

A CMOS Bio-Inspired 2-D Motion Direction Sensor Based on a Direction Computation Method Derived From the Directionally Selective Ganglion Cells in the Retina

Wen-Chia Yang, *Student Member, IEEE*, Li-Ju Lin, Herming Chiueh, *Member, IEEE*, and Chung-Yu Wu, *Fellow, IEEE*

Abstract—A CMOS bio-inspired motion direction sensor structure and its associated computation method are proposed. Both method and structure with excitation-inhibition operation are derived from the directionally selective ganglion cells (DSGCs) in the retina to mimic their functions. Edge-number normalization for direction calculation and pseudo-random tessellation (PRT) structure for pixel layout arrangement are also proposed to enhance the accuracy of the computation. An experimental chip based on the proposed method and structure has been designed, fabricated, and measured. The chip comprised 32×32 pixels with a pixel size of $63 \times 63 \mu\text{m}^2$ and a fill factor of 12.8%. The total chip size is $3.3 \times 4.2 \text{ mm}^2$ and the power consumption is 9.9 mW in the dark and 21 mW at a maximum clock rate of 10 MHz with 3.3-V power supply. The fabricated chip has been measured with different moving patterns, and a computation error of less than 11 degrees has been accomplished. This verifies the correct functions of the proposed motion direction sensor. With the capability of real-time motion detection and processing under low power dissipation, the proposed sensor is feasible for many applications.

Index Terms—Directionally selective ganglion cell, direction sensor, retinal chip, vision chip.

I. INTRODUCTION

THE retina, which is viewed as an extension of the brain [1], not only senses the incident scene of the environment but also interprets the scene in a format called visual language for the visual cortex in the brain [2], [3]. It also extracts the temporal and the spatial features of the scene for further analysis in the brain. One of the most essential features is motion direction. It is found that a kind of retinal cell called the directionally selective ganglion cell (DSGC) is related to motion direction detection

[4]. In recent research, the structure and mechanism of the cell have been understood. Based upon preliminary results [5], an operational model of the DSGC has been established. The model is suitable for real-time estimation of motion directions of an incident scene.

A DSGC model similar to [5] was developed and realized in a direction-selective silicon retina chip whose output has the characteristic of the DSGC [6]. The chip was tested with simple patterns to verify the DSGC functions. Further algorithms of direction computation for arbitrary patterns have not been developed.

Other hardware implemented motion sensors have been proposed [7]–[11]. They adopt the correlation-based algorithms inspired by biological models. The concept of a correlation-based sensor is to find the correlation between adjacent pixels over an interval of time and then determine the velocity and the direction of motion. Although its performance for velocity detection is good, the accuracy of 2-D direction detection is not verified when complex patterns are used. Because the range of the correlations is limited to neighboring pixels, these chips may suffer from the well-known aperture problem [12], [13] when calculating motion directions of complex patterns.

A stand-alone CMOS motion sensor based on retinal-processing circuits is proposed to realize a compact and real-time motion detection system [14]. The chip is implemented with a modified correlation-based algorithm and a pixel-level correlator to detect motion direction and speed. It has a high accuracy of direction and speed and can be operated over a wide range of speeds. However, only a stripe pattern perpendicular to the motion direction was tested. Moreover, the aperture problem that may degrade the accuracy of direction computation has not yet been solved in the proposed chip. It also has a higher power dissipation of 120 mW.

This paper proposes a direction computation structure based on the DSGC model to extract the motion direction of scenes. The proposed structure is realized in CMOS technology, and a prototype chip is fabricated and tested. With the capability of real-time motion detection and processing with low power dissipation, the proposed sensor possesses many potential applications such as target tracking systems, wireless optical mice, pointing devices, optical remote controllers, and digital image stabilization systems.

In Section II, the directional computation structure is described. The techniques of edge-number normalization and

Manuscript received January 27, 2011; revised April 24, 2011 and May 16, 2011; accepted May 25, 2011. Date of publication June 07, 2011; date of current version November 02, 2011. This work was supported in part by National Science Council (NSC), R.O.C., under project 100-2220-E-009-022 and in part by “Aim for the Top University Plan” of the National Chiao Tung University and Ministry of Education, Taiwan. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Paul C. P. Chao.

W.-C. Yang and C.-Y. Wu are with the Department of Electronics Engineering and the Institute of Electronics, National Chiao Tung University, Hsinchu 300, Taiwan (e-mail: peterwu@mail.nctu.edu.tw).

L.-J. Lin is with the Biomimetic Systems Research Center, National Chiao Tung University, Hsinchu 300, Taiwan.

