

# Elaborated Reichardt detectors

Jan P. H. van Santen and George Sperling

Department of Psychology, New York University, 6 Washington Place, New York, New York 10003

Received July 9, 1984; accepted October 12, 1984

The elaborated Reichardt detector (ERD) proposed by van Santen and Sperling [J. Opt. Soc. Am. A 1, 451 (1984)], based on Reichardt's motion detector [Z. Naturforsch. Teil B 12, 447 (1957)], is an opponent system of two mirror-image subunits. Each subunit receives inputs from two spatiotemporal filters (receptive fields), multiplies the filter outputs, and temporally integrates the product. Subunit outputs are algebraically subtracted to yield ERD output. ERD's can correctly indicate direction of motion of drifting sine waves of any spatial and temporal frequency. Here we prove that with a careful choice of either temporal or spatial filters, the subunits can themselves become quite similar or equivalent to the whole ERD; with suitably chosen filters, the ERD is equivalent to an elaborated version of a motion detector proposed by Watson and Ahumada [NASA Tech. Memo. 84352 (1983)]; and for every choice of filters, the ERD is fully equivalent to the detector proposed by Adelson and Bergen [J. Opt. Soc. Am. A 2, 284-299 (1985)]. Some equivalences between the motion detection (in  $x, t$ ) by ERD's and spatial pattern detection (in  $x, y$ ) are demonstrated. The responses of the ERD and its variants to drifting sinusoidal gratings, to other sinusoidally modulated stimuli (on-off gratings, counterphase flicker), and to combinations of sinusoids are derived and compared with data. ERD responses to two-frame motion displays are derived, and several new experimental predictions are tested experimentally. It is demonstrated that a system containing ERD's of various sizes can solve the correspondence problem in two-frame motion of random-bar stimuli and shows the predicted phase dependencies when confronted with displays composed of triple sinusoids combined either in amplitude modulation phase or in quasi-frequency modulation phase. Finally, it is shown that, while the ERD may in some instances give larger responses to nonrigid than to rigid displacements, the subunits (and hence the ERD) are especially well behaved with continuous movement of rigid or smoothly deforming objects.

## 1. INTRODUCTION

Recent theoretical work on motion detectors in human vision has shown a remarkable convergence. Current theories assume that a motion detector involves the comparison of the outputs of two spatiotemporal filters.<sup>1-7</sup> This general idea was originally formulated by Reichardt<sup>7</sup> in the context of experiments on insects. In Reichardt's original version, the detector had no spatial filters: it tapped the visual field at two *points*. Reichardt's detector formed the basis for a model proposed by van Santen and Sperling,<sup>4</sup> the *elaborated Reichardt model*, that included spatial input filters and also a linking hypothesis relating the combined response of many detectors to a measure of human performance. (In our terminology, a *detector* refers to the elementary motion-detecting unit; a *model* refers to the whole system that combines outputs of many detectors into a prediction of performance.) The elaborated Reichardt *model* was shown to explain several old and new phenomena in human motion perception. Its basic component was an elaborated Reichardt *detector* (original Reichardt detector plus spatial receptive fields), and it had some enticing properties. For a wide choice of spatial and temporal filters, the elaborated Reichardt detector (ERD) was shown to be well behaved from a Fourier-analytic point of view. That is, the detector indicated the correct direction of motion for a drifting sinusoidal grating of any spatial or temporal frequency.

In this paper, we present further theoretical and empirical results on several versions of the ERD—including ERD subunits, special cases, and generalizations. Section 2 outlines Reichardt's original detector and the elaborated Reichardt detectors. Section 3 derives the response of various ERD's to any display. We analyze the relationship between ERD's

and the detectors recently proposed by Watson and Ahumada<sup>5</sup> and Adelson and Bergen.<sup>6</sup> Coming from a Fourier-analytic tradition, their detectors are superficially quite different from the ERD although they share with it the general architecture (two spatiotemporal input filters and a comparison operation). We prove that these apparently different detectors actually are, or can be elaborated to become, equivalent to the ERD's. Section 4 outlines some equivalences and differences between one-dimensional motion detection (in  $x, t$  space-time) and two-dimensional pattern detection (in  $x, y$  space). This section also discusses receptive fields, windowing, frequency analysis, and related theoretical issues. Section 5 applies the general results derived in Section 3 to particular, commonly used displays; temporally modulated, drifting sinusoidal gratings, combinations of gratings, two-frame displays, and displays of rigid-object translation. Various empirical phenomena that can and cannot be explained by the ERD are considered. We reject claims<sup>8-10</sup> that the ERD can explain certain phenomena involving spatial resolution in *nonmoving*, sinusoidally modulated stimuli; we report predictions and results with four paradigms involving two-flash stimuli.

## 2. ELABORATED REICHARDT DETECTOR

### A. Original Reichardt Detector

*Subunits.* We begin the discussion of the ERD by explaining the detector as originally formulated by Reichardt.<sup>7</sup> As in previous work,<sup>1-4</sup> to simplify the exposition, we consider a version of the ERD that has been stripped of several temporal filters. In Subsection. 2.D we consider these additional filters.

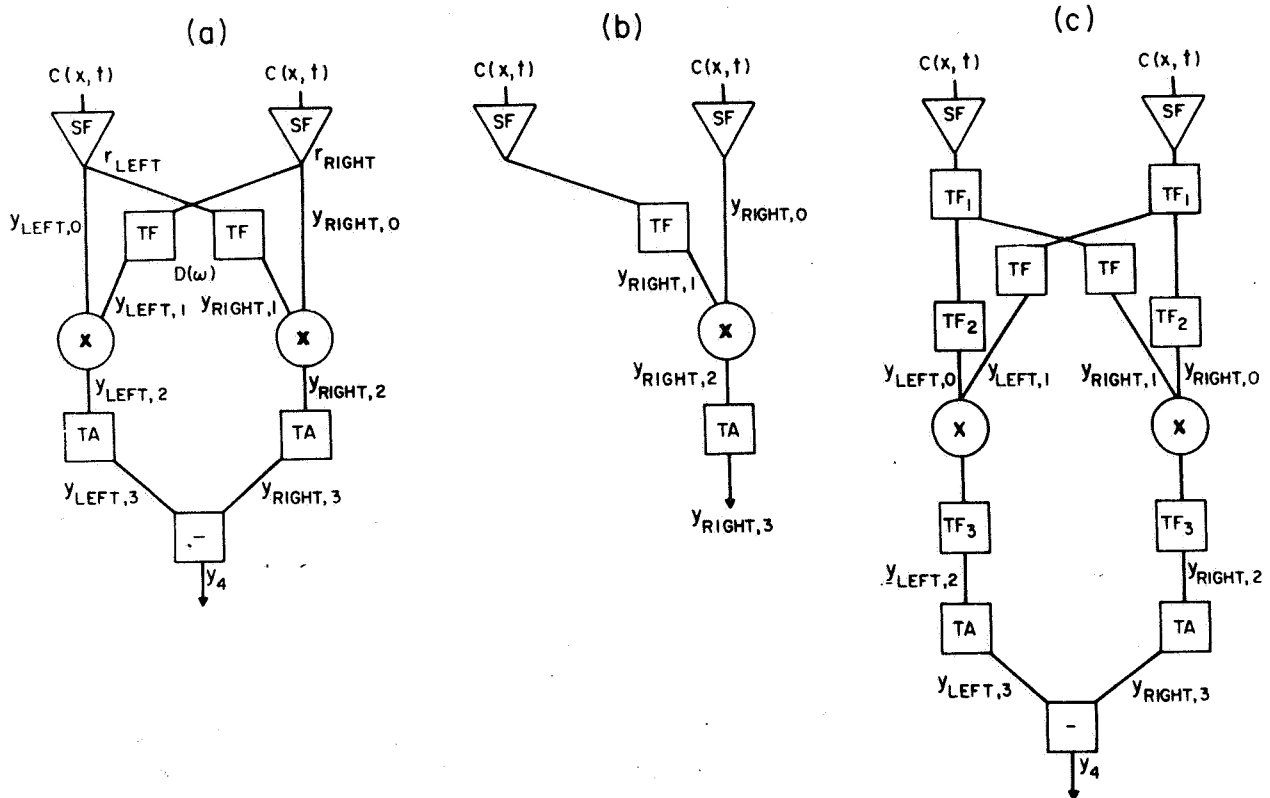


Fig. 1. (a) The ERD. The input is a luminance pattern with contrast  $c(x, t)$ ; it is sampled by linear spatial filters (receptive fields, SF's) with spatial responses  $r_{left}$  and  $r_{right}$  centered at locations  $x_{left}$  and  $x_{right}$ ;  $y_{i,H}$  ( $H = left, right$ ) represents the signal at the various stages  $i$  for the left and right subunits. TF indicates a linear, time-invariant filter with Fourier transform  $D(\omega)$ ,  $\times$  indicates a multiplication unit, TA indicates a temporal integration operation, and  $-$  indicates a unit that subtracts its left from its right input. (b) The right subunit of an ERD. (c) The ERD of (a) with three pairs of additional temporal filters ( $TF_1, TF_2, TF_3$ ) that leave its operation essentially unchanged.

Figure 1(a) depicts the ERD; the original Reichardt detector was similar except that it lacked the spatial filters marked SF and that the temporal integration filter TA was restricted to infinite time averaging. Reichardt assumed that motion detectors are composed of two component subunits tuned to motion in opposite directions (left and right). Subunit outputs cannot be directly accessed by subsequent processes, such as decision and response processes; rather, the response of the left subunit is subtracted from that of the right subunit to yield the detector's response. It follows that, if the output of the right subunit exceeds the output of the left subunit, detector response is positive, indicating rightward motion; likewise, if the output of the left subunit exceeds the output of the right subunit, detector response is negative, indicating leftward motion.

The subunits that make up a detector share two input channels that sample the visual field at two points in space. The input to each subunit thus consists of the temporal luminance patterns at its input points. That is, any dynamic display can be thought of as a spatiotemporal luminance pattern, i.e., a pattern in which luminance varies as a function of spatial location  $x$  and of time  $t$ . (We restrict the discussion to luminance patterns that vary in only one spatial dimension,  $x$ .) At a fixed point in space, a spatiotemporal luminance pattern produces a temporal luminance pattern. This temporal luminance pattern is a sine wave for a drifting sinusoidal grating; for stroboscopic motion, it is a *comb* function.

Subunits of a Reichardt detector operate on the basis of the delay-and-compare principle. That is, subunits detect motion

by delaying the temporal luminance pattern in one input channel and comparing this delayed pattern with the non-delayed pattern in the other channel. Subunit output reflects how close this match is.

In Reichardt detectors, the delay operation is implemented by an arbitrary, linear, time-invariant, temporal filter—the delay filter. Such a filter does not merely delay its input; it may alter both the phase and the amplitude of an input sine wave. However, it does not change the basic sinusoidal wave shape. (In general, a linear, time-invariant filter may alter the shape of any other, nonsinusoidal wave form.) The modulation transfer function of the delay filter is  $D(\omega)$ ;  $Re$  and  $Im$  denote its real and imaginary parts, respectively. The filter's amplification factor,  $|D(\omega)|$ , and its phase delay,  $\tau = \tan^{-1} \{-Im[D(\omega)]/Re[D(\omega)]\}$ , both depend on the temporal frequency  $\omega$  of the input sine wave.

The comparison operation is implemented by multiplication followed by infinite time averaging. That is, the detector computes the correlation between the undelayed temporal modulation pattern in one input channel and the delayed temporal modulation pattern from the other.

*Physiological basis of multiplication.* The question of whether multiplication is a plausible neural operation has aroused much speculation, because of the variety of ways in which multiplication can be realized. Let the direct and delayed inputs to the multiplier component of the Reichardt detector be designated as  $x$  and  $y$ , respectively. While the obvious physiological explanation of multiplication would be a form of multiplicative interneuronal excitation, the Rei-

Reichardt detector has an equivalent form in which  $x$  is multiplied by  $1 - y$  instead of by  $y$ . This equivalent form derives from shunting inhibition<sup>11,12</sup> in which  $x$  is divided by  $1 + y$ . Multiplication is a reasonable approximation to various forms of inhibition because the first term in a Taylor series expansion of a variety of nonlinear combination rules contains the product  $xy$ .<sup>13,14</sup> It follows that multiplication is quite compatible with current physiological knowledge.

*Responses of the original Reichardt detector.* For displays with sinusoidal temporal modulation (such as drifting gratings and counterphase gratings), the original Reichardt detector's response is directly proportional to  $\sin(\varphi)$ , where  $\varphi$  is the temporal phase difference between temporal modulation patterns in the two input channels. Obviously, the response is maximal when  $\varphi = \pi/2$ .

*Spatial and temporal aliasing.* For any particular sine-wave grating, the temporal phase difference between inputs to a particular detector will depend on the distance between its input channels and on the spatial frequency of the grating. Because the response depends on  $\sin(\varphi)$ , an original Reichardt detector has a spatial aliasing problem; for certain spatial frequencies of a drifting grating, the response always signals the wrong direction. Mathematically, note that  $\varphi = 2\pi df\Delta x$ , where  $\Delta x$  denotes the distance between input channels,  $d$  denotes the direction in which the grating drifts (+1 for rightward motion, -1 for leftward motion), and  $f$  denotes the spatial frequency of the grating. When the distance between input channels is between one-half and one spatial period, the sign of  $\sin(\varphi)$  is opposite that of  $d$ , thus signaling the incorrect direction of motion.

Some choices of the temporal filter TF lead to incorrect direction responses. In Subsection 5.A we consider what properties the spatial receptive fields and the temporal filters must have to prevent spatial and/or temporal aliasing.

## B. Elaborated Reichardt Detector

*Absence of spatial and temporal aliasing.* The simplifying assumption that the input channels to the Reichardt detector have point-shaped spatial receptive fields may have been appropriate for the insect facet eyes for which the original Reichardt detector was developed, but it is clearly inappropriate for human vision. Even for insects, the point input assumption was abandoned by Fermi and Reichardt.<sup>15</sup> First, in the human visual system, to the extent that the detector is meant to reflect the activity of neurons, the assumption of point receptive fields is untenable. Second, human observers do not have the aliasing problem of the original Reichardt detector, discussed above. The aliasing problem is eliminated in the elaborated Reichardt detector by substituting spatial-frequency-selective receptive fields, marked SF in Fig. 1(a), for the point inputs. More generally, van Santen and Sperling<sup>4</sup> showed how to choose spatial and temporal filters (SF and TF) so that the sign of detector output was correct for drifting sinusoids of any spatial and temporal frequency and included this nonaliasing feature as an assumption of their ERD.

*Generalization of infinite time averaging; temporal linking assumption.* In the present paper, we use as our focal model a more general version of the ERD than previously<sup>4</sup> used, namely, a version in which the infinite time-averaging component is replaced with an optional, arbitrary linear temporal filter. As we shall see below, to obtain simple mathematical

results in certain applications, some form of time averaging (but not infinite time averaging) has to be reintroduced as an auxiliary assumption, but there are many applications in which even this weaker form of averaging is not needed. This generalization brings the ERD into closer correspondence with the visual system, and yet general properties can be proved that obtain for wide classes of stimuli and ERD's.

Temporal integration is essential to provide the temporal linking assumption. That is, to generate predictions for direction-discrimination experiments, we need to specify how a time-varying output of the comparison operation is mapped into a single real number—strength—which can be translated into the probability of a correct direction-discrimination response. More specifically, we require that positive values of strength map into probabilities of seeing rightward motion that exceed 0.5, while negative values map to probabilities of seeing leftward motion that exceed 0.5.

The question of how much integration occurs at the level of the detector and how much at higher levels is a difficult one to decide; for specificity, we have chosen to let integration—when it is needed to generate a prediction—always take place inside the detector. In this way the voting rule, i.e., the rule by which outputs from multiple detectors are combined, can be kept quite general.

*Voting rule.* The critical distinction between the elaborated Reichardt model and the original or elaborated Reichardt detector is the presence of a voting rule in the model. To be useful, psychophysical predictions must be robust with respect to voting rules (maximum rule, strength addition, etc).<sup>16</sup>

## C. Subunits of the Elaborated Reichardt Detector

The ERD, like the original Reichardt detector, assumes that only the final response (the difference between the subunit responses) is used. There is a good reason for ignoring responses at the subunit level, since, in general, neither the sign nor the magnitude of these responses is meaningful. However, when we make specific assumptions about either the temporal filter TF or about the receptive fields SF, subunit responses may become not only better behaved but in fact equivalent to responses at the detector level, either for a restricted class of displays or for all displays. Figure 1(b) depicts a subunit of an ERD. Two special subunit detectors are considered.

