

A Biologically Inspired Analog IC for Visual Collision Detection

Reid R. Harrison, *Member, IEEE*

Abstract—We have designed and tested a single-chip analog VLSI sensor that detects imminent collisions by measuring radially expanding optic flow. The design of the chip is based on a model proposed to explain leg-extension behavior in flies during landing approaches. We evaluated a detailed version of this model in simulation using a library of 50 test movies taken through a fisheye lens. The algorithm was evaluated on its ability to distinguish movies ending in collisions from movies in which no collision occurred. This biologically inspired algorithm is capable of 94% correct performance in this task using an ultra-low-resolution (132-pixel) image as input. A new elementary motion detector (EMD) circuit was developed to measure optic flow on a CMOS focal-plane sensor. This EMD circuit models the bandpass nature of large monopolar cells (LMCs) immediately postsynaptic to photoreceptors in the fly visual system as well as a saturating multiplication operation proposed for Reichart-type motion detectors. A 16×16 array of two-dimensional motion detectors was fabricated in a standard $0.5\text{-}\mu\text{m}$ CMOS process. The chip consumes $140\ \mu\text{W}$ of power from a 5 V supply. With the addition of wide-angle optics, the sensor is able to detect collisions 100–400 ms before impact in complex, real-world scenes.

Index Terms—CMOS imager, collision detection, Gilbert multiplier, insect vision, neuromorphic systems, optic flow, smart sensor.

I. INTRODUCTION

THE VISUAL detection of imminent collisions is an ability possessed by animals ranging from insects [1]–[3] to birds [4] to humans. The speed with which rapidly approaching objects are detected suggests that the neural circuitry responsible for this sense is hardwired and operates beneath “high-level cognition” as a type of visual reflex. An algorithm providing this capability would be valuable in the design of autonomous robots using vision for navigation, and could also have applications in the area of automobile safety. In embedded applications such as these, the collision-detection algorithm must be instantiated as a small, low-power sensor capable of visually sensing a dynamic scene and ultimately producing a timely warning signal that can be used for collision avoidance.

Collision-detection algorithms have been implemented on robotic platforms using a traditional charged-couple device (CCD) imager and CPU to perform the sensing and real-time computation. Duchon and colleagues implemented an optic-flow-based collision-detection algorithm on a large

mobile robot with a powerful CPU that successfully navigated through a large laboratory using only a wide-angle video camera [5]. Blanchard and Verschure implemented Rind’s collision-detection algorithm on a small tabletop robot using three remote Pentium II PC’s (two 450 and one 333 MHz) to perform the real-time image processing [6], [7]. More recently, this a variant of this algorithm was used in preliminary experiments to guide an autonomous blimp, again using remote image processing on a PC [8].

Starting with the pioneering work of Mead in the 1980s [9], researchers have developed custom VLSI implementations of visual processing algorithms in the hopes of creating smaller, micropower sensors. In 1994, Abbott and colleagues presented an analog CMOS chip with 60 photodetectors and differentiators to perform edge enhancement [10]. A method for using the chip as a collision detector was presented, but explicit tests were not performed. As the chip had only a single one-dimensional (1-D) array of photodiodes, limited performance in general scenes would be expected. In 1996, Indiveri and colleagues presented an analog CMOS “time-to contact” chip [11], [12]. The chip used two concentric rings of 12 photodetectors each to create a ring containing 12 radially oriented motion detectors. The chip successfully produced estimates of collision times for high-contrast concentric rings that repeatedly triggered the motion detectors. The algorithm was not evaluated on real-world scenes. Also in 1996, Ancona and colleagues proposed a system built with two custom VLSI chips—a CMOS image sensor and a digital ASIC—and a microcontroller with additional RAM chips [13]. This multi-chip system used the same algorithm as in [11], which used sparse sampling of the image to compute image velocity about a contour.

Our goal in this work was to develop a hardware implementation of a visual collision-detection algorithm in the form of a single-chip “smart sensor” consuming less than one milliwatt of power. Unlike previous silicon implementations, our chip was designed to integrate information from every pixel in a dense two-dimensional (2-D) imager array in the hope that this approach will yield a more robust sensor. In the pursuit of robustness we also evaluated our algorithm using a library of movies recorded in complex, real-world scenes.

In this paper, we first present the biologically inspired algorithm that was used in both simulation and silicon. We next describe how the algorithm was evaluated and refined in simulation using a library of movies. Then we present the implementation of the algorithm in a standard CMOS VLSI process. Finally, we show experimental results from this single-chip visual sensor.

Manuscript received July 23, 2004; revised February 7, 2005. This work was supported in part by the Naval Air Warfare Center, China Lake, CA. This paper was recommended by Associate Editor T. S. Lande.

The author is with the Department of Electrical and Computer Engineering, University of Utah, Salt Lake City, UT 84112 USA (e-mail: harrison@ece.utah.edu).

Digital Object Identifier 10.1109/TCSI.2005.853517

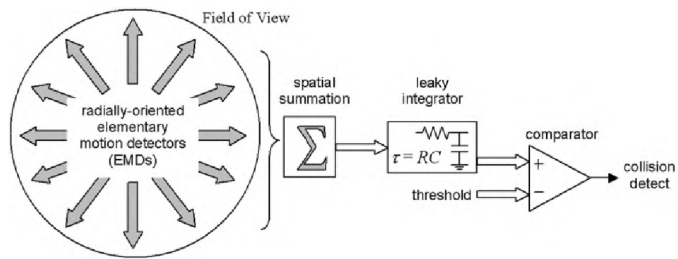


Fig. 1. Diagram of collision detection algorithm. The responses of radially oriented motion detectors are integrated in space and time, and compared against a fixed threshold.

II. MOTION DETECTOR MODELS

A. Models of Collision Detection in Animals

While several models have been proposed to explain collision detection [1]–[3], [14], some involve relatively high-level concepts such as identifying an approaching object in a dynamic image and measuring its size (e.g., [2]). The model proposed by Borst and Bahde in 1988 to explain the landing behavior of flies [1] is particularly amenable to low-level hardware implementation. As flies approach a large stationary object, they extend their legs to a landing position a few hundred milliseconds before contact. The response is visually mediated, and the leg extension timing relative to contact time varies depending on the visual structure of the landing site. Highly patterned objects with many high-contrast edges elicit earlier landing responses than visually sparse objects having large featureless regions.

The model developed by Borst and Bahde to explain this behavior employs a radially oriented array of motion detectors centered in the direction of flight (see Fig. 1). As the animal approaches a static object, an expansive optic flow field is produced on the retina. A wide angle field of view is useful since optic flow in the direction of flight will be zero. The response of this radial array of motion detectors is summed and then passed through a leaky integrator (a low-pass filter). If this response exceeds a fixed threshold, an imminent collision is detected and the animal can take evasive action or prepare for a landing. This expansive optic flow model has recently been used to explain landing and collision avoidance responses in the fruit fly [3]. All optic-flow-based collision detection algorithms require an initial processing stage that estimates motion in an image.

B. Motion Detector Architecture

Motion detection mechanisms in flies have been studied for nearly 50 years, starting with the so-called Reichardt model of motion detection proposed in 1956 [15]. Much experimental work has reinforced the validity of this delay-and-correlate elementary motion detector (EMD) and expanded the original model to an elaborated version shown in Fig. 2 [16]–[18]. As shown in this figure, two adjacent photoreceptors convert light intensity into an electrical signal. After photoreception in the retina, the mean light intensity needs to be subtracted. Eliminating the potentially large dc light level—which contains no motion information—makes better use of the limited-dynamic-range channels that follow. This dc level elimination is accomplished in the large monopolar cells (LMCs), which are directly postsynaptic to photoreceptors in the fly.

