A Neuromorphic aVLSI Model of Global Motion Processing in the Fly

Shih-Chii Liu

Abstract—Flies orientate themselves quickly in an unstructured environment through motion information computed from their low-resolution compound eyes. The fly visual system is an example of a robust motion system that works in a natural environment. In this paper, we describe a low-power analog very large scale integration chip that models motion computation in the fly. The architecture of this motion chip closely follows the anatomical layout of the fly visual layers. The output of the chip corresponds to the responses of the wide-field direction-selective cells in the final layer of the visual system. The silicon chip has a one-dimensional array of 37 elementary motion detectors (EMDs) each of which provides local motion information. The EMD outputs are aggregated in a nonlinear way to produce a motion output that is independent of the stimulus size and contrast. We employed various circuit techniques to ensure robust motion computation in each processing stage. Results from the circuit fabricated in a 1.2-μm CMOS technology are compared with the responses of the direction-selective cells.

Index Terms—analog integrated circuits, biological cells, biological system modeling, motion analysis, neural network hardware, very-large-scale integration.

I. INTRODUCTION

C O N V E N T I O N A L motion sensors usually work well only under predefined conditions. Computer vision techniques use serially-scanned high-resolution charge coupled device (CCD) imagers and frame-based processing, and require a fast computational engine. Neuromorphic chips offer a variety of low-power, single-chip solutions for computing the motion of stimuli with a wide range of contrasts and speeds, under a large dynamic range of lighting conditions [1], [2], [9], [5], [12], [17], [21], [23], [24]. These solutions are based on algorithms that model the motion computation in biological systems.

In this work, we describe a low-power, analog very-large-scale-integrated (aVLSI) motion sensor that models the motion computation believed to be performed in the fly visual system. We use the fly visual system as our model since flies are capable of navigating rapidly in an unstructured environment through motion information computed from a robot without being hampered by CCD imagers that require a high scanout rate for real-time interaction.

The architecture of this chip closely follows the anatomical layout of the processing layers in the fly visual system. The details of the anatomy and physiology of the fly visual system is described in Section II. The global motion output is generated from the nonlinear aggregation of 37 elementary motion detectors (EMDs). The circuits of the different layers in the architecture are described in Sections III and IV. We employed different circuit techniques to reduce the effects of background intensity and mismatches between layers of processing within each pixel. These techniques are described in Section V. Measured results from a chip fabricated in a 1.2-μm CMOS technology are presented in Sections VI and VII. We compare the chip results with the responses of the wide-field direction-selective cells in the fly.

II. BIOLOGICAL BACKGROUND

Each of the two compound eyes of the fly consists of 3000–4000 ommatidia (or “little eyes”), each with its own lens. The compound eye has a resolution of about 1.5° and is equivalent to a 60 × 60 imager (depending on the species). Thus, the fly does not need high spatial resolution information to compute optical flow information needed for navigation. The visual system of the fly consists of four layers: the retina, the lamina, the medulla, and the lobula complex. The retinotopic arrangement of the pixels in the first three layers of the visual system makes it easy to map the structure of the system to silicon.

The photoreceptors in the retina transduce the incoming light into a voltage signal. Each photoreceptor employs local gain control by adapting to the background intensity, thus ensuring that the receptor has a high transient gain to changes in intensity and a low dc gain. The receptor primarily codes contrast information. The laminar monopolar cells (LMCs) perform a temporal bandpass filtering of the receptor signals without retaining any information about the background intensity [16], [18]. The LMC output is an amplified and inverted version of the receptor signal. Very little is known about the functionality of the cells in the medulla since the small size of the cells make them difficult to record.

