

NN 127

Silicon Retina with Correlation-Based, Velocity-Tuned Pixels[†]

Tobias Delbrück
139-74 Caltech
Pasadena CA 91125
818-356-6986
tdelbruck@caltech.edu

Abstract—I report the first functional two-dimensional silicon retina that computes a complete set of local direction-selective outputs. The chip motion computation uses unidirectional delay lines as tuned filters for moving edges. Photoreceptors detect local changes in image intensity, and the outputs from these photoreceptors are coupled into the delay line, where they propagate with a particular speed in one direction. If the velocity of the moving edges matches that of the delay line, then the signal on the delay line is reinforced. The output of each pixel is the power in the delay line signal, computed within each pixel. This power computation provides the essential nonlinearity for velocity-selectivity. The delay line architecture differs from the usual pairwise correlation models in that motion information is aggregated over an extended spatiotemporal range. As a result, the detectors are sensitive to motion over a wide range of spatial frequencies.

[†] A brief abstract reporting preliminary results appeared as T. Delbrück, “A silicon network for motion discrimination that uses spatiotemporal interpolation,” *Soc. of Neuroscience Abstracts*, no. 143.8, p. 344, 1991.

The author has just graduated from the Computation and Neural Systems Program at the California Institute of Technology, Pasadena CA 91125.

I have designed and tested functional one- and two-dimensional silicon retinas with direction-selective, velocity-tuned pixels. A velocity-selective detector requires only a single delay element and nonlinearity for each tuned velocity, and is sensitive to both light and dark contrasts. The use of adaptive photoreceptors and compact circuits makes for a well-conditioned input signal and small circuit offsets, resulting in robust operation. All circuits work in subthreshold, resulting in low power consumption. Pixels with three hexagonal directions of motion selectivity are approximately $(225 \mu\text{m})^2$ area in a 2- μm CMOS technology, and consume less than 5 μW of power.

I. ANALOG HARDWARE MOTION COMPUTATION

In this introduction, I place the new chip in the context of previous approaches to analog-hardware visual motion computation. Essentially, there are two approaches: Those that use spatiotemporal image *gradients*, and those that use spatiotemporal image *correlation*. Hardware systems that utilize each approach have been previously built. The chip reported here falls into the second category.

Tanner and Mead built the first analog two-dimensional optical-flow chip [35][36][30]. The Tanner chip computes a *single* motion vector for the *global* two-dimensional optical flow, using a gradient-based scheme (for references, see [28]). The motion vector is computed cooperatively by a network of circuits, each circuit using local spatiotemporal derivative information. The circuits reach a consensus on the global image motion vector, using the overall constraint that the total derivative of the image intensity is zero. This constraint allows the chip to solve the aperture problem, which arises because a *local* measurement cannot detect motion *along* an edge [28].

The two-dimensional direction-selective chip built by Benson and Delbrück [5] performs a different and much simpler computation that is based on biological studies. The pixel outputs respond selectively to local rightward motion, and the motion computation is feedforward, rather than cooperative. The architecture is based on the idea of null-inhibition direction selectivity, developed by Barlow and Levick [4] to explain the direction selective responses found in rabbit retinas. In a null-inhibition model, the pixel output is inhibited by edges coming from the null-direction. In the preferred direction, the pixel output in response to an edge is not inhibited, because the inhibition arrives after the excitation.

I distinguish the *value*-encoded global velocity-vector output of the Tanner chip, from the *place*-encoded, direction-selective, velocity-tuned outputs of the Benson–Delbrück chip. The computations are qualitatively different, because one encodes the pattern velocity as a vector with magnitude and direction, and the other encodes the motion of features by the location of active outputs. The Tanner chip does a

higher-level computation, integrating spatiotemporal image information from the entire moving image in a cooperative calculation, whereas the Benson–Delbrück chip simply responds selectively to rightward versus leftward motion.

The operation of the Tanner chip is not robust, perhaps because of the demands of the mathematically sophisticated motion algorithm it implements. High-contrast images, with sharp edges, are required if the chip is to produce a direction-selective output, and the precision of the velocity measurement is poor. The reason for this nonideal performance is the combination of inherent susceptibility of gradient-based schemes to temporal and spatial noise, low-sensitivity receptors, and large circuits with poor offset characteristics. The strongest signals are derived from places in the image with sharp gradients and high contrasts, but gradient-based schemes require images with smoothly varying contrast if they are to obtain reliable estimates of spatial and temporal derivatives. The performance of the chip is compromised severely by these conflicting requirements.

The operation of the Benson–Delbrück chip is robust in comparison with that of the Tanner chip, because the Benson–Delbrück chip, unlike the Tanner chip, responds correctly to real input scenes of moderate contrast. There are several reasons for this performance difference. The algorithm is much simpler, the circuits are more compact and hence less offset prone, and a well-conditioned input is generated by the use of adaptive photoreceptors. However, this comparison is arguably pointless, given the major differences in computation performed by the two chips.

The null-inhibition architecture has the advantage of a large degree of direction-selectivity and hence, inherently robust operation. The disadvantage is that it is difficult to produce a response-versus-speed tuning that is other than a simple lowpass determined by the pixel spacing, so construction of a set of detectors tuned to different speeds is not straightforward. Moreover, this architecture does not deal with the aperture problem, and, in fact, knows nothing of global motion constraints.

The chip described here has place-encoded outputs that encode information about components of local edge motion along the detector directions, like the Benson–Delbrück chip. It is also like the Benson–Delbrück chip in that the motion computation does not deal with solving the aperture problem. It is like the Tanner chip in that the computation integrates information over an extended spatiotemporal region. Our chip is the first functional analog device that computes a full set of local direction-selective outputs in two dimensions. It is also the first reported chip that uses correlation-based *analog* computation to produce direction-selective, velocity-tuned outputs. (Reference [22] reports a correlation-based one-dimensional motion circuit that correlates digital pulses, produced by image intensity changes, in analog time.)

I omit discussion of other analog motion chips that have been fabricated only as one-dimensional circuits [22][23]. For the most part these architectures cannot be fabricated practically in a two-dimensional architecture, because they are not parsimonious with wire. I also omit discussion of analog change-detecting circuits [8][13], digital motion circuits [2], and discrete-component correlation-detector systems [16][32].

The remainder of this paper is organized as follows. Section II is a heuristic description of correlation motion detectors, and of the delay line motion architecture in one and two dimensions. Section III describes a mathematical model of the circuit operation in the time and frequency domains, and investigates the model to understand the limiting behavior, and the differences between the new model and previous pairwise correlation detectors. Section IV is a description of the circuits used in the chip. Section V is about qualitative and semiquantitative measurements on the two-dimensional motion chip, and quantitative measurement on the one-dimensional chip. Section VI and Section VII conclude the paper with a short discussion of future directions of this work and a summary of the results.

II. CORRELATION-BASED MOTION DETECTORS

The simplest correlation-based motion detector is shown in Fig. 1(a). A pair of spatially separated, feature-detecting cells feeds into a nonlinearity that detects coincidence of the two inputs. One of the inputs to the nonlinearity is delayed. When the motion is matched to the detector spatial separation and delay, the output becomes large. The nonlinearity, although often formulated as a multiplicative operation, can be any operation that detects *coincidence* of the inputs. Hence, the correlation detector may linearly combine the spatiotemporally separated inputs, and then use a threshold element, or any other expansive (faster-than-linear) nonlinearity.

The correlation detector was conceived by Hassenstein and Reichardt during experiments on a beetle visual system [18]. Their model, and later modifications worked out through experiments on the fly visual system, are more complicated than the model in Fig. 1(a) [33]. The major elaboration is the addition of a complementary direction selectivity. The output from the Hassenstein-Reichardt model is the difference between two outputs tuned to complementary directions. This differencing operation removes the common-mode signal.

A. Extension of the Simple Correlation Detector

My motion architecture extends the pairwise correlation model in Fig. 1(a), by using a unidirectional *delay line* as a tuned filter for a particular velocity, as shown in Fig. 1(b). The input to the delay line comes

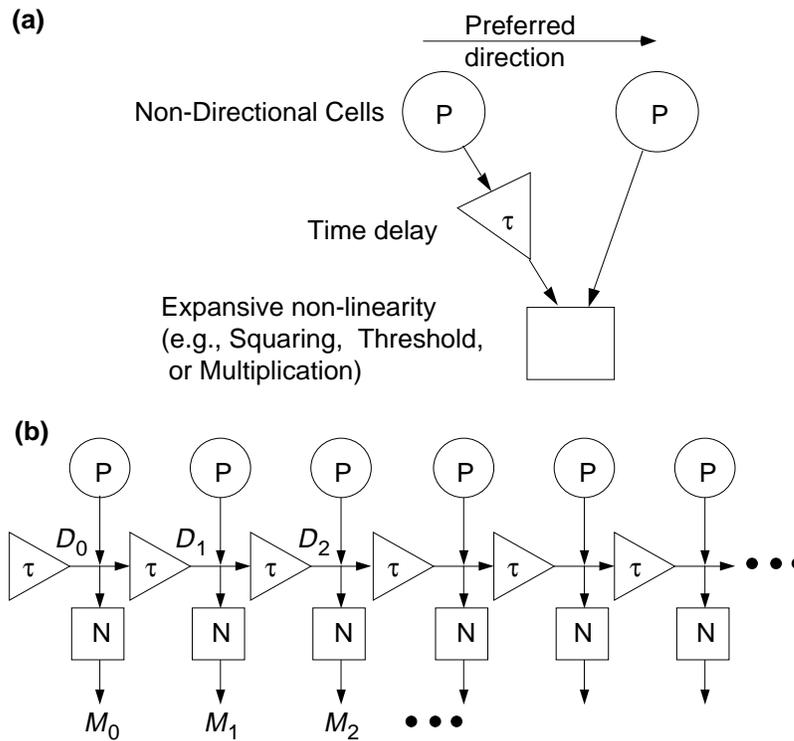


Fig. 1. Correlation-based motion-detectors. **(a)** Pairwise correlation detector. Nondirectional feature-detecting cells feed into an output nonlinear element that measures coincidence between the spatiotemporally separated inputs. The exact nature of the nonlinearity is not important, as long as it is expansive. **(b)** Extended correlation detector. Nondirectional cells (P) couple activity into a unidirectional delay line with delay τ per stage. The delay line signal (D_n) is large when the image motion is matched to the delay line speed. The expansive output nonlinearities (N) compute a velocity selective signal M_n that is monotonic in the magnitude of the delay line signal.

from photoreceptors that respond to local intensity change. The photoreceptor outputs are coupled capacitively into the delay line. The delay line is composed of low-pass filters. An edge passing over a photoreceptor creates a traveling signal in the delay line that spreads and decays with time. If the velocity of the edge is matched to that of the delay line, the successive inputs to the delay line pile up in synchrony. If the input velocity is not matched to the delay line velocity, the successive inputs arrive in asynchrony and do not pile up.