*Variant 1. Subunits with constant  $\pi/2$  temporal phase delay.* Subunits whose temporal delay filter delays all temporal frequencies by a quarter of a temporal cycle are called  $\pi/2$  temporal phase-delay subunits. A filter with this  $\pi/2$  phase-delay property has a modulation transfer function  $D(\omega)$  of the form  $-i|D(\omega)|$ . Although this filter is not realizable, Watson and Ahumada,<sup>5</sup> who propose it, point out that there are realizable filters that approximately have the  $\pi/2$  phase-delay property over a large range of temporal frequencies.

*Variant 2. Subunits with constant  $\pi/2$  receptive-field phase difference and equal receptive-field spectral powers.* Another important special case consists of subunits whose input receptive fields have the following property: For every spatial frequency, the spatial positions of a sinusoidal grating that maximize the responses of the receptive fields differ by one quarter of a spatial cycle. These subunits will be called  $\pi/2$  spatial phase shift subunits. Formally, let the receptive fields be denoted  $r_H$ , where  $H = \text{left, right}$ . Let  $f$  denote

spatial frequency. Then the Fourier transforms,  $R_H$ , of these receptive fields have the property that  $R_{\text{left}}(f) = -i \sin(f) \times R_{\text{right}}(f)$ . A special case consist of odd and even Gabor functions.<sup>17,18</sup> Simple cells in monkey area 17 apparently have the property of a  $\pi/2$  spatial phase difference.<sup>19</sup>

**D. Additional Temporal Filters**

The basic behavior of the ERD does not change when more temporal filters are added. Figure 1(c) shows an ERD that contains three additional pairs of filters  $TF_1$ ,  $TF_2$ , and  $TF_3$  in the input channels, in the direct paths between the input channels and the multipliers, and after the multipliers. It is easy to prove that any such multifilter ERD has the same output as an ERD with only one pair of additional filters directly in the input channels ( $TF_1$ ). Thus filters inside the detector do not create more generality and can be ignored. For mathematical convenience, except where issues of model equivalence are discussed, we shall also ignore the additional temporal filters in the input channels. The model depicted in Fig. 1(a) is the basis for most of our mathematical derivations.

**E. Auxiliary Assumptions**

These assumptions are not an integral part of the model and are needed only for some of the derivations (see Table 1).

(1) The temporal filter TA (Fig. 1) takes the special form of temporal integration over an appropriately chosen finite

**Table 1. Restrictive Auxiliary Assumptions Used in Derivations of Principal Results<sup>a</sup>**

Result	Auxiliary Assumption			
	TI	B	S	R
Multiplier output [Eq. (4)]				
Subunit output; ERD output (temporal-frequency segregation) [Eqs. (7), (8)]	X			
Equivalence of ERD and $n/2$ temporal-phase subunit [Eqs. (9), (10)]	X			
Equivalence of ERD and $n/2$ spatial-phase subunit [Eqs. (11), (12), (36)–(39)]	X			X
Equivalence of ERD and elaborated W-A detector	X			
Equivalence of ERD and Adelson-Bergen formulation [Eqs. (13)–(18)]				
Response to drifting gratings [Eqs. (19)–(24)]	X			
Response to counterphase and on-off gratings [Eqs. (25)–(27)]	X			
General two-frame response [Eqs. (29)–(33)]				X
Two-frame spatial responses [Figs. 9–13]		X	X	
Two-frame sine-wave response [Eqs. (34)–(35)]		X	X	
Two-frame sine-wave response [Appendix A]				X

<sup>a</sup> TI, temporal integration; B, balanced receptive fields; S, separability; R, rigid continuous motion displays. X indicates that the assumption was used.

interval. This window includes display onset and extends after display termination until multiplier output has decayed to insignificance. For periodic displays, integration is assumed to extend over precisely  $n$  ( $n = 1, 2, \dots$ ) periods. For periodic displays, this assumption seems more reasonable than the infinite time averaging in Reichardt's original model and in the earlier version of the ERD.<sup>1-4</sup> However, for periodic displays, averaging over one time period and infinite time averaging are precisely equivalent.

(2) Receptive fields are balanced, that is, their spatial integrals are zero because their on and off areas exactly cancel each other. Examples include differences of (appropriately scaled) Gaussians and all antisymmetric receptive fields.

(3) Spatial and temporal filters in the input channels are assumed to satisfy the separability hypothesis.<sup>20,21</sup> This hypothesis states that the response of an input channel can be obtained by calculating the spatial and temporal responses of a channel separately and then multiplying them, i.e., by correlating the (spatiotemporally modulated) input pattern with the receptive field of a channel and then convolving the (temporally modulated) receptive field output with the impulse response function of the channel's temporal filter.

**3. RESPONSE CHARACTERISTICS OF THE ELABORATED REICHARDT DETECTOR AND ITS VARIANTS**

**A. Elaborated Reichardt Detector**

Because ERD's are composed mainly of linear filters, their responses can best be derived by a Fourier approach.

Let  $c(x, t)$  denote the display (the input to the Reichardt detector). We assume that  $c(x, t)$  has a Fourier transform. Mathematically, this means primarily that the integral  $\iint |c(x, t)| dx dt$  converges or, in concrete terms, that the display is on only during a finite time interval and has a finite spatial extent. In practice, of course, this must always be so. We define the Fourier transform  $C(f, \omega)$  as

$$C(f, \omega) = \iint c(x, t) \exp[-2\pi i(fx + \omega t)] dx dt.$$

(Lower-case letters are used for spatiotemporal, real-valued functions and capital letters for their Fourier transforms.) The space-time function  $c(x, t)$  can be recovered from  $C(f, \omega)$  using the standard formula for the inverse Fourier transform:

$$c(x, t) = \iint C(f, \omega) \exp[2\pi i(fx + \omega t)] d\omega df.$$

For simplicity, we proceed under the assumption that any additional temporal filters [ $TF_1$ ; see Fig. 1(c)] are separable from the input spatial filters (SF's). However, we state here without proof that, with the exception of the two-frame results reported in Section 5, separability is not needed for any other results. Thus let receptive fields of an ERD be denoted  $r_H$  and their Fourier transforms by  $R_H$ , where  $H = \text{left, right}$ . Then the output  $y_{H,0}(t)$  of receptive field  $H$  is given by

$$\begin{aligned} y_{H,0}(t) &= \int r_H(x) c(x, t) dx \\ &= \iint C(f, \omega) \exp(2\pi i \omega t) R_H^*(f) d\omega df, \end{aligned} \quad (1)$$

and the Fourier transform of  $y_{H,0}$  is

$$H_{H,0}(\omega) = \int C(f, \omega) R^*(f) df. \quad (2)$$

Here, superscript \* indicates the complex conjugate. Let the delay filter have Fourier transform  $D(\omega)$ . Then the output from the delay filter is given by

$$y_{H,1}(t) = \int D(\omega) Y_{H,0}(\omega) \exp(2\pi i \omega t) d\omega. \quad (3)$$

After the multiplication operation, we have

$$y_{H,2}(t) = \left[ \int Y_{H,0}(\omega) \exp(2\pi i \omega t) d\omega \right] \times \left[ \int D(\nu) Y_{H',0}(\nu) \exp(2\pi i \nu t) d\nu \right]. \quad (4)$$

We now proceed using the auxiliary assumption of temporal integration of  $y_{H,2}(t)$ . Suppose that two arbitrary functions,  $g_1(\tau)$  and  $g_2(\tau)$ , are zero for all  $\tau < T_1$  and  $\tau > T_2$ . Thus the window of integration is  $[T_1, T_2]$ . Then the time integral  $TI$  of the product at time  $t$  is given by

$$\begin{aligned} TI_t(g_1, g_2) &= \int_{T_1}^t g_1(\tau) g_2(\tau) d\tau \quad (t \geq T_1) \\ &= \int_{-\infty}^{\infty} g_1(\tau) g_2(\tau) d\tau \quad (t \geq T_2). \end{aligned} \quad (5)$$

Using Parseval's identities, it immediately follows that

$$TI_t(g_1, g_2) = \int G_1(\omega) G_2^*(\omega) d\omega. \quad (6)$$

Applying these results to Eq. (4), it follows that subunit output is simply given by

$$y_{H,3} = \int Y_{H,0}^*(\omega) Y_{H',0}(\omega) D(\omega) d\omega, \quad (7)$$

where the  $Y$ 's denote the Fourier transforms. Finally, detector output is given by

$$y_4 = \int [Y_{right,0}^*(\omega) Y_{l,0}(\omega) - Y_{left,0}^*(\omega) Y_{r,0}(\omega)] D(\omega) d\omega. \quad (8a)$$

Because outputs are considered only after the observation interval is over and the internal response  $y_2$  has been temporally integrated, both subunit outputs [Eq. (7)] and detector output [Eq. (8a)] are time independent (see Section 4). This is indicated by the absence of  $t$  from the expressions for these outputs in Eqs. (7) and (8a).

*Segregation of temporal frequencies.* ERD output,  $y_4$ , can be considered as the integral over  $\omega$  of the responses to the temporal frequency components of the input display,  $c(x, t)$ . Let  $c_{\omega_0}(x, t) d\omega$  denote all those Fourier components of  $c(x, t)$  whose temporal frequencies are between  $\omega_0$  and  $\omega_0 + d\omega$ . Let  $y_4(\omega_0) d\omega$  denote the response to  $c_{\omega_0}(x, t) d\omega$ . It is easy to show that detector response  $y_4(\omega_0)$  to  $c_{\omega_0}(x, t)$  is given by

$$y_4(\omega_0) = [Y_{r,0}^*(\omega_0) Y_{l,0}(\omega_0) - Y_{l,0}^*(\omega_0) Y_{r,0}(\omega_0)] D(\omega_0). \quad (8b)$$

From Eqs. (8a) and (8b) it immediately follows that  $y_4 = \int y_4(\omega_0) d\omega_0$ ; the total response is the integral of the responses to the temporal frequency components. (The corresponding

property for periodic displays was derived by van Santen and Sperling.<sup>4</sup>)

### B. $\pi/2$ Temporal Phase-Delay Subunits

When  $D(\omega)$  is of the form  $-i|D(\omega)|$ , Eq. (7) reduces to

$$y_{H,3} = -i \int Y_{H,0}^*(\omega) Y_{H',0}(\omega) |D(\omega)| d\omega. \quad (9)$$

Note that  $y_{right,3} = -y_{left,3}$ , from which it follows that

$$y_4 = 2y_{right,3} = -2y_{left,3}. \quad (10)$$

Thus the output of a  $\pi/2$  phase-delay subunit is always equivalent to ERD output in the sense that they are directly proportional to each other.

### C. $\pi/2$ Spatial Phase-Shift Subunits

For this subunit, output is given by the rather uninviting equations

$$y_{left,3} = -i \iiint C^*(f, \omega) \times \text{sign}(f) C(g, \omega) R_{right}(f) R_{right}^*(g) D(\omega) d\omega df dg \quad (11)$$

and

$$y_{right,3} = +i \iiint C^*(g, \omega) \times \text{sign}(f) C(f, \omega) R_{right}(g) R_{right}^*(f) D(\omega) d\omega df dg. \quad (12)$$

In general, for this subunit detector, output is not equivalent to ERD output, as will be shown in Subsection 5.A by a counterexample. However, equivalence between subunit output and ERD output holds for certain classes of displays, which include the class of rigidly translating objects (see Subsection 5.B).

### D. Elaborated Watson-Ahumada Detector

*A linear subunit.* Watson and Ahumada<sup>5</sup> propose a motion subunit that corresponds—in structure but not in components—to an ERD subunit [Fig. 2(a)]. We discuss here how a reasonable elaboration of their detector is equivalent to a special case of the ERD.

Basically, Watson and Ahumada (W-A) replace the ERD's multiplier with an adder, resulting in a linear detector. To generate direction selectivity, they cleverly use phase-independent  $\pi/2$  spatial and temporal filters. Specifically, they assume that the spatial filters SF, the receptive fields, and temporal delay filters TF have the  $\pi/2$  property, independent of frequency, with the added assumption that the temporal filter has unit gain, that is,  $|D(\omega)| = 1$  for all  $\omega$ . Their subunit requires an additional pure delay filter (indicated by delay in the figure), because of the basically noncausal nature of the temporal filter; this delay can be ignored in the present discussion. For drifting sine-wave displays, the output of a W-A subunit is zero when the display moves in the nonpreferred direction and is a temporal sine wave with the same frequency as the display when the display moves in the preferred direction. The reason for this behavior is that sine-wave stimuli moving in the nonpreferred direction generate temporal sine waves within the subunit that exactly cancel each other in the adder. This, in turn, is due to the phase invariance (as a function of frequency) of the spatial and temporal filters.

Although the W-A subunit exhibits a form of direction selectivity, to convert the subunit into a motion detector re-

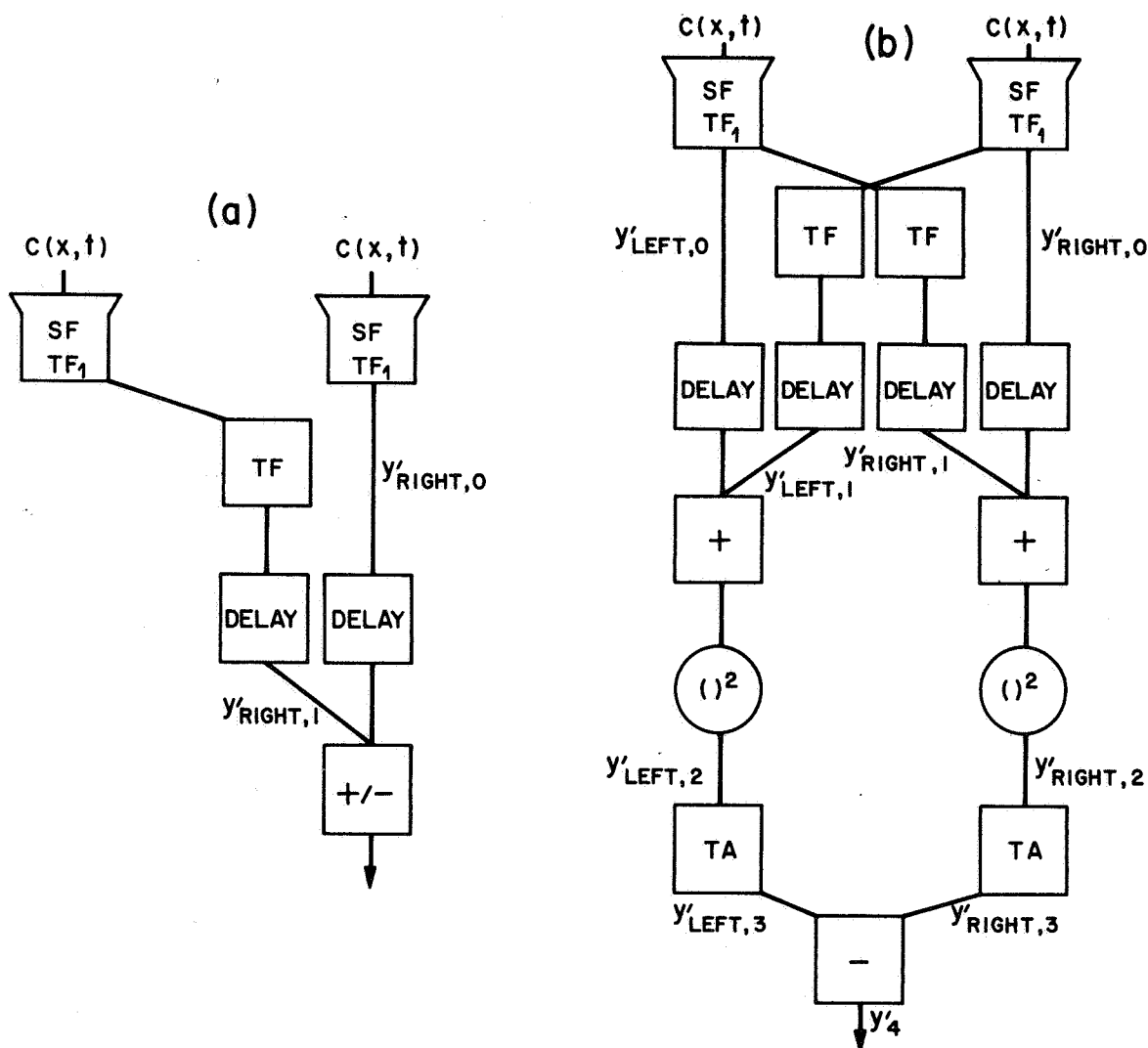


Fig. 2. Watson and Ahumada's<sup>5</sup> motion subunit and the elaborated W-A detector. (a) Their original linear subunit. SF,  $TF_1$  indicate separable spatial and temporal input filters, TF indicates linear, temporal delay filters, DELAY indicates an absolute delay inserted to simplify the description of the (noncausal) linear filters, +/- indicates that the operation is addition or subtraction according to whether the subunit signals motion to the right or left, respectively. (b) An elaborated W-A detector that includes squaring, temporal filtering TA and the subtraction of outputs of two mirror-image W-A subunits and is equivalent to an ERD.

quires additional components to convert the subunit's time-varying output into a single real number—the motion strength. That is, making a detector requires implementing a linking assumption. Although the W-A subunit is linear, it cannot be combined with a linear temporal linking assumption, such as time averaging or temporal integration, as the following example illustrates. Consider two drifting gratings that have the same temporal and spatial frequency; both move to the right and are 180 deg out of phase. Their sum is identically zero. Consider a W-A subunit that is tuned to rightward motion and an appropriate linking assumption so that it gives positive responses to either sine wave. If the output from this detector were transformed linearly into a strength, then the response to this sum should be equal to the sum of the separate responses to the component sine waves because the subunit itself is linear. But that would imply that the response to a zero display is positive, which is absurd. It follows that either the temporal linking assumption must be nonlinear or some nonlinear component must be added to the model. The problem is that, in distinguishing the nonzero

from the zero response (preferred from nonpreferred direction), it is the magnitude (a nonlinear computation) of the response that is critical.