There are 50–60 wide-field direction-selective cells in the lobula plate. Most of them respond to moving stimuli with graded responses. The remaining cells in the lobula plate respond with spikes. The outputs of all the direction-selective cells provide information about the translational motion along and rotational motion about different axes in the visual field.
This information controls the fly’s motor reflexes during flight. The responses of these cells can be modeled by the spatial integration of the outputs of local (EMDs). The first EMD model, Fig. 1(a), was first proposed by Hassenstein and Reichardt in 1956 [14] and the spatial integration of the outputs of the EMDs by a wide-field direction-selective cell is shown in Fig. 1(b). Hassenstein and Reichardt’s model was based on the optomotor response of the weevil Chlorophanus, but it also reproduces the responses of the wide-field cells. An EMD receives inputs from two neighboring pixels (or photorece-ptors). One of the inputs is delayed through a lowpass filter. This delayed signal is multiplied with the input at the adjacent pixel. The output of the left multiplier of the EMD is sensitive to stimuli moving from left to right, while the output of the right multiplier is sensitive to stimuli moving from right to left. The two outputs are subtracted to remove the common-mode or flicker dependence on the input.

To compute the EMD output, $R$ in Fig. 1(a), we assume that the input to a receptor is a sine wave $I + \Delta I \sin(\omega t + \theta)$ and the input to the neighboring receptor is $I + \Delta I \sin(\omega t + \theta + (2\pi\alpha/\lambda))$. Here $I$ is the mean intensity level, $\omega$ is the temporal frequency, $\lambda$ is the spatial wavelength, $\theta$ is the spatial phase of the sinewave, and $\alpha$ is the pixel spacing. The output $R$ can then be expressed as

$$R = \Delta I^2 \frac{\omega \tau}{1 + \omega^2 \tau^2} \sin\left(\frac{2\pi\alpha}{\lambda}\right) d$$

where $d$ is +1 for the preferred direction and is −1 for the null direction. Notice that the output is dependent on the input temporal frequency $\omega$ and the time constant $\tau$ of the lowpass filter. The dependence of $R$ on the square of $\Delta I$ only holds true in the responses of the biological cells for contrasts below 0.3. For higher contrasts, the response saturates.

III. aVLSI ELEMENTARY MOTION DETECTOR

The architecture of one of the silicon EMDs is shown in Fig. 2. The circuitry comprising the layers in this architecture is described in this section. Unless otherwise mentioned, the circuits are operated in the subthreshold domain.

A. Photoreceptor

The adaptive photoreceptor circuit first described by Delbrück [6] and modified by Liu [19] in Fig. 3 transduces via a photodiode, the incoming light $E$ into a voltage signal $V_{\text{ph}}$. The output $V_{\text{ph}}$ codes the logarithm of the light intensity $E$ if $M_5$ is in subthreshold. Changes in the local intensity $dE$ are amplified by the transient gain of the circuit $A$, which is determined by the ratio of the capacitors, that is, $A = (C_1 + C_2)/C_1$. The circuit adapts to the background intensity at a time constant that depends on the capacitors $C_1$ and $C_2$, and the conductance of the adaptive element $M_1$. The conductance of $M_1$ is set by its gate voltage $V_{\text{g}}$, which is derived from the source follower circuit on the right of Fig. 3. The details of the circuit are dis-
Fig. 2. Block diagram of the architecture of the silicon EMD for computing motion. The light input goes into a photoreceptor circuit with local gain control. The receptor output is then bandpass filtered by the LMC circuit. The output of the LMC circuit splits into two pathways. In one pathway, the signal is delayed through a low-pass filter. ON and OFF transient signals are generated from the delayed and undelayed signals. Subsequently, the ON and OFF delayed signals in each pixel are correlated (or multiplied) with the undelayed signals from the adjacent pixel. The two outputs in the figure are sensitive to a stimulus moving from left to right.

Fig. 3. Circuit diagram of photoreceptor circuit. The feedback consists of a resistor implemented by a pFET transistor $M_1$. The conductance of the resistor is controlled by the external bias $V_{in}$. By changing the value of $V_{in}$, we can change the conductance of $M_1$. Hence, also the receptor's adaptation time constant. The transient change in $V_{ph}$ in response to a step change in the intensity, $(dE/E)$ can be described by

$$dV_{ph} = U_T \frac{C_1 + C_2}{\kappa C_1} \frac{dE}{E}$$

where $U_T$ is the thermal voltage and $\kappa$ is the gate efficiency in driving the transistor. The output of the circuit responds primarily to the signal contrast and has a dc operating point that adapts to the background intensity.

