Computing Motion Using Analog VLSI Vision Chips: An Experimental Comparison Among Four Approaches

Christof Koch[1], Andrew Moore[1], Wyeth Bair[1],
Timothy Horiuchi[1], Brooks Bishofberger[1] and John Lazzaro[2]

[1]Computation and Neural Systems Program, 216-76, California Institute of Technology, Pasadena, CA 91125 USA.
[2]University of Colorado at Boulder, Campus Box 425, Boulder, CO 80309

Abstract

We have designed, built and tested a number of analog CMOS VLSI circuits for computing 1-D motion from the time-varying intensity values provided by an array of on-chip phototransistors. We present experimental data for three such circuits and discuss their relative performance. One circuit approximates the correlation model, one the gradient model, while a third chip uses resistive grids to compute zero-crossings to be tracked over time by a separate digital processor. All circuits integrate image acquisition with image processing functions and compute velocity in real time. Finally, for comparison, we also describe the performance of a simple motion algorithm using off-the-shelf components.

1 Introduction

There exist two broad categories of algorithms for recovering the optical flow field underlying the time-varying intensity patterns falling onto a retina or camera (for an overview, see [38]).[1] *Token-based* or *long-range* schemes identify low or high-level tokens in every image, such as edges, corners or missiles, and track these from frame to frame [37] [24]. They lead to an estimate of motion at sparse locations in the image. The key problem associated with these schemes is the *correspondence* problem, i.e. matching up the tokens in frame i with the same tokens in frame $i+1$. The difficulty associated with the correspondence problem decreases with sampling interval Δt between frames and with number of tokens.

The second class of motion algorithms uses the intensity or a linear function of the intensity at every location to compute the optical flow field throughout the image. During the last decade there has been an increasing interest in these *intensity-based* or *short-range* methods (for a review see [11] and [32]). The two main approaches that have been proposed for determining the optical flow are the *correlation, second-order or spatio-temporal energy* methods [8] [30] [1] [40] [43] [33] [3] [5] and the *differential methods* [29] [6] [15] [36] [10] [44] [39] [42]. Common to all correlation methods is that the intensity $I(x,y,t)$ is passed through a linear spatio-temporal filter and multiplied with a delayed version of the filtered intensity from a neighboring receptor [30]. The output of these methods is a quadratic functional from which velocity or speed has to be extracted. A number of population coding schemes have recently been proposed for this purpose [9] [7]. Gradient methods, on the other hand, exploit the relationship between the velocity and the ratio of the temporal to the spatial derivative. These methods yield a direct estimate of the optical flow field. However, they require evaluation of first or second-order spatial and temporal derivatives of the image intensities.

In two dimensions, most intensity-based methods are plagued by the aperture problem in either its strong or its weak form [24] [32]. For instance, the standard gradient scheme only estimates the velocity component perpendicular to the local image gradient [15]. Thus, in these cases computing optical flow is an ill-posed problem, since no unique solution exist [31]. In other cases (e.g. [33] [39]), the optical flow is merely ill-conditioned, that is, it depends on the exact knowledge of the initial values. In most cases a regularization step is imposed so that the final optical flow obeys some sort of smoothness constraint.

Common to all these methods is a large associated computational overhead, preventing real-time machine vision applications within most industrial, military or deep space/planetary settings except on anything but large, costly and power-hungry computers. Special-purpose hardware for computing optical flow in real-time becomes therefore a very attractive possibility. Here at Caltech, the laboratories of Carver Mead as well as ours have focused on a special class of such vision systems, analog, non-clocked CMOS VLSI circuits with on-chip photoreceptor arrays [27] [16] (for an alternative approach see [14]). A number of working chips, integrating image acquisition with different early vision algorithms, such as filtering, edge detection, binocular stereo and surface interpolation, have been designed, fabricated (via the government silicon

[1]Given the topic of this paper, we make no distinction here between the optical flow field induced by the time-varying image intensities and the underlying 2-D velocity field, a purely geometrical concept [13] [41].

TH0390-5/91/0000/0312/$01.00 © 1991 IEEE

foundry MOSIS) and successfully tested (for an up-to-date overview, see [25]). We here discuss three different analog circuits for computing the 1-D optical flow associated with the on-chip 1-D photoreceptor array. The next section discusses a chip approximating a Reichardt correlation algorithm, while section 3 presents data from a mixed analog-digital circuit. This system, tracking thresholded zero-crossings, bears similarity to the Marr & Ullman [24] scheme of computing velocity along zero-crossings of the $\nabla^2 G$ operator. Section 4 describes the performance of the Tanner-Mead motion detector circuit [35] in a novel mode of operation. Finally, in section 5, we compare the performance of these three analog, non-clocked chips with that of a simple system built out of a 1-D CCD imager and a programmable microprocessor.

1.1 Data acquisition and circuit design

When testing the performance of our different motion chips, we tried to directly compare their output under the same test conditions, in particular using the same stimulus and speed as well as background intensity. Accordingly, we built a conveyor belt system using an electric motor; belts with square wave gratings of various contrasts and spatial frequencies could be moved in view of the chips, with velocities that ranged over more than an order of magnitude. Moving stripe patterns were imaged onto the silicon surface using a narrow aperture lens directly positioned onto the chip. However, we did not achieve our initial goal of comparing all of the chips under identical operational conditions. This was mainly due to the fact that the different circuits have different optimal operating characteristics (e.g., some operate best under very low light conditions while others require higher light intensities).

All the data shown in this article is based upon measured data from working chips and not from circuit simulations. The chips were implemented with a standard 2.0 μm CMOS process available through the MOSIS silicon foundry.

Finally, we would like to mention that most of our circuits are directly or indirectly inspired by biological counterparts in the motion pathway of flies, rabbits or primates. In fact, it has been our experience that thinking about biological motion estimation systems (e.g. [17] [42]) leads to the design of more robust electronic circuits, while thinking about machine vision systems leads to a better understanding of the problems — such as gain-control, limitations in the precision of components — that any biological vision system must face.

2 A Pulse-Coded Correlation Circuit

The circuit discussed in this section was directly inspired by the correlation model as well as by the computational architecture found in the auditory system of owls [19]. We designed an analog VLSI chip that contains a large array of velocity-tuned units that correlate two events in time, using a delay-line structure [12]. In building motion detection systems using correlation methods, a typical approach with a clocked system measures the image shift over a fixed sampling time, while a dedicated analog hardware approach lends itself to the measurement of image shift time over a fixed distance. The latter is a local computation that gracefully scales to different velocity ranges without suffering from the problems of extended interconnection. It is this local property that we are using to compute motion.

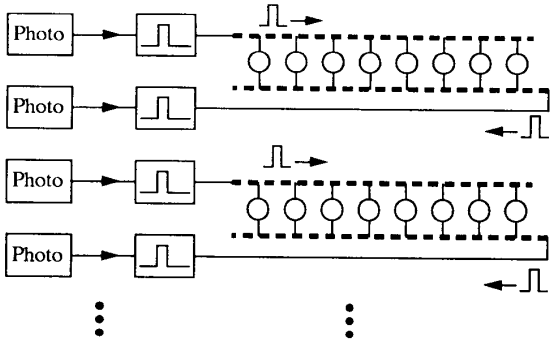

Figure 1. Block diagram of the pulse-coded correlation chip, showing two motion detector units. Rising light intensity signals at the photoreceptors are converted into pulses and sent down the delay line. Velocity is determined by the location where two pulses meet. The axon delay-line is drawn as a heavy dashed line and correlators are drawn as circles. See the text for explanation.

