Winner-Take-All-Based Visual Motion Sensors

Erhan Özalevli, *Student Member, IEEE*, Paul Hasler, *Senior Member, IEEE*, and Charles M. Higgins, *Senior Member, IEEE*

***Abstract*—We present a novel analog VLSI implementation of visual motion computation based on the lateral inhibition and positive feedback mechanisms that are inherent in the hysteretic winner-take-all circuit. By use of an input-dependent bias current and threshold mechanism, the circuit resets itself to prepare for another motion computation. This implementation was inspired by the Barlow–Levick model of direction selectivity in the rabbit retina. Each pixel uses 33 transistors and two small capacitors to detect the direction of motion and can be altered with the addition of six more transistors to measure the interpixel transit time. Simulation results and measurements from fabricated VLSI designs are presented to show the operation of the circuits.**

***Index Terms*—Analog VLSI, motion, pixel parallel, vision chips, VLSI design.**

I. Introduction

USING commercial off-the-shelf parts, it is highly challenging to endow a small autonomous robot with a powerful, efficient real-time vision system. Custom VLSI vision sensors are an alternative approach that can simultaneously achieve low power consumption and high computational throughput. The principles and mechanisms of biological visual systems have been utilized for nearly two decades to create reliable, robust implementations of focal plane VLSI vision sensors [1]. In order to achieve maximum spatial resolution and maintain a reasonable fill factor, it is vital to perform this computation with a minimum pixel transistor and capacitor count. We have designed a novel highly compact analog VLSI implementation of visual motion computation based on the winner-take-all (WTA) circuit.

The WTA circuit, originally designed by Lazzaro *et al.* [2], takes an arbitrary number of current inputs and continuously produces a large voltage output for the largest input, and small voltage outputs for the rest. This circuit achieves the effect of lateral inhibition between local computational elements by forcing some transistors into the triode region. The WTA circuit has been elaborated to add hysteresis [3], adding positive feedback so that the winning input is "latched" until another input greatly exceeds it. We utilize the lateral inhibition and positive feedback properties of the hysteretic WTA (HWTA) circuit in motion computation. In addition, we introduce an input-dependent bias current and threshold mechanism that serve to reset the circuit for another cycle of motion detection.

This new implementation was inspired by the mechanisms underlying direction selectivity in the retina of the rabbit, first modeled by Barlow and Levick [4], which have inspired previous analog VLSI implementations of visual motion sensors [5]. The Barlow–Levick (BL) visual motion detection model is based on the relative timing of excitation and inhibition from two neighboring photoreceptors. Other authors have employed similar mechanisms to implement time-of-travel sensors [6], [7] but none have previously utilized the HWTA circuit which allows us to do this quite compactly.

After describing the computational model and its implementation, we present both simulation results and measured experimental data from fabricated VLSI designs to demonstrate the function of the circuit in detecting the local direction and speed of visual stimuli.

Manuscript received April 20, 2004; revised July 22, 2004, October 28, 2004, and March 27, 2005. This work was supported by the Office of Naval Research under agreement N68936-00-2-0002. This paper was recommended by Associate Editor T. S. Rosing.

E. Özalevli is with the Department of Electrical & Computer Engineering, University of Arizona, Tucson, AZ 85721 USA, and also with the Department of Electrical & Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA.

P. Hasler is with the Department of Electrical & Computer Engeering, Georgia Institute of Technology, Atlanta, GA 30332 USA.

C. M. Higgins is with the Department of Electrical & Computer Engineering, University of Arizona, Tucson, AZ 85721 USA (e-mail: higgins@neuromorph.ece.arizona.edu).