It is well known that the time-averaged output of a linear system cannot be direction or speed dependent. In a linear system, we can alter the time order of the input without changing the time-averaged output,

because the response to a sum of inputs is the same as the sum of the responses to each input individually. Hence, the *average* signal on our delay line is not velocity selective. The *amplitude* of the signal clearly is velocity selective. Measuring amplitude is a nonlinear operation. We can use any nonlinearity that measures some metric of the amplitude of the delay line signal, and the result will be direction and speed selective. On my chip, I use a compact circuit that computes an approximate squaring function of the delay line signal. Hence, the nonlinearity approximately measures the filter output power.

The nonlinear operation is an even function of its input, so the delay line activities caused by edges of light and dark contrasts are both detected. A single nonlinearity, and a single delay line, are sufficient to compute a direction-selective, velocity-tuned response for either sign of image contrast.

The delay line architecture is robust against noise, because a particular output integrates information from an extended spatiotemporal range of the moving image. This feature is shared with the human visual system [31].

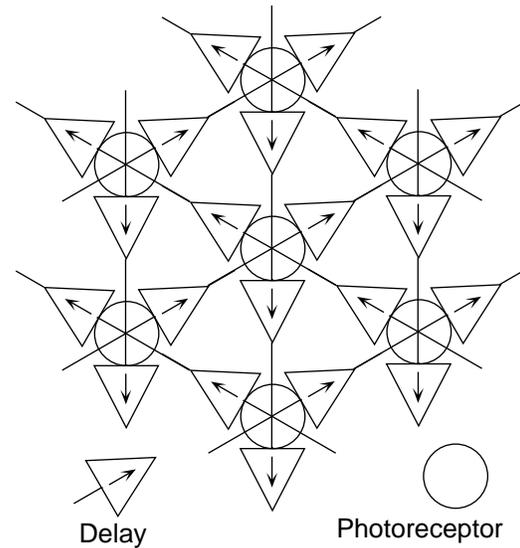
B. Two-Dimensional Architecture

In the two-dimensional chip, I use a hexagonal architecture that encodes any possible direction of motion using three directions (Fig. 2). Each pixel computes the response to three principal directions of motion. Since the three directions are nonorthogonal, any direction can be represented by weighted combinations of the three *nonnegative* outputs. This basis set may be familiar from the representation of color space by a set of three primary colors. In an orthogonal architecture, we would need four principal directions, each with nonnegative output, to represent an arbitrary direction of motion, or, equivalently, two bidirectional outputs.

Each pixel contains a single photoreceptor, three delay elements, and three output nonlinearities. The three outputs are scanned out and are displayed on a monitor as three different colors, through use of a scanning frame, as described in [29]. The Reichardt detector uses differencing of complementary directions to reduce common-mode responses. In the two-dimensional motion chip, I use an analogous scheme to increase the contrast between the three channels of output that is computed serially by the video driver circuits. This linear differencing operation does not do any additional information processing, so I will not discuss it further.

The aperture problem arises in any two-dimensional motion computation. A local edge-based measurement can measure only the component of edge motion orthogonal to an edge, and cannot measure motion along the edge. Another and more general way of stating this fact is that local measurements *constrain* the

Fig. 2. Two-dimensional architecture of the motion chip. Pixels are arranged in a hexagonal array. Each pixel contains a single photoreceptor, three delay elements, and three output nonlinearities. The delay lines run in the three principal directions shown by the arrows in the delay elements.



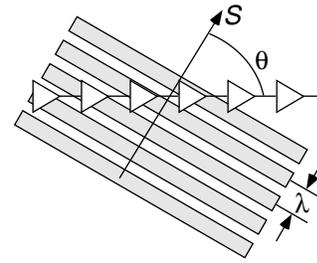
estimate of the actual pattern movement. An array of detectors measures a set of motion vectors, each member of which is *consistent* with the movement of the pattern as a whole. The different detectors, however, may disagree in their assumptions about the component of motion *along* their local edges. The global motion vector is the vector that most closely agrees with the set of constraints imposed by the local measurements. The Tanner chip ingeniously solves the constraint problem in a cooperative manner.

The delay line motion architecture does no global computation to reconcile local measurements, so the best that it can do, in principle, is the computation of the local motion vectors. In the present architecture, these quantities are computed such that they are consistent with the observed motion, but there is no built-in assumption that the measured motion is only orthogonal to the edge. The nonintuitive feature is that the motion that is consistent with the observed set of outputs may not be orthogonal to the edges. The reason is that a detector is sensitive to only spatiotemporal correlation along its preferred direction. An edge moving at a speed S in a direction oblique to the preferred direction, with angle θ , causes apparent motion along the preferred direction with speed $S/\cos\theta$ (Fig. 3).

III. THEORETICAL ANALYSIS OF MOTION CIRCUIT

To understand the properties of our model, and to derive results that can be verified experimentally, we shall analyze the one-dimensional motion circuit in both the time and frequency domains. The time-domain analysis is more intuitive. The frequency-domain analysis is more useful for comparison with measurement,

Fig. 3. Apparent motion and wavelength effects. A grating pattern with wavelength λ , moving at speed S and angle θ over motion detector with preferred direction to the right causes an apparent motion along the array at the increased speed $S/\cos\theta$. The apparent wavelength seen by the array is also increased to $\lambda/\cos\theta$. The “apparent”



motion can equally well be caused by true motion to the right at velocity $S/\cos\theta$ of edges that are oriented at angle θ ; there is no way, in principle, that the local measurement can distinguish the two.

and for understanding the response of the system to variations in stimulus parameters, such as velocity and spatial frequency. The nonlinearity is a simple feedforward operation done on the delay line signal, so our analysis will concentrate on the linear delay line. Some of the symbols used in this section are shown in Fig. 1

A. Time-Domain Analysis

We shall compute the delay line response to a single moving bright edge. To do this calculation, we compute the transfer function from photoreceptor to delay line, and the transfer function of an n -stage delay line. Combining these two transfer functions, we obtain the complete transfer function between the input and the delay line output n stages later. Using the transfer function, we compute the step response. The response to a moving edge is a sum over the temporally shifted step responses produced by the edge at successive receptors.

The photoreceptor outputs are coupled capacitively into the follower-integrator delay line section. Hence, the transfer function from receptor to delay line is the single high-pass filter

$$H_{in}(s) = \frac{\tau s}{\tau s + 1} \quad (1)$$

where s is the Laplace transform variable and τ is the time constant of each first-order section [30]. (It may be helpful to refer to the circuit diagram in Fig. 9). The delay line—a chain of n first-order, buffered, low-pass filters—has the transfer function

$$H_n(s) = \frac{1}{(\tau s + 1)^n} \quad (2)$$

We obtain the transfer function from photoreceptor input 0 to delay line output n by multiplying H_n and H_{in} . By computing the inverse Laplace transform, we find that the time response at the n th delay line tap to a bright intensity step at the first photoreceptor at time $t = 0$ is

$$r_n(t) = \frac{1}{n!} \left(\frac{t}{\tau} \right)^n e^{-t/\tau} \quad (3)$$

for $t > 0$. For $n = 0$, this function is a step-increase from 0 to 1, followed by a decaying exponential. For $n = 1$, this function is an initially-linear increase from 0, followed by the same decaying exponential. In general, (3) represents a wave packet that travels along the delay line, one stage per time τ . Asymptotic analysis shows that as n grows large, the response approaches a Gaussian shape with constant area. The width of the Gaussian broadens as \sqrt{n} , and the height decreases as $1/\sqrt{n}$. Note that the stimulus is a *step*-input at the photoreceptor that is high-pass filtered by the capacitive coupling to the delay line.

A motion stimulus causes a succession of inputs to the delay line, where the time delay ΔT between excitation of successive stages is

$$\Delta T = \frac{L}{S} \quad (4)$$

where L is the spatial separation between detectors, and S is the velocity of the moving pattern. Since the delay line is a linear system, the response to a moving bright edge is the superposition of the responses to the edge passing over all the contributing inputs. We define $R_n(t)$ to be the delay line response at tap n , in response to a moving bright edge that passes over the first input at time $t = 0$. It is easy to compute $R_n(t)$:

$$\begin{aligned} R_0(t) &= r_0(t) \\ R_1(t) &= r_1(t) + r_0(t - \Delta T) \\ R_2(t) &= r_2(t) + r_1(t - \Delta T) + r_0(t - 2\Delta T) \\ &\dots \\ R_n(t) &= \sum_{k=0}^n r_k(t - (n-k)\Delta T) \end{aligned} \quad (5)$$

We have assumed that either the delay line or the motion starts at tap 0. Fig. 4(a, c), show a graphical representation of the *linear* delay line response to a bright bar moving over the delay line at the optimum velocity and at one half of the optimum velocity. The bright edge causes positive activity, and the dark edge causes negative activity. Both signs of edge-contrast information travel along the delay line with a fixed speed and