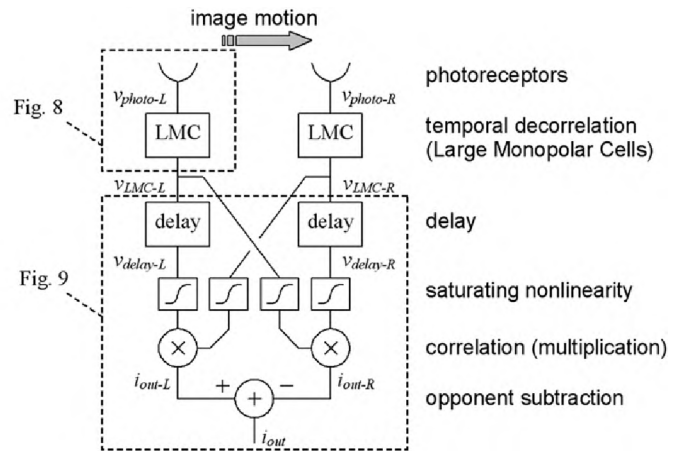


Fig. 2. Elaborated delay-and-correlate EMD.

Suppressing dc illumination and enhancing ac components of photoreceptor signals is a common theme in many biological visual systems. LMCs in the fly exhibit transient biphasic impulse responses approximately 40–200 ms in duration [19], [20]. In the frequency domain, this can be seen as a bandpass filtering operation that attenuates dc signals while amplifying signals in the 2–40 Hz range [20], [21]. In the lateral geniculate nucleus of cats, “lagged” and “nonlagged” cells exhibit transient biphasic impulse responses 200–300 ms in duration and act as bandpass filters amplifying signals in the 1–10 Hz range [22]. (Of course, the large-signal behaviors of these cells show significant nonlinearity. Nevertheless, the cells suppress dc signals and pass ac signals within their bandwidth, so they act as bandpass filters for small signals.) This filtering has recently been explained in terms of temporal decorrelation, and can be seen as way of removing redundant information from the photoreceptor signal before further processing [20], [23].

After this “transient enhancement,” or temporal decorrelation, the signals are delayed using the phase lag of a low-pass filter. While not a true time delay, the low-pass filter matches data from animal experiments and makes the Reichardt EMD equivalent to a simplified version of the oriented spatiotemporal energy filter proposed by Adelson and Bergen [24] with no explicit spatial filtering. Before correlating the adjacent delayed and nondelayed signals, a saturating static nonlinearity is applied to each channel. Without such a nonlinearity, the delay-and-correlate EMD exhibits a quadratic dependence on image contrast. In fly tangential neurons, motion responses show a quadratic dependence only at very low contrasts, then quickly become largely independent of image contrast for contrasts above 30%. Egelhaaf and Borst proposed the presence of this nonlinearity in the biological EMD to explain this contrast independence [25]. Functionally, it is necessary to prevent high-contrast edges from dominating the summed output of the EMD array.

After correlation, opponent subtraction produces a strong directionally selective signal that is taken as the output of the EMD. Unlike algorithms that find and track features in an image, the delay-and-correlate EMD does not measure true image velocity independent of the spatial structure of the image. However, recent work has shown that for natural scenes, these

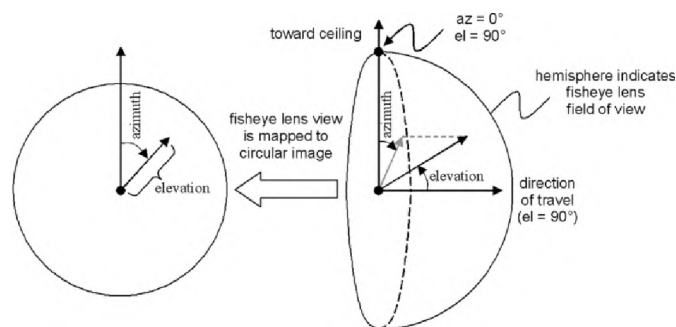


Fig. 3. Polar coordinate system used for all fisheye lens images.

Reichardt EMDs give reliable estimates of image velocity [26]. This reliability is improved by the addition of LMC bandpass filters and saturating nonlinearities. Experiments using earlier versions of silicon EMDs have demonstrated the ability of delay-and-correlate motion detectors to work reliably at very low signal-to-noise ratios [27].

III. SIMULATION EXPERIMENTS

A. Collection of Dataset

Toward the goal of developing a collision detection algorithm that works robustly in the presence of complex scenes, we compiled a database of real-world movies that were used to evaluate algorithm performance. We took digital movies using a digital camera (Nikon Coolpix 995) capable of capturing 70 high-quality (low compression) 320×240 JPEG images at a rate of 30 frames/second. We used a fisheye lens converter (Nikon FC-E8) to capture a complete 180° visual hemifield in front of the camera.

The fisheye lens mapped the forward visual hemifield onto a circle 240 pixels in diameter on the image plane (see Fig. 3). An object directly in front of the camera mapped to the center of this circular image. We defined this direction as zero elevation angle. An object directly above the camera (e.g., on the ceiling) mapped to the top of the circle. We defined this direction as 90° elevation, 0° azimuth. An object directly below the camera (e.g., on the floor) mapped to the bottom of the circle and was defined to be 90° elevation, 180° azimuth. Using this polar coordinate system, an object at the same height as the camera but lying to the left of center in its field of view would have an azimuth of $+90^\circ$ and an elevation of perhaps 45° .

The camera was mounted on a small custom-built motorized vehicle which could be moved in controlled trajectories across the floor of our laboratory. The camera was oriented so that the lens faced forward. The center of the lens was 11 cm above the floor, and was placed forward so that the vehicle itself was not visible in the fisheye field of view. The vehicle moved at a constant velocity of 31 cm/s.

We collected 50 movies of two types: “crash” and “no crash.” The 25 “crash” movies ended with the camera colliding with a fixed object: a wall, a book, a table leg, or some other object commonly encountered in a university lab. (To protect the fisheye lens, these movies were taken backward, with the camera first touching the obstacle and then the vehicle backing away. The frames were later reversed.) The 25 “no

crash” movies were taken in situations where the camera never collided with an obstacle. For example, the camera was driven under tables, across open areas of floor, or other noncollision trajectories. In two of the 25 “no crash” movies, the vehicle was modified to turn in place instead of move forward. While all movies were recorded in our $8.2 \text{ m} \times 7.9 \text{ m}$ lab, we made an effort to change the location and direction of trajectory in all 50 movies, and to include challenging collision detection situations, such as approaching bare walls. Note that the wall-floor and wall-ceiling boundaries were visible in the fisheye lens up to the point of collision. A typical “wide-angle” lens having a field of view of perhaps $\pm 35^\circ$ cannot see these important visual cues once it is within a certain distance of the wall.

As mentioned above, each raw movie consisted of 70 frames. Because the camera could not be synchronized precisely with vehicle motion, there were typically a few “wasted” frames in each image taken before the vehicle began moving. Each movie was analyzed to determine the onset of vehicle motion, and the excess frames were discarded. To maintain a uniform length for all 50 movies, we chose to trim each movie to 60 frames, which represented two seconds of motion and 62 cm of distance traveled.