2.1 System architecture

Fig. 1 shows the conceptual design of the detector and the organization of the chip. The image is directly focussed onto a one-dimensional array of 28 on-chip hysteretic photoreceptors spaced 50μm apart [4]. These photoreceptors enhance temporal changes in the incident light intensity. While configured like a follower, the circuit has its highest gain at higher frequencies. Additionally, the circuit has a compressive gain function for the amplitude of the signal, making it responsive to both small and large signals. Each photoreceptor is connected to a half-wave rectifying neuron circuit [21] that fires a single digital pulse of constant duration when it receives a quickly rising (but not falling) light-intensity signal. The duration of the pulses can be adjusted from approximately 1 $msec$ to 0.08 $msec$.

This rising light intensity signal is interpreted as a moving edge in the image passing over the photoreceptor. This signal is the image feature to be correlated. Note that from a computational point of view, we can use either the rising or the following intensity values, corresponding to an ON or to an OFF edge, as the feature to be correlated with. Due to the faster turn-on characteristics of the photoreceptor, however, a rising signal was chosen. Each neuron circuit is connected to an axon circuit [27] that propagates the pulse down its length. By orienting the axons in alternating propagation directions, as shown in Fig. 1, any two adjacent receptors generate pulses that will "race" toward each

other and meet at some point along the axon. Correlators between the two opposing axons detect when the two pulses pass each other, indicating the detection of a specific time difference. The width of the pulse in the axon circuits, which is adjustable, determines the pulse propagation rate down the line; the propagation rate determines the detectable velocity range.

2.1.1 Computing coincidences

To compute motion, the system measures the time that a feature takes to travel from one photoreceptor to one of its neighbors. By placing two delay lines in parallel that propagate signals in opposing directions, a temporal difference in signal start times from opposite ends will appear as a difference in the location where the two signals will meet. Between the axons, correlation units perform a logical AND with the axon signals on both sides. If pulses enter adjacent axons with zero difference in start times (i.e. infinite velocity), they meet in the center and activate a correlator in the center of the axon. If the time difference is small (i.e. the velocity is large), correlations occur near the center. As the time difference increases, correlations occur further out toward the edges. The left and right halves of the axon represent different directions of motion. At the chip level, when a single stimulus (e.g. a step edge) is passed over the length of the photoreceptor array with a constant velocity, a specific subset of correlators are activated that all represent the same velocity.

The major sources of error in the computation are related to fabrication offsets and noise. Component non-uniformities in the axon cause the pulses to be of slightly different durations which changes the propagation speeds at each location in the axon. This can shift the resulting correlation position, with the accumulated errors being the largest at longer time differences. Given the circuit component variability, we compute these coincidences for 27 pairs of opposing axons. A current summing line is connected to each of these correlators and is passed to a winner-take-all circuit [22] as one of the competing time delay channels. The winner of the winner-take-all computation corresponds to the bin that is receiving the largest number of correlated inputs. The output of the winner-take-all is scanned off the chip using an external input clock. Because the frequency of matches affects the confidence of the data, scenes that are denser in edges provide more confident data as well as a quicker response.

Note that the scheme we use to compute the velocity—by estimating the coincidence event receiving the maximal amount of support—approximates the "ridge" strategy Grzywacz and Yuille [7] advocate to compute velocity from a population of spatiotemporally oriented receptive fields.

2.1.2 Single versus bursting mode

The circuit described thus far uses a single pulse to indicate a passing edge. Due to the statistical nature of this system, a large number of samples are needed to make a confident statement of the detected time difference. By externally increasing the amplitude of the signal passed to the neuron during each event, the neuron can fire multiple pulses in quick succession. With an increased number of pulses traveling down the axon, the number of correlations increases, but with a decrease in accuracy, due to the multiple incorrect matches. The incorrect correlations are not random, however, but cluster around the correct velocity. The end result is a net decrease in resolution in order to achieve increased confidence in the final output.

The chip output is the measured time difference of two events in multiples of τ, the time-delay of a single axonal section. The final velocity v is given by $const/\Delta t$, where Δt corresponds to the signed time difference (measured in seconds/pixel). We set $const = 1$. Due to this inverse relationship, we expect to obtain the highest velocity resolution for slow speeds. However, due to the relatively small number of correlations at slower speeds, the signal-to-noise ratio will decrease. This will be less troublesome as larger photoreceptor arrays are implemented. The variable resolution in this computation can be an acceptable feature for control of robotic motion systems since higher velocity motions are often coarse, with fine control needed only at the lower velocities.

2.2 Performance

We fabricated the circuit shown in Fig. 1 using a double polysilicon $2\mu m$ process on a MOSIS Tiny Chip die containing about 8000 transistors. The chip has 17 velocity channels (8 channels in each direction as well as a center channel), and an input array of 28 photoreceptors. The voltages from the winner-take-all circuit are scanned out sequentially by on-chip scanners, the only clocked circuitry on the chip.

In testing the chip, gratings of varying spatial frequencies and natural images from newspaper photos and advertisements were mounted on a rotating drum in front of the lens. Although the most stable data was collected using the gratings, both image sources provided satisfactory data. Fig. 2(a) shows the winning time interval channel vs. actual time delay. The response is linear as expected. In Fig. 2(b), the data from Fig. 2(a) is converted into a measured velocity vs. input velocity plot. As described above, at the lower velocities, correlations occur at a lower rate, thus some of the lowest velocity channels do not respond. This is interpreted as zero velocity. Increasing the number of parallel photoreceptor channels will improve this situation. The circuit has been shown to measure, with varied settings of the axonal unit time constant, velocities from about 50 pixels/sec to over 1150 pixels/sec. Any given setting will measure a range of velocities just over one order of magnitude.

The response time for the chip depends strongly upon the type of stimuli used and the velocity range detected. The number of correlations per second determines the confidence in the channel, which determines the strength with which the channel drives the winner-take-all circuit. For faster velocities or stimuli with more edges, the chip will respond more quickly than for the slower speeds or sparser stimuli. While strongly dependent upon many parameter settings and

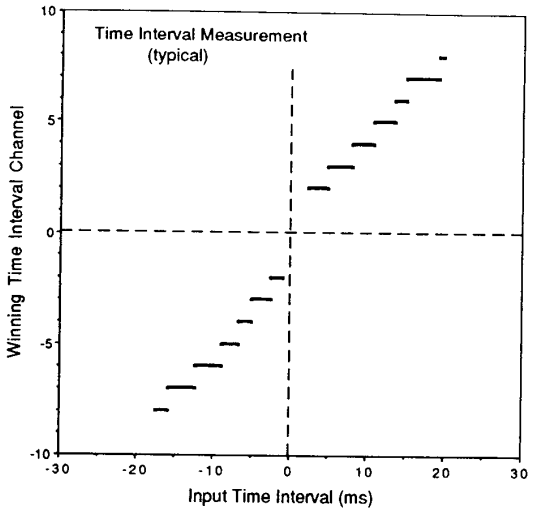

Figure 2a. Plot of winning output channel vs. input time interval. Each output channel represents a difference in the axon start times of 2τ. The highest velocities correspond to the shortest time interval. The horizontal shift in the negative velocities is believed to be due to a propagation delay in the long signal line on the chip. Its effect can be seen in the slightly different slopes in Fig. 2b.