H. Chiueh is with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu 300, Taiwan (e-mail: chiueh@soclab.org).

Digital Object Identifier 10.1109/JSEN.2011.2158642

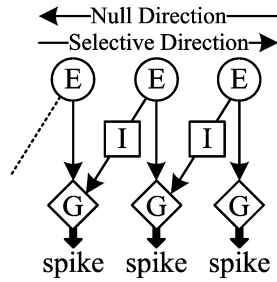


Fig. 1. Model of directionally selective ganglion cell (DSGC) in the retina.

pseudo-random tessellation adopted in the sensor design are demonstrated in detail. The overall simulation results of the proposed structure are also presented to verify its correct functions. In Section III, both the design and operation of the chip implemented to realize the proposed structure are detailed. A test chip with the sensing array size of 32×32 is designed and fabricated with a standard $0.35\text{-}\mu\text{m}$ CMOS technology. The experimental results of the chip are presented and discussed in Section IV. Finally, the conclusion is drawn in the last section.

II. DIRECTION COMPUTATION STRUCTURE

A. The Model of Directionally Selective Ganglion Cell

The DSGC model as proposed in [5] is shown in Fig. 1 where the cell responds most strongly when the stimulus of scenes is moving in a predetermined direction called the selective direction. The stimulus moving in the other direction called the null direction produces weaker or even no response. In Fig. 1, the upper ball labeled E denotes the excitatory input, which can be viewed as the transient input to the cell. The box labeled I sends an inhibitory signal to the neighboring cell in the null direction when it is activated into the “on” state by the excitatory input offered by E. The diamond labeled G at the bottom is the DSGC, which collects both the excitatory and the inhibitory inputs from E and I blocks. The DSGC determines whether to fire a spike depending on both of its excitatory and inhibitory inputs. Only if E is “on” and I is “off” will the cell fire a spike output.

Two operational examples of the DSGC model are demonstrated in Fig. 2. In Fig. 2(a), a stimulus bar moves in the selective direction. The stimulus triggers E_0 and then E_1 . After E_1 is activated, I_0 is activated and I_1 remains “off”. As G_1 receives the excitation from E_1 without being inhibited by I_1 , it fires a spike output. Although I_0 is activated, the generated inhibitory signal does not affect G_0 because G_0 has already fired a spike output in the previous state. In Fig. 2(b), the bar moves in the null direction. E_2 is triggered first and then I_1 is triggered to be in the “on” state. If I_1 maintains its “on” state until E_1 is triggered in the current movement, G_1 fires nothing because the excitation is suppressed by the inhibitory signal from I_1 . This state-maintaining behavior can be easily modeled as a low-pass filter or a delay. The time constant of the filter determines the range of the speeds that the cell can detect. Fig. 2 shows that the directional selectivity function of DSGC is realized by the model of Fig. 1 with opposite-directional interconnection structure and the excitation-inhibition operation.

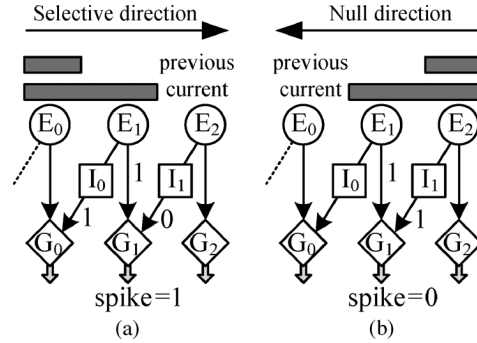


Fig. 2. DSGC model operation with a stimulus moving in different directions. Stimulus moves in (a) the selective direction and (b) the null direction.

B. Direction Computation Structure and Method

The model shown in Fig. 1 provides a simple and robust concept for motion direction extraction. Based upon the model, a direction computation structure that is suitable for silicon implementation is proposed. In the proposed structure, two major elements are adopted. Firstly, a binary imager named the retinal-processing circuit is used to simplify the processing procedure. Several structures of the retinal-processing circuit realizing a binary imager have been proposed [15]–[17] where the original photocurrent is compared with the spatially-smoothed one to obtain the binary output. No complex analog-processing circuit, ADC, or multibit processor is needed in the pixel level, which is an advantage that helps to shrink pixel size and decrease power consumption. Secondly, the low-pass filter in the inhibition path is replaced by the digital delay element, a D flip-flop, so that the delay time can be tuned accurately.