Digital Object Identifier 10.1109/TCSII.2006.875378

II. Computational Model

Barlow and Levick [4] modeled the direction selectivity of rabbit retinal ganglion cells, which fire strongly when a visual stimulus is moved across their receptive field in the "preferred" direction, and are inhibited by motion in the opposite direction. One formulation of their model is shown in Fig. 1(a). When a stimulus moves to the right (the preferred direction of unit R), excitation from photoreceptor PR1 arrives immediately, whereas inhibition from the neighboring photoreceptor PR2 arrives later. This causes the unit to be strongly activated. For a stimulus moving to the left, inhibition from PR2 arrives before, or at about the same time as, excitation from PR1 and the unit does not activate. Thus, the activation of each unit is specific to a particular direction of motion, dependent on which of the two photoreceptors was activated first.

A practical realization of this idea is shown in Fig. 1(b). The signals from two neighboring photoreceptors are high-pass filtered to remove sustained responses which do not contribute to motion computation. A half-wave rectification operation makes each pathway sensitive only to increases in illumination. Since an increase in one photoreceptor followed by a decrease in another photoreceptor does not necessarily indicate visual motion, we process only increases in illumination. The HWTA circuit, shown in Fig. 1(b) as two units which inhibit each other and also have positive self-feedback, detects the order of arrival of increases in illumination at the two photoreceptors and thus computes the direction of motion. These two units might correspond to two directionally selective retinal ganglion cells in the rabbit.

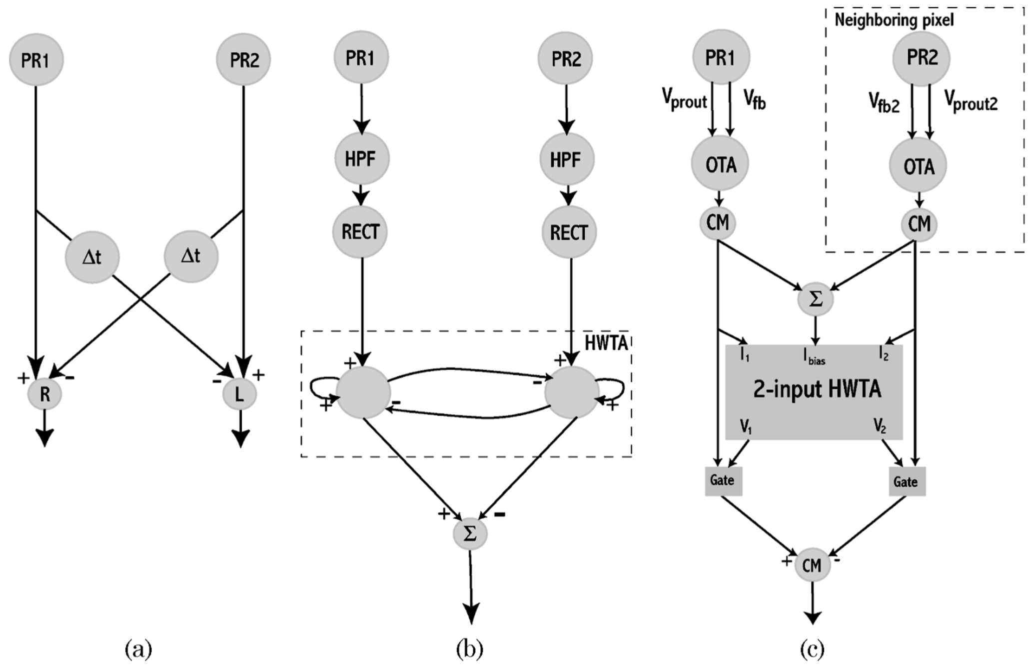

Fig. 1. Computational models of motion detection, and a diagram of their hardware implementation. PR1 and PR2 represent neighboring photoreceptors. (a) BL model for motion computation in rabbit retinal ganglion cells. Δt indicates a time delay, and $+$ and $-$ respectively represent excitation and inhibition. R and L indicate retinal ganglion cells sensitive respectively to rightward and leftward motion. If PR1 is stimulated before PR2 (that is, a rightward moving stimulus), excitation at unit R arrives before inhibition and the unit is activated. For leftward stimuli, inhibition arrives about the same time as excitation and the unit does not activate. (b) Computational model of HWTA-based motion sensor. HPF indicates a high-pass filter, RECT a half-wave rectification, and Σ a sum. The units inside the dashed box represent a model of the HWTA circuit with lateral inhibition and positive self-feedback. (c) Circuit block diagram of HWTA-based motion sensor. PR indicates an adaptive photoreceptor, OTA a 5-transistor transconductance amplifier, CM a current mirror, and Σ a sum. HWTA and gating circuits are described in the text.