B. Lamina

The circuitry in the second layer of the chip models the responses of the cells in the laminar layer. The response of the LMC cell is an amplified and inverted version of the input from the receptor. The dc dependence of the LMC output on the background intensity is zero. To implement this function, we use an inverting differentiator with an adaptive element $M_2$ between the output and the input as shown in Fig. 4(a). The transient gain in this circuit is set by the ratio of capacitors $(C_3 + C_4)/(C_4)$, i.e.,

$$\frac{dV_{MC}}{dV_{ph}} = \frac{C_3 + C_4}{C_4}.$$  (3)

The output returns to its dc level that is set by the current in the bias transistor $M_2$. The time constant that the circuit takes to return to steady state is set by the capacitances and the conductance of $M_3$. The conductance of $M_3$ is set by a source follower.
D. Transient Current Generation

The fourth layer models possible properties of cells in the medulla and lobula complex of the fly visual system. Each pixel has two hysteretic differentiator circuits [17] (see Fig. 6), whose inputs come from $V_{\text{LMC}}$ and $V_{\text{DMC}}$, respectively. In this circuit, we generate both an ON transient current $I_{\text{on}}$ and an OFF transient current $I_{\text{off}}$ from temporal changes in $V_{\text{LMC}}$ and $V_{\text{DMC}}$. The currents $I_{\text{on}}$ and $I_{\text{off}}$ are given by

\[ I_{\text{on}} = C_5 \frac{dV_{fb}}{dt}, \quad I_{\text{off}} = -C_5 \frac{dV_{fb}}{dt}. \tag{4} \]

Only one current is on at a time, while the other is essentially zero. Using (2) and (3), and assuming that $(dV_{fb}/dt) = (dV_{bmc}/dt)$, we get

\[ I_{\text{on}} = C_5 \frac{C_3 + C_4}{C_4} \frac{C_1 + C_2}{C_1} U_T \frac{dE}{dt}, \tag{5} \]
\[ I_{\text{off}} = -C_5 \frac{C_3 + C_4}{C_4} \frac{C_1 + C_2}{C_1} U_T \frac{dE}{dt}. \tag{6} \]

The voltages $V_{\text{on}}$ and $V_{\text{off}}$ from the differentiator circuit in one pixel are correlated with the voltages $V_{\text{on2}}$ and $V_{\text{off2}}$ from the differentiator circuit in the adjacent pixel using current correlators [4] (see Fig. 7). The current correlator, which is analogous to the multiplier in the Reichardt model, consists of two transistors in series. The output of the correlator consisting of $M_1$ and $M_2$ is

\[ I_{\text{corr}} = \frac{I_{\text{on}} I_{\text{on2}}}{I_{\text{on}} + I_{\text{on2}}} \tag{7} \]

The currents from two of the correlators are summed together to create $I_{\text{ref}}$, the output current for the preferred direction. The currents from the other two correlators are summed together to create $I_{\text{null}}$, the output current for the null direction. These currents feed into the aggregating circuit, which is described in the following section.

IV. NONLINEAR INTEGRATION OF EMDs

The wide-field cells produce a global motion output from the aggregation of the outputs of the EMDs. A major concern in aggregating the outputs is that we have to ensure that the global motion measure is not confounded with the number of EMDs that are stimulated by the external environment. Images of the natural environment have roughly a $1/f$ distribution in their spatial power spectra with most of the power concentrated in the low spatial frequencies (that is, natural images have very few high contrast edges). Hence, many of the EMDs will not be stimulated when a fly is in a natural environment. If we simply sum the EMD outputs, then the motion output will increase as more EMDs are stimulated by the environment.