We converted the color images to grayscale and reduced the resolution to polar-oriented pixels each subtending 1° of elevation angle and 1.4° – 2.8° azimuth angle. (Pixels in a polar image become smaller toward the center of the image, so we periodically reduced the number of pixels in each concentric “ring” to maintain a reasonable pixel size and thus approximate the uniform pixel layout in typical imager arrays.) Fig. 4(a) shows one image from a “crash” movie where the camera is approaching two filing cabinets. The image consists of 9000 pixels covering a full 180° view.

Using this dataset, we simulated the collision detection algorithm from Fig. 1 in MATLAB using a time step equivalent to the frame time of $1/30$ second. We implemented the LMC filter from Fig. 2 using an FIR filter with coefficients of -0.5 , $+1$, and -0.5 to give a biphasic response to a step change in light intensity with a similar time course to biological LMCs. This is roughly equivalent to a continuous-time bandpass filter having a center frequency of 15 Hz and a quality factor Q around two. We used a first-order IIR filter to implement the low-pass filter in the EMD ($\tau = 25$ ms, which corresponds closely to biological data from insects [18]) and the leaky integrator (also a low-pass filter) at the output ($\tau = 150$ ms). The saturating nonlinearity was implemented as a tanh function. EMDs were arranged radially between adjacent pixels.

B. Initial Simulation Results

Fig. 4 shows the various stages of processing in our algorithm for one frame from a “crash” movie. Fig. 4(a) shows the image taken by the camera after conversion to grayscale and polar pixelization. This represents the photoreceptor input to the EMD array. Fig. 4(b) shows the output of the LMC bandpass filter stage after the saturating nonlinearity. The edges of the filing cabinets, which are expanding in the field of view, are clearly enhanced by the LMC units. Fig. 4(c) shows the output of the delay (low-pass filter) stage after the saturating nonlinearity. Because the low-pass filter time constant is on the order of the frame

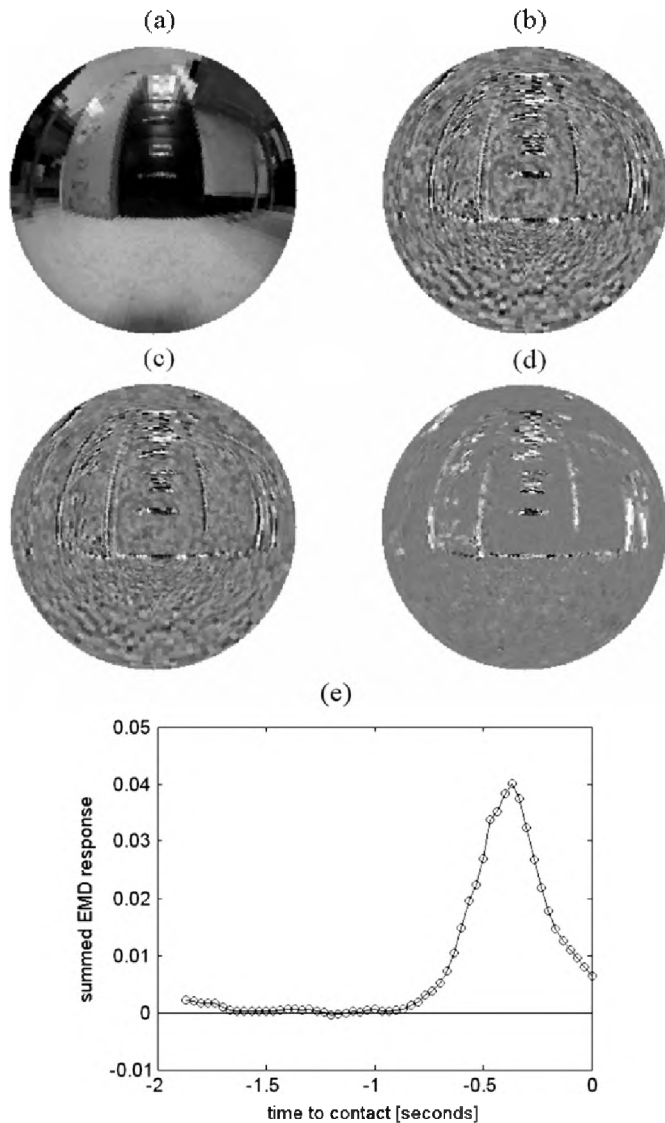


Fig. 4. Images from various levels of motion processing for a typical collision movie. (a) Photoreceptor. (b) LMC. (c) Delay. (d) Radial EMD opponent output. (e) Response after spatial summation and leaky integrator.

rate, the differences between Fig. 4(b) and (c) are subtle. Fig. 4(d) shows the opponent output of the radially oriented EMDs, where white represents a positive response and black represents a negative response. While more light pixels are present than dark pixels, it is clear that individual EMDs sampled at one instant in time are rather unreliable motion sensors, as has been observed in biological models [26], [28].

Fig. 4(e) shows the output of the collision detector after summing the EMD responses over the entire visual field and passing this signal through a leaky integrator. The response builds to a peak around 400 ms before contact with the obstacle, and then the response dies away before collision. Clearly, a simple threshold operation could be used to produce a warning several hundred milliseconds before collision.

Fig. 5 summarizes the performance of the collision detection algorithm on the 50-movie test set. Fig. 5(a) shows the superimposed algorithm responses for all 25 “crash” movies, and Fig. 5(b) shows the responses for all 25 “no crash” movies. Since

the purpose of the algorithm is to distinguish between imminent collisions and noncollision situations, we would like to set a threshold level that would be exceeded by every “crash” movie and not be reached by any “no crash” movie. As Fig. 5(c) illustrates, this ideal situation is impossible to obtain given the performance of our algorithm. There is no threshold level that results in zero false positives (reporting an imminent crash in a “no crash” situation) and zero false negatives (failing to detect an imminent collision). The data in Fig. 5(c) illustrate that the performance of the algorithm may be optimized by setting the threshold value to 17; this results in one false positive and three false negatives out of 50 trials, for a total of 92% correct.

We evaluated this data using the receiver operating characteristic (ROC) concept [29]. An ROC plot is a parametric plot of false positive rate versus false negative rate where the varied parameter is the detection threshold. Fig. 5(d) shows the ROC plot for the simulation data. The area underneath the ROC curve represents an estimate of the upper limit of the probability that the algorithm will make a correct decision. Note that a perfect collision detector would have an ROC curve that followed the left and upper edge of the square (encompassing an area of 1.0, or 100%), while a system behaving randomly would have an ROC curve of a straight line from the bottom-left corner to the upper-right corner of the square (encompassing an area of 0.5, or 50%). The area under the ROC curve (A_{ROC}) in Fig. 5(d) is 0.984; note that this value is overly optimistic when compared to the 92% correct detection rate observed in Fig. 5(c). In our simulations, A_{ROC} correlated well with observed performance so we used this single value to rate the overall performance of our algorithm in the 50-movie dataset. Algorithm performance was observed to be relatively insensitive to EMD parameters.

The simulation paradigm presented here allowed us to easily test alternative models for possible hardware implementation. For example, eliminating the saturating nonlinearities shown in Fig. 2 reduced A_{ROC} from 0.984 to 0.954 due to the increased contrast dependence. However, strengthening the nonlinearity so that a binarization of the signal was performed before correlation also decreased A_{ROC} (to 0.962), indicating that an analog signal should be preserved.