In order to synthesize a single output that represents the direction of motion in its sign rather than two individual directionally selective units, we subtract the outputs of these two units.

III. Circuit Implementation

A block diagram of the circuit implementation of a pixel in the HWTA-based motion sensor is shown in Fig. 1(c). Each pixel consists of four major circuit blocks: an adaptive photoreceptor, an operation transconductance amplifier (OTA), an HWTA, and a gating circuit used to produce the final current output. Additionally, two current mirrors are employed as half-wave rectifiers and another is used to reverse the direction of a current for subtraction via Kirchoff's current law (KCL).

An adaptive photoreceptor [8], shown in Fig. 2, is used in order to transduce the local light intensity into electrical signals. This circuit adapts to the mean light intensity level on slow time scales and provides a high gain for transient signals. We utilize two voltage signals from this circuit: V_{prout}, which has high gain to transient changes in contrast and a bandpass frequency response, and V_{fb} which mainly reflects the mean light intensity and has a low-pass frequency response.

A 5-transistor OTA is used to subtract V_{prout} from V_{fb} in order to remove the adapted illumination level of the photoreceptor from its output. This accomplishes an operation similar to high-pass filtering and also converts the signal to a current. The saturating current output of the OTA is used along with a high photoreceptor transient gain to provide a stereotyped magnitude of current signals to be processed by the rest of the circuit for most stimulus conditions.

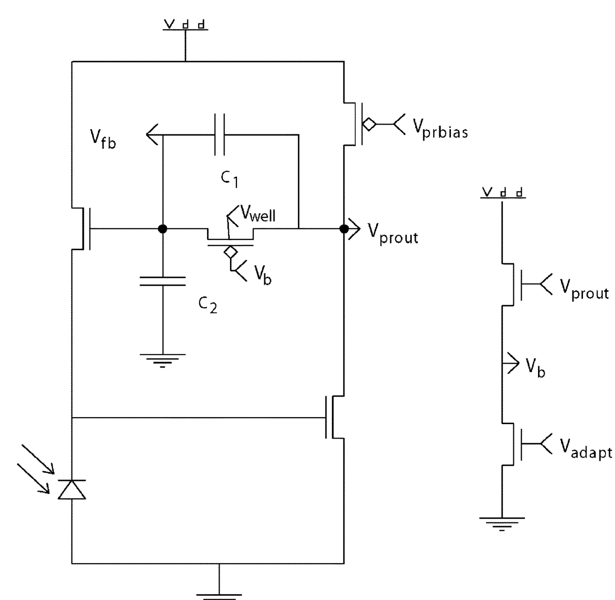

Fig. 2. Adaptive photoreceptor circuit [8]. The output V_{prout} responds strongly to transient changes in contrast, and the output V_{fb} represents the mean light intensity. V_{prbias} controls the frequency response and power consumption, and V_{adapt} and V_{well} the adaptation time constant.