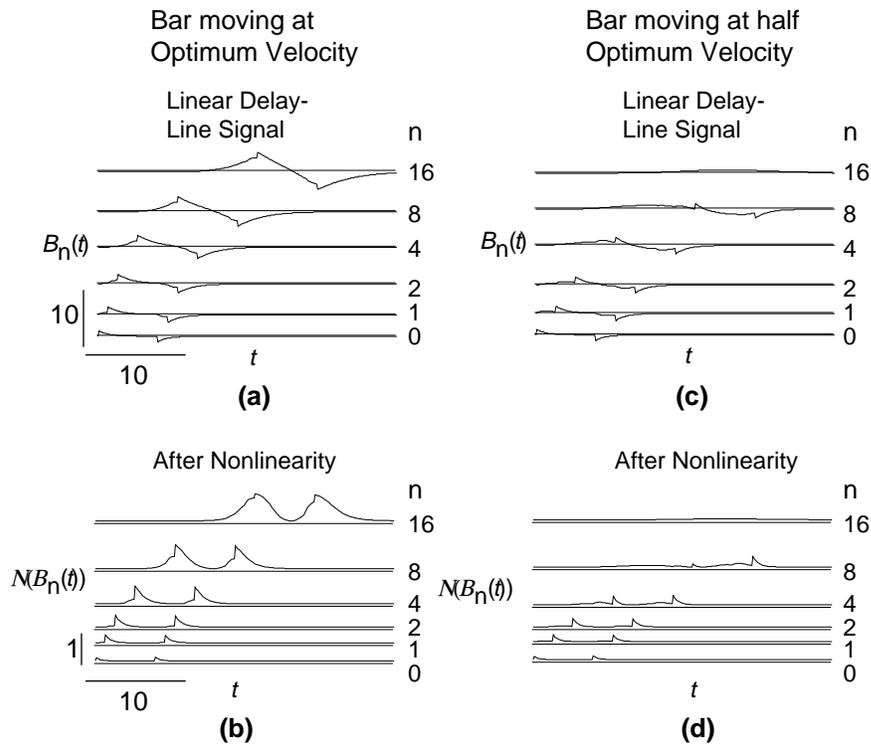
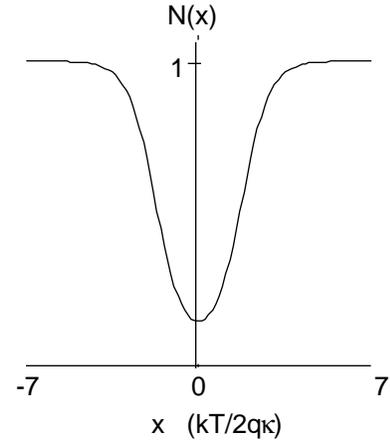


Fig. 4. Theoretical response of delay line to moving bar for two velocities. A bright bar of width 6 pixels, traveling to the right over the delay line, passes over the first tap at time $t = 0$. The mathematical form of the delay line response to the moving bar at the n th tap is $B_n(t) = R_n(t) - R_n\left(t - \frac{6}{S}\right)$, where R is the response to a single bright edge, computed in the text as (5), and S is the speed in pixels/time. Each trace shows the response at a particular tap as a function of time. Tap number is shown to right of trace. The linear delay line signals $B_n(t)$ are shown in **(a)** and **(c)**, and the result of passing the linear delay line signal through the antibump nonlinearity (Fig. 5), $N[B_n(t)]$, is shown in **(b)** and **(d)**. In (a) and (b), the edge travels at the natural speed of the delay line. In (c) and (d), the edge travels at one-half the speed of the delay line. When the bar motion is matched to the delay line speed, the successive inputs pile up, in response either to the bright or the dark edge. The nonlinearity rectifies $B_n(t)$, so both edge contrasts appear as a positive output. When the motion is not matched, the successive inputs arrive in asynchrony with the delay line signal, and the resulting output is smaller.

Fig. 5. The antibump output nonlinearity (6).

The input is a differential voltage x shown in natural input units, and the output is a current, shown here normalized to 1.



direction, though the activity has opposite polarity for the two edges. When the edge motion is matched to the speed of the delay line, the activity on the delay line piles up maximally. We can see how, in this linear system, the *integrated* response is unaffected by the stimulus velocity, even though the activity has larger peak magnitude when the motion is matched to the delay line speed. The longer the motion continues, the larger the peak activity. Saturating nonlinearities eventually limit the response of the delay line.

B. Antibump Output Nonlinearity

The antibump output-nonlinearity-circuit used in the motion chips computes the function

$$N(x) = \frac{I_b}{1 + \frac{w}{4} \operatorname{sech}^2 x} \quad (6)$$

where I_b is the bias current, x is the input voltage, in units of approximately $\frac{2kT}{q\kappa} \approx 70$ mV, and $w \approx 20$ is a geometrical layout parameter [11]. The output is a current. The shape of this function is shown in Fig. 5. The output is parabolic for small input, and saturates at the adjustable bias current. The width of the valley region in subthreshold is determined by layout geometry. The valley width can be increased by using an above-threshold bias current.

The output from the motion circuit is the nonlinearity (6) applied to the signal on the delay line. In response to the moving edge signal $R_n(t)$ on the delay line, the velocity-tuned outputs are $N(R_n(t))$. Fig. 4(c, d) show the response to the same moving bright bar after the nonlinearity is applied. We can see

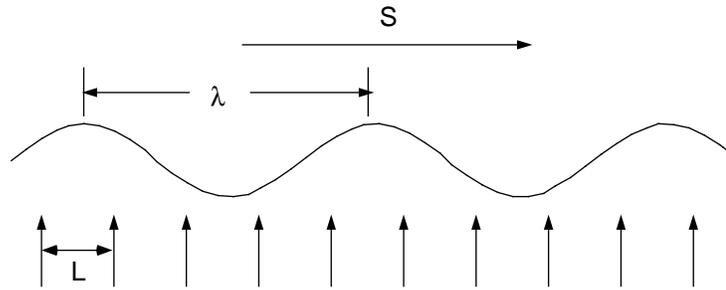


Fig. 6. Sinusoidal grating pattern moving over motion array. The grating moves at speed S . Pixel locations are shown as arrows. The pattern has wavelength λ , and the pixel spacing is L . The temporal frequency at each pixel is $\omega = 2\pi S/\lambda$. The pattern phase shift between pixels is $\phi = 2\pi L/\lambda$.

how the nonlinearity makes both the peak and integrated responses strongly velocity selective, and also how the even characteristic of the nonlinearity results in equivalent detection of both bright and dark edges.

C. Frequency-Domain Analysis

In the frequency domain, the stimulus is a one-dimensional sinusoidal variation in contrast, with wavelength λ and velocity S (Fig. 6). If the velocity is constant, then each input to the motion array is a simple sinusoidal function of time with the same temporal frequency,

$$\omega = \frac{2\pi S}{\lambda} \tag{7}$$

at each input. The signal at a particular tap of the delay line is a sum over all the contributing inputs. Each input is first phase shifted by the spatial separation, and then phase shifted again by the filtering in the low-pass delay line stages. The delay line signal is a pure sinusoidal function of time, with the same temporal frequency at each tap, multiplied by the transfer function for that tap.

The pattern phase shift of the input signal between each pixel, due to the pixel spacing and the sinusoidal spatial variation of image contrast, is

$$\phi = \frac{2\pi L}{\lambda} \tag{8}$$

Each stage of the delay line is separated by a low-pass filter, with transfer function

$$\frac{1}{i\omega\tau + 1} \quad (9)$$

Each input is a single high-pass filter, with transfer function

$$\frac{i\omega\tau}{i\omega\tau + 1} \quad (10)$$

We obtain the complete transfer function for tap n by summing the inputs from taps 0 through n , resulting in

$$\begin{aligned} H_n(\omega) &= \frac{i\omega\tau}{i\omega\tau + 1} \left(1 + \frac{e^{i\phi}}{i\omega\tau + 1} + \frac{e^{i2\phi}}{(i\omega\tau + 1)^2} + \dots \right) \\ &= \frac{i\omega\tau}{i\omega\tau + 1} \sum_{k=0}^n \frac{e^{ik\phi}}{(i\omega\tau + 1)^k} \end{aligned} \quad (11)$$

where we have factored out the high-pass filter common to each stage. The summation in (11) is a geometric series, so we can write the transfer function explicitly for an arbitrary number of stages:

$$H_n(\omega) = \frac{i\omega\tau}{i\omega\tau + 1} \frac{1 - \left(\frac{e^{i\phi}}{i\omega\tau + 1} \right)^{n+1}}{1 - \frac{e^{i\phi}}{i\omega\tau + 1}} \quad (12)$$

This expression represents a phasor that rotates about the complex origin at a frequency ω . The physically measurable part of the expression is the projection of (12) onto any axis in the complex plane, and is much more complicated than is (12).

The average signal on the delay line in response to a sinusoidal input pattern is zero, regardless of the velocity. The length of the phasor, however, is strongly direction and velocity selective. To obtain a response that is tuned for velocity requires a nonlinear operation that measures the length—on some metric—of the phasor. The delay line signal is fed into the antibump nonlinearity, which computes a squaring operation for small input signals. In explicit mathematical form, the time-dependent velocity-tuned output is

$$O_n(t, \omega) = N\left(\left| H_n(\omega) e^{i\omega t} \right| \right), \quad (13)$$

where N is the antibump nonlinearity in (6). Hence, the operation in (13) is proportional to the squared absolute transfer function, for low-contrast stimuli.

We can learn about the behavior of the system by studying the amplitude of the linear transfer function (12)—that is, the length of the phasor. We shall examine the case of a seven-input system. In Fig. 7, I have plotted the phasor at the seventh tap, $H_6(\omega)$, in response to the motion of a sinusoidal pattern in the preferred and null directions. The speed of the motion, in each case, is near the optimum speed for the preferred direction. Each phasor is shown as a number of adjoining line segments; each of these components is the contribution of one pixel input to the complete phasor.

When the grating moves in the preferred direction at the optimum speed, the phasor is as long as it can be, given an input pattern of a given amplitude and wavelength, because the components all point in nearly the same direction. They all point in the same direction because, for each input, the total pattern-phase-shift and the total low-pass-filtering-phase-shift match. There is also an overall rotation of the entire phasor, due to the high-pass filtering at each input to the delay line.

