R. (1965) Contrast sensitivity of the human eye for square-wave gratings. *J. opt. Soc. Am.* **55**, 1086–1091.

Thomas J. P. (1970) Model of the function of receptive fields in human vision. *Psychol. Rev.* **77**, 121–134.

Tolhurst D. J. (1975a) Reaction times in the detection of gratings by human observers: A probabilistic mechanism. *Vision Res.* **15**, 1143–1149.

Tolhurst D. J. (1975b) Sustained and transient channels in human vision. *Vision Res.* **15**, 1151–1155.

Tolhurst D. J. and Dealy R. S. (1975) The detection and identification of lines and edges. *Vision Res.* **15**, 1367–1372.

Note added in proof Since this article went to press a paper has appeared in which several of the ideas presented here are also expressed (Rashbass, C. (1976) Unification of two contrasting models of the visual increment threshold. *Vision Res.* **16**, pp. 1281–1283).