*Nonlinear elaborations of the Watson-Ahumada subunit.* In their original formulation, W-A<sup>5</sup> did not provide a linking hypothesis. Van Santen and Sperling<sup>4</sup> suggested that the most reasonable linking assumption would be squaring followed by temporal integration and subtraction. When a subunit's output is squared, temporal integration over an integral number of cycles yields the power. (In fact, in plotting the power spectrum of their subunit, W-A applied precisely this squaring transformation to its time-varying output.) Since the W-A subunit responds only to movement in the preferred direction, we supply a mirror-image subunit for motion in the opposite direction and subtract the oppositely tuned subunit outputs to provide the output of an elaborated W-A detector [Fig. 2(b)]. It is easy to show that this elaborated W-A detector is fully equivalent to an elaborated Reichardt detector with  $\pi/2$  temporal delay and  $\pi/2$  spatial phase-shift filters. (The proof follows the same steps

as those used in the next section to evaluate Adelson and Bergen's detector.) It is really quite remarkable that a detector composed of subunits based on an additive comparison operation is equivalent to an ERD based on multiplicative comparison. The equivalence comes about because the squaring (multiplication) used to compute power in the additive model ultimately results in the equivalent operations being carried out in a different order.

Recently, W-A<sup>22</sup> themselves have elaborated their subunit with a different principle—frequency counting. This nonlinear operation may be efficient for velocity estimation—indeed, it has been successfully used in computer image processing to estimate velocity in spatially bandpass-filtered displays.<sup>23</sup> For efficient discrimination, however, squaring and time averaging have seldom been surpassed, have been proved to be theoretically optimal under many conditions, and are the most plausible computation for the direction discrimination task. It is worth noting, however, that while ERD's detect motion and discriminate direction of motion, in isolation they cannot reliably estimate velocity, an issue considered in Subsection 5.B.4.

#### E. Reformulation of the Elaborated Reichardt Detector: Adelson and Bergen's Detector

Adelson and Bergen<sup>24</sup> propose a motion detector (A-B detector), which on analysis (see below) turns out to be fully equivalent to the elaborated Reichardt model; it performs the same computations in a different sequence. The A-B detector differs from the Reichardt detector [Fig. 1(a)] in the way the functions  $y_{H,0}$  and  $y_{H,1}$  are combined at the subunit level. In Fig. 3, the A-B combination rule is broken down into its arithmetic components. Formally (using double primes to distinguish the A-B formulation from the standard ERD formulation)  $y''_{H,0}(t) = y_{H,0}(t)$  and  $y''_{H,1}(t) = y_{H,1}(t)$ . In the standard formulation, subunit output before temporal integration,  $y_{H,2}(t)$ , is given by

$$y_{H,2}(t) = y_{H,0}(t)y_{H,1}(t); \quad (13)$$

in the A-B formulation,  $y''_{H,2}(t)$  is given by

$$y''_{H,2}(t) = [y_{H,0}(t) + y_{H,1}(t)]^2 + [y_{H',0}(t) - y_{H',1}(t)]^2. \quad (14)$$

Equation (14) can be rewritten as

$$y''_{H,2}(t) = 2[y_{H,2}(t) - y_{H',2}(t)] + k(t), \quad (15)$$

where  $k(t)$  is defined as

$$k(t) = y_{\text{right},0}{}^2(t) + y_{\text{right},1}{}^2(t) + y_{\text{left},0}{}^2(t) + y_{\text{left},1}{}^2(t). \quad (16)$$

After temporal filtering

$$\begin{aligned} y''_{H,3}(t) &= 2[y_{H,3}(t) - y_{H',3}(t)] + K(t) \\ &= 2y_4(t) + K(t), \end{aligned} \quad (17)$$

where  $K(t)$  is  $k(t)$  after filtering. From Eq. (17) it follows that

$$\begin{aligned} y''_4(t) &= y''_{\text{right},3}(t) - y''_{\text{left},3}(t) \\ &= 2y_4(t) \end{aligned} \quad (18)$$

For all displays, the output  $y''_4(t)$  of the A-B detector differs from the ERD only by a multiplicative constant. For this reason, the formulation of A-B is appropriately called a reformulation of the ERD.

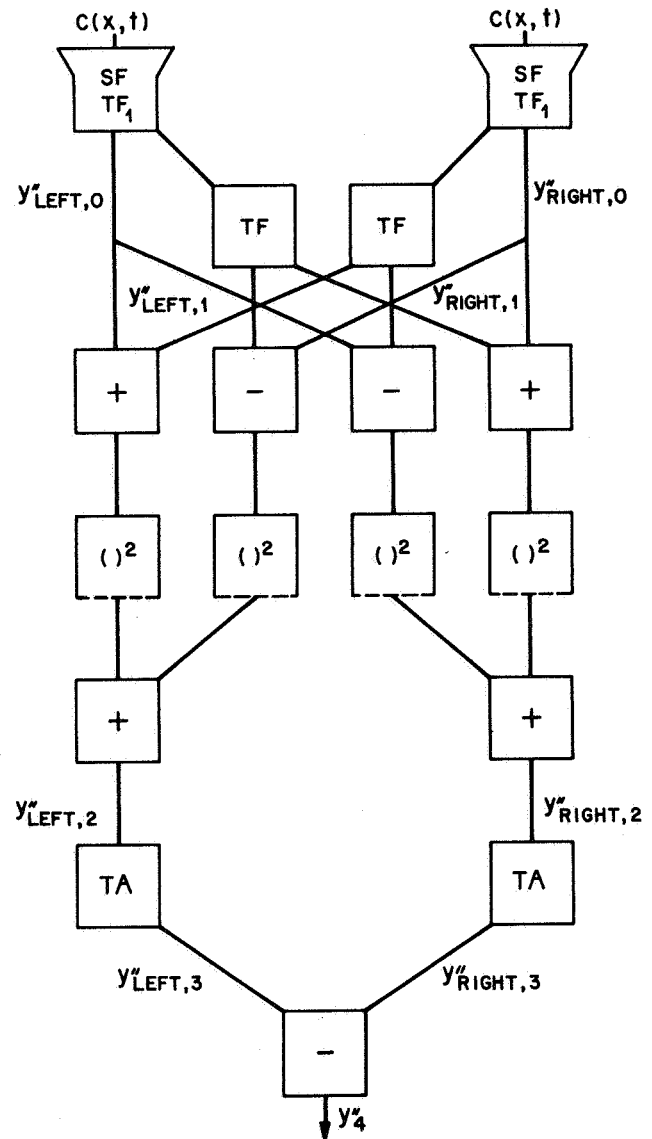


Fig. 3. Adelson and Bergen's<sup>24</sup> motion detector. Notation as in Figs. 1 and 2. We have provided the temporal filters TA. The A-B detector is equivalent to an ERD [Fig. 1(c)].

*Subunits of the Adelson-Bergen formulation.* The A-B reformulation of the ERD combines the inputs [Eq. (14)] in a more complicated way than the standard formulation. With  $\pi/2$  spatial phase shift and  $\pi/2$  temporal phase delay filters, A-B subunits have the following direction-selective response: For drifting sinusoidal gratings, subunit response is zero for movement in the nonpreferred direction (as in the W-A detector model) and is equal to  $K(t)$  for movement in the preferred direction. In the standard formulation, with  $\pi/2$  spatial phase shift and  $\pi/2$  temporal phase delay filters, subunit output is equivalent to detector output and is equal in magnitude but opposite in sign for these two drifting gratings. The nonnegativity of subunit output in the A-B formulation is consistent with the intuitive notion that direction-selective channels measure the spatiotemporal energy in the spatiotemporal frequency spectrum of a display, a notion that formed the impetus for A-B's work.

*Generalizations of the Adelson-Bergen formulation.* The



A-B formulation describes a detector that is a member of a larger class of detectors, which A-B call spatiotemporal energy models. These detectors have in common that  $y''_{H,2}(t)$  is of the form  $A_H^2(t) + B_H^2(t)$ , where  $A_{\text{left}}(t)$ ,  $B_{\text{left}}(t)$ ,  $A_{\text{right}}(t)$ , and  $B_{\text{right}}(t)$  are the outputs of four spatiotemporal filters. Equation (14) represents the special case in which  $A_H(t) = y_{H,0}(t) + y_{H,1}(t)$  and  $B_H(t) = y_{H',0}(t) - y_{H',1}(t)$ . It is generally not possible to decompose four arbitrary functions  $A_H(t)$ ,  $B_H(t)$  into four functions  $y_{H,j}(t)$  under the restriction that the latter functions share the temporal and spatial filters as specified by the ERD. However, if one drops the latter restriction and allows the four functions  $y_{H,j}(t)$  to be the outputs from four arbitrarily chosen spatiotemporal filters, then it is easy to show that the functions  $A_H(t)$ ,  $B_H(t)$  can always be cast in the form of Eq. (14). Such a model would be a generalized form of the ERD, in which subunits do not share receptive fields and temporal filters. Thus, if one gives the ERD the same freedom in choosing arbitrary spatiotemporal input filters, the resulting class of models is equivalent to the class of spatiotemporal energy models.

#### F. The General Fourier Analyzer

Because it is so simple, the general Fourier analyzer (GFA) is a kind of null hypothesis against which the more complicated motion theories must be justified. The GFA assumes that the psychophysical response to a motion display consists of the independent responses elicited by each spatiotemporal Fourier component of the display. For quantitative comparisons, it is usual to assume, additionally, that the response amplitude of the GFA to isolated Fourier components matches that of the system to which it is being compared, e.g., the ERD or the visual system. The GFA definitely is not a Reichardt variant. Nor is it even a detector until an explicit linking assumption is made, usually that magnitude (or power) is detected.

The GFA is an extension to vision of Helmholtz's model of the cochlea,<sup>25</sup> which asserted that the cochlea performed a Fourier analysis of the auditory stimulus. The virtue of the GFA is that it is an off-the-shelf model, widely known, and often useful. It is incomplete in that it makes no assumptions about how responses to various components might combine or about the spatiotemporal window within which the Fourier analysis is performed—the same difficulties that beset it in the auditory domain.

A Fourier analysis of the stimulus is often revealing. For example, a checkerboard has its fundamental Fourier components along the diagonals, not parallel to the edges of the checks. Ives<sup>26</sup> used these diagonal Fourier components as the basis of his acuity grating, and they influence even complex perceptual properties.<sup>27</sup>

The GFA is most useful in cases in which one Fourier component dominates the stimulus, and we will make extensive use of it as a comparison reference for the ERD. The only complication considered here for the GFA is that its magnitude of response to any Fourier component matches the assumed magnitude of the visual system's response to that component viewed in isolation. Once stimuli become complex, and the GFA has to be elaborated with windows and with combination rules, it no longer is off the shelf, nor simple, nor familiar; it is merely a general computational procedure for analyzing motion responses in the Fourier domain.

#### 4. ELABORATED REICHARDT DETECTOR IN RELATION TO MOTION DETECTORS IN GENERAL

In this section, we consider several issues that pertain to all motion stimuli and detectors: local versus global detection (windowing), detectors versus models, late versus early nonlinearities, representation of one-dimensional motion versus representation of two-dimensional spatial patterns ( $x$ - $y$  versus  $x$ - $t$  representations), receptive fields of motion detectors, the Fourier representation of common stimuli, and the relation of the general Fourier analyzer to the ERD and other detectors under consideration.

*Local versus global detection (windowing).* In the spatial domain, the ERD is local because it has receptive fields that confine most of its input to some local area. Receptive fields define the spatial window within which a detector detects. In the temporal domain, the general ERD also is local in so far as the TA filter [Fig. 1(a)] is. Under the assumption of infinite time averaging for the TA filter, the ERD is completely global in the temporal domain, as it is under the analogous assumption that it integrates the entire response to a temporally confined stimulus [Eq. (5)].

In contrast to the elaborated Reichardt detector, the general Fourier analyzer does not explicitly have windows in either the spatial or the temporal domain and therefore in the ultimate global model. The GFA is applicable only to spatially and temporally periodic and stationary displays and to other cases where windowing can be neglected.

*Models, channels, and detectors.* First, because they are local, detectors are assumed to be replicated at many locations, i.e., to be dense in visual space. Second, in each small neighborhood, different sizes of detectors are assumed to occur. A class of detectors of a single size is called a spatial frequency channel. Third, in each spatial location and for each size, there are detectors having different primary orientations; these detectors form orientation channels. Fourth, when the assumption of infinite time averaging is relaxed, detectors' responses are time varying.

The algorithm for combining the information from all these many detectors and moments in time is called a voting rule. Typical examples are: choose the detector having the maximum output during the trial interval; add the detector outputs according to an *a priori* weighting function and make the decision according to whether the sum exceeds a criterion value. We<sup>3</sup> were able to show that for many stimuli the choice between these two voting rules had no influence on the predicted psychophysical response of the elaborated Reichardt model. In general, optimum voting rules depend on the nature of noise in the model. In this paper, it will usually be possible to evaluate detectors without invoking voting rules. Voting rules, however, are central to perceptual theories of motion.