In contrast, when the pattern moves in the null direction, the phasors sum destructively, and the resultant phasor curls up on itself. The complete phasor is short, no matter how many inputs are combined, because the components always approximately cancel one another. The length of the phasor, as a function of the speed of the motion, has wiggles, corresponding to the periodic behavior of the destructive cancellation. These wiggles can be seen in the experimental data shown in Fig. 15 and Fig. 17.

Many of the correlation-based direction-selective models in the literature correspond to a two-input ($n = 1$) version of this system. Hence, it is interesting to compare the case of a two-input system with an sys-

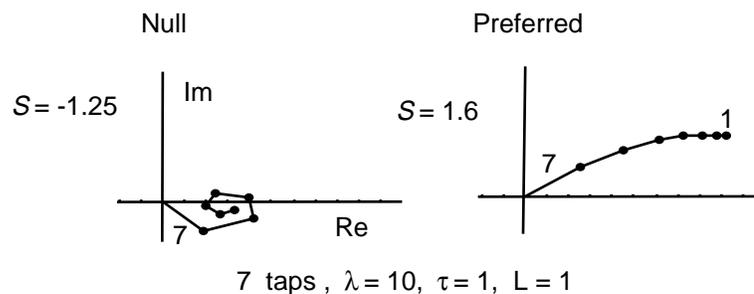


Fig. 7. Phasor summation properties for a seven-stage system. $H_6(2\pi S/\lambda)$ is plotted on the complex plane. The symbols are defined in Fig. 6. The contributions of individual taps are shown as component line segments; taps 1 and 7 are labeled. The optimum velocity is near $L/\tau = 1$.

tem that is effectively infinitely long. We shall first derive the mathematical behavior for each system, and then contrast the two systems.

Long System

The explicit form of the transfer function for an infinite number of stages is given by (12) with $n \rightarrow \infty$:

$$H_{\infty}(\omega) = \frac{1}{1 + \frac{1 - e^{i\phi}}{i\omega\tau}} \quad (14)$$

(Note that this transfer function includes the high-pass input to the delay-line.) By examining the behavior of the n -dependent term in the numerator of (12), we can determine that this limiting form is accurate as long as the following condition holds:

$$n > \left(\frac{\lambda}{L}\right)^2 \quad (15)$$

This restriction seems severe, but examination of the measured tuning curves shows that, in practice, there are only scale differences in the responses after the first few taps. We shall examine the behavior of (14) to understand how it is affected by spatial wavelength.

Since the numerator of (14) is 1, we can find the maximum value of the response by minimizing the absolute value of the denominator. This simple computation reveals that the maximum value of $|H_{\infty}(\omega)|$ occurs at

$$\omega_{\max} = \frac{2(1 - \cos \phi)}{\tau \sin \phi} \quad (16)$$

corresponding to the speed

$$S_{\max} = \frac{\omega_{\max} \lambda}{2\pi} = \frac{\lambda(1 - \cos \phi)}{\pi\tau \sin \phi} \quad (17)$$

For long wavelengths, $\phi = 2\pi L/\lambda$ is small, and we find the peak tuning occurs at the frequency

$$\omega_{\max} = \frac{\phi}{\tau} \quad (18)$$

corresponding to the speed

$$S_{\max} = \frac{L}{\tau} \quad (19)$$

which is reassuring, since this optimum speed corresponds with the intuitive idea that optimum motion is motion matched to the delay-line speed. The optimum speed is invariant with wavelength, for wavelengths longer than a few pixel spacings. This result is also surprising, because correlation models do not generally display this invariance.

We can see how this behavior comes about by examining the limiting case of a long wavelength. As we *increase* the wavelength λ , we *decrease* the pattern phase shift ϕ between pixels. (The pattern phase shift is the phase shift of the input, between adjacent pixels, just due to the wavelength of the pattern). The response maximum does not move, because the decrease in ϕ is compensated by a *decrease* in low-pass phase lag due to decreased frequency ω in the low-pass filter.

The first-order, low-pass phase lag cannot become greater than 90° . Eventually, as we decrease wavelength, the pattern phase shift decreases faster than does the low-pass phase lag. To compensate, the pattern must move faster to obtain an increased phase lag from the low-pass filter. Hence, the shorter the wavelength, the higher the optimum speed. In fact, when $\lambda = 2L$, the optimum velocity approaches infinity. This limit corresponds to the Nyquist sampling criterion—the spatial frequency pattern is sampled twice per cycle by the array. The trend toward higher optimum speed with shorter wavelength is visible in the experimentally measured velocity tuning curves in Fig. 17.

At the optimum velocity, in the limit $\lambda \gg L$, the squared magnitude of the transfer function takes the value

$$|H_\infty(\omega)|_{S_{\max}}^2 = \left(\frac{\lambda}{\pi L}\right)^2 \quad (20)$$

Hence, the squared magnitude of the response at the optimum speed goes as the square of the wavelength, independent of the optimum speed. At low speeds, $\omega \ll 1/\tau$, and the squared magnitude of the transfer function is given by

$$|H_\infty(\omega)|_{S \ll L/\tau}^2 = \left(\frac{S}{L/\tau}\right)^2 \quad (21)$$

Hence, for low speeds, the squared magnitude of the response is proportional to the square of the speed and invariant to wavelength. (The apparently conflicting results in (19), (20) and (21) are actually not conflicting, because the assumption that leads to (21) breaks down at S_{\max} .)

Two-Input System

As stated earlier, the two-input system is similar to existing correlation-based direction-selective detectors. Let us now consider the behavior of a two-input system in the same way we just did for the effectively-infinite delay line system. For two inputs, (12) becomes

$$H_1(\omega) = \frac{i\omega\tau}{i\omega\tau + 1} \left(1 + \frac{e^{i\phi}}{i\omega\tau + 1} \right) \quad (22)$$

We will form a particular instantiation of the pairwise Reichardt-type correlation detector by taking the difference between the squared magnitude of two pairwise detectors tuned to opposite directions, resulting in the following expression.

$$|H_1(\omega)|^2 - |H_1(-\omega)|^2 = 4 \left(\frac{\omega\tau}{\omega^2\tau^2 + 1} \right)^2 \omega\tau \sin\phi \quad (23)$$

We can think of this detector as the difference between two time-averaged pairwise detectors, like the one shown in Fig. 1(a), with opposite direction selectivities. The common-mode terms have canceled. The tuning is separable into a temporal and a spatial product. Hence, we can immediately deduce that the optimum response occurs at a particular temporal frequency, and not at a particular velocity. The pairwise detector is not a velocity-selective filter, but rather is a temporal-frequency-selective filter whose response is modulated by the geometrical interference term $\sin\phi$, familiar from the fly literature (see, for example, [7]). We can find the optimum temporal frequency ω_{\max} for the two-input system in the same way that we found S_{\max} for the infinite delay line. Independent of wavelength, the maximum of (23) occurs at the temporal frequency

$$\omega_{\max} = \sqrt{3}/\tau \quad (24)$$

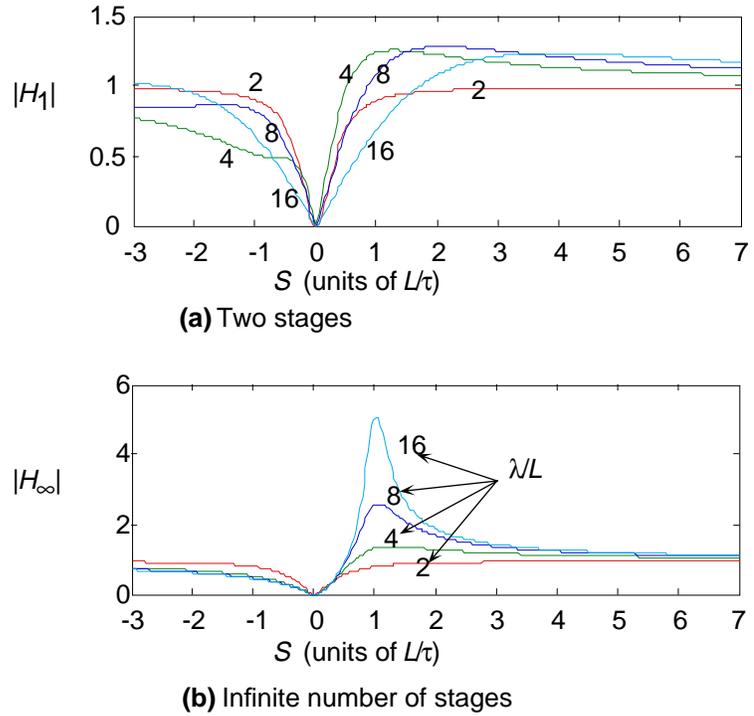
For low speeds, the temporal frequency is small compared with $1/\tau$, and we compute the response (23) as approximately

$$|H_1(\omega)|^2 - |H_1(-\omega)|^2 = 4(\omega\tau)^3 \sin\phi. \quad (25)$$

Hence, the response to slow motion depends on the wavelength through the $\sin\phi$ term, and is cubic in the speed.

Fig. 8 compares the theoretical transfer function amplitudes for the two-input system and the long system cases. Note that in this figure, we compare $|H_{\infty}|$ with $|H_1|$, rather than with the Reichardt detector in (23). In the two-input system, the optimum speed is proportional to wavelength (in accordance with (24)), but the amplitude of the response at the optimum speed is relatively invariant to wavelength. In the long sys-

Fig. 8. Theoretical plots of the magnitude of the transfer function for **(a)** two-stage system, compared with **(b)** infinite delay line, for various wavelength stimuli, shown next to curves. Parameters: $L/\tau = 1$.



tem, the optimum speed and the low-speed response are invariant to wavelength, but the amplitude of the response at the optimum speed is proportional to wavelength. We can also see these characteristics in the measured tuning curves shown in Fig. 17. Table 1 lists these theoretical results. In summary, the delay-line architecture differs from the pairwise detectors in that it is sensitive to stimuli with low as well as high spatial frequencies. The pairwise detector, because of the locality of the computation, cannot be very selective for motion of patterns with long wavelengths. The delay-line detector, in contrast, aggregates information over a spatial range corresponding to an arbitrarily-long wavelength, and hence can retain direction and speed selectivity even for long-wavelength patterns.