The output current of the OTA is then processed through a current mirror, passing only a single direction of current and thus accomplishing a half-wave rectification operation. The half-wave rectified currents from two neighboring pixels are then input to a two-input HWTA circuit. The underlying idea behind the hysteretic WTA circuit operation is to find the

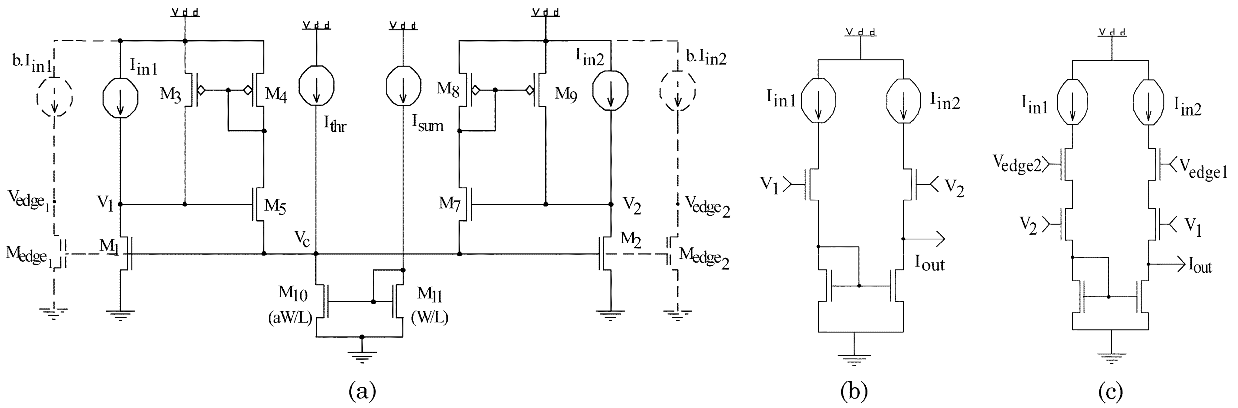

Fig. 3. HWTA-based motion detection circuit and its extended version. (a) Original HWTA circuit as used in motion computation (without dashed portions). $I_{\rm in1}$ and $I_{\rm in2}$ are the half-wave rectified output currents of the OTA circuits from two neighboring pixels, $I_{\rm thr}$ is a bias current that sets the minimum input current required to enable the circuit, and $I_{\rm sum}$ represents the sum of the input currents ($I_{\rm sum} = I_{\rm in1} + I_{\rm in2}$). V_1 and V_2 are the winner/loser outputs of the circuit. These outputs encode the direction of stimulus motion. The dashed portions of the circuit represent the additional elements used to implement the extended version of the HWTA-based sensor. Transistors $M_{\rm edge1}$ and $M_{\rm edge2}$ allow production of a voltage signal whose pulsewidth is proportional to the stimulus interpixel transit time. (b) Gating circuit of the original HWTA-based sensor. $I_{\rm in1}$ and $I_{\rm in2}$ are gated by winner/loser voltages V_1 and V_2 to determine the sign of the output current. (c) Gating circuit of the extended version of the HWTA-based sensor. Together with V_1 and V_2, $V_{\rm edge1}$ and $V_{\rm edge2}$ are used to gate the input currents.

maximum of the rectified currents after their sum exceeds a predetermined threshold. This maximum represents the pixel that was activated first, and thus the direction from which the visual motion signal arrived.

The HWTA circuit is shown in Fig. 3(a). The constant bias current usually employed in this circuit is replaced by $I_{\rm sum}$, the sum of the two input currents $I_{\rm in1}$ and $I_{\rm in2}$. These input currents represent the half-wave rectified OTA current outputs from two neighboring pixels. All of these currents are provided by duplicating output transistors of the rectifying current mirrors. The sum current is then mirrored (with gain a, explained below) to bias the HWTA circuit. A threshold current $I_{\rm thr}$ is subtracted from this sum current, ensuring that the circuit computes the direction of motion only if the sum of the input signals is bigger than the threshold. When the sum is smaller than the threshold level, the threshold current pulls up node V_c to reset the winning response so that the circuit can make a new decision for the next visual stimulus. The operational assumption for this circuit is that the input currents will both become small between the passing of subsequent image features.