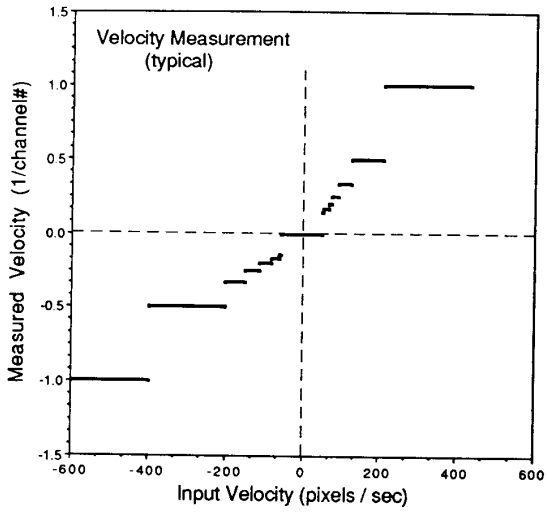

Figure 2b. Plot of measured velocity vs. input velocity using data from Figure 2a. Each output channel represents a specific velocity which is related to the inverse of the difference in axon start times. This relationship maps a large range of velocities into the short time interval channels, giving a coarse resolution at higher speeds. We only show the response of the chip up to about ± 500 pixels/sec, since our conveyor belt system fails to move at higher speeds.

conditions, typical response times run from 0.5 *sec* to 2 *sec*.

The performance under differing light levels depends primarily upon the ability of the photoreceptor and feature extraction circuit to deliver reliable feature detection signals. As described before, the hysteretic photoreceptor is extremely sensitive to both large and small changes in the intensity and allows the chip to operate at quite low light levels. Usable data was obtained with DC illumination from 1 mW/m^2 up to 1000 mW/m^2 over various gain settings of the coupling circuit between the photoreceptor and axon circuits.[2] With any particular gain setting it is possible to operate reliably over slightly greater than one order of magnitude of light intensity. The limiting factor for illumination is at the higher end where the dc level of the photoreceptor begins to reduce the amount of signal that is coupled into the neuron circuit.

2.3 Summary

Our implementation of the correlation model shows promise due to its relative robustness to light levels and contrast. Some of the issues to be discussed include flicker sensitivity, noise, velocity range, and possible design expansion.

The first and most limiting aspect of this particular circuit is the feature extraction circuit. The hysteretic photoreceptor is intended to enhance the temporal changes in the light signal and thus detect edges. This circuit is extremely sensitive and allowed the operation of the circuit to continue down into very low light levels and low contrast stimuli. However, temporally differentiating circuits in hardware are troublesome due to noise amplification, and in our circuit this manifests itself in the form of flicker sensitivity. Under fluorescent and some AC incandescent lighting, all of the photoreceptors' circuits fire consistently at $120Hz$, indicating infinite velocity and thus making it unusable in such lighting. A modification of the circuit to include a more sophisticated feature extraction stage would eliminate this problem.

The statistical nature of the computation allows the system to perform successfully in the presence of noise as well as to produce a usable confidence level measure. By summing votes for specific velocities across the chip and by using the burst-mode described above, it is possible to obtain a strong signal above the noise. If a different method for extracting the detected time difference were used in the place of the winner-take-all circuit, the current levels in each of the summing lines would provide a confidence level for each particular channel. It is interesting to note that despite the apparent loss of resolution caused by bursting, the confidence level measure can provide additional information to allow interpolation between the discretized velocity outputs.

[2]Note that the solar-constant is about 1400 W/m^2 while a value of 1 mW/m^2 corresponds to candle-light illumination.

A natural next step in developing motion detection circuits is the design of a 2-D array of motion detection units in order to integrate motion over an array. It should be remembered, however, that this particular circuit exploits the second spatial dimension on the silicon to represent time, making it necessary to use three dimensions to build a similarly designed 2-D motion detector.

3 Motion from Zero-crossings

This system computes motion by using an analog chip to localize zero-crossings and a digital microprocessor to track the zero-crossings. It approximates the scheme proposed by Marr & Ullman [24], but without their use of X and Y cells. The analog chip is a one-dimensional 64 pixel device which exploits on-chip photoreceptors and the natural filtering properties of resistive networks to implement an edge-detection scheme similar to the Difference of Gaussians (DOG) operator proposed by Marr and Hildreth [23]. The chip localizes the zero-crossings associated with the difference of two exponential weighting functions, and reports the locations of only those zero-crossings which have a slope greater than an adjustable threshold. A conventional digital microprocessor receives the locations of the zero-crossings from the analog chip and tracks their displacements over time to compute velocity.

3.1 The analog VLSI zero-crossing chip

Similar to a DOG, our chip takes the difference of two filtered versions of the input light intensity, but we avoid the difficulties associated with implementing Gaussian kernels in silicon and filter with first-order resistive networks instead. In these networks, each node is connected to an input data voltage via a conductance G and to its neighbors via resistances R. The characteristic length, corresponding to the standard deviation σ of a Gaussian, of the resulting filter function is given by $\lambda = 1/\sqrt{RG}$. The Green's function of the resistive network, a symmetrical decaying exponential, i.e. $e^{-|x|/\lambda}$, differs somewhat from the Gaussian filter, in particular at $x = 0$. Two resistive networks with different values of λ, achieved by using different resistances, then implement a difference-of-exponentials, or DOE, operator. This filter has some similarities to a $\nabla^2 G$ operator; for instance, the output of this DOE operator to a constant input is zero (for more details see [2]). The rounded peak of the Gaussian makes the DOG look like a "Mexican-hat", while the pointed peak of the decaying exponential makes the operator implemented by our chip appear more like a pointed "Witch-hat". After filtering the input image with the DOE operator, the chip localizes zero-crossings which ideally correspond to edges in the image and object boundaries in the scene. The entire process, from imaging to edge detection, occurs on-chip in four stages of analog circuitry: photoreceptors capture incoming light, a pair of 1-D resistive networks smooth the input image, transconductance amplifiers subtract the smoothed images, and mixed analog and digital circuitry localizes and thresholds the zero-crossings. Fig. 3 shows a block diagram of the first three stages of this processing, which is described in more detail below.

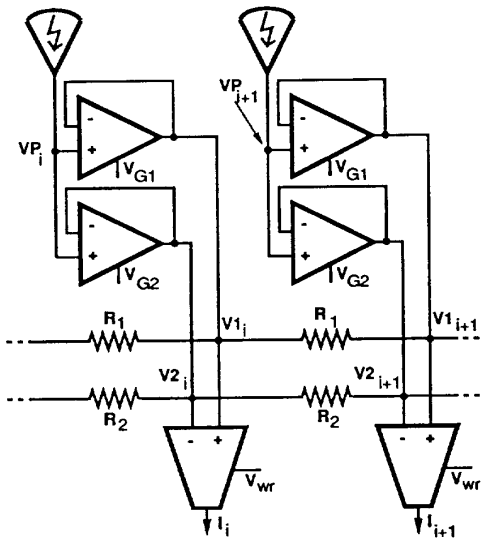

Figure 3. Zero-crossing chip circuit diagram. Logarithmic photoreceptors encode light intensity as voltages, VP, which are reported to the nodes of two resistive networks via transconductance amplifiers connected as followers. The voltage biases VG set the conductances. The network resistances, $R1$ and $R2$, are implemented as saturating resistors and are externally adjustable from voltage biases. The filtered images are subtracted by wide-range transconductance amplifiers which output a current, I, proportional to the voltage difference across their inputs.