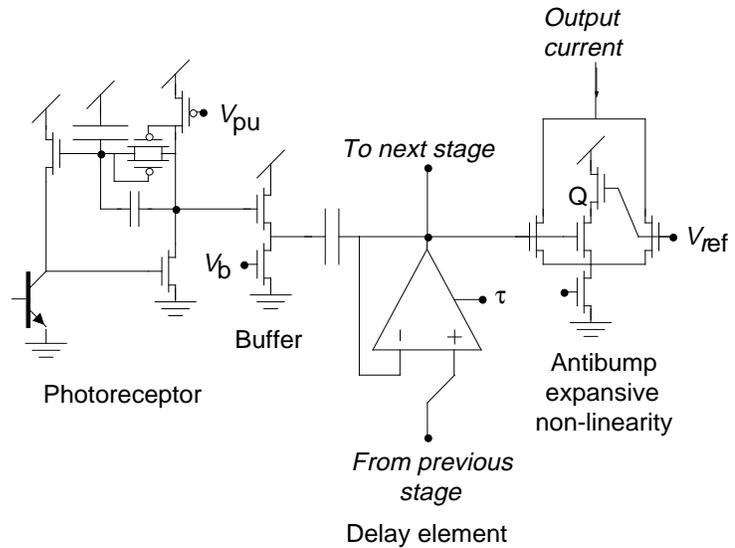
	Pairwise detector: $ H_1(\omega) ^2 - H_1(-\omega) ^2$	Delay-line detector: $ H_\infty ^2$
Location of optimum response	$\omega = \frac{\sqrt{3}}{\tau}$	$S = \frac{L}{\tau}$ (for $\lambda \gg L$)
Low-speed response	$(\omega\tau)^3 \sin\phi$	$\left(\frac{S}{L/\tau}\right)^2$ (for $\lambda \gg L$)
Response at optimum speed	$\propto \sin\phi$	$\propto \left(\frac{\lambda}{L}\right)^2$ (for $\lambda \gg L$)

Table 1. Comparison between pairwise Reichardt-type detector and delay-line detector.

Symbols are defined in Fig. 6.

Fig. 9.

Circuitry for a single stage of the motion circuit. An adaptive photoreceptor, with a contrast-sensitive gain of about 1 V/decade, feeds into a source-follower buffer, whose output couples capacitively into the follower integrator. The velocity-selective output current is computed by the antibump



expansive-nonlinearity circuit. When the delay line signal is pushed away from its equilibrium point (V_{ref}) by activity on the delay line, the output current increases, because the center leg of the antibump circuit is turned off. For additional directions of selectivity, only the Buffer, Delay element, and Antibump nonlinearity must be duplicated—all directions share the same photoreceptor input.

IV. THE MOTION CIRCUIT

The circuit for a single stage of the motion circuit is shown in Fig. 9. The output from an adaptive, high-gain, logarithmic photoreceptor circuit is coupled capacitively into a delay line stage. The delay line consists of a line of buffered, first-order, low-pass filters, implemented on the chips with follower integrators. The output nonlinearity is implemented with an antibump circuit.

The adaptive photoreceptor is essential for the robust operation of the circuit, because it provides well-conditioned input to the delay line over a wide range of lighting conditions. The receptor circuit has been reported previously [10][13][27] and is described in detail in [12]. In brief outline, the operation of the photoreceptor circuit goes as follows. The input leg of the receptor, consisting of the phototransistor and the source-follower feedback transistor, forms a signal voltage that is logarithmic in the intensity. The feedback circuit amplifies the logarithmic signal, and centers the operating point of the output signal around the history of the signal. The time-averaged output signal is stored on the feedback capacitor and serves as the ref-

erence around which the response is computed. The pair of diode-connected transistors act as a resistor-like adaptive element with monotonic I - V characteristic. The use of adaptation makes for a receptor with simultaneous high sensitivity and wide dynamic range.

The output from the adaptive photoreceptor connects to the input of a source follower amplifier that buffers the output of the receptor. Different directions of selectivity use separate buffers that share the common photoreceptor input. The source follower isolates different directions in the case of a motion network with more than a single direction of selectivity. A separate source-follower is used for each direction. The source-follower output can follow a decreasing input signal only as fast as the bias current can discharge the output. It can follow increasing signals at an arbitrary rate, determined by the input voltage. By biasing the follower strongly (i.e., turning up V_b), we ensure that the source-follower output can follow accurately both increasing and decreasing outputs from the photoreceptor.

The follower-integrator uses a simple five-transistor transconductance amplifier [30]. These simple transconductance amplifiers suffer from systematic offset of several mV, due to Early-effect drain conductance in the differential pair and in the current mirror. Over a 25-stage delay line, the offset can be more than 100 mV. Since we compute the output of the pixel relative to a fixed reference voltage V_{ref} that is the input to the first stage of the delay line, this offset is important. We set V_{ref} near V_{dd} , so that the differential pair and the current mirror are both balanced. The speed of the delay line, and hence of the velocity tuning, is set by the bias τ of the transconductance amplifier.

The voltage on the delay line connects to the input of an antibump circuit [11]. The other input to the antibump circuit connects to the reference voltage V_{ref} . If any activity on the delay line pushes the voltage away from V_{ref} , either in the positive (bright-edge) or in the negative (dark-edge) direction, the output current becomes larger, thus computing the power-like measurement of the linear delay line signal. Details of the antibump circuit operation are given in [11]. In brief, the antibump circuit works because the center leg of the circuit, consisting of the series-connected transistors marked with Q in Fig. 9, turns off when the differential input voltage is sufficiently large, forcing the bias current to flow through the outer legs of the circuit.

The delay line architecture for this circuit was inspired by an earlier architecture for sound synthesis of visual images [30]. The SeeHear chip sees an image, through use of on-chip photodetectors, and converts the image into an auditory equivalent of the visual image. This SeeHear chip used the same scheme of photodetector signals coupling into linear delay lines. The delay lines serve to synthesize the interaural delay cues that drive the auditory-localization system in the brain.

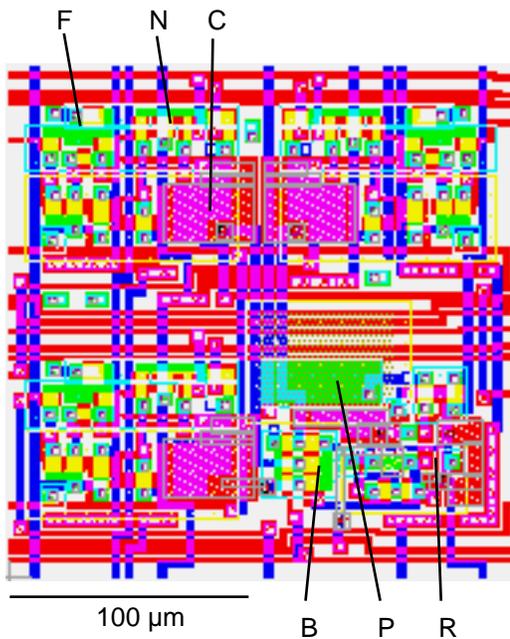


Fig. 10. Layout of pixel of two-dimensional motion circuit with hexagonal architecture. Pixel has one photoreceptor, three follower integrators, and three antibump nonlinearities. Lines indicate circuit components. Keys: F = follower, N = antibump nonlinearity, C = capacitor, B = buffer, P = phototransistor, R = photoreceptor. Scale bar shows dimensions. This pixel is arranged in the hexagonal architecture shown in Fig. 2.

The layout for one pixel from the two-dimensional motion chip is shown in Fig. 10. Most of the pixel area is covered with wire and capacitance, rather than with transistor. The amount of interpixel wiring is minimal compared with the amount of wire used for routing bias signals into the pixel or routing outputs from the pixel. The architecture is efficient, because these input and output signals are essential and occupy most of the pixel area.

V. EXPERIMENTAL RESULTS

In this section, I present experimental results from a two-dimensional motion chip with the hexagonal architecture, and from a one-dimensional motion chip with two opposing delay lines.

A. Results from Two-Dimensional Motion Circuit

Fig. 11 shows the real-time scanned output from the two-dimensional motion circuit in response to two moving patterns. Part (a) shows the response to a drifting square-wave grating. The buildup of activity away from the edge of the array in part (a) of the figure is evident as an increasing saturation of the color of the output. The longer the motion is visible to the array, the larger the response. Part (b) shows the response of the chip to a rotating spiral pattern that produces the illusion of expansive optical flow. The chip computes a

pseudolocal flow field. Motion that is not in a principal direction appears as a combination of the colors corresponding to the principal directions.

This chip consists of a 26 by 26 hexagonally arranged array of pixels. The power consumption of the core of the chip (all the analog computation done by the pixels) is 1.5 mW (dark) to 8 mW (light), depending on the brightness of the illumination, and hence on the raw pixel photocurrent. Hence, the power consumption of the analog computation, independent of photocurrent, is 1.5 mW, or 2.2 $\mu\text{W}/\text{pixel}$. About 80 % of this power is supplied to the antibump circuit, which I run at above-threshold bias to widen the transfer characteristic. The rest of the chip—consisting of the scanning frame, the clock driver, and the on-chip video amplifiers [29]—uses another 26 mW. The size of each pixel is 224 by 225 μm ; the entire chip fits on a 6.8- by 6.9-mm die and is fabricated in a 2- μm , double poly, *n*-well process available through MOSIS, the DARPA fabrication service [9].

Fig. 12 shows the directional tuning for the pixels in the two-dimensional chip. These plots show the average angular tuning for each of the directions represented by a pixel, for three different movement speeds. The responses are plotted in polar coordinates, and the distance from the origin is the magnitude of the average response. For speeds at or higher than the optimum velocity, the relative values of three outputs unambiguously determine the direction of the normal component of edge velocity. However, for speeds lower than the optimum, motion of a grating in a direction *oblique* to the detector orientation excites a detector more than motion at the same speed *along* the detector, resulting in responses with two peaks. This characteristic is an aspect of the aperture effect, since the delay-line detectors correlate image information along only the detector orientation.