C. Results With Reduced Field-of-View and Image Resolution

We observed that some of the “no crash” movies were triggering false positives when the camera moved past a nearby object, which produced outward optic flow in the periphery of the image. To improve performance, we limited the field of view from $\pm 90^\circ$ to $\pm 60^\circ$ to limit the effect of this optic flow caused by objects not in the vehicle’s path. This improved A_{ROC} to 0.994 while using a 5160-pixel image. With an efficient hardware implementation as the ultimate goal, we also reduced the overall resolution of the image as much as possible. Fig. 6(a) shows the image size and resolution used in the previous simulations, and is identical to Fig. 4(a). Fig. 6(b) shows the reduced image size and resolution used in the second simulation. The field of view was limited to an elevation of 60° , and the pixels were reduced to an elevation of 5° and an azimuth ranging from 11° to 22° . This resulted in an image containing only 132 pixels.

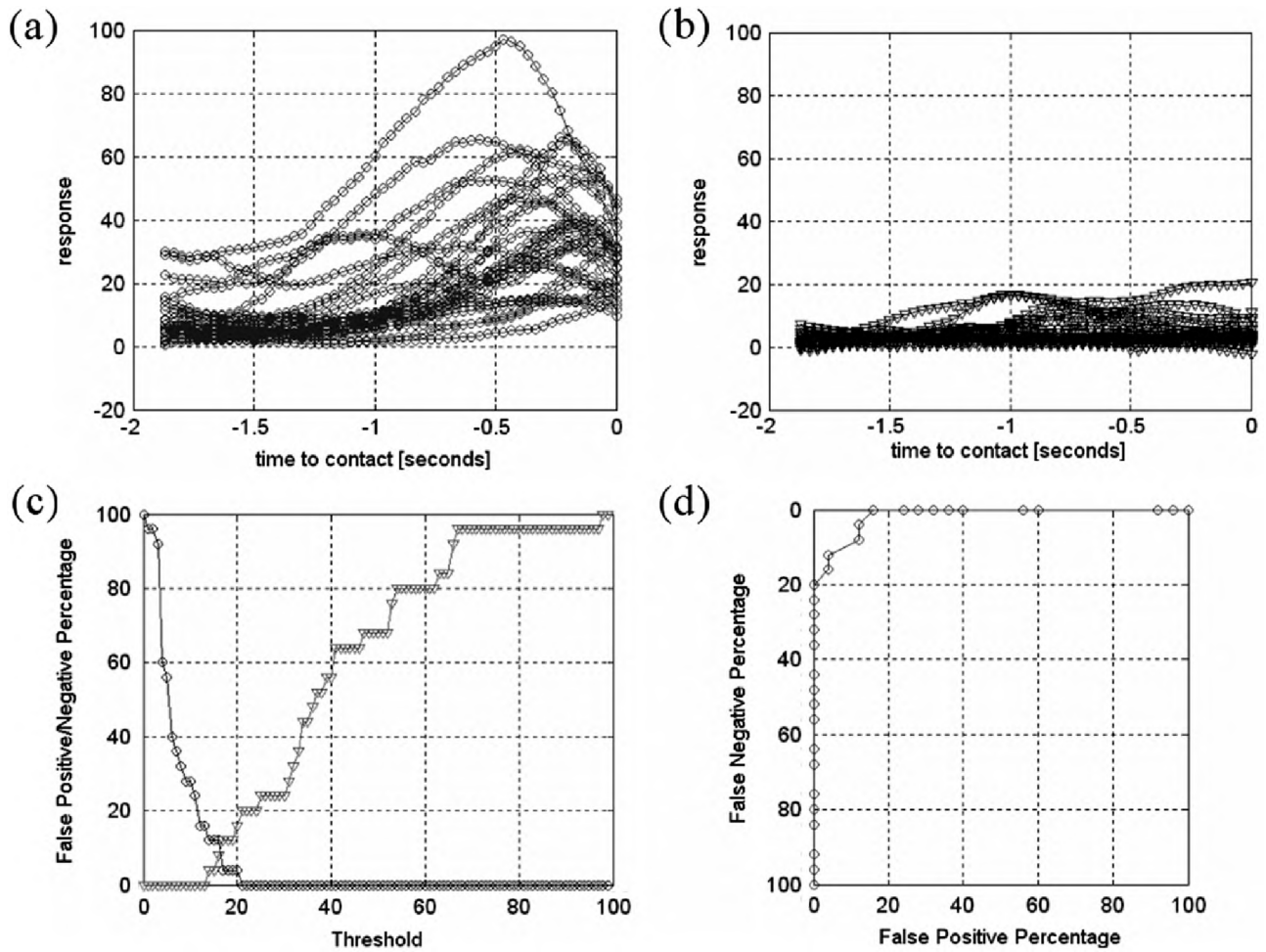


Fig. 5. Results of collision detection algorithm applied to all 50 movies at full resolution. (a) Output of collision detector for all 25 “crash” movies. (b) Output of collision detector for all 25 “no crash” movies. (c) False positives and false negatives as detection threshold is varied. (d) ROC curve. The area under the ROC curve is 0.984.

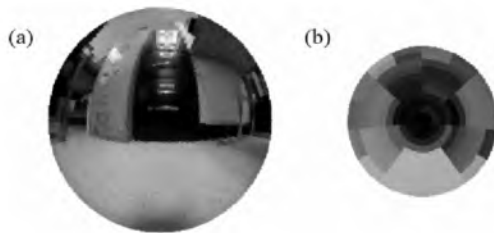


Fig. 6. (a) Image size used for initial simulations (9000 pixels). The field of view is 180° . (b) Reduced image used for later simulations (132 pixels). The field of view is 120° .

The simulation results using these angle-limited, reduced-resolution images are shown in Fig. 7. Note that Fig. 7 has the same format as Fig. 5. The algorithm performs with $A_{ROC} = 0.989$ across the entire dataset of 50 movies. While lower than the $A_{ROC} = 0.994$ achieved using a 120° field of view at full resolution, the algorithm performs surprisingly well considering the coarse images used. The best observed performance occurred with a threshold of 0.0029 [see Fig. 7(c)], giving two false positives and one false negative for a total of 94% correct. Further reduction of image resolution and/or field of view significantly lowered performance.

IV. INTEGRATED CIRCUIT IMPLEMENTATION

Our positive simulation results using 132-pixel images suggested that this collision detection algorithm was particularly amenable to implementation as a focal-plane analog VLSI smart sensor. Adding analog computation to an imaging chip typically results in large pixel size and a low fill factor. While this trade-off is undesirable for most image-recording applications, it can produce compact, low-power sensors suitable for use in small, power-limited autonomous systems [12], [27], [30].