The chip receives input from an array of photoreceptors spaced 100μm apart, encoding the logarithm of light intensity as a voltage. The set of voltages from the photoreceptors are reported to corresponding nodes of two resistive networks via transconductance amplifiers connected as followers. The followers' voltage biases can be adjusted off-chip to independently set the data conductances for each resistive network. The network resistors are implemented using saturating resistors developed by Mead [27]. Another pair of voltage biases allow independent off-chip adjustment of the resistances along the two resistive networks. The data conductance and network resistance values determine the space constant of the smoothing filter which each network implements. The sets of voltages along the networks represent the two filtered versions of the image. These filtered images are subtracted by wide-range transconductance amplifiers [27] which produce an output current proportional to the difference in voltage applied across their inputs. The array of currents produced by this circuitry corresponds to the result of applying the DOE operator to the input image.

The final stage of processing detects zero-crossings in the array of currents from the wide-range amplifiers and implements a threshold on the slope of those zero-crossings. Each pair of neighboring currents charge or discharge the inputs of an exclusive OR gate. The binary output of this gate indicates the presence or absence of a zero-crossing between two nodes. A second signal is generated by subtracting a threshold current from the magnitude of the difference between the neighboring currents mentioned above. If the charging current, representing the slope of the zero-crossing, is greater than the threshold current set by an off-chip bias voltage, then this signal charges a node to logical 1, otherwise, that node is discharged to logical 0. The conjunction of a zero-crossing and a steep slope causes the chip to report the existence of an edge at that location for any of 63 possible locations. The output can be thought of as a 63 bit word where each bit codes for the presence or absence of a zero-crossing at that particular location.

3.2 Data from the zero-crossing chip

Fig. 4 shows data taken from the zero-crossing chip. The input light profile is a bright bar. Oscilloscope traces show the filtered versions of the image from the nodes along the resistive networks. By setting the space constants of the networks differently, we have achieved varying amounts of smoothing. The difference of these two smoothed voltage traces is shown in Fig. 4c; arrows indicate the locations of two zero-crossings which the chip reports at the output. The reported zero-crossings accurately localize the positions of the edges in the image. Other zero-crossings were not reported because their slopes were less than the adjustable threshold. Thresholding allows for noise and imperfections in the circuitry and can be used to filter out weaker edges which are not relevant to the application.

Fig. 5 shows the response when two fingers are held 1 m in front of the lens and swept across the field of view. The fingers appear as bright regions against a darker background. The chip accurately localizes the four edges (two per finger) as indicated by the pulses below each voltage trace. As the fingers move quickly back and forth across the field of view of the chip, the image and the zero-crossings follow the object with no perceived delay. From sequences of frames like these, we can compute optical flow. Note that these are not successive frames, but are more representative of every 100-th frame that the motion detection system will receive (see below).

3.3 The microprocessor and motion detection

The motion detection system consists of a zero-crossing chip interfaced to a 12.5 MHz 80286 microprocessor-based single-board computer. The interface allows the microprocessor to receive 63-bit frames of zero-crossing data at just over 320 frames per second. As each new frame is read, the microprocessor updates the cumulative displacement of each zero-crossing and increments the number of frames over which that displacement has occurred. The system assumes that zero-crossings will not move more than

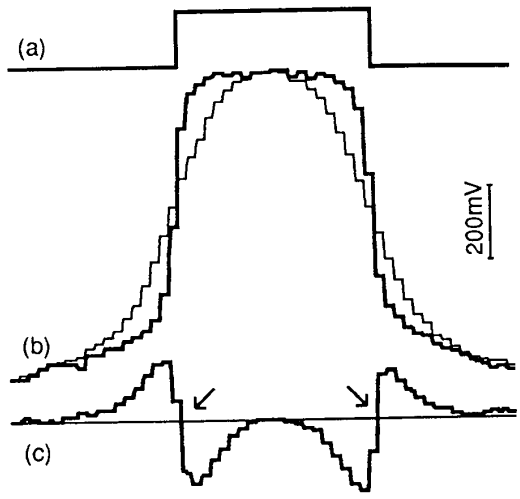

Figure 4. Measured response of the zero-crossing chip to a light bar stimulus. (a) Input light intensity, (b) voltage traces from the two resistive networks, and (c) difference of the voltage traces, corresponding to the image intensities convolved with a difference-of-two-exponentials (DOE) operator. The circuit correctly localizes the location of the two edges (arrows). The threshold suppresses zero-crossings with small magnitude slope.

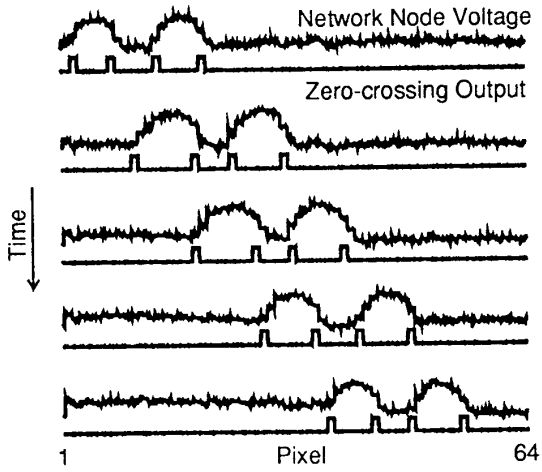

Figure 5. Zero-crossing chip response as two fingers are waved about 1 m in front of the lens. The upper traces show voltages from one resistive network; the lower traces show positions of zero-crossings reported by the chip.

2 pixels per frame. With our optics, this assumption is violated only at velocities in excess of approximately 700 degrees per second.

After tracking zero-crossings for a fixed number of frames, their individual velocities are computed in pixels per frame. These velocities are averaged for all zero-crossings which have been tracked for longer than a fixed number of frames. For the data shown here, an average full-field velocity is reported every second. Fig. 6 shows the average and standard deviation of the reported velocity over a 1 *min* period for a range of input velocities.

3.4 Performance analysis

Fig. 6 shows the output of the system for image velocities ranging from 0 to 450 pixels per second at two different light levels. The error bars show the standard deviation of the output velocity. Over most of this range, the standard deviation was less than four percent of the average value. Image velocity was limited by the lens and stimulus. The data shown for 10 W/m^2 is representative of the system response for light levels of 1 W/m^2 and higher. Below 1 W/m^2, the zero-crossing chip was unable to localize higher velocity zero-crossings. We believe this is due to R-C time constants associated with the circuitry of the analog chip. Also, as seen in Fig. 6, the reported velocity is less than the image velocity but remains linear. At lower light levels, zero-crossings due to offsets are more prevalent and introduce zeros into the average velocity computation, thus lowering the reported velocity. Such spurious zero-crossings can undermine the accuracy of the average velocity in more subtle ways as well. As light intensity drops, the linear range of output for this system becomes smaller around zero. Below 100 mW/m^2, the zero-crossing chip fails to detect edges, and the system cannot even detect direction of motion. Qualitatively, the useful range of operation for this system is from bright sunlight to dim indoor fluorescent or incandescent lighting. This range is achieved without changing gain or other parameters.

The zero-crossing chip fails at low light and contrast levels due to the small signal-to-offset ratio. Imperfections in the fabrication process cause many of the signals in the analog chip to be corrupted. The magnitude of this noise, called offsets, is a substantial fraction of the magnitude of the signal reported by the logarithmic photoreceptors. Although the logarithmic receptor allows operation over a wide range of lighting conditions, it compresses the range of voltages which are used to encode any particular scene and therefore decreases the signal-to-noise ratio. A hysteretic photoreceptor similar to the one used in the correlation chip described in the previous section would improve the signal-to-noise ratio, but would also increase sensitivity to lighting changes, and possibly compromise sensitivity to small velocities.