When $a \cdot I_{\rm sum} = a \cdot (I_{\rm in1} + I_{\rm in2})$ is less than $I_{\rm thr}$, the common node, V_c, is pulled up. V_c becomes high for input signals smaller than $I_{\rm thr}$, and since it controls the gates of M_1 and M_2, it causes output voltages V_1 and V_2 to become very close to zero.

When $a \cdot I_{\rm sum}$ becomes bigger than $I_{\rm thr}$, V_c goes low and transistors M_5 and M_7 supply currents to keep V_c at a reasonable level that is enough to drive M_1 and M_2. After that, the circuit starts to function as a WTA circuit and finds the maximum of two input signals. Depending on the ratio of the input signals, the sum current is shared hysteretically between M_1 and M_2. The hysteretic connection improves the winner selection of the input signals. When there is a small difference between the input signals, one of the output voltages goes lower and the other goes higher allowing for the decision of a winner. If it is determined, for example, that V_1 is bigger than V_2, the amount of current driven by M_5 becomes larger than the current driven by M_7. This effect hysteretically increases the difference between the output voltages V_1 and V_2. As a result, V_1 becomes high, indicating that $I_{\rm in1}$ is the winner, while V_2 goes low indicating that $I_{\rm in2}$ is the loser.

The current mirror providing $I_{\rm sum}$ must have gain $a > 1$. If $I_{\rm in1}$ is the first signal activated by a motion stimulus, after some time $I_{\rm in2}$ must become much bigger than $I_{\rm in1}$ as the stimulus passes. In this case, since the winner is still $I_{\rm in1}$, the hysteretic connection approximately supplies $a \cdot I_{\rm in2}$ to M_1. In order to continue to suppress $I_{\rm in2}$ by the comparison between the currents through M_1 and M_2, the hysteresis current $a \cdot I_{\rm in2}$ must be greater than the other input current $I_{\rm in2}$, and thus a must be greater than unity.

To provide a bidirectional current output whose sign represents the direction of motion of the visual stimulus, an output gating circuit is used [Fig. 3(b)]. The input currents $I_{\rm in1}$ and $I_{\rm in2}$ (provided using output transistors of the original rectifying current mirrors) are gated by the HWTA output voltages V_1 and V_2 and subtracted by a current mirror.

In the general case, once the HWTA circuit has decided the direction of motion, the output gating circuit allows the selected (winning) input current pulse to pass through. The sign of the output current pulse depends on the direction of the moving stimulus. The width of the output current pulses in response to a periodic stimulus is

$$T_{\rm pw} = C/f_t - D \tag{1}$$

where C is a constant reflecting the proportion of a stimulus period during which $I_{\rm in}$ is greater than $I_{\rm thr}$, and D is a fixed delay indicating the amount of time the HWTA circuit takes to make its decision. Thus, the pulsewidth output of this sensor encodes the temporal frequency of a periodic input stimulus, but not its speed.

In the case of a visual stimulus moving in a direction orthogonal to the sensor orientation and activating both photoreceptors at nearly the same time, the HWTA will choose the input