A. Photoreceptor and LMC Circuit Design

We adapted the EMD shown in Fig. 2 to a small, low-power CMOS integrated circuit. Fig. 8 shows a schematic of the photoreceptor and LMC bandpass filter. A $35\ \mu\text{m} \times 35\ \mu\text{m}$ well-substrate photodiode with diode-connected pMOS load operating in subthreshold converts the diode photocurrent into a voltage v_{photo} that is a logarithmic function of light intensity. The speed of this simple photoreceptor circuit is proportional to light intensity, so it is unsuitable for use in extremely low light levels. In dim environments, an active photoreceptor such as the type presented in [31] would be preferred. A pMOS source follower biased by $I_{SF} = 700\ \text{pA}$ buffers this signal so that the

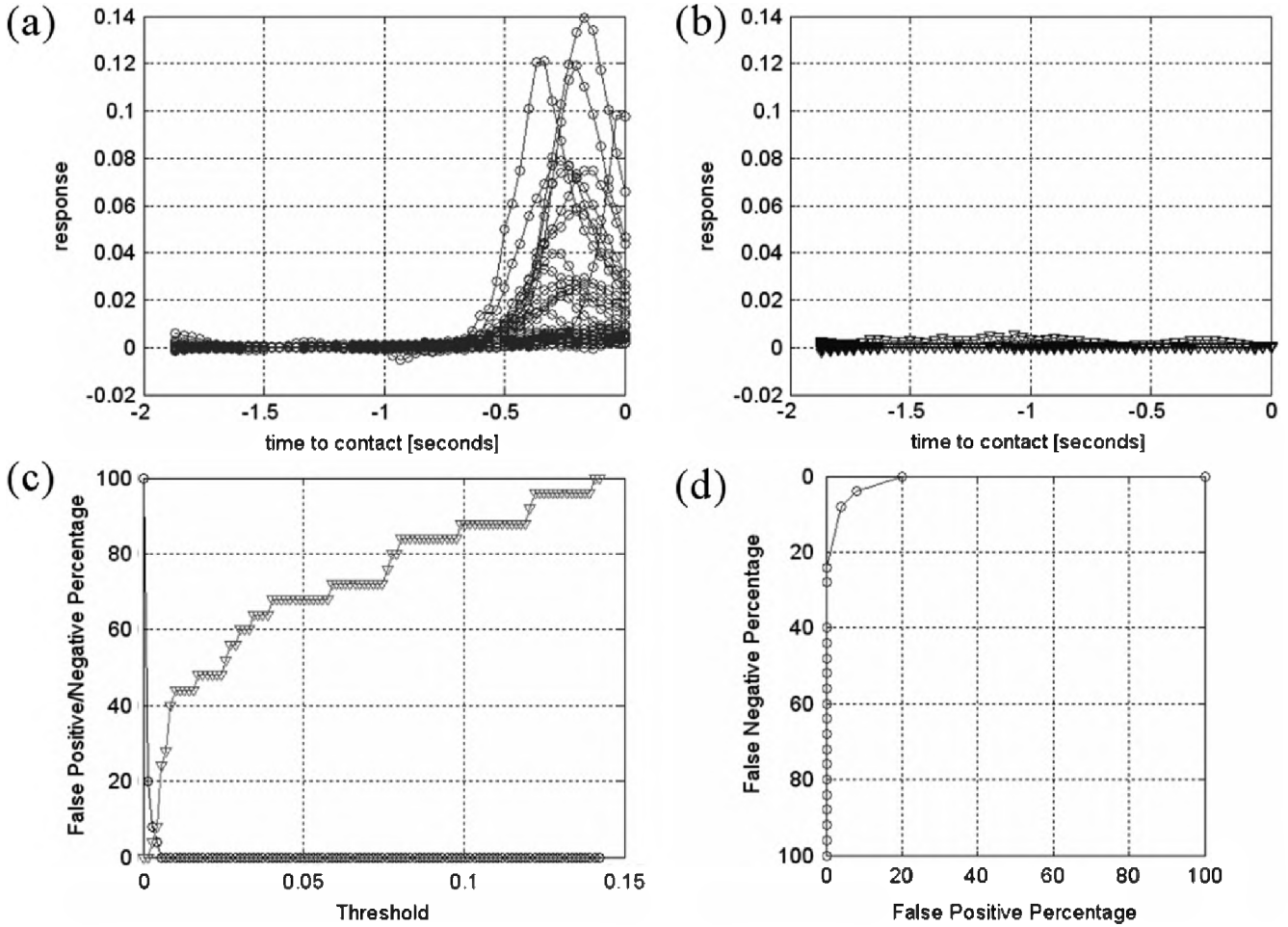


Fig. 7. Results of collision detection algorithm applied to all 50 movies at reduced resolution and field of view (132 pixels, $\pm 60^\circ$). (a) Output of collision detector for all 25 “crash” movies. (b) Output of collision detector for all 25 “no crash” movies. (c) False positives and false negatives as detection threshold is varied. (d) ROC curve. The area under the ROC curve is 0.989.

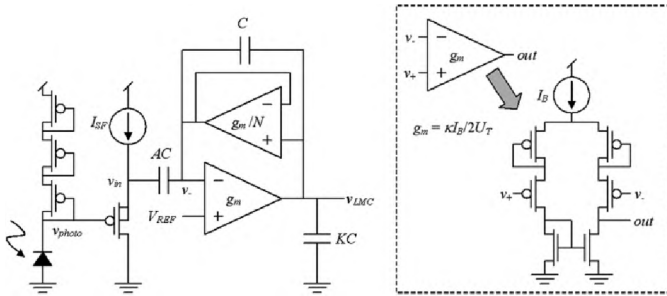


Fig. 8. Schematic of photoreceptor/LMC circuit. Detail of operational transconductance amplifier (OTA) shown in inset.

input capacitance of the LMC circuit does not load the photoreceptor. The time constant of the source follower is much shorter than the time constant of the LMC bandpass filter following it and thus does not significantly alter the time course of the photoreceptor signals.

The LMC bandpass filter consists of two operational transconductance amplifiers (OTAs) and three capacitors. The OTAs in the circuit are implemented with pMOS differential pairs using diode-connected transistors for source degeneration for extended linear range (see inset, Fig. 8) [32]. From the schematic shown in Fig. 8, we can write the following

equations using Kirchoff’s current law, setting $V_{REF} = 0$ for convenience

$$sKCv_{LMC} + sC(v_{LMC} - v_-) - g_m(0 - v_-) = 0 \quad (1)$$

$$sAC(v_- - v_{in}) + sC(v_- - v_{LMC}) - \frac{g_m}{N}(v_{LMC} - v_-) = 0. \quad (2)$$

From these two equations, the transfer function of the LMC circuit can be derived

$$\frac{v_{LMC}(s)}{v_{in}(s)} = -A \cdot \frac{N\tau_0 s \cdot (1 - \tau_0 s)}{(\tau_1 s)^2 + \frac{\tau_1 s}{Q} + 1} = -\frac{AN}{\beta} \cdot \frac{(1 - \frac{\tau_1}{\beta} s)}{\tau_1 s + \frac{1}{Q} + \frac{1}{\tau_1 s}} \quad (3)$$

where

$$\tau_0 = \frac{C}{g_m} \quad (4)$$

$$\beta = \sqrt{N(A+1)(K+1)} - N \approx \sqrt{NAK} \quad \text{if } A, K \gg 1 \quad (5)$$

$$\tau_1 = \beta\tau_0 \quad (6)$$

$$Q = \frac{\beta}{(K+N)}. \quad (7)$$

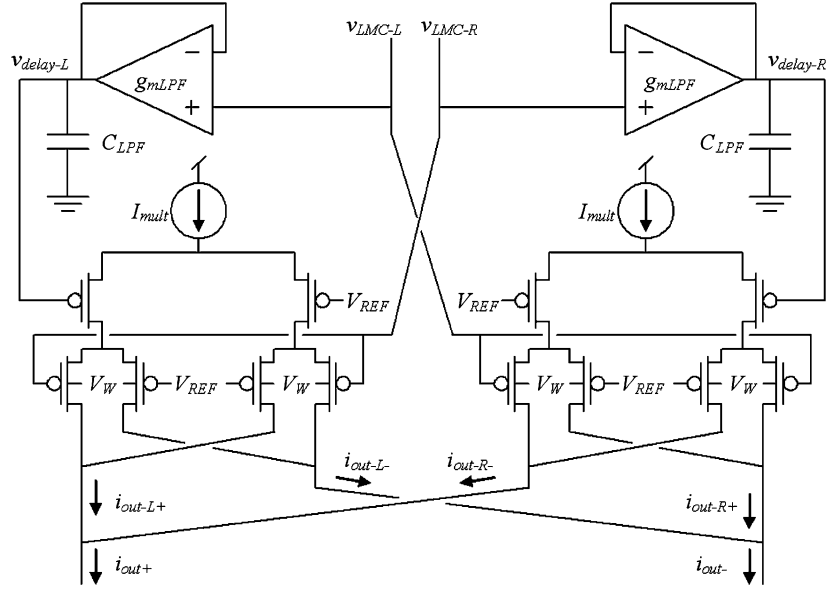


Fig. 9. Schematic of delay-and-correlate circuit. OTA-based g_m - C filters are used as low-pass filters. Subthreshold CMOS Gilbert multipliers are used for correlation.