B. Quantitative Results from a One-Dimensional Motion Circuit

I collected quantitative data primarily from a one-dimensional version of the motion circuit with two directions of selectivity. Fig. 13 shows the response from two taps of the one-dimensional circuit in response to a moving sinusoidal pattern. The buildup of direction selectivity is evident in the difference between the response of the early and late taps. A systematic asymmetry is evident in the lack of frequency doubling, except for the largest response. I am not certain of the origin of the asymmetry, although I suspect either the source-follower buffer or the photoreceptor.

Fig. 14 shows velocity-tuning curves for a number of taps from the same one-dimensional motion circuit, in response to the same moving pattern. The velocity tuning becomes sharper and more direction selective for later taps, although it appears that the tuning begins to approach a limiting form after about tap 15.

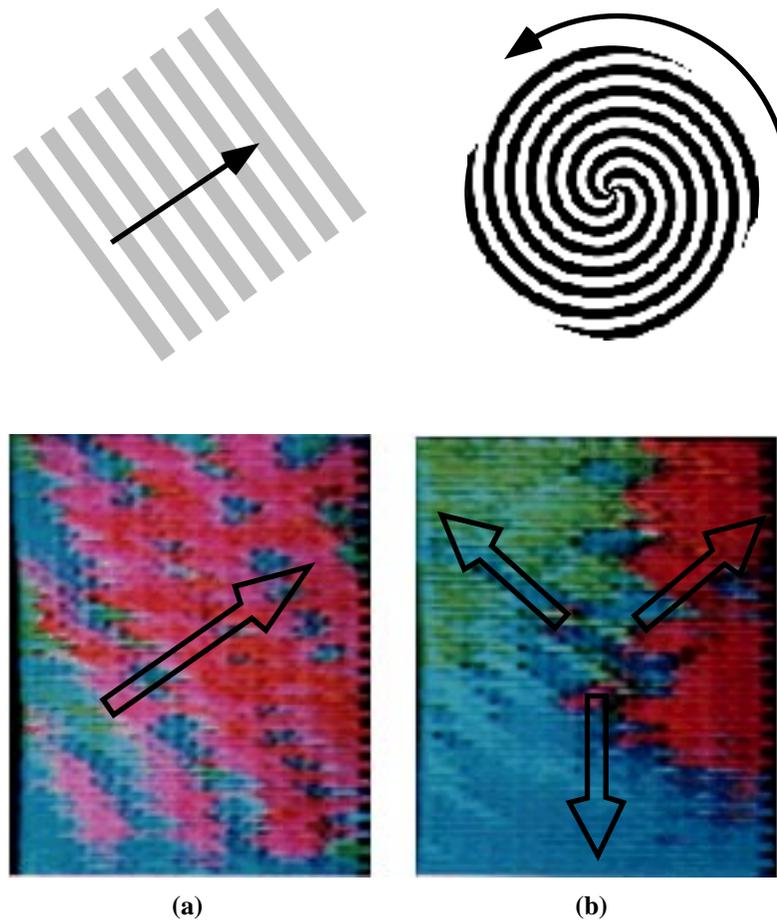
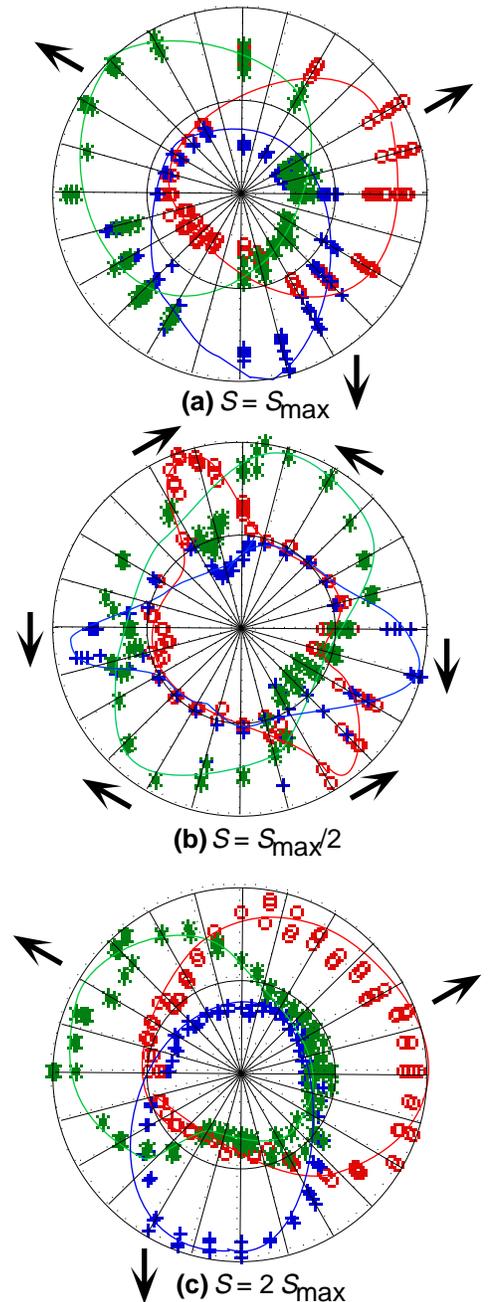


Fig. 11. Color photographs showing the video output of the two-dimensional motion chip to motion **(a)** of a square-wave grating pattern in a single principal direction (arrow), and **(b)** of a rotating spiral pattern that produces illusory expansive optical flow (arrows). The outputs of the chip encode motion downward as blue, motion upward and to the right as red, and motion upward and to the left as green. Encoding of intermediate directions is evident as mixing of the primary colors. The dynamic nature of the output of the chip makes direct photography difficult, so we produced these pictures by video taping the output from the color monitor, and then photographing the freeze-framed screen of the video-tape monitor.

Fig. 12. Directional-tuning of pixel in two-dimensional chip, for different stimulus speeds. Stimulus is sinusoidal grating pattern with wavelength approximately 10 pixel spacings. Each polar plot shows the response, measured as average video output voltage, of the three outputs from a pixel near the center of the chip. Each curve is labeled with an arrow showing the orientation of the delay line. I normalized each polar curve to the same maximum to adjust for monitor brightness corrections in the video circuits. The scatter in the points shows the noise in the 2 second averages. The solid curves are Bezier curves fitted by eye to the data points.

The stimuli in (a), (b), and (c) differ only in grating speed. In (a) the grating moves at the optimum speed, in (b), at half optimum speed, and in (c), at twice optimum speed. In (b), the apparent motion effect from Fig. 3 is evident as doubly peaked responses that are roughly perpendicular to orientation of delay line. In (c), the response does not display the same doubly-peaked effect, because oblique motion is faster than orthogonal motion, so oblique motion excites pixel less than does true motion.



The tuning curves shown in Fig. 14 are measurements of the peak-to-peak output voltage from the circuit. The theoretical form of this measurement can be computed from (13), plus the transfer function of the logarithmic current-sense amplifier used in the scanning frame [29]. In Fig. 15, I show fits to the data in Fig. 14. Only the tap number was varied among the theoretical curves; all other parameters (wavelength, spatial frequency, time constant, antibump circuit parameters, logarithmic current-sense amplifier parameters) were kept constant. The quality of the fits is remarkable, considering that the input to the system was from a grating pattern printed on a piece of paper.

A number of nonlinearities have been ignored in the analysis. The effect of saturating nonlinearities can be seen in Fig. 16, which compares the velocity tuning for low- and high-contrast grating patterns. The high-

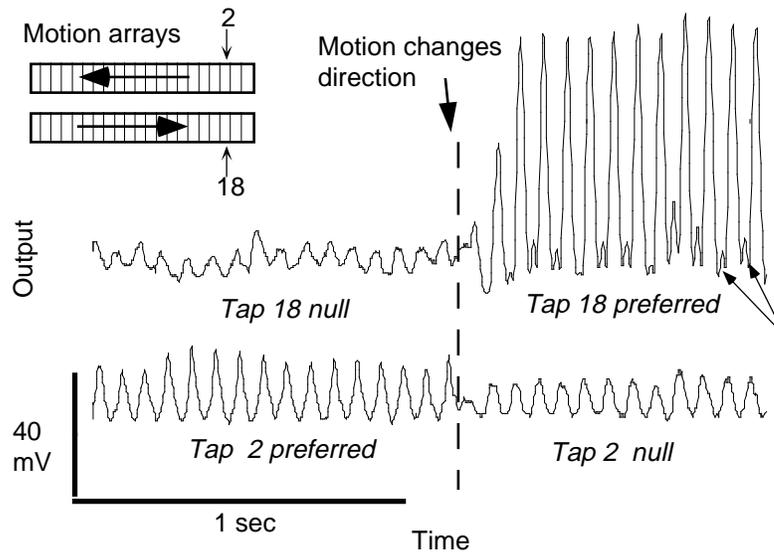
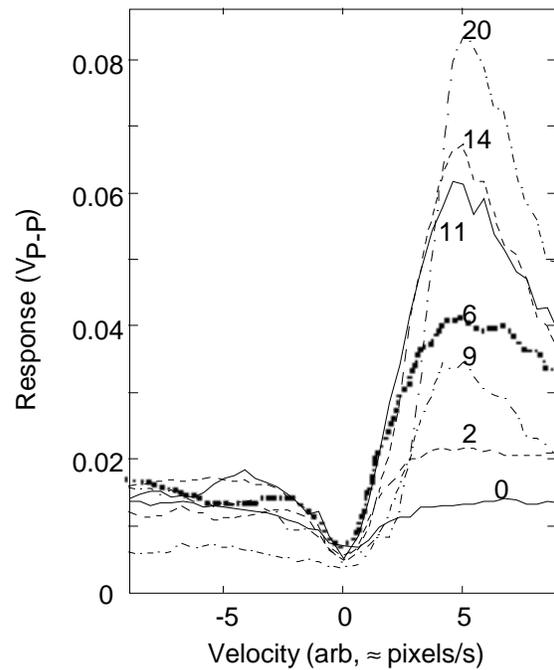


Fig. 13. Measured response from two taps of a one-dimensional motion circuit with both directions of sensitivity (see upper left). The stimulus is a moving low-contrast grating pattern. The motion network had a total of 21 taps. I simultaneously recorded tap 2 in one direction and tap 18 in the other direction (numbering starts at zero, as in the text), and I plot the output voltage from the logarithmic current-sense amplifier. The grating initially moved in the tap 2 preferred direction and in the tap 18 null direction; it then changed direction. After a slight latency, the response from tap 18 became much larger than that of tap 2. The lack of frequency doubling, except in the largest signals (arrows), arises from a systematic offset of unknown origin.