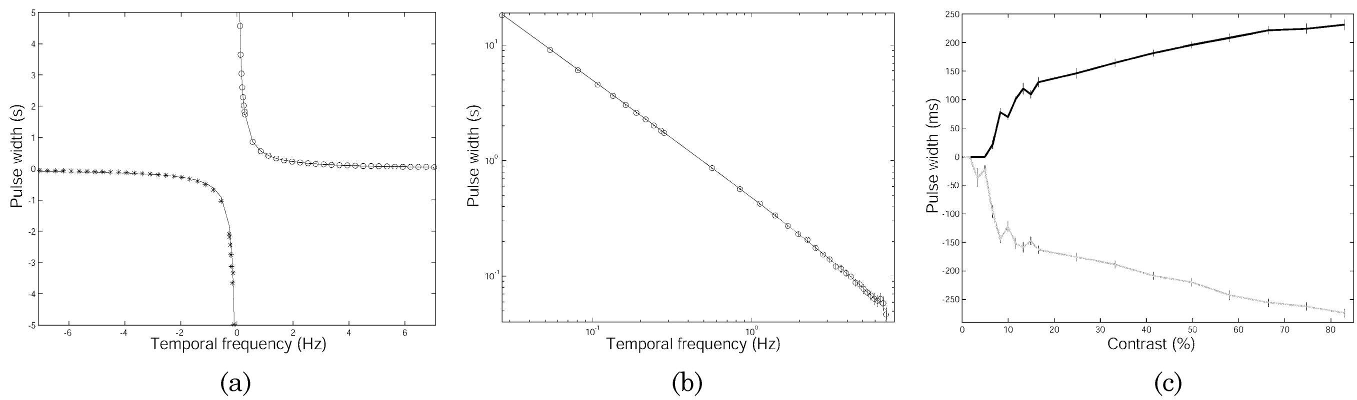

Fig. 4. Experimental measurements of the pulse output of the HWTA-based motion sensor to a moving square-wave stimulus of variable temporal frequency and contrast. Mean pulsewidth of motion output over 10 stimulus periods is shown. Temporal frequency data is shown with a theoretical fit to a constant multiple of inverse temporal frequency (see text). (a) Response to leftward (negative temporal frequency) and rightward (positive temporal frequency) moving gratings. Circles indicate the width of positive current pulses, asterisks negative current pulses. (b) Detail of temporal frequency response for rightward-moving gratings. Error bars (shown as vertical lines) represent standard deviation for each point. (c) Response to variable contrast moving gratings. Black line indicates width of positive current pulse responses to stimuli moving to the right, and gray line indicates width of negative current pulse responses to leftward stimuli. Error bars (shown as vertical lines) represent standard deviation for each point.

that comes first. If both inputs come at exactly the same time, transistor mismatch inside the HWTA circuit will determine the winner. Thus, in this version of the circuit there will be a motion output for a perfectly orthogonal stimulus (or equivalently for a flickering stimulus) that will be spatially random in sign due to spatial variations in transistor mismatch. This problem is fixed in the extended version of this circuit.

With a simple modification, this circuit can be extended to allow a pulsewidth coded measurement of stimulus speed. The extended HWTA circuit is shown in Fig. 3(a) (the addition to the original circuit is illustrated with dashed lines). The addition of transistors $M_{\rm edge1}$ and $M_{\rm edge2}$ produces signals $V_{\rm edge1}$ and $V_{\rm edge2}$ whose pulsewidth is proportional to the time taken for an image feature to pass from one photoreceptor to the next (the interpixel transit time). The addition of these transistors does not affect the operation of the HWTA circuit previously described.

The drain currents of the edge transistors are amplified versions of the input currents $I_{\rm in1}$ and $I_{\rm in2}$. The gain factor b must be chosen to be between $(a{+}1)$ and $(2a{+}1)$, as explained below. If we assume that $I_{\rm in1}$ is the first activated signal and $I_{\rm in2}$ is zero, the current through $M_{\rm edge1}$ will be $(a{+}1){\cdot}I_{\rm in1}{-}I_{\rm thr}$. The difference between the current $b \cdot I_{\rm in1}$ and the drain current of $M_{\rm edge1}$ causes $V_{\rm edge1}$ to go high. Thus, $V_{\rm edge1}$ rises when $I_{\rm in1}$ is activated. When $I_{\rm in2}$ is later activated, the drain current of $M_{\rm edge1}$ rises to $(a+1) \cdot I_{\rm in1} + a \cdot I_{\rm in2} - I_{\rm thr} \approx (2a+1) \cdot I_{\rm in1} - I_{\rm thr}$. Since this current is bigger than $b{\cdot}I_{\rm in1}$, $V_{\rm edge1}$ falls. Thus, if $I_{\rm thr}$ is small enough to be neglected, a value of b between $(a+1)$ and $(2a{+}1)$ will result in a $V_{\rm edge1}$ signal whose pulsewidth represents the interpixel transit time, which enables us to measure the speed of the visual motion stimulus.