Both physiological recordings [11] and optomotor experiments [22] have shown that the outputs of the wide-field cells are independent of the stimulus size except for small-sized stimuli. This gain control property can be modeled by the biophysical properties of the cell and its dendrites as shown by Borst [3] (see Fig. 8). The membrane potential $V_m$ of the cell is determined by the excitatory, inhibitory, and leak conductances $g_e$, $g_i$, and $g_{\text{leak}}$, and the corresponding reversal potentials are $E_e$, $E_i$, and $E_{\text{leak}}$.
Fig. 5. Response of inverting differentiator to a 30-mV p–p square wave for different values of $V_r$ and $V_{\text{resh}}$. Response of $V_{\text{mc}}$ and $V_p$ to different values of $V_{\text{resh}}$.

Fig. 6. Hysteretic differentiator circuit. This circuit generates transient currents $I_{\text{on}}$ and $I_{\text{off}}$ from changes in the output voltage $V_{\text{mc}}$ from the LMC circuit.

$E_r$, $E_{\text{leak}}$: The conductances $g_e$ and $g_i$ depend on the preferred and null motion outputs of an EMD, respectively. If the motion of the stimulus is in the preferred direction, then both $g_e$ and $g_i$ increase but $g_e > g_i$. The membrane potential can be written as

$$V_m = \frac{N E_e g_e + N E_i g_i + E_{\text{leak}} g_{\text{leak}}}{N g_e + N g_i + g_{\text{leak}}} \quad (8)$$

where $N$ is the number of activated EMDs. For large $N$ and assuming $E_{\text{leak}} = 0$, we get

$$V_m = E_e \frac{g_e}{g_e + g_i}. \quad (9)$$

In the equation above, we assume that $E_e = -E_i$. So for large $N$, the membrane potential only codes the motion independent of the number of activated EMDs.

We can map this model to the current-mode aggregator circuit shown in Fig. 9. Here, the outputs of the EMDs map to input currents (rather than conductances) and the reversal potential

$$E_e$$

sets the global bias current $I_b$. Each module consists of the transistors $M_1$ to $M_7$, which implement a Gilbert normalizer [10]. However, $M_1$, $M_6$, and $M_7$ are shared amongst all the modules so there are only four transistors in each module. The preferred ($I_{\text{pref}}$) and null ($I_{\text{null}}$) outputs of each silicon EMD are normalized with respect to the global bias current $I_b$. The input current $I_{\text{leak}}$ acts as the leak conductance and determines the number of inputs at which the output saturates. The output current of the left leg $I_{\text{pref}}$ of each module is given by

$$I_{\text{pref}} = I_b \frac{I_{\text{pref}}}{I_{\text{pref}} + I_{\text{null}} + I_{\text{leak}}}. \quad (10)$$

Authorized licensed use limited to: ETH BIBLIOTHEK ZURICH. Downloaded on June 07,2010 at 00:17:56 UTC from IEEE Xplore. Restrictions apply.
If \( N \) EMDs are activated and assuming that the preferred and null outputs are the same for all the \( N \) EMDs, then \( I_{\text{agpref}} \) is now

\[
I_{\text{agpref}} = I_b \frac{I_{\text{pref}}}{N I_{\text{pref}} + N I_{\text{null}} + I_{\text{leak}}}.
\]

(11)

The global preferred current, \( I_{\text{agpp}} \), is the sum of all \( I_{\text{agpref}} \)’s, i.e.,

\[
I_{\text{agpp}} = \sum_{i=1}^{N} I_{\text{agpref}} = I_b \frac{N I_{\text{pref}}}{I_{\text{pref}} + I_{\text{null}}}.
\]

(12)

assuming that \( N I_{\text{pref}} + N I_{\text{null}} \gg I_{\text{leak}} \). Similarly, the global null current, \( I_{\text{aggn}} \), is given by

\[
I_{\text{aggn}} = I_b \frac{I_{\text{null}}}{I_{\text{pref}} + I_{\text{null}}}.
\]

(13)