*One-dimensional motion stimuli.* In general, a monocular, achromatic motion stimulus is three dimensional. It is represented as a luminance  $c(x, y, t)$  that is a function of two spatial dimensions ( $x, y$ ) and time ( $t$ ). Here, however, we confine ourselves to motion stimuli that are functions of  $x$  and  $t$  only and do not vary in the  $y$  direction. For example, such stimuli are vertically oriented gratings or patterns; and ERD's are assumed to be oriented optimally for detecting these gratings. (Orientation selectivity, that is, off-angle detection



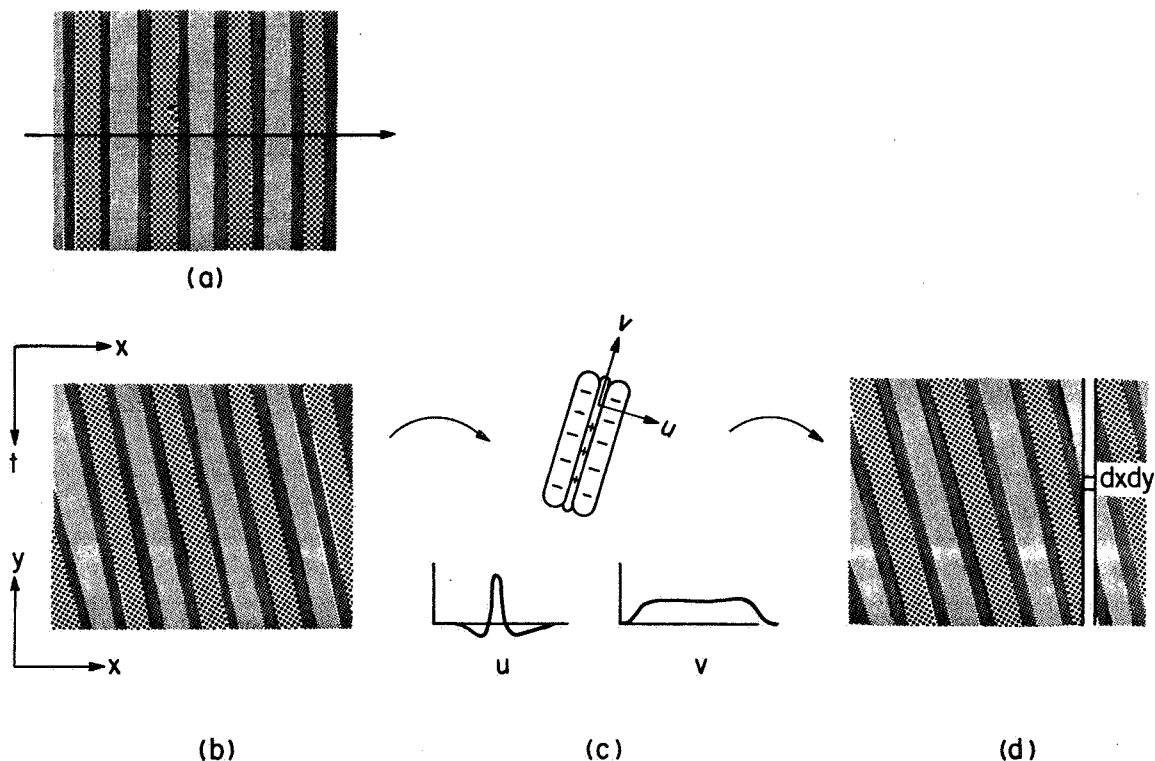


Fig. 4. (a) Artist's rendition of a single frame of a one-dimensional drifting sinusoidal grating. The various shadings represent stimulus intensities. The arrow indicates the direction of movement. (b) Two-dimensional  $(x, t)$  representation of (a). The horizontal dimension  $x$  represents the instantaneous spatial luminance pattern. The vertical dimension represents time. The slope inversely indicates the speed of movement. (b) Also can be interpreted as a static, diagonally oriented grating in two dimensions, that is, a grating pattern in  $(x, y)$  space. (c) A linear,  $u, v$  separable, receptive field (approximation to a simple cell in cortical area 17 that receives (b) as an input (spatial interpretation). The  $+$  and  $-$  symbols are a conventional notation used to indicate areas in which the response to a point of light is positive and negative, respectively. The receptive field is oriented along its primary axes  $u, v$ , and the local responsivity of the field (its impulse response) is indicated in the cross-sectional graphs,  $u, v$ . (d) Transformation of (b) via (c) onto a receptor field. The small square  $dx dy$  indicates the physical area occupied by a receptor located at  $x, y$  (spatial interpretation). In the motion interpretation, the vertical slit represents the physical domain of a motion receptor, that is, a motion receptor at location  $x$  responds to inputs from all times  $t$ .

by motion detectors, is a complicated issue that is beyond the scope of the present treatment.) Figure 4(a) illustrates a static view of a drifting sinusoidal grating; Fig. 4(b) illustrates the representation of the drifting sinusoid as a function of  $x$  and  $t$ .

*Receptive fields;  $(x, y)$  versus  $(x, t)$ .* It is obvious from Fig. 4 that the one-dimensional, drifting sinusoidal grating described as a function of  $x$  and  $t$  has the same representation as a static two-dimensional sinusoidal grating described as a function of  $x$  and  $y$ .<sup>5,6,28,29</sup> It is instructive to consider the similarities and differences between methods of detecting these two stimuli,  $c(x, y)$  and  $c(x, t)$ .

For detection of static  $x, y$  gratings, it is assumed that there are receptors with receptive fields of various sizes and orientations and that the receptors with the largest responses detect the grating or contribute most heavily to a weighted-detection response.<sup>30-32</sup> Each kind of receptor (a particular orientation and size) exists in many points within the visual field, and the stimulus is represented on a field of these receptors. If it is assumed that the receptors are approximately linear, at least in their responses to low-contrast, near-threshold stimuli, then the receptor can be regarded as a linear operator, and the transformation from stimulus field to receptor field is easily computed [Fig. 4(c)]. An especially useful property of linear systems is that input sine waves are transformed to output sine waves differing only in phase and amplitude but not in

frequency. Thus we know immediately that a stimulus  $x, y$  sinusoidal pattern [Fig. 4(b)] also is represented in any linear receptor field as a sinusoidal pattern, as illustrated in Fig. 4(d).

When the stimulus pattern in Fig. 4(b) is interpreted as a spatial  $x, y$  pattern, every point  $x, y$  in the receptor field of Fig. 4(d) represents the magnitude of response of a receptor located at  $x, y$ . When the pattern is interpreted as an  $x, t$  drifting grating, the point  $x, t$  in the receptor field represents the output at time  $t$  of the receptor located at point  $x$ . Different orientations in  $x, y$  correspond to different velocities and left-right directions in  $x, t$ .

Every small area  $dx dy$  in the  $x, y$  receptor field of Fig. 4(d) represents a different receptor. When interpreted as an  $x, t$  field, only  $x$  locations represent receptors; physically, a motion receptor occupies a whole vertical slit ( $dx$ ). A single-motion receptor transmits signals about all times  $t$ . Physically, the structures of the  $x, y$  and  $x, t$  fields are quite different. Functionally, they are quite similar. There are inevitable asymmetries between  $x, y$  and  $x, t$  because past and future are not so symmetrical as left and right or up and down. A motion receptor can respond only to past inputs; a spatial receptor can respond to inputs either above or below its central location. With respect to the set of receptors in physical  $x, y$ , the  $x$  and  $y$  dimensions are essentially symmetrical and arbitrary (i.e., they can be freely rotated). With respect to

motion receptors, the  $x, t$  dimensions are completely constrained and unsymmetrical. These asymmetries can be large, or they can be quite minimal; they depend on the specific nature of the  $x, y$  and  $x, t$  receptive fields. Thus, even though the physical natures of the  $x, y$  and  $x, t$  receptors are essentially different, the formal descriptions of their transformations are equivalent.

Throughout this section, we have been careful to refer to linear operators as receptors. When it comes to detection, a nonlinear operation, the same arguments transpose perfectly. Physically, the detection of patterns in  $x, y$  involves quite different structures than does the detection of patterns in  $x, t$ ; formally, the detection operations can be equivalent.

*Early and late nonlinear detectors.* A linear system is one in which superposition (and certain other technical conditions) obtain. Superposition holds if and only if for every choice of two inputs,  $a, b$  (with corresponding outputs,  $A, B$ ), the combined input  $a + b$  produces an output of exactly  $A + B$ . On the other hand, directionally selective motion detectors are inherently nonlinear (see Subsection 3.D). For any motion detector, even when the detector output is a continuously graded confidence level (rather than the intrinsically nonlinear categorical yes/no response), there are profound physical and logical reasons for detector nonlinearity.

The simplest kind of detector is one that is describable as a linear system right up to the last stage—a decision stage that, typically, computes the power of its input and/or determines whether it exceeds a criterion value.<sup>33</sup> This is a late nonlinear detector. A Reichardt detector becomes nonlinear at the point at which the left and right inputs are multiplied—an earlier nonlinearity. However, the elaborated W-A detector, which performs an equivalent overall computation to the ERD, is late nonlinear. A late-nonlinear motion detector allows one to regard the linear transformation carried out by the motion detector in  $x, t$  as being analogous to the linear receptive-field transformation carried out by the spatial detector in  $x, y$  (as in Fig. 4).

Nonlinear systems (such as the ERD) do not have transfer functions (linear receptive fields). But, because the ERD (for certain choices of its filters) has a late-nonlinear dual, the elaborated W-A model, one may reasonably ask: What is its receptive field? The answer is that the ERD's receptive field in  $x, t$  looks remarkably like the  $x, y$  receptive fields typically proposed for spatial detectors [Fig. 4(c)]. Different orientations of  $x, y$  receptive fields indicate optimal sensitivity to gratings with the corresponding orientations. Similarly, different orientations of the receptive field in  $x, t$  indicate optimal sensitivity to motion with the corresponding speeds and left-right directions. The receptive field of the ERD's late-nonlinear dual is computed before the last stage of temporal filtering, time averaging [TA, Fig. 2(b)]. Time averaging would stretch the receptive field vertically in Fig. 4. However, since time averaging is preceded by a nonlinear operation—squaring—the receptive field concept is no longer applicable at this stage.

*Fourier representations of commonly occurring functions.* What does it mean that receptive fields of spatial mechanisms in  $x, y$  are so similar to receptive fields of one-dimensional motion mechanisms in  $x, t$ ? Both receptive fields are designed to discover patterns, and these patterns are quite well described in terms of narrow bands of Fourier components. In both  $x, y$  pattern detection and  $x, t$  motion detection, a

major component of detection reduces to frequency analysis—the discovery and detection of Fourier components. In fact, the ERD's remarkable property of segregating temporal frequencies [Eqs. (8a) and (8b)] implies that it detects any two or more sine-wave patterns in  $x, t$  independently unless they happen to have the same temporal frequency. Therefore, it is useful at this point to consider the frequency analysis of the stimuli that will be used to test the ERD and its variants.

Figure 5 illustrates some two-dimensional functions, which can be interpreted as either  $x, y$  or  $x, t$ , and their Fourier transforms; the  $x, t$  interpretation is used here. Luminance cannot take negative values; therefore all contrast functions are added to an unchanging, uniform background. A uniform, unchanging field has zero spatial and temporal frequency; there its transform is a point at the origin in the middle of the panel. Because the transformation from a function to its Fourier transform is linear, the transform of any sum of functions is simply the sum of the transforms of the components. Thus every transform panel of Fig. 5 has a point at the origin representing the uniform background.

Functions a–e in Fig. 5 are periodic; one period of an infinite two-dimensional mosaic is shown. For specificity, the displayed area can be regarded as representing 1 deg of visual angle (dva)  $\times$  1 sec. Panel a illustrates a right-drifting sinusoid. The period of  $c(x, t)$  contains four cycles in  $x$  and one cycle in  $t$  [shown in  $C(f, \omega)$ ] and indicates a velocity [slope of  $C(f, \omega)$ ] of  $(1 \text{ cycle/sec})/(4 \text{ cycles/dva}) = 0.25 \text{ dva/sec}$ . Panel b illustrates a right-drifting sinusoid with double the spatial frequency and double the temporal frequency of a, and hence the same velocity. Panel c illustrates spatially uniform flicker; the spatial frequency is zero, the temporal frequency is 8 Hz. Rotating panel c 90 deg would yield a vertical sinusoid whose Fourier transform also would be rotated 90 deg, corresponding to a spatial frequency of 8 cycles/dva. Panel d illustrates a left-moving sinusoid, velocity 0.5 dva/sec.

The axes of the transforms in panels a–d are perpendicular to the stripes in  $c(x, t)$ , which is generally true for moving patterns. Every pair of points in  $C(f, \omega)$  that is symmetrically disposed around the origin represents a drifting sinusoidal component of  $c(x, t)$ . Panel e illustrates a multiplicity of components in a sampled (stroboscopic) motion display derived from d by briefly flashing the moving sinusoid eight times in each cycle. The motion path is ambiguous; three possible paths are shown. The Fourier transform also represents these paths; only the first three of many possible paths are shown. Note the perpendicularity of the motion path and the axis of the corresponding pair of points in the Fourier transform.

The functions in panels f–j are not periodic in  $x$ ; they are spatially confined; what is shown is all that is nonzero in  $x$ . Panel f illustrates a continuously present, stationary single line; its transform is a line rotated 90 deg. Panel g shows a continuously present, stationary random bar pattern. Its transform is restricted to  $C(f, 0)$ , but owing to  $c(x, t)$ 's limited spatial extent,  $C(f, 0)$  is not a constant but represents the statistical properties of  $c(x, t)$ .

The functions in h–j are spatial and temporally confined. Panel h shows a briefly flashed sinusoidal grating pattern. Its transform is not completely uniform in the vertical dimension because of the significantly nonzero duration of the flash.  $C(f, \omega)$  is not confined to just two spatial locations ( $\pm 4$  cycles/dva) because of the restricted spatial extent of  $c(x, t)$ . Panel i il-

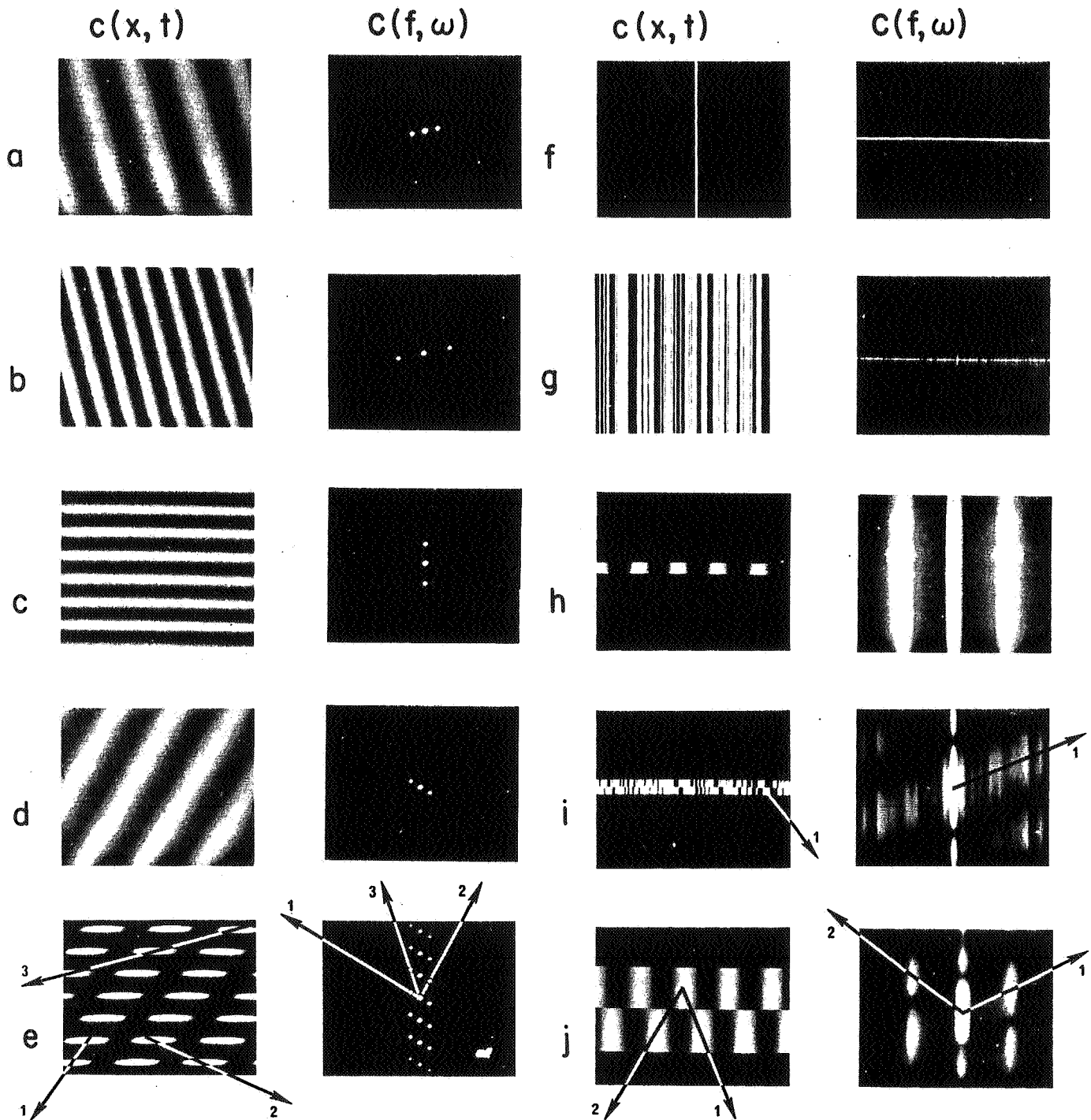


Fig. 5. Examples of two-dimensional functions considered in this paper and their Fourier transforms. For the functions  $c(x, t)$ , the horizontal dimension is  $x$ , the vertical is  $t$ , and the shading indicates the value of  $c(x, t)$ . For the Fourier transforms  $C(f, \omega)$ , the horizontal dimension is spatial frequency  $f$ ; the vertical dimension is temporal frequency  $\omega$ ; the shading indicates the magnitude of the complex number  $C(f, \omega)$ . The one-dimensional motion interpretation ( $x, t$ ) rather than the two-dimensional space ( $x, y$ ) interpretation is used in the following descriptions. Functions (a) to (e) are periodic; one period of an infinite two-dimensional mosaic is shown. (a) Right-drifting sinusoid. (b) Right-drifting sinusoid with double the spatial frequency and double the temporal frequency of (a) and hence the same velocity. (c) Spatially uniform, stationary flicker. (d) Left-moving sinusoid. (e) A sampled (stroboscopic) motion display derived from (d). The motion path is ambiguous; three possible paths (1, 2, 3) and their transforms are shown. The contrast functions (f-j) are not periodic in  $x$  but are spatially confined to the  $x$  interval shown; functions (h-j) are also temporally confined. (f) Continuously present, stationary single line. (g) Continuously present, stationary random bar pattern. (h) A briefly flashed sinusoidal grating pattern. (i) A two-flash presentation of a random bar pattern moved to the right. (j) A two-flash presentation of a sinusoidal grating pattern moved to the right.

illustrates a two-flash presentation of a random bar pattern in which the second flash is moved to the right of the first. The arrows (1) indicate the  $x, t$  movement and the many Fourier components consistent with this movement direction; inconsistent components are also apparent. Panel j illustrates a two-flash presentation of a sinusoidal grating pattern moved to the right; 1 (right) and 2 (left) represent the two most prominent movement interpretations and their Fourier transforms.