The output signal v_{LMC} is centered around V_{REF} , a dc voltage which was set to 1.0 V. We sized the capacitors in our circuit to give $A = 20$ and $K = 5$ (with $C = 70$ fF). The transconductance of the lower OTA was set by adjusting its bias current I_B

$$g_m = \frac{\kappa}{(\kappa + 1)} \cdot \frac{I_B}{2U_T} \quad (8)$$

where κ is the weak inversion slope (typically between 0.6 and 0.9) and U_T is the thermal voltage kT/q (approximately 26 mV at room temperature) [32]. We set the bias current in the upper OTA five times smaller to achieve $N = 5$.

As we see from (3), the LMC circuit acts as an ac-coupled bandpass filter centered at $f_1 = 1/2\pi\tau_1$, with a quality factor Q set to 2.5 by capacitor and current ratios. The circuit also has a zero at βf_1 , but since $\beta = 25$ in our circuit, the zero takes effect outside that passband and thus has little practical effect on the filter. We used a bias current of $I_B = 35$ pA in the lower OTA and 7 pA in the upper OTA to center the passband near 20 Hz, which was chosen because it lies in the range of LMC response measured in the fly. The LMC circuit presented here allows the designer to adjust the center frequency and Q factor to selectively amplify frequencies present in moving images.

B. Delay-and-Correlate Circuit Design

The LMC circuits from each photoreceptor pass their signals to the the delay-and-correlate circuit shown in Fig. 9. The delay is implemented as a first-order low-pass filter. The OTAs in this circuit used two diode-connected transistors in series for extended linear range. The time constant of this filter is given by

$$\tau_{LPF} = \frac{C_{LPF}}{g_{mLPF}}. \quad (9)$$

We used $C_{LPF} = 700$ fF and set τ_{LPF} to around 25 ms, which is in the range of biological motion detectors. This required a bias current of 9 pA for each OTA. Since current sources suffer from mismatch, particularly in subthreshold operation, we used rela-

tively large bias transistors for this circuit ($W/L = 4.8 \mu\text{m}/2.4 \mu\text{m}$ in a $0.5\text{-}\mu\text{m}$ process) to improve cross-chip matching.

We implemented the correlation function using a CMOS Gilbert multiplier operating in subthreshold [9], [33]. The output currents of the multipliers in Fig. 9 can be expressed as

$$\begin{aligned} i_{outL+} - i_{outL-} &= I_{mult} \tanh \frac{\kappa(v_{delay-L} - V_{REF})}{2U_T} \\ &\quad \times \tanh \frac{\kappa(v_{LMC-R} - V_{REF})}{2U_T} \end{aligned} \quad (10)$$

$$\begin{aligned} i_{outR+} - i_{outR-} &= I_{mult} \tanh \frac{\kappa(v_{delay-R} - V_{REF})}{2U_T} \\ &\quad \times \tanh \frac{\kappa(v_{LMC-L} - V_{REF})}{2U_T} \end{aligned} \quad (11)$$

For small differential input voltages, $\tanh(x) \approx x$ and the circuit acts as a linear multiplier. As the input signals grow larger, the tanh nonlinearity dominates and the circuit acts more like a digital exclusive-or gate. We use this inherent circuit nonlinearity as the desired saturating nonlinearity in our EMD model (see Fig. 2). The previous LMC circuit provides sufficient gain to ensure that we are usually operating well *outside* the linear range of the multipliers.

Traditional CMOS Gilbert multipliers require that the dc level of the upper differential input be shifted relative to the dc level of the lower differential input. This is required to keep all transistors in saturation. To avoid the cost in chip area, power consumption, and mismatch associated with level shifters, we introduce a novel circuit modification that allows both the upper and lower differential inputs to operate at the same dc level. As shown in Fig. 10, we lower the well potential of the lower pMOS transistors from V_{DD} to a dc voltage V_W . This lowered well voltage

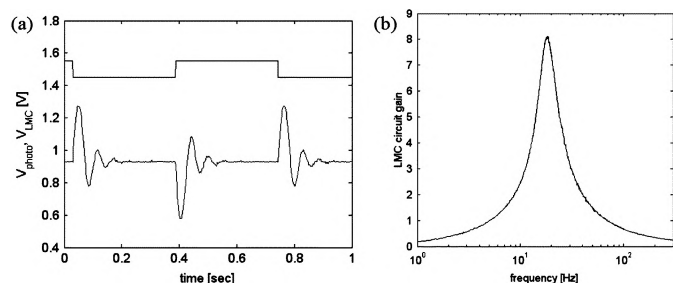


Fig. 12. Measurement of LMC circuit performance. (a) Step response of LMC circuit. (b) Frequency tuning of LMC circuit.

Fig. 11(b) shows the EMDs that were summed to produce the global radial motion signal. Diagonally oriented EMDs were derived from the sum of a horizontal and a vertical EMD. The center 4×4 pixels were ignored, as motion near the center of the field of view is typically very small in collision situations. We used custom-built ultra-wide-angle optics to give the chip a field of view ranging from $\pm 52^\circ$ at the sides to $\pm 74^\circ$ at the corners. The simulations presented in Section III showed that a field of view of at least $\pm 60^\circ$ was necessary for reasonable performance using this algorithm.

Before testing the array, we characterized an individual LMC circuit that was configured to have a voltage input v_{photo} provided externally using a function generator. We provided a 1.4 Hz, 100 mVpp square wave and observed the LMC circuit output [see Fig. 12(a)]. The LMC circuit exhibits a transient oscillatory step response similar to its biological counterpart. Using a spectrum analyzer, we measured the transfer function of the circuit [see Fig. 12(b)]. The LMC circuit acts as a bandpass filter centered at 19 Hz, with a measured Q of 2.3.

The entire collision detector chip consumed $140 \mu\text{W}$ of power. Most of this was consumed by peripheral biasing circuits; the 17×17 pixel array used only $5.2 \mu\text{W}$ (18 nW per pixel). To test the complete collision detection chip, we implemented the leaky integrator and comparator from Fig. 1 using off-chip components. In future implementations, these circuits could be built on chip using little power.