We obtain the final motion output \( I_{\text{mot}} \) by subtracting \( I_{\text{aggn}} \) from \( I_{\text{agpp}} \), so

\[
I_{\text{mot}} = I_{\text{agpp}} - I_{\text{aggn}} = I_b \frac{I_{\text{pref}} - I_{\text{null}}}{I_{\text{pref}} + I_{\text{null}}}.
\]

(14)

A nice property of spatially aggregating the outputs of the EMDs is that the temporal variation of the response of one EMD in response to a spatial grating input is removed in the final wide-field cell output. The reduction in the variation of the response occurs because the response of the EMDs to the moving pattern are spatially out of phase with each other. This phenomena has been demonstrated in the HS cell by Single, \textit{et al.} using calcium imaging [25].

V. MISMATCHES

Silicon hardware, like biological neural hardware, has mismatches between pixels or cells on the same substrate. The coefficient of variation (CV)—defined as the standard deviation over the mean—of the mismatches from the biological LMC neurons is in the order of five percent between recorded cells [16]. The CV as measured from a silicon imager fabricated in a 2-\( \mu \)m CMOS ORBIT process is about 6%–10% [19]. The mean and standard deviation for the CV measure is obtained from the peak-to-peak responses of the photoreceptor output of the individual imager pixels in response to an LED that was modulated by a 10-Hz square wave. The pixels’ mismatches need to be handled if comparisons or correlations between pixels are part of the motion computation algorithm.

We employed various circuit techniques to reduce the effects of mismatches and common-mode input factor in each processing stage. The ways we used to increase the robustness of the computation are as follows.

- We bandpass filtered the signal from the photoreceptor going into the LMC circuit so that there is no common-mode dependence on the LMC output.
- The dc level of the input from the LMC circuit to the hysteretic differentiator circuit affects the transient dynamics of the generation of the ON and OFF currents. Since the dc voltage of the LMC input to the hysteretic differentiator circuit is constant, the dynamics are not affected by the background light intensity.
- We have chosen to use transient current inputs instead of voltage inputs to a correlator circuit so that we would not have to deal with mismatches in the DC input levels at the multiplier. A previous hardware implementation of the correlation model, where voltage outputs from
neighboring pixels are fed into a multiplier circuit, failed to provide a robust direction-sensitive output. The signal was not robust because the input dc mismatches between the pixels creates different outputs at the multiplier. The use of transient currents ensures that there will be no output current (that is, zero motion signal) if there is no input change at the receptor output. The mismatches in the voltage multiplier could be reduced by layout and sizing techniques or nulled using tunnelling-injection structures [7] or UV techniques. Additional circuitry complexity and area would have been required to implement these features.

- We used current-mode circuits whenever possible (for example, in the nonlinear aggregator circuit) so that we would not have to deal with the limited input linear range of a transconductance amplifier. In addition, current-mode circuits are also usually more compact than the voltage-mode counterparts.

VI. EXPERIMENTAL RESULTS FROM AN aVLSI EMD

In these experiments, we measured the output of a single fabricated aVLSI EMD. We used the output current of the EMD to charge or discharge a capacitor on a node $V_{\text{mot}}$, which also has a leak conductance to a reference voltage $V_{\text{ref}}$ of 2.5 V. If the input to the chip moves in the preferred direction $V_{\text{mot}}$ increases above 2.5 V. If the input moves in the null direction $V_{\text{mot}}$ decreases below 2.5 V. The temporal responses of an EMD to a sinewave grating of contrast 0.2 moving in the preferred direction are shown in Fig. 10(a) and in the null direction in Fig. 10(b). The contrast described in the experiment results is the Michelson contrast $I_{\text{max}} - I_{\text{min}} / I_{\text{max}} + I_{\text{min}}$, where $I_{\text{max}}$ and $I_{\text{min}}$ are the maximum and minimum intensities respectively of the stimulus. Notice the frequency doubling of the output in these figures.