**5. DERIVATIONS OF ELABORATED REICHARDT DETECTOR AND SUBUNIT RESPONSES TO STANDARD DISPLAYS**

In this section, we use the general results obtained in previous sections to investigate the response of the ERD and two subunit variants to four important classes of displays: (1) simple drifting sine-wave gratings; (2) combinations of sine waves, such as counterphase and on-off gratings; (3) two-frame motion; and (4) rigid translations.

**A. Gratings with Sinusoidal Temporal Modulation**

The results below are based on the auxiliary assumption of temporal integration over a suitably chosen temporal interval (see Subsection 2.E). In terms of appropriate experiments to test these predictions, this means that one should present at least several temporal cycles of a display, viewed or temporally windowed in such a way that the psychophysical response is determined by the maximum modulation presented and not by the manner of onset or termination of the display.

**1. Responses to Drifting Sinusoidal Gratings**

A drifting sinusoidal grating is defined as

$$c(x, t) = L_0 + m \cos(2\pi f_0 x + 2\pi \omega_0 dt). \quad (19)$$

Here, the average luminance is  $L_0$ , the modulation amplitude is  $m$ , the temporal frequency is  $\omega_0 \geq 0$ , the spatial frequency is  $f_0 \geq 0$ , and the direction of movement is  $d(-1, 1)$ . Without loss of generality,  $L_0$  can be ignored.

For computing ERD responses to sinusoids, it is more convenient to switch to polar notation. As before, for any  $G(\omega)$  we define its magnitude as  $|G(\omega)| \equiv \{\text{Im}^2[G(\omega)] + \text{Re}^2[G(\omega)]\}^{1/2}$  and its phase as  $\tau_\omega \equiv \tan^{-1}\{-\text{Im}[G(\omega)]/\text{Re}[G(\omega)]\}$ .

Let  $\tau_{\omega_0}$ ,  $\rho_{\text{left},f_0}$  and  $\rho_{\text{right},f_0}$  denote the phase angles of  $D(\omega_0)$ ,  $R_{\text{left}}(f_0)$ , and  $R_{\text{right}}(f_0)$ , respectively. Subunit output is given by

$$y_{H,3} = m^2 |D(\omega_0)| |R_H(f_0)| |R_{H'}(f_0)| \times \cos[\tau_{\omega_0} - d(\rho_{H,f_0} - \rho_{H',f_0})]. \quad (20)$$

Obviously,  $y_{\text{left},3} = -y_{\text{right},3}$  for both subunit variants 1 and 2. Finally, ERD output  $y_4$  can be written as

$$y_4 = m^2 |D(\omega_0)| |R_{\text{left}}(f_0)| |R_{\text{right}}(f_0)| d \times \sin(\tau_{\omega_0}) \sin(\rho_{\text{right},f_0} - \rho_{\text{left},f_0}). \quad (21)$$

This is a direct generalization of Eqs. (13) and (15) in Ref. 4.

**Nonaliasing.** The conditions for nonaliasing of an ERD are easily derived from Eq. (21). Nonaliasing requires, for all  $\omega_0$  and  $f_0$ ,

$$|D(\omega_0)| d \sin(\tau_{\omega_0}) \geq 0 \quad (22)$$

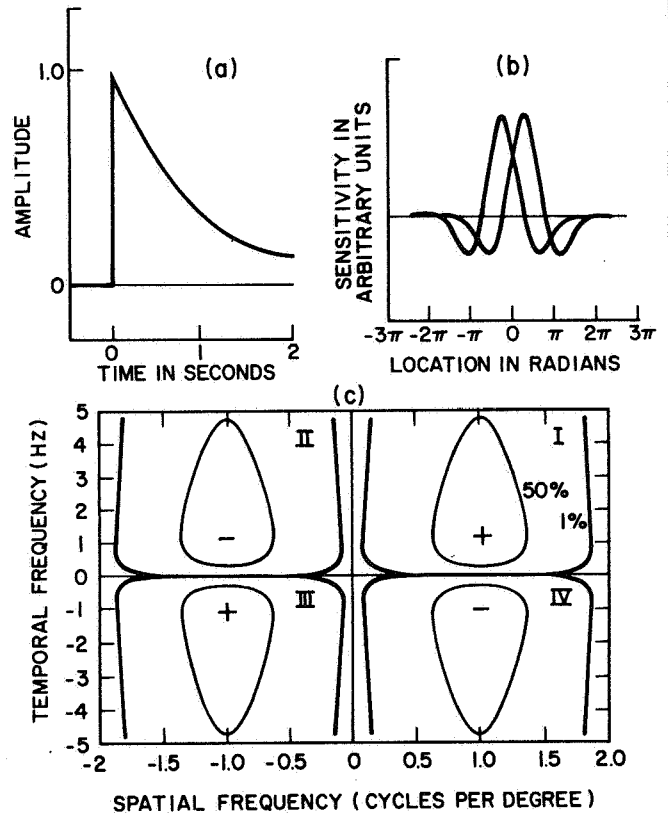


Fig. 6. Frequency responses of an ERD. (a) Impulse response function of the temporal delay filter (TF, Fig. 1), a first-order low-pass filter with time constant of 1 Hz. (b) Spatial filters (left and right receptive fields) of the ERD. The abscissa indicates the spatial location of a line stimulus input; the horizontal line indicates zero response; the ordinate indicates output amplitude in response to this stimulus. The one-dimensional receptive fields are even Gabor functions, separated by one-quarter cycle of the optimal spatial frequency for the individual receptive fields [Eq. (24)]. (c) Response of this ERD to drifting sinusoidal gratings as a function of their spatial and temporal frequency. Two contour lines are shown, representing the 50 and 1% contours of the absolute value of the response relative to the maximum response. The locations where the actual maxima occur are indicated by the + and - signs. These signs also indicate the sign of ERD output within each quadrant.

and

$$|R_{\text{left}}(f_0)| |R_{\text{right}}(f_0)| \sin(\rho_{\text{right},f_0} - \rho_{\text{left},f_0}) \geq 0. \quad (23)$$

Conditions (22) and (23) are easy to satisfy. When the temporal filter has a  $\pi/2$  delay, condition (22) always is satisfied, and when a detector has receptive fields with a  $\pi/2$  phase shift, condition (23) always is satisfied. However,  $\pi/2$  phase shifts are sufficient, not necessary. There is a wide range of temporal and spatial filters that an ERD can have and still function properly. In fact, the great advantage of the ERD over detectors composed of only a single subunit is precisely the ERD's ability to perform well without narrow constraints on its filters.

Figure 6 illustrates responses of a nonaliasing ERD with simple (nonoptimal) filters. The temporal filter is a first-order low-pass filter with impulse response function  $\text{TF}(t) = e^{-t}$  [Fig. 6(b)]. This type of filter does not have the  $\pi/2$  temporal phase-delay property. Its phase delay is always positive and increases asymptotically to  $\pi/2$ .

The receptive field filters are two displaced, symmetric "Mexican-hat" functions given by

$$r_H(x) = W(x - x_H)\cos[(x - x_H)f_0] \quad H = \text{left, right.} \quad (24)$$

The weighting function  $W$  is Gaussian,  $W(x) = \exp(-x^2/4\sigma^2)$ , where  $\sigma = 1.41$ ,  $x_{\text{left}} = -\pi/2$ ,  $x_{\text{right}} = \pi/2$ , and  $f_0 = 1$ . These displaced, but otherwise identical, receptive fields do not have the  $\pi/2$  spatial-phase-shift property. In fact, the phase shift is directly proportional to spatial frequency and thus varies between zero and infinity. It should be noted, however, that at the spatial frequency that maximizes the individual receptive field responses the phase shift is approximately  $\pi/2$ .

Figure 6(c) shows contour plots of those spatial and temporal frequencies at which the absolute value of the response  $|y_d|$  is 1 and 50% from the peak response amplitude. Figure 6(c) shows that despite the casually selected spatial and temporal filters, the detector behaves properly, giving significant responses that always have the correct sign (no aliasing). Other filter choices yield comparably efficient ERD's.

Figure 6 represents the response  $|y_d|$  of the ERD. If it were desired only to represent response amplitude, Fig. 6 could have been drawn in a single quadrant. An alternative interpretation of Fig. 6 is as representing the magnitude of the Fourier transform of a linear system that exhibits precisely the same magnitude of response to Fourier components as does the ERD. The original, unelaborated W-A model [before squaring and time averaging [Fig. 2(a)]] is one such ERD dual. Assuming a particular dual uniquely constrains the other quadrants of the Fourier transform (except for sign). Other motion detectors (such as a GFA) may also have the same response magnitudes, and they may or may not exhibit the same phase relations as the ERD dual. Without other constraints, the response data of Fig. 6(c) do not imply a unique motion detector.

The representation in Figure 6(c) of ERD response magnitude enables one to compute ERD responses graphically to more complex stimuli. For example, a simple sinusoid traveling to the left is represented by symmetric points in quadrants 1 and 3. The ERD response in both of these quadrants is positive, the response to these points add [Eqs. (8)], and so the ERD's response is positive. A simple sinusoid traveling to the right is represented by points in quadrants 2 and 4; responses in both of these quadrants are negative; therefore the ERD's response is negative. When two stimuli are presented simultaneously, and their temporal frequencies are different, the responses simply add; for movements in opposite directions, the responses cancel completely with appropriately balanced stimuli. The property that all responses in quad-

**Table 2. Response of the Elaborated Reichardt Detector to Combinations of Drifting Sinusoids as a Function of Their Differences in Spatial and/or Temporal Frequency<sup>a</sup>**

Temporal Frequency	Spatial Frequency	
	Different	Same
Different	$i_s$	$i_s$
Same	$i_s$	$d$

<sup>a</sup>  $i_s$  indicates independent combination with an ERD of component's powers according to a sum rule.  $i_s$  indicates independent combination between different-sized ERD's according to an arbitrary voting rule.  $d$  indicates dependent, nonlinear combination within an ERD.

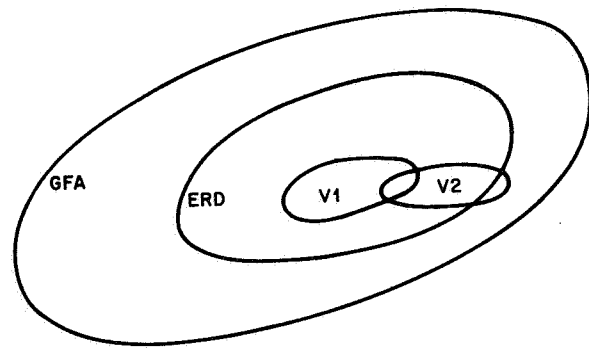


Fig. 7. The range of responses to combinations of drifting sinusoids of: the GFA (with arbitrary combination rules), the ERD, and two ERD subunits ( $V1 = \pi/2$  temporal phase-delay subunit;  $V2 = \pi/2$  spatial phase-shift subunit).

rants 1 and 3 have signs opposite all responses in quadrants 2 and 4 is a graphic representation of the fact that this ERD is nonaliasing, that is, it classifies every simple, drifting sinusoidal grating correctly with respect to left-right direction of motion.

*Combination properties of the elaborated Reichardt detector.* Drifting sinusoids are too simple for critical tests of motion detectors; stimuli composed of combinations of two or more drifting sinusoids are needed. Three kinds of sinusoidal grating combinations are significant for the ERD (see Table 2).

(1) When drifting gratings have different temporal frequencies, the ERD has a remarkable additivity property: The response of the ERD (and its variants) to a stimulus that is the sum of sinusoids is simply the sum of the ERD's responses to the isolated components. This is independent combination according to a sum rule.

(2) Independent combination still results when components of a compound drifting grating have the same temporal frequency but widely different spatial frequencies. Because of the spatial-frequency selectivity of the ERD, the component gratings stimulate different ERD's and therefore still combine independently according to the voting rule for combining independent detector responses. In cases 1 and 2, ERD's behave essentially like the simplest of general Fourier analyzers—independent responses to the component stimuli. One powerful criterion for independent combination is phase independence—the response is the same independent of the phases between stimulus components.

(3) When both temporal and spatial frequencies of component grating are similar, the ERD predicts profound dependencies. However, that is just where the windowing assumptions of the ERD become important. Nearby frequencies (space or time) imply beats, i.e., quite different things happen locally (during brief intervals or within spatial neighborhoods), and only with infinite temporal windows do these beat episodes disappear (into perfect frequency analysis). We show in Sec. 5.B.3 that the particular way in which spatial components interact perceptually is uniquely predicted by the ERD (and its variants). However we do not deal here with temporal frequency resolution (e.g., temporal beats).

The GFA does not distinguish stimuli with beats from any other stimuli. Failures of frequency resolution (beats) require

elaboration of a GFA; any particular set of phase relations would not be an obvious extension of a GFA.

*Comparison of the responses of the elaborated Reichardt detector and the two subunit variants to combinations of sinusoids.* The relations among the various detectors under consideration is best illustrated by a Venn diagram (Fig. 7). Each point represents an input-output pair, that is, a particular kind of response of a detector to an input consisting, for specificity, of a pair of sinusoids. The enclosed areas represent the range of achievable input-output pairs (over the range of parameters) for each of four kinds of detectors: GFA, ERD, Variant 1 (ERD subunit with  $\pi/2$  temporal-phase-shift filters), and Variant 2 (ERD subunit with  $\pi/2$  spatial-phase-shift filters). For the GFA, any voting rule (not merely addition) is admissible for combining Fourier components, so this GFA is the most general model. The ERD, which combines components as described in the previous section (and in Table 2) is much more specific. Variant 1 is always equivalent to a full ERD. Variant 2 is equivalent to an ERD for stimuli in which each temporal frequency is associated with at most one spatial frequency but can give slightly different responses to other complex stimuli.

2. Counterphase and On-Off Gratings

An extended counterphase grating is described by

$$c(x, t) = L_0 + m \cos(2\pi\omega_0 t) \cos(2\pi f_0 x). \quad (25)$$

Ignoring  $L_0$ , the Fourier transform of  $c$  is

$$C(f, \omega) = m, \quad f = \pm f_0 \text{ and } \omega = \pm \omega_0 \\ = 0, \quad \text{elsewhere} \quad (26)$$

The counterphase grating and its Fourier transform are illustrated in Figs. 8(a), 8(b), and 8(d). Because of the symmetry of the Fourier transform of the counterphase grating—one point falls in each of the four quadrants of the ERD cloverleaf (Fig. 6)—it immediately suggests that the ERD

response will be zero, as in the case of two sine waves moving in opposite directions. This is easily proved. In polar notation, subunit output is

$$y_{H,3} = m^2 |D(\omega_0)| |R_{left}(f_0)| \\ \times \cos(\tau_{\omega_0}) \cos(\rho_{left,f_0}) \cos(\rho_{right,f_0}). \quad (27)$$

Obviously,  $y_{left,3} = y_{right,3}$ , so  $y_4$  is always zero.

Equation (27) is an example of a case in which the output from a  $\pi/2$  spatial-phase-shift subunit (Variant 2) is not equivalent to ERD output (Subsection 3.C). When the difference between  $\rho_{right,f_0}$  and  $\rho_{left,f_0}$  is  $\pi/2$  (and ERD output is zero), the product  $\cos(\rho_{left,f_0}) \cos(\rho_{right,f_0})$  (which determines subunit output) is generally nonzero. For counterphase gratings, the response of  $\pi/2$  spatial subunits may or may not be zero; ERD response always is zero.

An on-off temporally modulated spatial grating is defined by Eq. (28) and illustrated by Figs. 8(a), 8(c), and 8(e).

$$c(x, t) = L_0 + m [\cos(2\pi\omega_0 t) + 1] \cos(2\pi d f_0 x). \quad (28)$$

This on-off display is equivalent to a counterphase display plus a stationary sinusoidal grating. Since an ERD has zero response to stationary stimuli, the ERD response to the on-off modulated grating will be the same as to the counterphase grating, i.e., zero. Analytically, it is easy to show that the counterphase Eq. (27) also describes subunit responses to on-off gratings.