In the extended circuit, both the voltage outputs of the hysteretic WTA circuit (V_1 and V_2) and the edge signals ($V_{\rm edge1}$ and $V_{\rm edge2}$) are used to gate the current output as illustrated in Fig. 3(c). Gating with both edge and "winner" signals ensures that the output both reflects the sign of motion and has a pulsewidth proportional to the interpixel transit time. Because winner signal V_1 for the local photoreceptor is used to gate the current from the neighboring photoreceptor, this gating removes responses to single-photoreceptor, orthogonal, or flickering stimuli by requiring that both photoreceptors must be stimulated in sequence.

IV. Results

The HWTA circuit shown in Fig. 3 (not including dashed portions) was fabricated through MOSIS in a 1.6-μm process, in a 10×11 array on a 2.1 mm $\times$ 2.1 mm die. In this section, we present characterizations of the pulse output of this fabricated circuit in response to moving square-wave grating stimuli presented using a liquid crystal display (LCD) screen. We also present a simulation of the extended circuit (including dashed portions in Fig. 3) to demonstrate sensitivity to stimulus speed. In simulations and on the fabricated chip, the current gain factors a and b were 2 and 4, respectively.

Fig. 4(a) shows the results of varying the temporal frequency of a 50% duty cycle square-wave grating stimulus with the spatial frequency held constant at 0.5 cycles per pixel. Fig. 4(a) shows that leftward-moving stimuli result in negative current pulses, and rightward-moving stimuli result in positive pulses. As explained in the previous section, the width of these pulses can be well fit by a multiple of inverse temporal frequency minus a fixed delay: $T_{\rm pw} = 0.5/f_t - 20$ ms, where the constants were empirically determined. The multiplier 0.5 is consistent with the 50% duty cycle of the stimulus. Fig. 4(b) shows detail of the temporal frequency response over 2.5 orders of magnitude in frequency. The error bars on this graph show the standard deviation of the pulsewidth, which is very small on this scale, less than 1% of the mean for the larger pulsewidths and rising to around 10% for the smaller widths. On a log–log plot, the graph is largely linear except at high frequencies, in which case the delay term becomes significant.

Fig. 4(c) shows the results of varying the contrast of a square-wave grating stimulus with fixed spatial frequency (0.5 cycles/pixel) and fixed temporal frequency (2.2 Hz). The saturation of response for high contrast levels is mostly due to

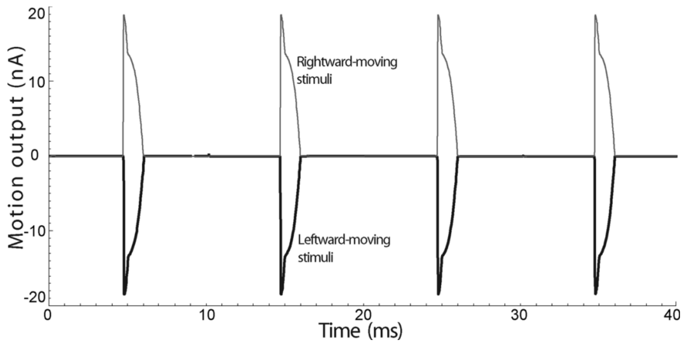

Fig. 5. Simulated current pulse output of the extended version of the HWTA-based sensor to 100-Hz square-wave stimuli that are set to have 1-ms delay between peaks at the two photoreceptors. Using the edge signals V_{edge1} and V_{edge2}, these outputs are active only during the transition of an image feature from one photoreceptor to the next, and thus are close to 1 ms in width.

differential input saturation of the OTA. As the contrast of the grating decreases, the amplitude of the photoreceptor response gets smaller. This slightly decreases the output pulsewidth because the sum of the two input signals to the HWTA circuit go below the fixed threshold sooner. Still, the sensor is capable of detecting the direction of stimulus motion to less than 5% contrast.