We plotted the dependence of the aVLSI EMD on the temporal frequency of the input by presenting to the chip a fixed spatial frequency sinewave moving at different velocities. We measured the temporal frequency of the input as seen by the photoreceptor circuit and the average EMD output for each velocity. The results are shown in Fig. 11(a). The upper curve corresponds to the stimulus moving in the preferred direction and the lower curve is obtained from a stimulus moving in the null direction. The shape of the response is similar to that measured in the HS cell of the fly as shown in Fig. 11(b). This figure was adapted from the work of Hausen [15]. The output of this wide-field cell can be predicted by the EMD model as shown by Egelhaaf and Borst [8]. The range of frequencies for which there is a response extends from about 1 Hz to about 25 Hz.

We also measured the direction selectivity of the output to a square-wave grating moving at different orientations. The response of the chip can be approximated by a cosinc function as shown in Fig. 12.

VII. EXPERIMENTAL RESULTS FROM AN aVLSI HS CELL

The figures described in this section are taken from the output of an aVLSI HS cell, which aggregates the outputs of 37 EMDs.
using the circuit shown in Fig. 9. For experimental measurements, the outputs of the aggregator circuit $I_{agg}$ and $I_{agg1}$ are used to charge and discharge a capacitor, respectively, on a node $V_{ref}$, which also has a leak conductor to a reference voltage $V_{ref}$. These currents could also be read individually.

A simple summation of the EMDs outputs will lead to an HS output that increases as more EMDs are stimulated. In Fig. 13, we showed the effect of linearly adding the motion outputs of the individual EMDs. The current outputs of the EMDs are used to charge or discharge a capacitor on a common node. The temporal variation in the output is not present but the output saturates at the power rails if enough EMDs are stimulated. Under this condition, it would not be possible to measure the motion quantitatively from the output. Even if we summed the preferred and null output currents of the EMDs to create a global preferred current and a global null current, these global currents increase with the number of stimulated EMDs.

By using nonlinear aggregation of the EMD outputs, we can prevent saturation of the output at the rails. The transient response of the aggregator circuit to an input sinewave grating that starts moving from rest to 14 Hz in the preferred direction and in the null direction is shown in Fig. 14. The large transient response at the start of the motion is not always present at every trial and depends on the output states of the individual blocks in the EMD.

The temporal frequency response of the silicon HS cell was measured for three different contrasts (Fig. 15). The curves for the two higher contrasts, 0.58 and 0.92, almost overlap, thus showing that aggregator output is independent of contrast at least above 0.58. This contrast saturation property is necessary if we do not want the motion output to depend on the signal contrast. Egelhaaf and Borst have shown physiologically that the response of the biological HS cell saturates for contrasts above...
motion output from the nonlinear aggregation of 37 elementary motion detectors so that the motion output is independent of stimulus area and contrast. We employed various circuit techniques to generate a robust motion signal by reducing the effects of background intensity and mismatches on the motion computation. This motion sensor differs from some of the other motion circuits, for example, by Sarpeshkar et al. [24] and Etienne-Cummings et al. [9], as the entire computation up the motion output stage is performed on analog signals. In the implementations mentioned above, a pulse which codes the presence of an “edge” in the scene is first created before the motion of this “edge” is computed. Our circuit implementation differs from Harrison and Koch’s implementation in 1998 [12], and includes the contrast normalization circuit, the laminar circuits, and the nonlinear aggregation response of the wide-field cells.

We have used the outputs of these motion sensors as inputs to a controller system on a robot [20]. The system exhibits two behaviors similar to those of the fly; optomotor response and fixation. This system allows us to evaluate models of visuo-motor controllers in working systems. The interaction of these continuous-time sensors with the motor system provides a powerful method of exploring real-time interaction of the robot with its environment.

ACKNOWLEDGMENT

The author would like to thank R. Douglas for supporting this paper, the MOSIS foundation for fabricating this circuit, and A. M. Whatley and T. Delbrück for proofreading and critical reading of this paper’s manuscript.

REFERENCES