3. No Spatial Frequency Doubling

Kulikowski<sup>8,9</sup> and Murray *et al.*<sup>10</sup> studied the effects of the type of temporal modulation (drifting versus counterphase versus on-off) on the highest resolvable spatial frequency of these gratings (resolution limit). They concluded that, for a range of experimental conditions, the resolution limit is twice as high for on-off gratings as for counterphase gratings, a phenomenon they called spatial-frequency doubling. They assert that “[the property of the Reichardt detector that] ...

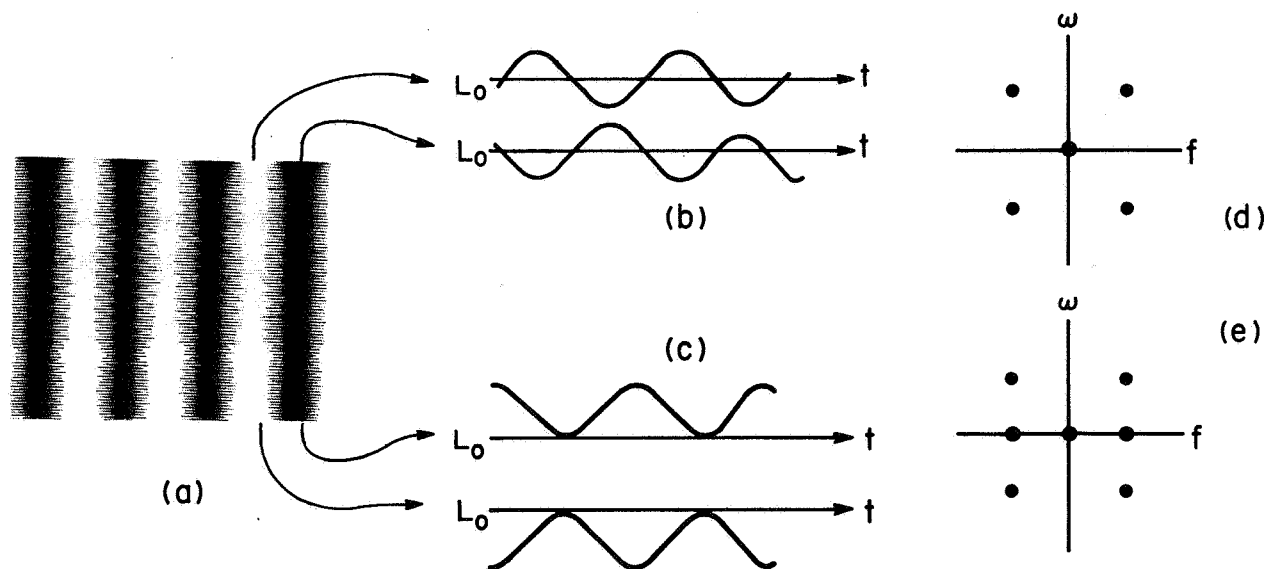


Fig. 8. Fourier representations of counterphase and on-off gratings. (a) Freeze-frame  $x, y$  representation of a grating. (b) Counterphase grating; temporal modulation of luminance (ordinate) of adjacent bars around a mean luminance  $L_0$  as a function of time (abscissa). Ordinate: luminance; abscissa, time. (c) On-off modulated grating. (d) Fourier transform of counterphase and (e) on-off modulated grating. Note that these two stimuli differ only by the presence of a static sinusoidal grating shown in (e).



signals of two adjacent inputs are cross-correlated . . . [leads to] . . . halving of the spatial resolving power . . . [p. 151]."<sup>10</sup> We now demonstrate that this assertion is incorrect.

On-off and counterphase grating stimuli differ only in the presence of a stationary grating (Fig. 3), and the ERD does not respond to stationary stimuli of any kind nor do they affect the ERD's responses to moving stimuli. Therefore finding different responses to on-off and counterphase gratings cannot be explained by any reasonable version of the ERD. In particular, as noted above, detector output to both display types is zero. Second, if one allows for the possibility that observations in these experiments are mediated by subunits, we have just shown that at the subunit level the responses to the two display types are identical. Third, if one allows for the possibility that appears to be implicit in the claims by Kulikowski and his co-workers, namely, that responses to on-off gratings are made at the input channel level while responses to counterphase gratings are made at the subunit level, the ERD still makes no predictions about resolution limit ratios. The reason is that subunit output depends on the receptive field distance and receptive field shapes, whereas, obviously, responses at the input channel level depend only on receptive field shape. Thus changes in receptive field distance have an effect on the subunit response to counterphase gratings but not on the input channel response to on-off gratings. Hence there can be no *a priori* relationship between these two response. To put it simply, because the ERD calculates a temporal, not a spatial, cross correlation, spatial-frequency doubling cannot be explained by the ERD.

## B. Two-Frame Displays

Historically, apparent motion has been most studied with displays consisting of two successively flashed frames. Most work on random-dot motion employs this two-frame paradigm. Since it is widely assumed that, under proper spatiotemporal conditions, two-frame displays are processed primarily by simple, low-level motion detectors,<sup>34</sup> it is important to establish that the ERD applies to the psychophysics of two-frame displays. We first derive some general results for multiple-frame displays, then specialize these results to two-frame displays, and finally consider some particular two-frame displays.

A multiple-frame display  $c(x, t)$  has the general form

$$c(x, t) = \sum_{j=1}^N c_j(x) m_j(t). \quad (29)$$

Here,  $c_j(x)$  is the picture presented as the  $j$ th frame, and  $m_j(t)$  denotes the intensity time course of that frame. For example, the same frame may flash four times at intervals of 20 msec, or it may be illuminated continuously for 80 msec. In the first case,  $m_j(t)$  is a comb with four 20-msec teeth; in the second case,  $m_j(t)$  takes on a constant value during an 80-msec interval and is zero outside of this interval.

The Fourier transform of a multiframe display  $c(x, t)$  is given by the sum of the transforms of the  $N$ -component displays:

$$C(f, \omega) = \sum_{j=1}^N C_j(f) M_j(\omega). \quad (30)$$

Here,  $C_j$  and  $M_j$  are the Fourier transforms of  $c_j$  and  $m_j$ , respectively [Eq. (1a)]. We introduce the following term, which

is useful for understanding multiple-frame motion. Analogously to Eqs. (1) and (2), let

$$r_{H,j,c}(x) = \int c_j(x') r_H(x' - x) dx' \quad (31a)$$

$$= \int C_j(f) R_H^*(f) df. \quad (31b)$$

In words,  $r_{H,j,c}$  denotes the response magnitude of receptive field  $H$  at location  $x$  for frame  $c_j$ ; it is the spatial component of  $y_{H,0}$ . The second equality (31b) follows from the standard Parseval identities for Fourier transforms. We will omit the subscript  $c$  when this does not lead to misinterpretations.

To derive a general equation for ERD output  $y_4$  in response to multiframe displays, we substitute the particular form of input display [Eq. (31b)] into the defining equations for the ERD, Eqs. (3b) and (8a), simplify, and obtain

$$y_4(x) = \sum_{1 \leq j < k \leq N} [R_{\text{right},j}(x) R_{\text{left},k}(x) - R_{\text{right},k}(x) R_{\text{left},j}(x)] \times \left\{ \int D(\omega) [M_j^*(\omega) M_k(\omega) - M_k^*(\omega) M_j(\omega)] d\omega \right\}. \quad (32)$$

Because of the concern with an ERD's response as a function of its spatial location  $x$  within the display,  $y_4$  is here explicitly written as a function of  $x$ . The expression  $[R_{\text{right},j}(x) R_{\text{left},k}(x) - R_{\text{right},k}(x) R_{\text{left},j}(x)]$  occurs sufficiently often to deserve a symbol of its own, which will be  $\Delta_{jk}(x)$ , referring to the fact that the expression is a  $(2 \times 2)$  determinant.

What is astonishing about the ERD's response to multiframe displays [Eq. (32)] is the full separability of ERD output as a sum of products of spatial ( $x$ ) and temporal ( $\omega$ ) factors. The spatial factors, which depend only on the spatial characteristics of the display and the ERD parameters, are all grouped into one term, and the temporal factors are grouped into a second term. Separability has important implications for experimental tests of the ERD. For example, whenever the temporal parameters of different displays are held constant [i.e., all displays have the same intensity time functions  $m_j(t)$ ], differences in ERD responses to these displays depend only on differences in the factors  $\Delta_{jk}(x)$ . For two-frame displays, Eq. (32) reduces to

$$y_4(x) = \Delta_{12}(x) \int D(\omega) [M_1^*(\omega) M_2(\omega) - M_2^*(\omega) M_1(\omega)] d\omega. \quad (33)$$

In the remainder of this section, we derive two-frame predictions based on  $\Delta_{12}(x)$  in Eq. (33), and we examine new experimental evidence for them. We first consider spatial random bar patterns, then composite sine waves, and finally simple sine waves.

### 1. Elaborated Reichardt Detector (ERD) Computations and Psychophysical Observations of Two-Frame Random-Bar Patterns: Size-Specific ERD Channels Resolve the Correspondence Problem

A typical two-frame random-bar motion stimulus is illustrated in Fig. 9. Both frames may be regarded as extended stimuli that are viewed through a fixed, stationary window; the second frame is the same as the first frame, except for a rigid rightward displacement. The lower part of Fig. 9 depicts values of the term  $\Delta_{12}(x)$  as a function of  $x$ -axis location; the different



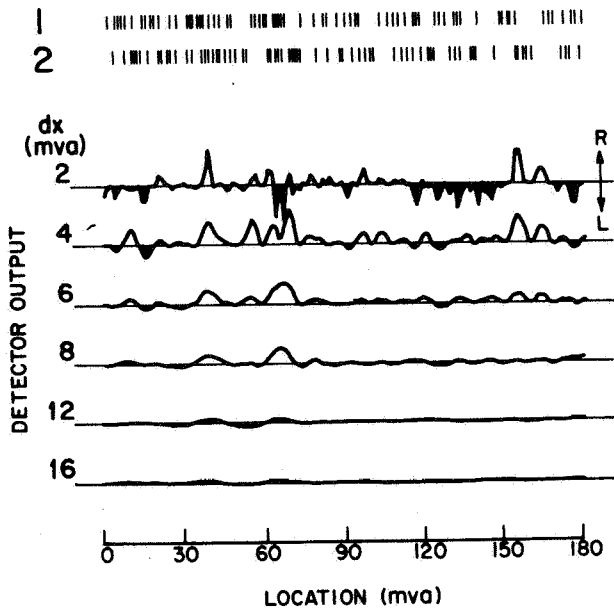


Fig. 9. The response to two-frame random bar motion of ERD's with the same spatial filters as in Fig. 6. A representative display (frames 1, 2) is shown at the top of the figure. Average distance between bars is 2.5 min of visual angle (mva). The between-frame displacement is 6 mva rightward. ERD responses are given as a function of their  $x$  location; up indicates rightward response (R), and down indicates leftward response (L). Zero-response reference lines are drawn for each detector size to highlight the sign—the indicated direction—of detector outputs. Incorrect direction responses (leftward) are indicated by dark shading. Different lines represent different sizes of proportionally scaled ERD's; the label indicates their size in terms of the distance  $\Delta x$  between their receptive field centers.

functions represent ERD's of different sizes. When the output of an ERD of a given size is below the reference line for that size, the ERD indicates leftward motion, and rightward motion otherwise. This computation uses detectors with the spatial receptive fields illustrated in Fig. 6(a) and described by Eqs. (24).

The results from the computation of two-frame bar motion in typical displays show the following. The smallest ERD's ( $\Delta x = 2$  mva) do not consistently indicate a particular direction of motion. Larger ERD's ( $\Delta x = 4, 6, 8$  mva) consistently indicate motion correctly. These output patterns are consistent with the perceptions of these two-frame random bar displays. For the display of Fig. 9, observers perceive a multitude of bars that jump in both directions, while at the same time one observing a patternless, rightward undercurrent.

With the assistance of David H. Parish in our laboratory at New York University, we made related theoretical and psychophysical observations using bandpass-filtered random-bar stimuli. (1) We verified that ERD's indeed are selectively activated by input stimuli whose passbands correspond to the sizes of their receptive fields; (2) the perceived ambiguity or nonambiguity of the bandpass-filtered two-frame displays corresponded to the computed ERD responses in the activated band. There was good agreement between the ERD responses and our perceptual observations. Others<sup>35,36</sup> have found similarly expected results with bandpass filtered motion stimuli.

In attempting to arrive at a simple algorithm for computing the motion in a two-frame random-bar display, it is not ob-

vious which particular bar in frame 2 should be taken to correspond to a given bar in frame 1, especially when there is some noise or deformation in the second frame. This is the correspondence problem. The two-frame random-bar simulation of the ERD shows that an analysis, based on calculating the term  $\Delta_{12}(x)$  resolves—or at least vastly simplifies—the correspondence problem for these patterns. It should be emphasized that the frame-to-frame displacement of six units is three times the average interbar distance, thus making it all but impossible to apply a nearest-neighbor algorithm.

A two-dimensional spatiotemporal Fourier analysis of the two-frame random bar display [Fig. 5(i)] clearly shows the characteristic spectrum of sampled rigid motion. These are the Fourier components detected by the ERD in its resolution of the direction of motion. But the ERD's do not need to make use of the fact that, in this display, motion consists of rigid displacement. As we shall see below, the ERD is indifferent to object rigidity insofar as computing its output is concerned.

## 2. Elaborated Reichardt Detector Computations and Psychophysical Observations of Double-Component Sine-Wave Patterns: Rigid versus Nonrigid Motion

Figures 10 and 11 illustrate displays that each consist of two sine-wave components with spatial frequencies of  $f$  and  $3f$ , where  $f = 0.75$  cycle/dva. Both figures illustrate the same location  $\times$  size computations for ERD responses to these displays as Fig. 9 showed for responses to random-bar patterns.

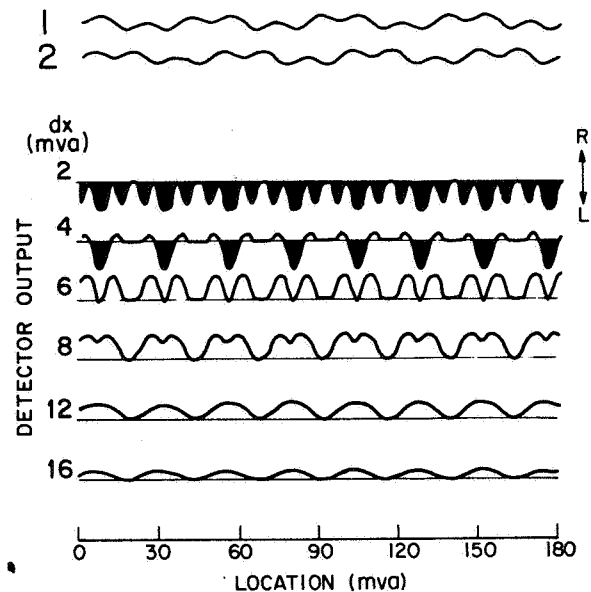


Fig. 10. Two-frame two-component sinusoidal stimuli in rigid motion. The two frames are shown at top; they consist of superimposed sine waves whose spatial frequencies have a ratio of 3:1. The displacement is  $\pi/4$  of the lower frequency, resulting in a rigid displacement of the entire pattern. Responses as a function of spatial location  $x$  and ERD size  $\Delta x$  (left-center to right-center distance of scaled receptive fields) are shown for ERD's with the same parameters as in Figs. 6 and 9. Despite the rigid displacement, small ERD's ( $\Delta x \leq 4$  mva) indicate leftward motion, and large ERD's ( $\Delta x \geq 6$  mva) indicate rightward motion, resulting in an ambiguous overall response.

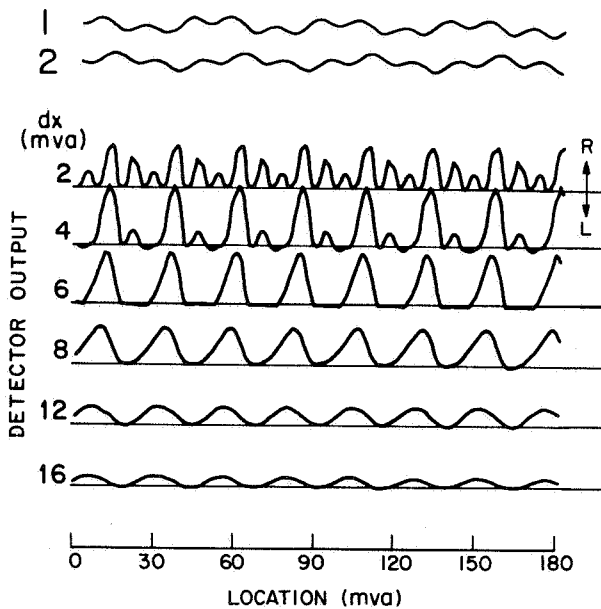


Fig. 11. Two-frame nonrigid motion. Same as Fig. 10, except that sine-wave components are each displaced by one quarter of their respective cycles. In this nonrigid displacement, the luminance pattern itself changes from a peaks-subtract to a peaks-add pattern. Nevertheless, ERD's of all sizes correctly indicate rightward motion.