Another limitation in the performance of the zero-crossing chip is the photoreceptor response time. The measured response time of the chip to the appearance of a detectable discontinuity in light intensity varies from about 100 μsec in bright indoor illumination to about 10 $msec$ in a dark room, and these response

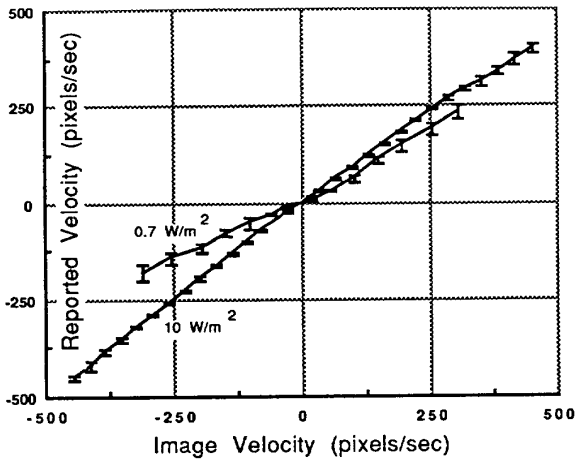

Figure 6. Zero-crossing motion detection system output for two light intensities. At higher light intensities (about 10 W/m^2 and above) the output is linear and accurate over a large range of velocities. At lower intensities, the zero-crossing chip cannot localize fast edges, and lower signal-to-offset ratios introduce spurious zero-crossings that compromise accuracy (error bars show standard deviation).

times seem to be dominated by the phototransistor.

Finally, spatial and temporal aliasing limit the performance of this system. As the spatial frequency of features increases, zero-crossings appear closer together and the correspondence problem arises. This is a function of the environment, the lens, and the photoreceptor spacing on the chip. Interfacing the zero-crossing chip to a digital computer requires clocking the output from the chip. In theory, this causes temporal aliasing at higher velocities, but the slow time response of the photoreceptors cause the system to fail before temporal aliasing is noticed.

4 Computing Motion from Derivatives of Intensity

The Tanner and Mead optical motion detector circuit [34] [35] [36] implements the gradient model, in which the ratio of the temporal derivative to spatial derivative is computed. The circuit was originally tested with idealized inputs over a limited range of light levels. Our goal was to put a lens on the chip and to demonstrate a robust response to a wide range of real-world moving scenes. As developed below, a multiplication rather than a ratio of derivatives is used to achieve this goal with this circuit.

4.1 Modes of operation of the Tanner-Mead motion detection circuit

The gradient model of motion estimation relies on the assumption that the total derivative of intensity over time does not change. If we assign a unique velocity v to any given location and assume that the

intensity profile $I(x,t)$ translates rigidly in the plane, we have in one dimension:

$$\frac{dI}{dt} = I_x v + I_t = 0, \tag{1}$$

where I_x and I_t are the spatial and temporal intensity derivatives. This can be written as

$$v = -\frac{I_t}{I_x}. \tag{2}$$

In analog circuitry, no simple way exists to directly divide two signed variables, i.e., to directly evaluate $z = x/y$. Instead, we use a multiplication—easy to implement—within a feedback circuit, such that the overall circuit minimizes $(z \cdot y - x)^2$. At the minimum, $z \cdot y = x$ and the signed variable z corresponds to the desired ratio. The original Tanner-Mead motion detector is more complicated since it computes this ratio for a two-dimensional input. In one dimension, however, the above simplification applies, and the analogous one-dimensional circuit computes the value of v that minimizes

$$\sum_i \left(I(i)_x v + I(i)_t\right)^2, \tag{3}$$

where the sum is taken over all locations i. The underlying assumption is that the velocity v is constant over the patch containing all locations i. When this functional is minimized, the velocity is the correct ratio of derivatives. The associated Euler-Lagrange equation that the circuit is computing is

$$\sum_i \left(I(i)_x v + I(i)_t\right) I(i)_x = E \tag{4}$$

When the term $E \mapsto 0$, eq. 3 is minimized. The second multiplication by I_x eliminates the singularity that would otherwise result when $I_x = 0$, i.e., when there is no detail in the scene. This was a practical point in the design that also has elegant mathematical underpinnings.

Two constants in the actual chip scale the additive terms of this expression: K_1 weights the time derivative I_t, while K_2 weights the error signal fed back as velocity. In terms of these parameters, the equation for the signal E is

$$\sum_i \left(e^{K_2} I(i)_x v + e^{K_1} I(i)_t\right) I(i)_x = E \tag{5}$$

These exponential factors originate in the exponential dependency of the source-drain current on the gate voltage in the subthreshold domain we operate our device in. For $e^{K_1} \gg e^{K_2}$, eq. 5 reduces to a multiplication of the spatial and the temporal derivatives (the associated gate voltages are $K_1 < 0.5$Volt and $K_2 \approx 0.8$Volt). The feedback loop is then broken, since the value fed back for v has no effect on the error signal. While the division mode of operation has been demonstrated successfully [34] [35] [36] for idealized inputs, the multiplication mode of operation gives more robust output for general stimuli and so it is the mode of operation explored here.

4.2 Performance

The motion detection circuit operating as a multiplier was tested for moving stripes of various spatial frequencies at different light levels. We fabricated and tested chips with a single one-dimensional motion detection element analogous to the original Tanner-Mead 2D element as well as chips with 30 such elements that compute an aggregate output signal by combining the error signals of all 30 elements. Figure 7 shows the response of the single-element circuit (containing two, spatially offset photoreceptors) to moving, high contrast stripes. At the top the output of the two photoreceptors can be seen. The output shown in the bottom part of the figure is essentially zero for the times when the stripe edge does not fall on the photoreceptors. The noisy character of the response is due to circuitry developed for the division mode of operation; besides introducing this artifact, it does not affect the performance of the chip in the multiplication mode.

In the multiplication mode, with a stimulus consisting of high contrast black-and-white stripes, the circuit produces an output signal which is linear in velocity (over a limited velocity range) when integrated over time. From eq. 5, the output signal is $E = \sum_i I(i)_t \cdot I(i)_x$ in this mode. Each stripe edge produces a single pulse with the correct sign of the velocity; the value of E integrated over time indicates the number of stripes that have gone by in that time. The faster the pattern moves, the greater the integrated value of E. For chips with a single element, integration is achieved off-chip via an RC circuit with a long time constant (1−3 seconds). For multi-element chips, the aggregation process itself performs some temporal integration; further integration off-chip may be used as well.

The response of the chip varies with light level for two reasons. First, the photoreceptor response itself varies with light level. The AC response increases and the DC level drops as the light level increases. Second, when the AC signal from the photoreceptor is larger than a certain level, the amplifiers saturate and the time derivative computation is distorted such that I_t is reported as artificially high. The consequence of this is shown in Fig. 8. Here, the response of a single element is plotted for a range of velocities at high, medium and low light levels. The response is relatively linear at a given illumination, and drops off (but remains linear) as the illumination is decreased. This drop-off is due to distortion in the time derivative computation [28]. The chip does respond, however, over a range of two and a half orders of magnitude of illumination. This is due to the compressive logarithmic nonlinearity of the photoreceptors [26] [27]. These data are for high contrast stripes—the chip responds poorly to targets of medium or low contrast.