Fig. 14. The velocity tuning for a number of taps of a one-dimensional motion circuit. Plotted are the peak-to-peak outputs from the logarithmic current-sense amplifier fed from the antibump nonlinearity at each tap. The tap number is shown next to each curve. The gain of the logarithmic sense amplifier is approximately 100 mV per decade of current. Circuit offsets are visible as an inversion of the expected order between taps 6 and 9.



contrast grating pattern saturates the differential inputs in the follower integrators or in the antibump nonlinearity, causing a shift of the peak response toward a lower velocity, and a widening of the peak.

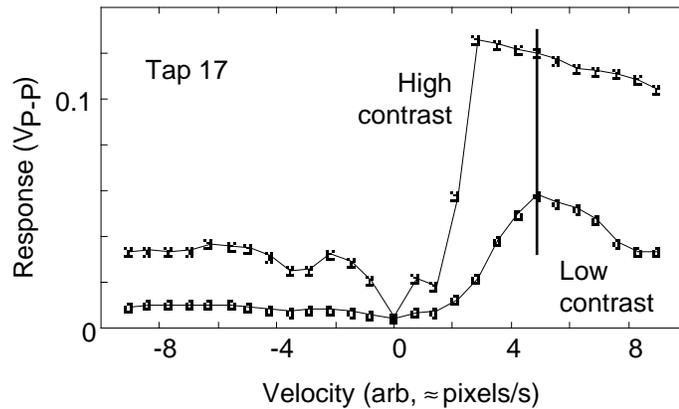
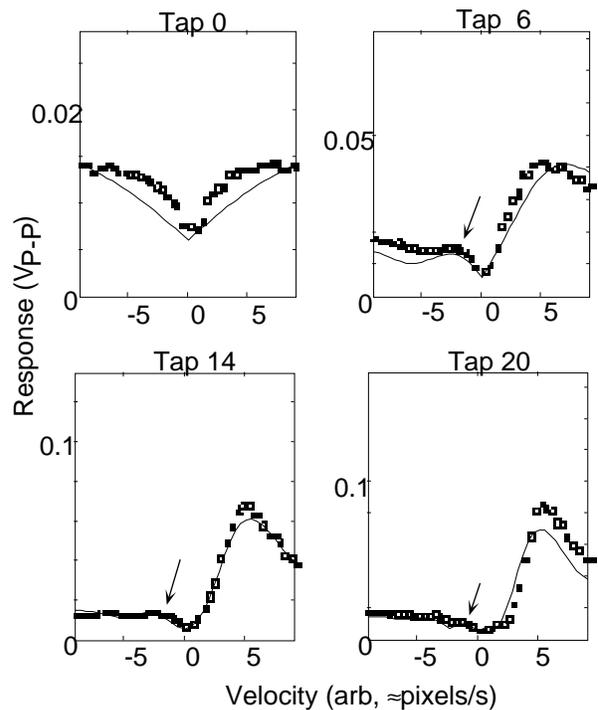


Fig. 16. Effect of saturating nonlinearities on motion-circuit response. The data in the top curve were collected with the use of a high-contrast ($\approx 50\%$) grating; those in the lower curve were collected with the use of a low-contrast ($\approx 10\%$) grating.

Fig. 17 shows the velocity tuning of several taps of the network for different wavelength stimuli. For the late taps, the peak response occurs at a constant speed, relatively invariant to wavelength. For tap 1, the optimum velocity is a strong function of wavelength. For all the taps, the primary effect of varying the wavelength of the stimulus is a simple scaling of the amplitude of the response with wavelength. The speed of the optimum response shifts toward higher speeds as the wavelength is decreased. All these effects are consistent with the theoretical analysis presented earlier. The scaling of tap 1 response amplitude with wavelength, however, is inconsistent. I think that this result is due to a defocused image, which decreases the effective modulation depth of the image more for shorter wavelengths, and hence results in a smaller response for shorter wavelength.

Fig. 15. Velocity tuning with theoretical fits. Each plot shows measured velocity-tuning data as squares, with theoretical fits as solid curves. Response plotted is the peak-to-peak output voltage from a logarithmic current-sense amplifier that senses the antibump output current. Theoretical curves are derived from (14), plus the transfer function for the sense amplifier. The tap number is shown at the top of each curve and corresponds to the numbering used in the theory. Note the vertical scale—the response grows with tap number, as in Fig. 14. Arrows point out wiggles in null direction response that are fitted by theory. All parameters in the model are identical for each theoretical curve, except for the tap number.



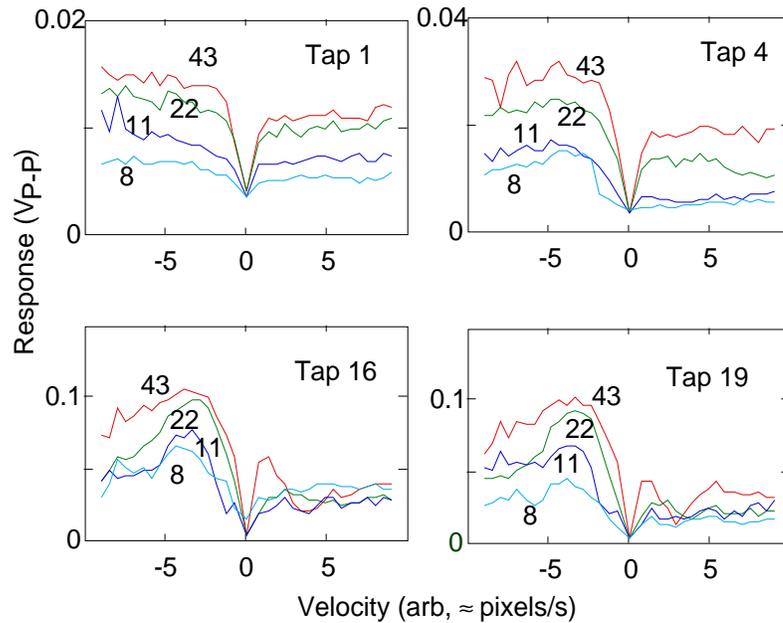


Fig. 17. Effect of changing the stimulus wavelength. Each plot shows the measured output of a particular tap of the one-dimensional motion chip as a function of the stimulus velocity, in response to a moving sinusoidal grating pattern. In each graph, four plots are shown, one for each wavelength of the stimulus. The numbers on each plot show the wavelength, in pixels, of the grating.

VI. DISCUSSION

The inspiration for the spatiotemporal integration properties of the chip's motion architecture came from psychophysical observations of human motion perception [31]. I am aware that cortex is far more complex than my simple model. The important thing to point out, however, is that the robust functioning of the chip is at least partially due to the spatiotemporal aggregation, while I know of at least three other attempts to build functional analog correlation detectors that have only marginal performance. The spatiotemporal aggregation in the chip also has the consequence that it allows the computation to see motion over a wide

range of spatial frequencies, as opposed to the short spatial scales that a pairwise detector is capable of discriminating.

A given combination of the three pixel outputs can correspond to a range of possible pattern velocities of the actual scene. This characteristic is an expression of the aperture effect, and is true for *any* local motion detector. The three spatiotemporal correlations computed in each pixel are valid computations in the sense that the information they report is consistent with a possible motion of a one-dimensional pattern. It can turn out that the output is not consistent with a motion orthogonal to the edge, as seen in Fig. 12(b). However, there is absolutely no reason why motion must be orthogonal to an edge. At first sight, the apparent-motion effect suggests why cortical direction-selective cells are orientation tuned to edges orthogonal to the preferred direction, and why fly motion-sensitive cells are sensitive to the temporal frequency more than to image speed [14][20]. The argument goes that, in cortex, the orientation tuning removes the ambiguity about edge motion by ensuring that the computation measures only those motion components that are orthogonal to the edge. The same argument may be made in flies [19][39]. However, I believe that it is not settled whether these arguments are computationally valid or only opportunistically based on particular aspects of experimental observations.

The receptive field of the delay line motion detector resembles a simplified Adelson-Bergen spatiotemporal energy model [1]. Different spatial locations in the receptive field have different time delays to get to the output of the detector. The spatial receptive field is a long, skinny, rectangle starting at the location of the pixel and extending back along the delay line. The antibump output nonlinearity is similar to the squaring nonlinearity used in the Adelson-Bergen detectors. My detectors share receptive field with each other, through the common shared signal on the delay line.

One approach to performing full computation of the image motion is to plaster the input space with tuning curves. In cortex, there are spatial- and temporal-frequency tuned neurons with many different tuning curves [3]. The ensemble of broadly tuned, imprecise responses, that covers the whole space of inputs, can be combined to form a precise estimate of optical flow [17][21][24][37]. In comparison with some models of cortical motion computation, the number of velocity tunings in my motion chip is tiny—only three curves are used to cover the entire space of inputs. Models of cortex often plaster the space with dozens or hundreds of different tuning curves. Can we build systems with even tens of tuning curves? Not on one chip! This approach would require a number of motion chips because of the large number of required tuning curves. The problem of integrating multiple massively parallel analog chips is now being addressed [26].

It would be good to automatically adjust the tuning of the motion detector to match the input image, in analogy with other adaptive sensory processing like the automatic gain control used in the photoreceptors.