In the displays of Fig. 10, both sinusoidal components are displaced, between frames, by the same amount of  $f\pi/2$  rad, i.e., by one quarter of the spatial cycle of the  $f$  component. This results in rigid motion. However, the ERD's give conflicting outputs. Small detectors tuned to the  $3f$  component indicate leftward motion; large detectors tuned to the  $f$  component indicate rightward motion. In Fig. 11, the same sine-wave components are used. However, this time the components are each displaced by one quarter of their respective spatial cycles, i.e., by  $f\pi/2$  and  $f\pi/6$  rad. This results in nonrigid motion. All detectors indicate the same direction of motion, even though the spatial pattern changes entirely from frame 1 to frame 2 (from a peaks-subtract pattern to a peaks-add pattern). Informal psychophysical observations showed that the rigid display is indeed perceived as ambiguous, with different parts of the display moving in different directions, while the nonrigid display appears to have a consistent direction of motion and, if anything, looks rigid.

The  $f + 3f$  demonstrations show that the ERD is not inherently sensitive to object rigidity. The predictive success of the random-bar and the  $f + 3f$  simulations make it seem as though the ERD decomposes a display into its spatial sine-wave components. In particular, computations of ERD responses to two-frame random bar patterns (organized into size-specific channels) suggest the feasibility of simple spatial frequency analysis as a solution to the correspondence problem. Even though in this case spatial analysis does seem to suffice, we have demonstrated previously,<sup>4</sup> and shall demonstrate again below, that spatial-frequency analysis, by itself, fails to predict motion responses; spatiotemporal analysis is required. Specifically, we<sup>4</sup> demonstrated displays in which the ERD correctly predicted the direction of perceived motion, and a detector that tracks sinusoidal components from frame to frame made incorrect predictions. The ERD does indeed perform a spatiotemporal-frequency analysis of a  $x, t$  stimulus

display but only when the temporal-frequency components are different or, if not, when the spatial-frequency components are sufficiently different to stimulate nonoverlapping ERD populations (see Table 2).

### 3. Elaborated Reichardt Detector (ERD) Computations and Psychophysical Observations of Triple-Component Sine-Wave Patterns: The ERD Does Not Perform Spatial Sine-Wave Analysis

Figures 12 and 13 show the results of computations for displays consisting of three sine-wave components, having spatial frequencies of  $2f$ ,  $3f$ , and  $4f$  and amplitudes of 0.5, 1.0, and 0.5. These three components can be combined in two ways. The first, in Fig. 11, results in an amplitude-modulated grating with a carrier grating of frequency  $3f$  and a modulation frequency  $f$ . The second, in Fig. 12, differs from an amplitude-modulated grating only in that the  $4f$  component is phase shifted by  $\pi$ , or, equivalently, is sign reversed. The resulting pattern is usually called a quasi-frequency-modulated grating. In both displays, the displacement is rigid and identical (slightly smaller than one half of the cycle of the central  $3f$  component). Yet the pattern of detector outputs is quite different for the two displays. The quasi-frequency-modulated grating gives rise to conflicting detector outputs in most sizes of channels; the amplitude-modulated grating does not. Again, these patterns of detector outputs are consistent with informal psychophysical observations.

The triple-component computations show that, even though the ERD is indifferent about the preservation of some local detail as exemplified in the rigid/nonrigid display simulations, it does not invariably perform a global spatiotemporal sine-wave analysis of the display (like the GFA). Phase relations between spatiotemporal-frequency components are critical, both for the ERD's and for perception, indicating that the component spatial frequencies in these displays are not perceived independently by ERD's or the visual system. (Of

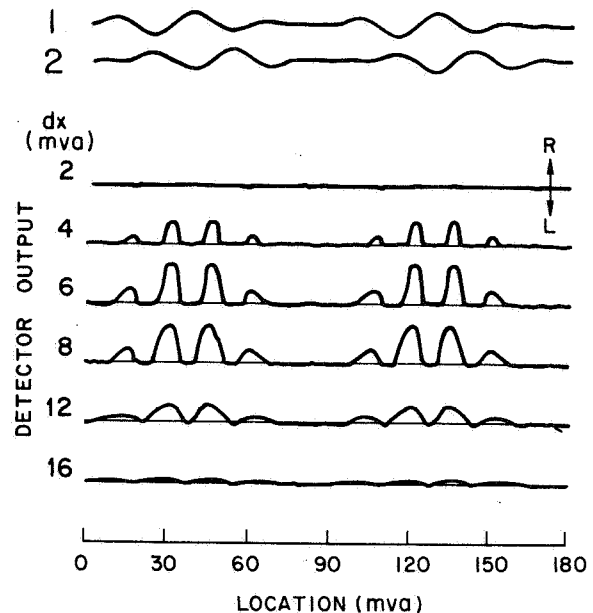


Fig. 12. Two-frame, three-component sinusoids in an amplitude-modulated grating displaced rigidly in the right. Parameters as in Figs. 6 and 9; ERD's of all sizes and at all locations have nonnegative responses, indicating unambiguous rightward motion.

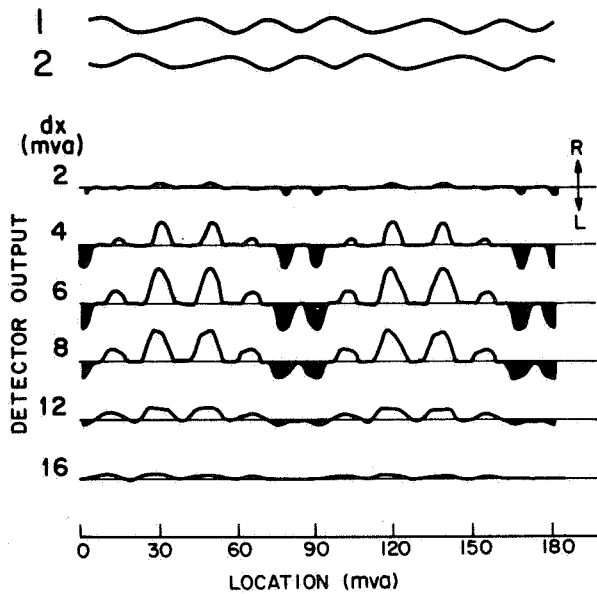


Fig. 13. Two-frame, three-component sinusoids in a quasi-frequency-modulated grating displaced rigidly to the right. Although this grating contains exactly the same spatial sine-wave components with the same respective amplitudes as in Fig. 12, ERD responses are quite different. Except for the largest size ( $\Delta x = 16$  mva), all detectors have outputs of different signs depending on location, indicating an ambiguous overall response.

course, the observed phase dependencies could be incorporated into the voting-combination rules of an elaborated GFA, but this would merely be an *ad hoc* translation of ERD time-domain computations to the frequency domain.)

**Conclusions for location  $\times$  size computations.** Taken together, these computations can be interpreted to mean that, in two-frame displays, the ERD's perform a form of primitive feature matching, where the features are defined in terms of whatever happens to match the spatial receptive fields. This matching procedure is sufficiently sophisticated to resolve the correspondence problem in the same way as the visual system for the classes of displays considered here.

4. Two-Frame Sine-Wave Displays

Two-frame motion of spatial sinusoids was illustrated in Fig. 5(j). Some new predictions can be derived analytically for these displays. Let  $c(x, t)$  be a two-frame sine-wave display, that is,

$$c(x, t) = [L_0 + A_1 \sin(2\pi fx - \varphi_1)]m_1(t) + [L_0 + A_2 \sin(2\pi fx - \varphi_2)]m_2(t). \quad (34)$$

It can be shown that the term  $\Delta_{12}(x)$  is given by

$$\Delta_{12}(x) = A_1 A_2 \sin(\varphi_1 - \varphi_2) = A_1 A_2 \sin(\varphi). \quad (35)$$

Here,  $\varphi$  is the spatial phase difference between the two frames. Equation (35) assumes balanced receptive fields; Appendix A derives predictions when the assumption is not made.

In Eq. (35), the dependence of ERD output on location  $x$  has disappeared; ERD output depends only on the product of the modulation amplitudes and the sine of the phase shift. This is an extremely powerful prediction, because it must hold no matter what shapes the receptive fields have or what temporal filters the ERD has.

Logically speaking, for rightward motion,  $0 < \varphi_1 - \varphi_2 < \pi$ , while for leftward motion,  $0 < \varphi_2 - \varphi_1 < \pi$ . That is, for two-frame sine-wave displays it makes sense to define the direction of motion as that corresponding to the shortest phase shift. Since the sign of ERD output indicates direction of motion, Eq. (35) shows that  $\Delta_{12}$  always indicates the correct direction.

**Equivalence for ERD's of  $\pi/2 \pm \vartheta$  in two-frame sinusoids.** Equation (35) implies several powerful, testable predictions. First, ERD output is identical for any two values of  $\varphi$  that are symmetric around  $\pi/2$ , such as  $4/15\pi$  (48 deg) and  $11/15\pi$  (132 deg), and is maximal when  $\varphi = \pi/2$ . Again, this is true for any receptive fields. In particular, it is not necessary for this prediction that spatial receptive fields be shifted by  $\pi/2$ , as in subunit Variant 2. Thus all ERD's of all sizes have precisely the same response to these widely different displays.

A strong implication of Eq. (36) is that, provided their phases are symmetric around  $\pi/2$ , no ERD can distinguish between two two-frame motion displays that differ in velocity even by an order of magnitude. This prediction does not carry over to visual perception. For example, the 48- and 132-deg spatial-phase-shift displays described above are perceptually quite different, with the large phase shift giving rise to a clearly visible change and the small phase shift giving the impression of hardly any change at all. Thus velocity discrimination in this experiment cannot be based entirely on temporally integrated outputs from an ERD but must involve other computations, possibly in other channels or possibly in direct access to output of a kind of subunit that differs from ERD outputs in these two-frame displays. (When subunit filters are the special  $\pi/2$  filters, then subunits are equivalent or almost so to ERD's; when filters are not so chosen, subunit output is generally not very meaningful. Velocity discrimination is inherently difficult for ERD's.)

In formal psychophysical tests of the symmetry prediction with respect to direction (not velocity) discrimination, four

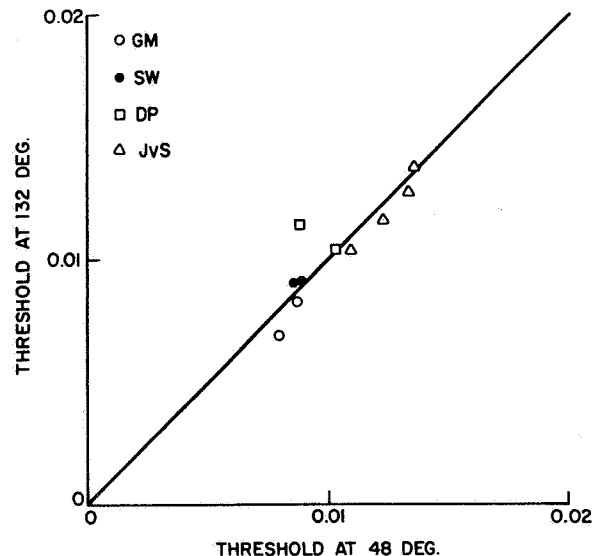


Fig. 14. Data from a two-frame direction-discrimination experiment with sinusoidal grating stimuli. The coordinates represent contrast thresholds for 71% correct direction discrimination with frame-to-frame displacements of 48 and 132 deg. The straight line represents exactly equal thresholds. Data for four subjects are shown; each point represents  $2 \times 500$  trials in an interleaved-staircases procedure.

subjects viewed displays  $1 \times 2$  deg in viewing angle. Two frames were exposed consecutively, each for 80 msec. Threshold contrasts for left-right motion discrimination were determined by interleaved up-down staircases. Figure 14 shows typical results for one experiment with spatial sinusoids, 1.5 cycle per degree of which were shifted either 48 or 132 deg between frames. Statistically equal contrast thresholds for motion-direction discrimination were observed with spatial phase shifts between frames of  $\pi/2 - \varphi$  and  $\pi/2 + \varphi$ , as predicted.<sup>4</sup> Even though the 48- and 132-deg displays appear to move slowly and rapidly, respectively, motion-direction discrimination is the same for both. The same symmetry in motion-direction discrimination was observed with other spatial frequencies and phase shifts and with Gaussian as well as square spatially windowed presentations. Similar symmetry between  $\pi/2 - \varphi$  and  $\pi/2 + \varphi$  was reported by Nakayama and Silverman,<sup>37</sup> who also present compelling data to support their contention that (1) motion-direction discrimination and (2) the detection of moving stimuli (by their subjects) involve radically different computations. The identical response of ERD's to two-frame sinusoids with different velocities suggest a third computation that is yet to be discovered.

*Three more two-frame sinusoid predictions.* Second, with the appropriate linking assumptions<sup>16</sup> it follows that contrast threshold should depend on the square root of the reciprocal of  $\sin(\varphi)$ . Here, contrast threshold is defined as the value of  $A$  ( $A_1 = A_2$ ) at which criterion performance is reached. The difficulty in testing this prediction is that the response of the human motion discrimination system to contrast amplitude saturates when contrast reaches only 2%.<sup>37</sup> This extreme saturation of responses to contrast makes testing quantitative predictions based on different contrast responses extremely problematical.

Third, it should not make a difference whether the amplitude of the first frame exceeds that of the second or vice versa, as long as the product of the amplitudes remains constant. This prediction carries over into psychophysical predictions without special linking assumptions and is borne out by our psychophysical observations.

Fourth, more generally, when  $\varphi$  is fixed, and  $A_1$  and  $A_2$  are covaryed such that their product remains constant, ERD output remains constant. Again, with appropriate linking assumptions,<sup>16</sup> it follows that this constant-product rule should also be found psychophysically. Again, the nonlinearity of the motion system's contrast response makes testing this prediction problematical.

### C. Rigid, Continuous Motion

There is a paradox in the ERD vis à vis object rigidity. The previous section presents a situation in which an array of ERD's detects direction of motion better in a nonrigid than in a rigid but otherwise identical display. However, for rigid, continuous translations, subunits are well behaved in that less-restrictive assumptions about the filters have to be made to achieve equivalence between subunits and the ERD (Subsections 3.B and 3.C). Specifically, for rigid displays the  $\pi/2$  spatial phase-delay subunit is always equivalent to a full ERD.

Let  $c(x, t)$  be a rigid display, i.e.,

$$c(x, t) = \gamma(x - vt). \quad (36)$$

Here,  $v$  is velocity and  $\gamma(x)$  is the rigidly displacing spatial pattern. The Fourier transform is given by

$$C(f, \omega) = \Gamma(f)\delta(f + \omega/v), \quad (37)$$

where  $\Gamma(f)$  is the Fourier transform of  $\gamma(x)$ . It can be shown that

$$\begin{aligned} y_{\text{left},3} &= -i \int \Gamma^*(f)\sin(f)\Gamma(f)R_{\text{right}}(f)R_{\text{right}}^*(f)D(-fv)df \\ &= -y_{\text{right},3}. \end{aligned} \quad (38)$$

The same result can be shown for the larger class of displays whose Fourier transforms are of the form

$$C(f, \omega) = \Gamma(f)\delta[f + H(\omega)], \quad (39)$$

where  $H(\omega)$  is the transform of an arbitrary movement function. For rigid displays,  $H(\omega) = \omega/v$ . This class includes objects that undergo smooth shape transformations during motion.

Equations (38) and (39) indicate that  $\pi/2$  spatial-phase-shift subunits of the ERD are equivalent to the ERD with respect to rigid displacements of arbitrary spatial patterns and suggest that the full ERD, which tends to cancel out defects of its subunits, would be well behaved even if the spatial and temporal filters did not exactly satisfy the  $\pi/2$  temporal and spatial phase conditions of the subunits [Variants 1 and 2, Eqs. (2.1) and (2.2)].