We were unable to test the chip using the recorded movies (see Section III) due to noise feedthrough from the ~ 70 Hz refresh rate of monitors. Instead, we tested the chip by moving a black square with a width of 12.8 cm toward the lens at various speeds. In keeping with notation from [2], we denote object size by the half-width $s = 6.4$ cm. Fig. 13(a) shows the measured response of the chip when the black square approaches at a velocity v of 17 cm/s. The response peaks approximately 230 ms before contact, then rises again during the last 150 ms. Careful observation of the experiment revealed that the lens cast a shadow on the target at distances of less than 3 cm. As the target approached the lens this shadow expanded, yielding an additional expanding visual stimulus. To eliminate this artifact, we truncated the data for distances less than 3 cm from contact. Fig. 13(b) shows experiments performed with velocities ranging from 12 cm/s (peaking sooner) to 24 cm/s (peaking later).

We also mounted two collision detector chips on a small robotic platform (see Fig. 14). The two sensors looked slightly to the left and right of the nominal direction of travel. If an

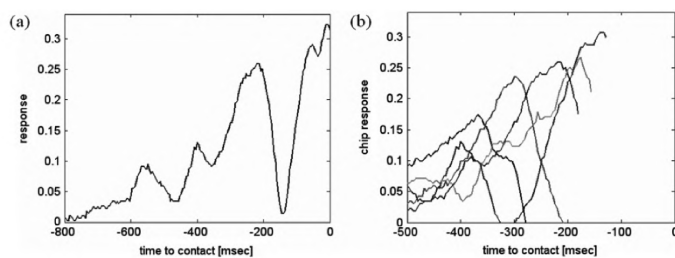


Fig. 13. (a) Chip response to approaching dark square ($s = 6.4$ cm) approaching at $v = 17$ cm/s. (b) Chip response for dark square ($s = 6.4$ cm) approaching at five different velocities: 12 cm/s, 14 cm/s, 17 cm/s, 19 cm/s, and 24 cm/s. Shadow artifacts ($d < 3$ cm) have been removed.

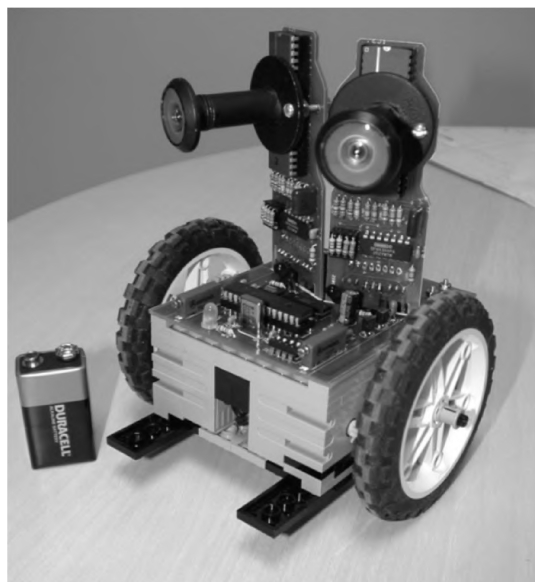


Fig. 14. Collision avoidance robot with two collision detector chips. Custom ultra-wide-angle optics are mounted over the chips, which are mounted in 40-pin DIP packages.

imminent collision was detected on the left, the robot would stop and turn to the right. If an imminent collision was detected on the right, the robot would stop and turn to the left. This simple implementation of the model presented in [3] was able to navigate through a cluttered laboratory with few collisions. A video of this robot is available at <http://www.ece.utah.edu/~harrison/robot/>.

VI. SUMMARY AND CONCLUSIONS

As shown in Figs. 5(a), 7(a), and 13, the time course of the collision detector output signal has a distinct shape, peaking before collision and then collapsing. This is similar to the activity patterns observed in LGMD neuron in locusts [2] and η neurons in pigeons [4] during collisions. While more complex models positing the measurement of true image velocity and object size have been used to explain this peculiar time course [2], we observe that a simple model integrating the output of a radial EMD array gives qualitatively similar responses.

In both the locust and the pigeon, the time at which the response peaks (relative to the time of collision) is observed to be a linear function of s/v , where s is the radius of the object the observer is colliding with and v is the approach velocity. As shown in Fig.

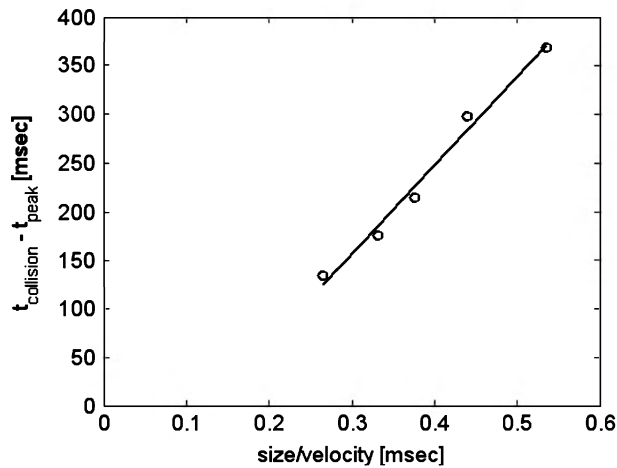


Fig. 15. Peak time versus size-to-velocity ratio in silicon collision detector. Data is taken from Fig. 13(b).

15, the data from our silicon system also exhibits a linear relationship between time of response peak and s/v . We observed similar results in simulations using synthetic images of an approaching black disc on a white background (data not shown).

Furthermore, we observed that the slope of this line, while relatively insensitive to the EMD parameters, is strongly affected by the field of view used by the algorithm. Using this empirical relationship (data not shown), we observed that the timing of η neurons in pigeons [4] corresponded to a receptive field 6.6° in diameter and the timing of the LGMD neuron in three species of locusts [2] corresponded to receptive fields between 38° and 48° in diameter using the algorithm shown in Fig. 1. Additional neurophysiological studies are needed to determine if receptive fields of these sizes are actually present in the visual systems of these animals.

We have demonstrated that this model of collision detection can be implemented in a small, single-chip sensor dissipating only $140 \mu\text{W}$ of power. Digital implementations of this algorithm would result in easier programmability at the likely expense of greater power consumption. By recording a library of “crash” and “no crash” real-world movies, we were able to evaluate and refine the collision-detection algorithm before the chip was designed. This approach likely saved us several design-fabricate-test cycles and also increased the robustness of the resulting sensor. While the algorithm was quantitatively evaluated in simulation using complex scenes, the chip was quantitatively evaluated using only high-contrast squares. We were unable to test the chip using the recorded movies due to noise feedthrough from the ~ 70 Hz refresh rate of monitors. However, qualitative tests using two chips mounted on a small robotic platform demonstrated a high degree of robustness in the sensors as the robot wandered through a cluttered room. More extensive and quantitative testing of the chip on mobile platforms will better characterize its performance.

ACKNOWLEDGMENT

The author wishes to thank M. Love for assistance in collecting the movie library. Ms. Love was supported by the University of Utah Engineering Scholars Program.