Fig. 9 shows the output of a chip with 30 elements computing an aggregate velocity. A motor is driving a wheel with stripes on its rim in view of the chip; the motor is powered with a current from a function

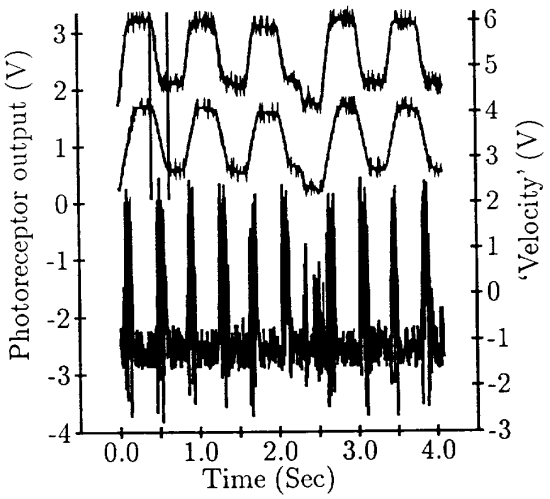

Figure 7. Measured response of a single derivative multiplication element to moving stripes. The top two traces are photoreceptor outputs (one trace has been moved down 1.5 Volt for clarity) and the bottom trace is the product of the temporal and spatial derivative

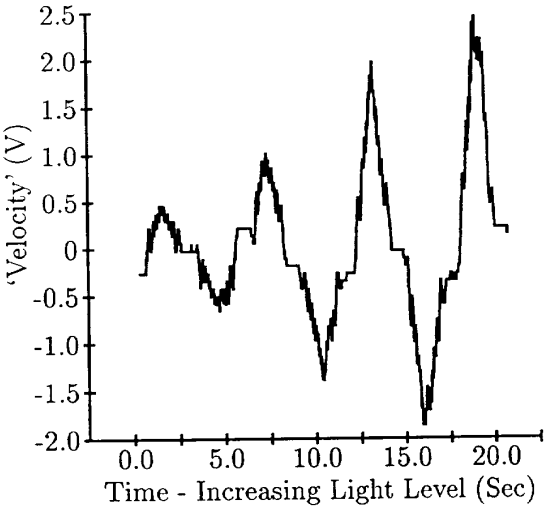

Figure 9. The response of a 30 element motion detection chip for a striped pattern whose velocity varies as a triangle wave. During the course of the measurement period (20 *sec*), a bright light source was brought closer and closer to the chip. This results in the chip

generator set to produce slow triangle waves. The circuit is again linear in the velocity of the pattern—both positive and negative phases of the triangle wave are reproduced well. The common aggregation line serves to temporally integrate the outputs of the elements, smoothing over the noise evident in Fig. 7 (no off-chip temporal integration is performed here). This figure also demonstrates the unwanted side effect that the chip reports higher velocities as the light intensity increases. Increasing the photoreceptor AC output is achieved by bringing the lamp closer and closer to the wheel over a twenty second period (Fig. 9). Mirroring this, the peak velocity output increases in time.[3]

In this mode of operation, the chip evaluates an expression of the form $I_t \cdot I_x$ and responds with a signed pulse to a moving edge. Thus, for our square grating stimulus operating at a fixed light level, doubling the spatial frequency of a striped pattern doubles the pulse rate. A low-pass filtered version of the output, then, should be linear in spatial frequency. Fig. 10 shows this for high-contrast striped patterns at three spatial frequencies and a range of velocities. The relative spatial frequencies of the patterns with 1/4, 1/2 and 1 inch wide stripes are 4, 2 and 1 respectively. The slope roughly doubles from the 1 inch to 1/2 inch patterns and from the 1/2 inch to 1/4 inch patterns. This data was taken from a multi-element chip and so its output is the aggregate (sum) of the error signals from its elements. The summation saturates fastest for the highest frequency, later for the middle frequency, and not at all for the lowest frequency pattern at the highest available test velocity.

4.3 Summary

The principal problems with this motion circuit are the variation in the output with light level and contrast and the very fact that this mode of operation must be used when the chip is used on arbitrary images. Variation of the chip output with spatial frequency derives from its mode of operation.

These problems could be solved in two ways. First, the range of the differential pairs throughout the circuit should be widened. Second, a more sophisticated photodetector should be incorporated in this circuit. DC level clamping and AC gain control at the photoreceptor level would eliminate the variation in velocity output with light level and would extend the sensitivity to lower contrasts. Indeed, with a better input, this circuit would be able to operate as originally intended and spatial frequency dependency would be eliminated.

5 A Fully Digital System

In order to compare our analog circuits against their digital counterparts as well as to be able to quickly test vision algorithms using the reliability of a CCD system, a completely digital circuit was built incorporating a linear 256 pixel element as well as a small microprocessor designed for real-time use. We use a Harris RTX2001A microprocessor operating at 8 MHz. It includes 8K of RAM memory and executes FORTH directly. The system computes a 1-D field velocity based upon a simple correlation method. The CCD (Fairchild Linear CCD) retrieves image data at a maximum rate of 2800 images/second with greater than 12 bits of accuracy. Each pixel is sent through an 8-bit A/D converter, which updates the processor memory at a maximum rate of 2000 images/second. With this structure, the processor is able to access the image residing in memory as rapidly as possible.

[3]Note that in Fig. 9 the dead zones in the velocity output are due to the motor threshold.

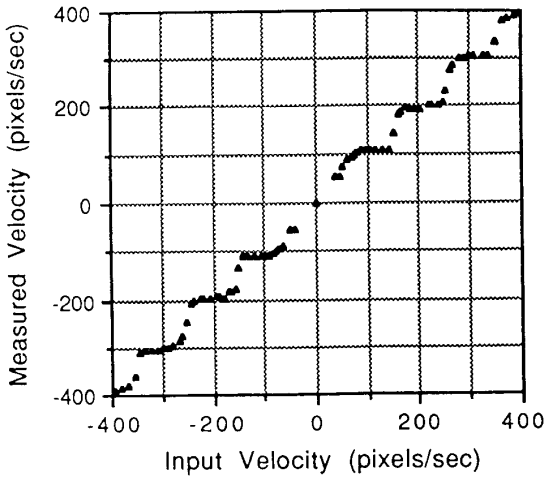

Figure 11. The output of our digital system, using a one-dimensional, commercial 256 pixel CCD camera, is demonstrated here. The reported velocity remain constant as long as the image intensity is above 120 mW/m^2. The waviness of the output is due to the fact that the minimal reported shift is 1 pixel. Interpolation would allow sub-pixel accuracy.

The global image velocity is estimated by first storing two images sampled 10 $msec$ apart. These two images are then subtracted from each other and the absolute value of this difference, summed over the entire 1-D image, is computed. We term this the error associated with a 0 pixel shift. The same operation is also performed when the second image is shifted by $\pm 1, \pm 2, \pm 3$ and ± 4 pixels with respect to the first image. The computer then finds the shift corresponding to the smallest error signal. In other words, it approximates a correlation model in a manner reminiscent of the algorithm of Bülthoff *et al.* [3]. The microprocessor then retains the second image, waits 10 $msec$, stores a new image, and performs the shift comparison again. In the interest of speeding up the algorithm, we did not perform any filtering on the image. In spite of this lack of pre-filtering, the algorithm performed remarkable well. Simple modifications of this algorithm enable the system to also compute the spatially varying optical flow field as well as time-to-contact.

The amplitude of the shift (from -4 to $+4$) is accrued over a period of one second and this sum (number of pixels moved/sec) is plotted against actual im-

[15] B.K.P. Horn and G. Schunck "Determining Optical Flow," *Artif. Intell.* **17:** 185–203, 1981.

[16] C. Koch, "Seeing chips: analog VLSI circuits for computer vision", *Neural Compt.* **1:** 184-200, 1989.