This strategy would use the available dynamic range more fully, and the adaptation state itself would give information about the image motion. Biological studies suggest that animals with cortex do not adjust the tuning curves, while insects do. (However, people have not looked very hard in cortex for short-time-scale adaptive behaviors.) The fly visual system has only a few tens of visual output neurons. These direction-selective, velocity-tuned neurons, which integrate information from the entire visual field, adjust their tuning curves in response to the scene velocity [6][25][34][38]. The adjustment takes the form of a variable time constant that is shorter for higher image velocities or temporal frequencies (which it is—velocity [34] or frequency [6]—is under contention). This scheme is attractive from a chip designer's point of view, because it potentially allows a single velocity-tuned unit to cover a large dynamic range and still retain high sensitivity, in a manner analogous to the adaptive photoreceptor circuits used in the pixels.

VII. SUMMARY

I have described the first functional two-dimensional analog VLSI implementation of a motion detector based on the 35-year-old Hassenstein-Reichardt correlation detector. The delay line extension to the usual pairwise correlation model has the novel functional property that it integrates information over an extended spatiotemporal region. The functionality of this chip is part of an accumulating body of evidence that we will eventually build complete, functional, neuromorphic visual systems.

ACKNOWLEDGMENTS

I thank the anonymous reviewers for constructive comments, Carver Mead, Martin Egelhaaf, Shih-Chii Liu, Buster Boahen, Misha Mahowald, Humbert Suarez, Rahul Sarpeshkar, and Rod Goodman for valuable discussion, and Lyn Dupré for style editing. I also thank the MOSIS fabrication service for making this type of exploratory chip design possible. This work was funded by the Office of Naval Research under grant number NAV N00014-89-J-1675 and by the California Competitive Technologies Program. The inspiration to build a correlating detector array came from Werner Reichardt's work on the fly visual system, and from Nicola Franceschini's work on recording responses to optical stimulation of single pairs of ommatidial receptors. The idea of extending the correlation model to more than a pairwise correlation was inspired by a talk given by Ken Nakayama at the Society of Neuroscience Annual meeting in Phoenix in 1989.

REFERENCES

- [1] E.H. Adelson and J.R. Bergen, "Spatiotemporal energy models for the perception of motion," *J. Opt. Soc. Am.*, vol. 2, pp. 284–299, 1985.
- [2] P. Ang, P.A. Ruetz, and D.Auld, "Video compression makes big gains," *IEEE Spectrum*, Oct 1991, pp. 16–19, 1991.
- [3] C. Baker, "Spatial- and temporal-frequency selectivity as a basis for velocity preference in cat striate cortex neurons," *Visual Neuroscience*, vol. 4, pp. 101–113, 1990.
- [4] H.B. Barlow and W.R. Levick, "The mechanism of directionally selective units in the rabbit's retina," *J. Physiol.*, vol. 178, pp. 477–504, 1965.
- [5] R. G. Benson and T. Delbrück, "Direction -selective silicon retina that uses null inhibition," in *Advances in Neural Information Processing Systems 4*, D.S. Touretzky, Ed., San Mateo, CA: Morgan Kaufmann, pp. 756–763, 1991.
- [6] A. Borst and M. Egelhaaf, "Temporal modulation of luminance adapts time constant of fly movement detectors," *Biol. Cybernetics*, vol. 56, pp. 209–215, 1987.
- [7] E. Buchner, "Elementary movement detectors in an insect visual system," *Biol. Cybernetics*, vol. 24, pp. 85–101, 1976.
- [8] C.P. Chong, C.A.T. Salama, and K.C. Smith, "Image-motion detection using analog VLSI," *Journal of Solid State Circuits*, vol. 27, pp. 93–96, 1992.
- [9] D. Cohen and G. Lewicki, "MOSIS—the ARPA silicon broker," in *Proceedings from the Second Caltech Conference on VLSI*, California Inst. of Tech., Pasadena CA, pp. 29–44, 1981.
- [10] T. Delbrück, "An electronic photoreceptor sensitive to small changes in intensity," in *Advances in Neural Information Processing Systems 1*, D.S. Touretzky, Ed., San Mateo, CA: Morgan Kaufmann, pp. 720–727, 1988.
- [11] T. Delbrück, "'Bump' circuits for computing similarity and dissimilarity of analog voltages," in *Proc. of International Joint Conference on Neural Networks*, vol. 1, pp. I-475–479, 1991.
- [12] T. Delbrück, *Investigations of Analog VLSI Visual Transduction and Motion Processing*, Ph.D. Thesis, Department of Computation and Neural Systems, California Institute of Technology, Pasadena CA 91125, 1993.
- [13] T. Delbrück and C.A. Mead, "Time-derivative adaptive silicon photoreceptor array," in *Proc. SPIE*, vol. 1541, pp. 92–99, 1991.
- [14] H. Eckert, "Functional properties of the H1-neurone in the third optic ganglion of the blowfly, *Phaenicia*," *J. Comp. Physiol.*, vol. 135, pp. 29–39, 1980.
- [15] M. Egelhaaf and W. Reichardt, "Dynamic response properties of movement detectors: theoretical analysis and electrophysiological investigation in the visual system of the fly," *Biol. Cybernetics*, vol. 56, pp. 69–87, 1987.

- [16] N. Franceschini, J. Pichon, C. Blanes, “Real-time visuomotor control: from flies to robots,” in *IEEE Fifth Intl. Conf. on Advanced Robotics, June 1991, Pisa, Italy*, pp. 91–95, 1991.
- [17] N.M. Grzywacz and A.L. Yuille, “A model for the estimate of local image velocity by cells in the visual cortex,” *Proc. Royal Soc. Lond.*, vol. B 239, pp. 129–161, 1990.
- [18] B. Hassenstein and W. Reichardt, “Systemtheoretische analyse der Zeit-, Reihenfolgen- und Vorzeichenbewertung bei der Bewegungsperzeption des Rüsselkäfers *Chlorophanus*,” *Z. Naturforsch.*, vol. 11b, pp. 513–524, 1956.
- [19] J.H. van Hateran, “Directional tuning curves, elementary movement detectors, and the estimation of the direction of visual movement,” *Vision Res.*, vol. 30, pp. 603–614, 1990.
- [20] K. Hausen, “Monokulare und binoculare Bewegungsauswertung in der Lobula plate der Fliege (Monocular and binocular computation of motion in the lobula plate of the fly),” in *Verh. Dtsch. Zool. Ges.*, Stuttgart: Gustav Fischer Verlag, pp. 49–70, 1981.
- [21] D.J. Heeger, “Optical flow using spatiotemporal features,” *Int. J. Computer Vision*, vol. 1, pp. 279–302, 1988.
- [22] T. Horiuchi, J. Lazzaro, A. Moore and C. Koch, “A delay line based motion detection chip,” in *Advances in Neural Information Processing Systems 3*, R. Lippman, J. Moody, D. Touretzky, Eds., San Mateo, CA: Morgan Kaufmann, pp. 406–412, 1991.
- [23] T. Horiuchi, W. Bair, B. Bishofberger, J. Lazzaro, J. and C. Koch, “Computing motion using analog VLSI Chips: an experimental comparison among different approaches,” *Int. J. Comp. Vision*, vol. 8, pp. 203–216, 1992.
- [24] S.R. Lehky and T.J. Sejnowski, “Neural model of stereoacuity and depth interpolation based on a distributed representation of stereo disparity,” *J. Neurosci.*, vol. 10, pp. 2281–2299, 1990.
- [25] T. Maddess, “Afterimage-like effects in the motion-sensitive neuron H1,” *Proc. R. Soc. Lond.*, vol. B-228, pp. 433–459, 1986.
- [26] M.A. Mahowald, *VLSI Analogs of Neuronal Visual Processing: A Synthesis of Form and Function*, Ph.D. Thesis, California Inst. of Tech., Dept. of Computation and Neural Systems, Pasadena CA, 1992.
- [27] M.A. Mahowald, “Silicon retina with adaptive photoreceptors,” in *Proc. SPIE/SPSE Symposium on Electronic Science and Technology: from Neurons to Chips*, vol. 1473, April 1991.
- [28] D. Marr, *Vision*, New York: W.H. Freeman & Co., 1982.
- [29] C.A. Mead and T. Delbrück, “Scanners for visualizing activity of analog VLSI circuitry,” *Analog Integrated Circuits and Signal Processing*, vol. 1, pp. 93–106, 1991.
- [30] C.A. Mead, *Analog VLSI and Neural Systems*. Reading, MA: Addison–Wesley, 1989.
- [31] K. Nakayama and G.H. Silverman, “Temporal and spatial characteristics of the upper displacement limit for motion in random dots,” *Vision Res.*, vol. 24, pp. 293–299, 1984.
- [32] J.M. Pichon, C. Blanes, and N. Franceschini, “Visual guidance of a mobile robot equipped with a network of self-motion sensors,” in *Proc. SPIE*, vol. 1195, pp. 44–53, 1989.

- [33] W. Reichardt, "Evaluation of optical motion information by movement detectors," *J. Comp. Physiol.*, vol. 161, pp. 533–547, 1987.
- [34] R.R. de Ruyter van Steveninck, W.H. Zaagman, and H.A.K. Mastebroek, "Adaptation of transient responses of a movement-sensitive neuron in the visual system of the blowfly *Calliphora erythrocephala*," *Biol. Cybernetics*, vol. 54, pp. 223–236, 1986.
- [35] J.E. Tanner and C. Mead, "An Integrated Analog Optical Motion Sensor," in *VLSI Signal Processing, II*, S.Y. Kung, Ed., New York: IEEE Press, pp. 59–76, 1986.
- [36] J.E. Tanner, *Integrated Optical Motion Detection*. Ph.D. Thesis, California Inst. of Tech., Dept. of Computer Science, Pasadena CA, 1986.
- [37] P.A. Viola, S.G. Lisberger and T.J. Sejnowski, "Recurrent eye tracking network using a distributed representation of image motion," in *Advances in Neural Information Processing Systems 4*, pp. 380–387, 1992.
- [38] W.H. Zaagman, H.A.K. Mastebroek, and R.R. Van Steveninck, "Adaptive strategies in fly vision: on their image-processing qualities," *IEEE Trans. Systems, Man, and Cybernetics*, vol. SMC–13, pp. 900–906, 1983.
- [39] J.M. Zanker, "On the directional sensitivity of motion detectors," *Biol. Cybernetics*, vol. 62, pp. 177–183, 1990.