REFERENCES

- [1] A. Borst and S. Bahde, “Visual information processing in the fly’s landing system,” *J. Comp. Physiol. A*, vol. 163, pp. 167–173, 1988.
- [2] F. Gabbiani, H. G. Krapp, and G. Laurent, “Computation of object approach by a wide-field, motion-sensitive neuron,” *J. Neurosci.*, vol. 19, pp. 1122–1141, 1999.
- [3] L. T. Tammero and M. H. Dickinson, “Collision-avoidance and landing responses are mediated by separate pathways in the fruit fly, *Drosophila melanogaster*,” *J. Exp. Biol.*, vol. 205, pp. 2785–2798, 2002.
- [4] H. Sun and B. J. Frost, “Computation of different optical variables of looming objects in pigeon nucleus rotundus neurons,” *Nature Neurosci.*, vol. 1, pp. 296–303, 1998.
- [5] A. P. Duchon, W. H. Warren, and L. P. Kaelbling, “Ecological robotics,” *Adapt. Behav.*, vol. 6, pp. 473–507, 1998.
- [6] M. Blanchard, P. F. M. J. Verschure, and F. C. Rind, “Using a mobile robot to study locust collision avoidance responses,” *Intl. J. Neural Syst.*, vol. 9, pp. 405–410, 1999.
- [7] M. Blanchard, F. C. Rind, and P. F. M. J. Verschure, “Collision avoidance using a model of the locust LGMD neuron,” *Robot. Autom. Syst.*, vol. 30, pp. 17–38, 2000.
- [8] S. Bermúdez i Badia and P. F. M. J. Verschure, “A collision avoidance model based on the lobula giant movement detector (LGMD) neuron of the locust,” in *Proc. Intl. Joint Conf. on Neural Networks*, 2004, pp. 1532–1537.
- [9] C. Mead, *Analog VLSI and Neural Systems*. Reading, MA: Addison-Wesley, 1989.
- [10] D. Abbott, A. Moini, A. Yakovlev, X. T. Nguyen, A. Blanksby, G. Kim, A. Bouzerdoum, R. E. Bogner, and K. Eshraghian, “A new VLSI smart sensor for collision avoidance inspired by insect vision,” in *Proc. SPIE*, vol. 2344, Boston, MA, Nov. 1994, pp. 105–115.
- [11] G. Indiveri, J. Kramer, and C. Koch, “Parallel analog VLSI architectures for computation of heading direction and time-to-contact,” in *Advances in Neural Information Processing Systems 8*, D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, Eds. Cambridge, MA: MIT Press, 1996.
- [12] R. Sarpeshkar, J. Kramer, G. Indiveri, and C. Koch, “Analog VLSI architectures for motion processing: from fundamental limits to system applications,” *Proc. IEEE*, vol. 84, no. 7, pp. 969–987, Jul. 1996.
- [13] N. Ancona, G. Creanza, D. Fiore, R. Tangorra, B. Dierickx, G. Meynants, and D. Scheffer, “Real-time, miniaturized optical sensor for motion estimation and time-to-crash detection,” in *Proc. SPIE*, vol. 2950, Berlin, Germany, 1996, pp. 75–85.
- [14] F. C. Rind and D. I. Bramwell, “Neural network based on the input organization of an identified neuron signaling impending collision,” *J. Neurophysiol.*, vol. 75, pp. 967–985, 1996.
- [15] B. Hassenstein and W. Reichardt, “Systemtheoretische Analyse der Zeit-, Reihenfolgen-, und Vorzeichenbewertung bei der Bewegungserperzeption des Rüsselkäfers *Chlorophanus*,” *Z. Naturforsch.*, vol. 11b, pp. 513–524, 1956.
- [16] W. Reichardt and T. Poggio, “Visual control of orientation behavior in the fly,” *Q. Rev. Biophys.*, vol. 9, pp. 311–375, 1976.
- [17] W. Reichardt and M. Egelhaaf, “Properties of individual movement detectors as derived from behavioral experiments on the visual system of the fly,” *Biol. Cybern.*, vol. 58, pp. 287–294, 1988.
- [18] M. Egelhaaf and A. Borst, “Transient and steady-state response properties of movement detectors,” *J. Opt. Soc. Amer. A*, vol. 6, pp. 116–127, 1989.
- [19] S. B. Laughlin, “Matching coding, circuits, cells, and molecules to signals—General principles of retinal design in the fly’s eye,” *Progr. Ret. Eye Res.*, vol. 13, pp. 165–196, 1994.
- [20] J. H. van Hateren, “Theoretical predictions of spatiotemporal receptive fields of fly LMCs, and experimental validation,” *J. Comp. Physiol. A*, vol. 171, pp. 157–170, 1992.
- [21] —, “Processing of natural time series of intensities by the visual system of the blowfly,” *Vision Res.*, vol. 37, pp. 3407–3416, 1997.
- [22] A. B. Saul and A. L. Humphrey, “Spatial and temporal response properties of lagged and nonlagged cells in cat lateral geniculate nucleus,” *J. Neurophysiol.*, vol. 64, pp. 206–224, 1990.
- [23] D. W. Dong and J. J. Atick, “Temporal decorrelation: a theory of lagged and nonlagged responses in the lateral geniculate nucleus,” *Network*, vol. 6, pp. 159–178, 1995.
- [24] E. H. Adelson and J. R. Bergen, “Spatiotemporal energy models for the perception of motion,” *J. Opt. Soc. Amer. A*, vol. 2, pp. 284–299, 1985.
- [25] M. Egelhaaf and A. Borst, “Transient and steady-state response properties of movement detectors,” *J. Opt. Soc. Amer. A*, vol. 6, pp. 116–127, 1989.

- [26] R. O. Dror, D. C. O'Carroll, and S. B. Laughlin, "Accuracy of velocity estimation by Reichardt correlators," *J. Opt. Soc. Amer. A*, vol. 18, pp. 241–252, 2001.
- [27] R. R. Harrison and C. Koch, "A robust analog VLSI Reichardt motion sensor," *Analog Integr. Circuits Signal Process.*, vol. 24, pp. 213–229, 2000.
- [28] W. Reichardt and M. Egelhaaf, "Properties of individual movement detectors as derived from behavioral experiments on the visual system of the fly," *Biol. Cybern.*, vol. 58, pp. 287–294, 1988.
- [29] D. M. Green and J. A. Swets, *Signals Detection Theory and Psychophysics*. New York: Wiley, 1966.
- [30] X. Arreguit, F. A. van Schaik, F. V. Bauduin, M. Bidiville, and E. Raeben, "A CMOS motion detector system for pointing devices," *IEEE J. Solid-State Circuits*, vol. 31, no. 12, pp. 1916–1921, Dec. 1996.
- [31] T. Delbrück, "Analog VLSI Phototransduction by Continuous-Time, Adaptive, Logarithmic Photoreceptor Circuits," California Institut. of Technol., Pasadena, Tech. Rep. CNS 30, 1994.
- [32] L. Watts, D. A. Kerns, R. F. Lyon, and C. A. Mead, "Improved implementation of the silicon cochlea," *IEEE J. Solid-State Circuits*, vol. 27, no. 5, pp. 692–700, May 1992.
- [33] B. Gilbert, "A precise four-quadrant multiplier with subnanosecond response," *IEEE J. Solid-State Circuits*, vol. 3, no. 12, pp. 365–373, Dec. 1968.



Reid R. Harrison (S'98–M'00) received the B.S. degree in electrical engineering from the University of Florida, Gainesville, in 1994 and the Ph.D. degree in computation and neural systems from the California Institute of Technology, Pasadena, in 2000.

He joined the University of Utah, Salt Lake City, in 2000, where he is now an Assistant Professor of Electrical and Computer Engineering and an Adjunct Assistant Professor of Bioengineering. He has over 30 refereed publications since 1999 in the fields of low-power analog and mixed-signal CMOS circuit

design, integrated electronics for neural interfaces and other biomedical devices, and hardware for biologically inspired computational systems.

Dr. Harrison organized the 2001 IEEE SSCTC Workshop on Low-Power Circuits in Arlington, VA. He received the National Science Foundation CAREER Award in 2002.