[17] C. Koch, H.T. Wang, B. Mathur, A. Hsu and H. Suarez, "Computing optical flow in resistive networks and the primate visual system," *Proc. IEEE Workshop on Visual Motion*, pp. 62-72, Irvine, March 1989.

[18] C. Koch, W. Bair, J.G. Harris, T. Horiuchi, A. Hsu and J. Luo, "Real-time computer vision and robotics using analog VLSI circuits," In: *Advances in Neural Information Processing Systems Vol. 2.*, D. Touretzky, ed., pp. 750-757, Morgan Kaufmann, San Mateo, CA, 1990.

[19] M. Konishi, "Centrally synthesized maps of sensory space," *Trends Neurosci.* **4**: 163-168, 1986.

[20] S. Laughlin, "The reliability of single neurons and circuit design: a case study." In: *The computing neuron*, R. Durbin, C. Miall and G. Mitchison, eds., pp. 322-336. Addison-Wesley, Menlo Park, CA, 1989.

[21] J. Lazzaro and C. Mead "Circuit models of sensory transduction in the cochlea." In: *Analog VLSI Implementations of Neural Networks*, C. Mead and M. Ismail, eds., pp. 85-101, Kluwer, Norwell, MA, 1989.

[22] J. Lazzaro, S. Ryckebusch, M.A. Mahowald and C. Mead "Winner-take-all networks of O(n) complexity." In: *Advances in Neural Information Processing Systems Vol. 1*, D. Tourestzky, ed., pp. 703-71, Morgan Kaufmann, San Mateo, CA, 1988.

[23] D. Marr and E. Hildreth "Theory of edge detection", *Proc. Roy. Soc. Lond. B* **207:** 187-217, 1980.

[24] D. Marr and S. Ullman "Directional selectivity and its use in early visual processing. *Proc. R. Soc. Lond. B* **211:** 151-180, 1981.

[25] B. Mathur and C. Koch, *Visual Information Processing: From Neurons to Chips*, editors, SPIE, 1991.

[26] C. Mead, "A sensitive electronic photoreceptor," in: 1985 Chapel Hill Conference on Very Large Scale Integration, H. Fuchs, ed., Computer Science Press, Chapel Hill, NC, 1985.

[27] C. Mead, *Analog VLSI and Neural Systems.* Reading, MA: Addison-Wesley, 1989.

[28] A. Moore and C. Koch, "A multiplication-based analog motion detection chip," *Visual Information Processing: From Neurons to Chips*, B. Mathur and C. Koch, editors, SPIE, 1991.

[29] H.H. Nagel "Analysis techniques for image sequences," *Proc. 4th Int. Joint Conf. Pattern Recognition*, Kyoto, Japan, Nov 1978.

[30] T. Poggio and W. Reichardt "Considerations on models of movement detection," *Kybern.* **13:** 223-227, 1973.

[31] T. Poggio, V. Torre and C. Koch, "Computational vision and regularization theory," *Nature* **317:** 314-319, 1985.

[32] T. Poggio, W. Yang and V. Torre, "Optical flow: computational properties and networks, biological and analog." In: *The computing neuron*, R. Durbin, C. Miall and G. Mitchison, eds., pp. 355-370. Addison-Wesley, Menlo Park, CA, 1989.

[33] W. Reichardt, R.W. Schlögel and M. Egelhaaf "Movement detectors of the correlation type provide sufficient information for local computation of 2-D velocity field," *Naturwiss.* **75:** 313-315, 1988.

[34] J. Tanner, *Integrated optical motion detection*, Ph.D. Thesis, Caltech Dept. of Comp. Sci., 1986.

[35] J. Tanner and C. Mead, "An integrated optical motion sensor," *VLSI Signal Processing II*, S-Y. Kung, R.E. Owen, and J.G. Nash, eds., 59-76, IEEE Press, NY, 1986.

[36] J. Tanner and C. Mead, "Optical motion sensor." In: *Analog VLSI and Neural Systems*, C. Mead, pp. 229-255, Addison-Wesley, Reading, MA, 1989.

[37] S. Ullman, *The interpretation of visual motion.* MIT Press, Cambridge, MA, 1979.

[38] S. Ullman, "Analysis of visual motion by biological and computer systems," *IEEE Computer*, **14:** 57-69, 1981.

[39] S. Uras, F. Girosi, A. Verri and V. Torre "A computational approach to motion perception," *Biol. Cybern.* **60:** 79–87, 1988.

[40] J.P.H. van Santen and G. Sperling "A temporal covariance model of motion perception," *J. opt. Soc. Am. A* **1:** 451-473, 1985.

[41] A. Verri and T. Poggio "Motion field and optical flow: qualitative properties," *IEEE PAMI* **11:** 490-498, 1989.

[42] H.T. Wang, B. Mathur and C. Koch "Computing optical flow in the primate visual system," *Neural Comp.* **1:** 92-103, 1989.

[43] A.B. Watson and A.J. Ahumada "Model of human visual-motion sensing," *J. opt. Soc. Am. A* **2:** 322-341, 1985.

[44] A. L. Yuille and N.M. Grzywacz "A computational theory for the perception of coherent visual motion," *Nature* **333:** 71-73, 1988.

7 Conclusion

The work reported here represents an effort over three years to build robust, analog motion sensors with on-chip photoreceptors. We have achieved moderate success in that we are able to compute the global velocity of a 1-D image in real-time.

We are continuing to port our vision chips onto small, highly mobile and autonomous vehicles (toy cars) in order to demonstrate their use as smart sensors in a real-life environment [18]. We are also continuing our quest for more robust motion circuits. Our next major goal, however, is the design of circuits enabling us to compute the 1-D optical flow, i.e. to estimate a velocity vector at different locations across the retina, in order to compute such quantities as the focus of expansion and time-to-contact [13].

Correlation methods are substantially more robust than methods requiring the evaluation of spatial and temporal derivatives and constant photoreceptor output. Of course, correlation methods appear to be used throughout the animal kingdom, from flies to humans [37] [11] [32] [5]. While gradient methods have been very popular in the computer vision community (e.g. [29] [6] [15] [10] [44] [39]), they require very accurate components, such as are available in 16 and 32 bit machines. However, given the limited accuracy we can achieve today in our MOSIS fabricated circuits (no more than $7-8$ bits and possibly less) and given the limited accuracy of neuronal components,[7] it seems unlikely that gradients methods are robust enough to yield the precise optical flow fields required for various vision tasks, such as computing the 3-D velocity field or finding structure-from-motion.

Acknowledgments

We thank Carver Mead for providing laboratory resources for the design and fabrication of the analog chips and Steve DeWeerth, John Tanner, and John Harris for their help throughout the years. Additional thanks to John Harris for critically editing this paper. We thank the ONR, the NSF, Rockwell International Science Center and Hughes Aircraft Corporation for financial support of VLSI research in our laboratory, the Systems Development Foundation, the ONR, and Hewlett-Packard for computing support in the laboratory of Carver Mead and to DARPA for MOSIS fabrication services. A.M. was supported by fellowships from the Parsons Foundation and the Pew Charitable Trust and by research assistantships from the ONR, the Joint Tactical Fusion Program, the Center for Research in Parallel Computation, and the Program in Advanced Technologies, which was sponsored by GM and TRW. W.B. is supported by a NSF Graduate Fellowship and performed some of the work described here at the Hughes Aircraft AI Center.