## 6. CONCLUSIONS

This paper provides new results on some variants of the elaborated Reichardt model. We have taken as our focal model a slightly more general version than that of van Santen and Sperling<sup>4</sup> in which infinite time averaging is replaced by arbitrary temporal filtering. For certain applications, we introduced one or more of the following auxiliary assumptions; finite temporal integration—a weaker assumption than infinite time averaging—net balancedness of excitatory and inhibitory areas of receptive fields, and separability of spatial and temporal input filters. Although the linear components of a model for one-dimensional motion perception in  $x, t$  are remarkably similar to the corresponding operations for pattern perception in  $x, y$ , the physical detectors are essentially different because a motion detector at location  $x$  must be concerned with events at all times  $t$ , whereas a pattern detector's interests are localized in both  $x$  and  $y$ .

### A. Summary of Results

The following theoretical results concerning the elaborated Reichardt detector's response to one-dimensional movement in  $x, t$  were derived:

1. Under the assumption of temporal integration, subunit output is equivalent to detector output (1) for all displays when the temporal filter delays all temporal frequencies by  $\pi/2$  or (2) for rigid object displays with arbitrary motion trajectory when the spatial filters have a constant spatial phase difference of  $\pi/2$ .

2. The W-A detector,<sup>5</sup> when elaborated into a Reichardt-type opponent scheme, is equivalent to a special case of the ERD.

3. For any choice of spatial and temporal filters, output

from the substantially more complicated Adelson-Bergen (A-B) detector<sup>24</sup> is equivalent to output from the ERD. At the subunit level, provided that both the temporal delay and spatial receptive field filters have the  $\pi/2$  property, A-B outputs differ from ERD outputs by an additive constant,  $K$ . The presence of  $K$  in the A-B variant guarantees nonnegativity of subunit output, while the absence of  $K$  in the ERD guarantees that the sign of subunit output indicates whether motion is to the left or to the right.

4. A wide variety of temporal and spatial filters suffices to guarantee that the temporally integrated ERD output will indicate the correct direction of motion for a drifting grating of any spatial or temporal frequency (nonaliasing).

5. For two-frame displays, temporal and spatial factors are completely separable in the ERD output. Holding temporal display parameters constant, ERD outputs depend simply on  $\Delta_{12}(x)$ , the  $2 \times 2$  determinant formed from the magnitudes of spatial responses the two input channels to each of the frames, at each spatial location  $x$ . Analysis of  $\Delta_{12}(x)$  for two-frame displays and comparing this with visual responses shows that

(a) For two-frame random bar patterns, the ERD solves the correspondence problem similarly to the way the visual system does. The ERD performs neither individual bar feature matching nor spatial sine-wave component matching. Rather, it matches local patterns that have shapes similar to its input receptive fields.

(b) For two-frame double-sinusoid motion, the ERD has a larger output when each sinusoidal component moves separately, nonrigidly, for its own optimal distance than when the components move together rigidly.

(c) For two-frame, triple-sinusoid rigid motion, the phase relations between the sinusoids are critical in determining ERD response (and visual) responses, indicating the essential nonlinearity of this motion computation.

(d) For two-frame sine-wave displays, output ERD's with balanced receptive fields are independent of spatial location  $x$  and given by the product of the frame contrast amplitudes and the sine of the displacement angle. Equal contrast thresholds were predicted for detection of a slow movement (displacement =  $\pi/2 - \varphi$ ) and the symmetric fast movement (displacement =  $\pi/2 + \varphi$ ),  $0 < \varphi < \pi/2$ . Similarly, when unequal contrast amplitudes were assigned to the first and second frames, equal thresholds were predicted independent of the assignment. Both predictions were psychophysically confirmed.

### B. Some Remaining Issues

First, the ERD does not explain certain phenomena that might appear to be within its scope, notably the reported doubling of threshold contrast in counterphase as compared with drifting gratings.<sup>38,39</sup> These are phenomena that involve detection of motion stimuli, not discrimination of their direction; the relation between detection and discrimination mechanisms still needs to be clarified. Second, an ERD does not compute the velocity of a display; velocity discrimination requires other information than is computable even from arrays of ERD outputs. Third, Nakayama and Silverman<sup>37</sup> observe a compressive nonlinearity in the input channels to motion detectors that complicates the mathematics and testing of ERD's and other theories. Fourth, the connection of the ERD with recordings made from direction-selective

single cells has not been treated here because it is quite unclear. Finally, the ERD has been elaborated only for tasks in which observers were required to report a single direction of motion over a temporal interval. To deal with time-varying changes in motion during the observation interval requires further elaboration. In spite of its limitations, our results establish the ERD is at once the simplest and the most powerful of the mechanisms proposed for the human discrimination of direction of motion.

### APPENDIX A: TWO-FRAME SINE-WAVE RESPONSE OF ELABORATED REICHARDT DETECTORS WITH UNBALANCED RECEPTIVE FIELDS

This appendix outlines the implications of unbalanced receptive fields on the response of ERD's to two-frame sine-wave displays. An unbalanced receptive field is one that has a nonzero output [Eq. (1)] in response to spatially uniform illumination. A two-frame sine-wave display is defined by Eq. (34):

$$c(x, t) = [L_0 + A_1 \sin(2\pi fx - \varphi_1)]m_1(t) + [L_0 + A_2 \sin(2\pi fx - \varphi_2)]m_2(t), \quad (A1)$$

We simplify the discussion by assuming that  $A_1 = A_2 = A$ . Also, we assume that the left and right receptive fields are equally unbalanced:  $U = \int r_{\text{left}}(x')dx' = \int r_{\text{right}}(x')dx'$ . Without loss of generality, we let  $\varphi_1 = 0$  and  $\varphi_2 = \varphi$ . Let  $x$  denote ERD location. It can be shown that Eq. (35) now becomes

$$\Delta_{12}(x) = \alpha A^2 \sin(\varphi) + \beta L_0 U A (1 - \cos \varphi)^{1/2} \cos(x - \gamma). \quad (A2)$$

Here,  $\alpha$ ,  $\beta$ , and  $\gamma$  are factors that depend on spatial frequency  $f$  and on receptive field parameters but not on  $x$  or  $\varphi$ .

Equation (A2) has the following consequences. First, the value of  $\Delta_{12}(x)$  varies with location,  $x$ .  $\Delta_{12}(x)$  varies sinusoidally with  $x$  around a mean value of  $\alpha A^2 \sin(\varphi)$ , which is the (location-independent) response for ERD's with balanced receptive fields [Eq. (35)].

Second, the effect of the spatial variation of ERD output on psychophysical responses depends on how responses are combined spatially. Specifically, the contribution from the spatially varying component may vanish if the combination rule weighs peaks and troughs equally. This happens, for example, if responses are integrated with respect to  $x$ . On the other hand, if there is no exact cancellation of peaks and troughs, then the spatial variation does have an influence. An example is given by the maximum rule, in which case the combined response becomes

$$\Delta_{12} = \alpha A^2 \sin \varphi + \beta L_0 U A (1 - \cos \varphi)^{1/2}. \quad (A3)$$

Third, if the spatially varying component had an effect, the effect would be (1) to introduce an asymmetry in the dependency on  $\varphi$  and (2) to complicate the relation between  $A$  and  $\varphi$ . As was mentioned in Section 5, Nakayama and Silverman<sup>37</sup> and the present authors have no systematic departures from symmetry of psychophysical responses around  $\pi/2 \pm \varphi$ . However, in these same experiments, significant departures were found from the rule that contrast threshold should depend on the square root of the reciprocal of  $\sin(\varphi)$ ; the ob-

served relationship was somewhere between square root and linear. The spatially varying component of the response cannot account for failure of the inverse-square-root prediction since a large spatially varying component would be expected to cause a significant symmetry between  $\pi/2 \pm \varphi$ .

The probable cause of the failure of the inverse-square-root prediction is that motion receptors may have a nonlinear contrast response. Is a nonlinear contrast response compatible with an ERD response that is symmetric in  $\varphi$  around  $\pi/2$ ? It can be shown that when the intensity transduction nonlinearity is antisymmetric—transduces positive and negative inputs with equal gains but opposite signs—and when ERD responses are spatially integrated by the voting rule of the model, then the ERD response is symmetric in  $\varphi$ . Thus, one can construct an ERD system that has precisely the observed relationship between  $\varphi$  and  $A$ .

### ACKNOWLEDGMENTS

This research was supported by U.S. Air Force Office of Scientific Research grant 80-0279. The authors wish to acknowledge useful discussions with E. Adelson, J. Bergen, and M. S. Landy.

### REFERENCES

1. J. P. H. van Santen and G. Sperling, "A new class of models of motion-sensitive units in human vision," presented at the 16th annual Meeting of the Society for Mathematical Psychology, Boulder, Colorado, 1983.
2. J. P. H. van Santen and G. Sperling, "A temporal covariance model of motion perception," *Invest. Ophthalmol. Vis. Sci. Suppl.* 24, 277 (1983).
3. J. P. H. van Santen and G. Sperling, "Applications of a Reichardt-type model of two-frame motion," *Invest. Ophthalmol. Vis. Sci. Suppl.* 25, 14 (1984).
4. J. P. H. van Santen and G. Sperling, "A temporal covariance model of motion perception," *J. Opt. Soc. Am. A* 1, 451-473 (1984).
5. A. B. Watson and A. J. Ahumada, "A look at motion in the frequency domain," NASA Tech. Memo. 84352 (1983).
6. E. H. Adelson and J. Bergen, *J. Opt. Soc. Am. A* 2, 284-299 (1985).
7. W. Reichardt, "Autokorrelationsauswertung als Funktionsprinzip des Zentralnervensystems," *Z. Naturforsch.* 12b, 447-457 (1957).
8. J. J. Kulikowski, "Pattern and movement detection in man and rabbit; separation and comparison of occipital potentials," *Vision Res.* 18, 183-189 (1978).
9. J. J. Kulikowski, "Spatial resolution for the detection of pattern and movement (real and apparent)," *Vision Res.* 18, 237-238 (1978).
10. I. Murray, F. McCana, and J. J. Kulikowski, "Contribution to two movement detecting mechanisms to central and peripheral vision," *Vision Res.* 23, 151-159 (1983).
11. G. G. Furman, "Comparison of models for subtractive and shunting lateral-inhibition in receptor-neuron fields," *Kybernetik* 2, 257-274 (1965).
12. G. Sperling and M. M. Sondhi, "Model for visual luminance discrimination and flicker detection," *J. Opt. Soc. Am.* 58, 1133-1145 (1968).
13. J. Thorson, "Small-signal analysis of a visual reflex in the locust," II. Frequency dependence," *Kybernetik* 3, 53-66 (1966).
14. V. Torre and T. Poggio, "A synaptic mechanism possibly underlying direction selectivity to motion," *Proc. Royal Soc. London Ser. B* 202, 409-416 (1978).
15. G. Fermi and W. Reichardt, "Optomotorische Reaktionen der Fliege," *Musca Domestica Kybernetik* 2, 15-28 (1963).
16. See Ref. 4, p. 456, Note 17.
17. D. Gabor, "Theory of communication," *J. Inst. Electr. Eng.* 93, 429-457 (1946).
18. S. Marcelja, "Mathematical description of the responses of simple cortical cells," *J. Opt. Soc. Am.* 70, 1297-1300 (1980).
19. D. A. Pollen and S. F. Ronner, "Phase relationships between adjacent simple cells in the visual cortex," *Science* 212, 1409-1411 (1981).
20. H. R. Wilson, "Spatiotemporal characterization of a transient mechanism in the human visual system," *Vision Res.* 20, 443-452 (1980).
21. P. Burt and G. Sperling, "Time, distance, and feature trade-offs in visual apparent motion," *Psychol. Rev.* 88, 171-195 (1981).
22. A. B. Watson and A. J. Ahumada, "A model of how humans sense image motion," *Invest. Ophthalmol. Vis. Sci.* 25, 14 (1984).
23. S. A. Rajalla, A. N. Riddle, and W. E. Snyder, "Application of the one-dimensional Fourier transform for tracking moving objects in noise environments," *Comput. Vision Graphics Image Process.* 21, 280-293 (1983).
24. E. H. Adelson and J. Bergen, "Motion channels based on spatio-temporal energy," *Invest. Ophthalmol. Vis. Sci. Suppl.* 25, 14 (A) (1984).
25. H. S. F. von Helmholtz, *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik* (Dover, New York, 1863); (reprint, 1954).
26. H. E. Ives, "A visual acuity test object," *Electron. World* 55, 939-940 (1910).
27. S. Anstis and C. W. Tyler, "Induced tilt from checkerboards: edged vs. Fourier components," *Invest. Ophthalmol. Vis. Sci. Suppl.* 21, 165 (1980).
28. G. Sperling and J. P. H. van Santen, "Models of motion perception," presented at the U.S. Air Force Office of Scientific Research Review of Sponsored Basic Research in Human Information Processing, Sarasota, Florida, 1982.
29. G. Sperling, "Theories of motion perception," presented at the 17th Annual Meeting of the Society for Mathematical Psychology, Chicago, Illinois, 1984.
30. N. Graham, "Spatial-frequency channels in the human visual system: detecting edges without edge detectors," in *Visual Coding and Adaptability*, C. S. Harris, ed. (Erlbaum/Halstead, Potomac, Md., 1980).
31. A. B. Watson, "Detection and recognition of simple spatial forms," in *Physical and Biological Processing of Images*, O. J. Braddick and A. C. Sleight, eds. (Springer-Verlag, Berlin, 1982), pp. 100-114.
32. H. R. Wilson, "Psychophysical evidence for spatial channels," in *Physical and Biological Processing of Images*, O. J. Braddick and A. C. Sleight, eds. (Springer-Verlag, Berlin, 1982), pp. 88-99.
33. G. Sperling, "Linear theory and the psychophysics of flicker," *Doc. Ophthalmol.* 18, 3-15 (1964).
34. O. Braddick, "A short-range process in apparent motion," *Vision Res.* 14, 519-529 (1974).
35. J. J. Chang and B. Julesz, "Displacement limits for spatial frequency filtered random-dot cinematograms in apparent motion," *Vision Res.* 23, 1379-1385 (1983).
36. K. Nakayama and G. H. Silverman, "Temporal and spatial properties of the upper displacement limit in random dots," *Vision Res.* 24, 293-300 (1984).
37. K. Nakayama and G. H. Silverman, "Detection and discrimination of sinusoidal grating displacements," *J. Opt. Soc. Am. A* 2, 267-274 (1985).
38. E. Levinson and R. Sekuler, "The independence of channels in human vision selective for direction of movement," *J. Physiol.* 250, 347-366 (1975).
39. A. B. Watson, P. G. Thompson, B. J. Murphy, and J. Nachmias, "Summation and discrimination of gratings moving in opposite directions," *Vision Res.* 20, 341-348 (1980).

**Jan P. H. van Santen**

Jan P. H. van Santen received the B.A. and M.A. degrees in psychology from the University of Amsterdam in 1972 and 1974 and the Ph.D. degree in mathematical psychology from the University of Michigan in 1979. He is a member of technical staff at the Computer Division of AT&T Bell Laboratories. Formerly, he held a postdoctoral fellowship at the Acoustical and Behavioral Research Center at Bell Laboratories and was an associate research scientist at the Human Information Processing Laboratory of the Department of Psychology, New York University, where the research for this paper was conducted. His current research interests include statistics, artificial intelligence, and visual perception.

**George Sperling**

George Sperling received the B.S. degree in mathematics from the University of Michigan, Ann Arbor in 1955, the M.A. degree from Columbia University in 1956, and the Ph.D. degree from Harvard University in 1959, both in experimental psychology. He is director of the Human Information Processing Laboratory at New York University, where since 1970 he has been Professor of Psychology. Since 1959, he has also been a member of the technical staff of the

Acoustical and Behavioral Research Center at AT&T Bell Laboratories, Murray Hill, New Jersey. Along with the above positions, he has been a visiting professor at New York University (1962-1963), Duke University (1964), Columbia University (1964-1965), the University of California at Los Angeles (1967-1968), University College, London (1969-1970), the University of Western Australia (1972), and the University of Washington (1977), and a visiting scholar at Stanford University (1984). He is the author of about 100 papers that deal mostly with mathematical and experimental analyses of human perception and information processing and with computer image processing. Professor Sperling was awarded the Gomberg Chemistry Prize and was elected to Phi Beta Kappa, Phi Eta Sigma, Phi Kappa Phi, and Sigma Xi while at the University of Michigan. He has been elected a Fellow of the American Psychological Association, the Optical Society of America, and the American Association for the Advancement of Science. He has been elected to the Society of Experimental Psychologists and to the Executive Boards of the Eastern Psychological Association, the Society for Computing in Psychology, and the Society for Mathematical Psychology, which he also served as chairman. He is a member of the Psychonomic Society and founder of the Annual Interdisciplinary Conference. In 1969 he was awarded a Guggenheim Fellowship. His current research deals with computational methods and models for human perception and information processing.