Fig. 5 shows a simulation of the extended version of the sensor in response to a 100-Hz square-wave grating set to have 1-ms delay between stimulation of successive photoreceptors. The current outputs show a pulsewidth close to 1 ms as expected from the speed of the stimulus. By gating the output of the original HWTA sensor with the edge signals V_{edge1} and V_{edge2}, the extended HWTA motion sensor encodes stimulus speed in its pulsewidth.

V. Discussion

We have described a visual motion algorithm inspired by direction selectivity in the rabbit retina which employs a hysteretic WTA circuit in a novel fashion. We have shown with experimental measurements of a fabricated VLSI chip that the direction of a moving visual stimulus can be measured using this circuit over more than two orders of magnitude in temporal frequency. Further, simulations show that pulsewidth encoded speed measurement can be added in an extended version of the circuit. The strong responses for both directions of motion and experimentally demonstrated sensitivity to low contrast caused by saturation at the transconductance amplifier stage makes this sensor a good candidate for biomimetic robotic applications.

TABLE I
Comparison of Pulse-Based Visual Motion Sensors

-	FT [7]	FTI [6]	HWTA-based
Trans. count	40	23	33
Cap. count	9	9	2
Area	$50{,}000\mu m^2$	$27{,}500\mu m^2$	$18{,}150\mu m^2$
Process	$2\mu m$	$2\mu m$	$1.6\mu m$
Max. pulse	140 msec	65 msec	>10 sec

The capacitors in a pixel-parallel motion computation circuit often drive the pixel area, and limit scaling to higher process resolutions. These capacitors are particularly important for generating very long pulses, and thus responses to very slow visual stimuli. The visual motion sensor described here uses the HWTA circuit to achieve these long pulses without using large capacitors. In Table I, we compare the HWTA-based motion sensor with two other pulse-based motion sensors. Taking the process resolution into account, the areas of facilitate-trigger-and-inhibit (FTI) [6] and HWTA-based sensors are very similar and both much smaller than the area of facilitate-and-trigger (FT) [7] sensor. However, in contrast to the FTI and FT sensors, there are only two capacitors used in the HWTA-based sensor. Despite this, the maximum pulsewidth of the HWTA-based sensor in Table I reveals the fact that it can detect stimuli at much lower speeds than the other two sensors.

References

[1] A. Moini, *Vision Chips*. Norwell, MA: Kluwer, 1999.
[2] J. Lazzaro, S. Ryckebusch, M. A. Mahowald, and C. A. Mead, "Winner-take-all networks of $O(n)$ complexity," *Adv. Neural Inf. Process. Syst.*, vol. 1, pp. 703–711, 1989.
[3] J. A. Starzyk and X. Fang, "CMOS current mode winner-take-all circuit with both excitatory and inhibitory feedback," *Electron. Lett.*, vol. 29, no. 10, pp. 908–910, May 1993.
[4] H. B. Barlow and W. R. Levick, "The mechanism of directionally selective units in rabbit's retina," *J. Physiol. (London)*, vol. 178, pp. 477–504, 1965.
[5] R. G. Benson and T. Delbrück, "Direction selective silicon retina that uses null inhibition," *Adv. Neural Inf. Process. Syst.*, vol. 3, pp. 756–763, 1991.
[6] J. Kramer, "Compact integrated motion sensor with three-pixel interaction," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, pp. 455–460, 1996.
[7] R. Sarpeshkar, W. Bair, and C. Koch, "Visual motion computation in analog VLSI using pulses," *Adv. Neural Inf. Process. Syst.*, vol. 5, pp. 406–412, 1993.
[8] S. C. Liu, "Silicon retina with adaptive filtering properties," *Analog Integr. Circuts Signal Process.*, vol. 18, pp. 243–254, 1999.