[7] Laughlin [20] estimates that the light-induced signal generated in fly photoreceptors and transmitted to their next interneuron, the large monopolar cell, has about 80 levels of discrimination. Thus, in a system where redundancy is stripped away and the signal is produced by an array of over 1200 synapses, responding in parallel to 6 highly correlated signals from identical photoreceptors, at best $6-7$ bits of information is available.

References

[1] E.H. Adelson and J.R. Bergen "Spatio-temporal energy models for the perception of motion," *J. opt. Soc. Am. A* **2:** 284-299, 1985.

[2] W. Bair, and C. Koch. "An Analog VLSI Chip for Finding Edges from Zero-crossings." In: *Advances in Neural Information Processing Systems Vol. 3* Touretzky, D.S., Lippman, R., eds., Morgan Kaufmann, San Mateo, CA, in press, 1991.

[3] H. H. Bülthoff, J.J. Little and T. Poggio " Parallel computation of motion: computation, psychophysics and physiology. *Nature* **337:** 549-553, 1989.

[4] T. Delbrück and C. Mead "An Electronic Photoreceptor Sensitive to Small Changes in Intensity." In: *Neural Information Processing Systems 1*, D. Touretzky, ed., pp. 720-727, 1989.

[5] M. Egelhaaf, A. Borst and W. Reichardt "The computational structure of a biological motion detection system," *J. Opt. Soc. Am. A.* **6:** 1070-1087, 1989.

[6] C.L. Fennema and W.B. Thompson "Velocity determination in scenes containing several moving objects," *Comp. Vis. Graph. Proc.* **9:** 301-315, 1979.

[7] N.M. Grzywacz and A. L. Yuille, "A model for the estimate of local image velocity by cells in the visual cortex," *Proc. R. Soc. Lond. B* **239:** 129-161, 1990.

[8] B. Hassenstein and W. Reichardt "Systemtheoretische Analyse der Zeit, Reihenfolgen, und Vorzeichenauswertung bei der Bewegungsperzepion des Rüsselkäfers *Chlorophanus*," *Z. Naturforsch.* **11b:** 513–524, 1956.

[9] J.D. Heeger, "Model for the extraction of image flow," *J. Opt. Soc. Am. A* **4:**1455-1471, 1987.

[10] E.C. Hildreth *The measurement of visual motion*, MIT Press, Cambridge, 1984.

[11] E. Hildreth and C. Koch "The analysis of visual motion: from computational theory to neuronal mechanisms," *Ann. Rev. Neurosci.* **10:** 477-533, 1987.

[12] T. Horiuchi, J. Lazzaro, A. Moore, C. Koch "A Delay-Line Based Motion Detection Chip," In: *Advances in Neural Information Processing Systems Vol. 3*, D. Touretzky, ed., Morgan Kaufmann, San Mateo, in press, 1991.

[13] B.K.P. Horn *Robot vision*, MIT Press, Cambridge, 1986.

[14] B.K.P. Horn, "Parallel networks for machine vision." *Artif. Intell. Lab. Memo* **No. 1071**. MIT, Cambridge, 1989.

age velocity in Fig 11. One can see the step-function where the calculated velocity is rounded to the nearest multiple of integer pixel shifts. In addition, the points between the steps are a result of the averaging effect of sampling for 1 *sec*. This simple algorithm can be altered to better demonstrate that the output is linear. By shifting the image further in each direction, each step will be shallower and the graph will better approximate a continuous line. Furthermore, we can approximate sub-pixel velocities by interpolating between steps by utilizing the computed error function. At any setting, the CCD provides valid output over light intensities differing by more than one order of magnitude, in DC or 120 Hz illumination. The system operates down to 120 mW/m^2, at which point features are no longer detected and the reported velocity drops sharply to zero. Preliminary tests show promising results from zero-crossing detection and from edge tracking schemes. As a result, we look forward to exploring these options and making comparisons between them and others in the near future.

6 Comparison

We have described four methods of motion measurement: pulse correlation on a single analog chip, digital tracking of zero-crossings provided by a single analog chip, derivative multiplication on a single analog chip, as well as digital autocorrelation of grey levels from a CCD imager. In this section, the four methods of motion measurement are compared.

6.1 Operating range

The four methods differ in their ability to measure motion as the lighting conditions are changed. The derivative multiplication chip operates over a range of two to three orders of magnitude of light intensity, but its output is, in general, not directly proportional to velocity and varies with light level. The pulse correlation chip operates over one order of magnitude of light level; this range may be adjusted, and the output does not vary within this range. The zero-crossing chip can provide good zero crossings to a digital system for tracking over two orders of magnitude of light level.[4] The CCD imager used in the grey-level autocorrelation system responds to light over a range of two orders of magnitude; the range of operation can be varied by varying the integration time of the CCD output scheme.[5] The zero-crossing tracking system–using logarithmic photoreceptors–is the most robust of the four in this respect, since it reports consistent velocities as the light level is varied over the greatest range. The pulse correlation chip uses hysteretic photoreceptors, which have very high gain for small changes in intensity. While this endows the chip with superior contrast sensitivity, it also responds strongly to the 120 Hz flicker seen in ordinary room lights. However, we could eliminate this sensitivity by appropriate on-chip filtering.

6.2 Spatial frequency dependency

All of the methods except the derivative multiplication chip report the same velocity as the spatial frequency of striped patterns is varied. The derivative multiplication chip output is linear in spatial frequency, as discussed in section 4.

6.3 Contrast sensitivity

The pulse correlation chip and the grey-level autocorrelation system have the best sensitivity to contrast. The pulse correlation chip owes its superior contrast sensitivity to its hysteretic photoreceptors, which have very high gain for small changes in intensity.[6] The digital grey level autocorrelation method owes its superior contrast sensitivity to the extremely low noise of the CCD that provides its input. The other two systems work best for high contrasts. There exists a trade-off between contrast and light level sensitivity, such that the systems that are more sensitive to low contrasts operate best only within a relatively small range of light levels.

6.4 Power consumption

The lowest-power method is the derivative multiplication chip. It requires only a single off-chip operational amplifier to give velocity readings. Total on-chip power consumption is $< 1mW$, most of which is used in the photoconversion stage. The zero-crossing chip used in the feature tracking method uses slightly more power, since its outputs are scanned off with digital shift registers. Of course, the tracking is performed with a PC which requires a large amount of power in comparison to the analog chips; correlation or tracking circuitry on the same chip as zero-crossing detection is possible and would tremendously reduce the complexity and power requirements of this method. The pulse correlation chip has the next greater power requirement, due to the quasi-digital nature of the delay lines, but it does not require any external components. Finally, the grey level autocorrelation method uses far more power than any of the individual analog chips, since it requires special clocking chips to drive the CCD, analog-to-digital converters, and a microprocessor, but this system still uses less power than the PC interfaced to the zero-crossing-chip.

6.5 Robustness

In terms of overall robustness, we have ranked the methods as follows: the grey-level autocorrelation method is most robust, followed in order of decreasing robustness by the zero-crossing tracking method, the pulse correlation chip, and the derivative multiplication chip. Interestingly, this roughly corresponds to their relative use of power. The CCD imager, a conventional, off the shelf component that forms the front end of the system, is not nearly as noisy as the photoreceptors we use on the analog chips. This is not surprising as many hundreds of man-years of engineering have been devoted to building very accurate CCD cameras.

[4] Note that these are the upper two decades of intensity that we tested. The zero-crossing chip works well outdoors, at image intensities not available in our laboratory.

[5] Note, however, that autogain, a common feature on CCD cameras, would increase this range by several orders of magnitude.

[6] However, the receptor also makes the chip sensitive to the 120 Hz flicker of ordinary room lights.