Fig. 3 shows the proposed pixel structure with interpixel connections. Each pixel contains a retinal-processing circuit as the imager, two registers for frame storage, one multiplexer, two NAND gates, and four direction selection (DS) units to mimic DSGC function. The retinal-processing circuit generates the binary image output. The binary outputs of two successive frames are sampled and stored in two D flip-flops in registers denoted CF and PF. CF denotes the output of the current frame, whereas PF denotes that of the previous frame. The interval of two consecutive sampled points of the registers can be accurately controlled by the clock signal labeled *im_load* in the figure. The non-inverting output of the registers CFQ and PFQ are selected by a multiplexer and sent to the four DS units as excitatory signals. The inverting outputs CFQB and PFQB are sent to the DS units of the four neighboring pixels as the inhibition signals for the excitation-inhibition operation. The four DS units receive excitatory and inhibitory signals and generate the outputs $d(+X)$, $d(+Y)$, $d(-X)$, and $d(-Y)$ in a pixel so that 2-D motion direction can be extracted. Each DS unit contains an AND gate and a multiplexer. With an excitation signal within the cell and an inhibition signal from one of its four neighboring cells as inputs, the AND gate can realize the excitation-inhibition function of DSGC.

The two NAND gates are used to generate \overline{on} and \overline{off} signals indicating the pixel’s status of being turned on or off. The status is determined according to the inverting and the non-inverting outputs of the two registers. The \overline{on} signal is used to control the

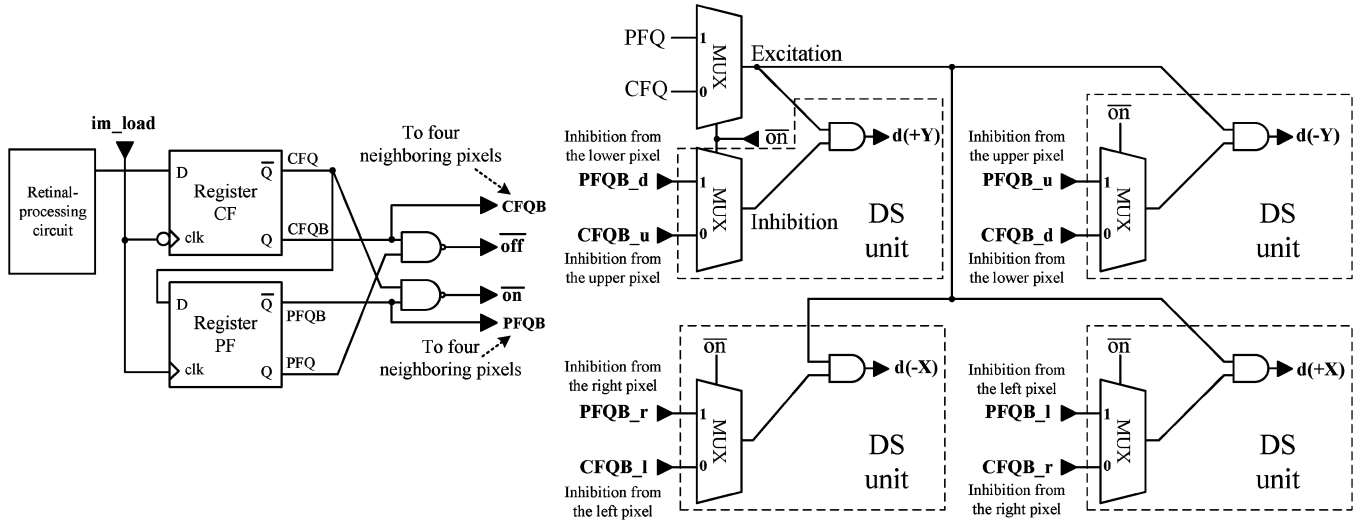


Fig. 3. Pixel structure adopted in the proposed direction computation structure.

multiplexer in the DS unit. If both CFQ and PFQ are 1 or 0, $\overline{\text{on}}$ and $\overline{\text{off}}$ is 1 indicating the stimulus on the pixel is not altered. If $\overline{\text{on}}$ signal is 0 (on is 1) and $\overline{\text{off}}$ is 1, it indicates that the pixel is turned on with CFQ = 1 and PFQ = 0. In this case, CFQ is selected as the excitation signal of the four DS units by a multiplexer. Signal CFQB from the neighboring pixel that can provide opposite directional inhibition is selected by a multiplexer in each DS unit. If the stimulus is moving along +Y direction, CFQB_u is 1. Thus, $d(+Y)$ is 1 and $d(-Y)$, $d(+X)$, $d(-X)$ are 0 to indicate the motion direction is +Y. If $\overline{\text{off}}$ signal is 0 ($\overline{\text{off}}$ is 1) and $\overline{\text{on}}$ is 1, they indicate that the pixel is turned off with CFQ = 0 and PFQ = 1. In this case, PFQ is selected as the excitation signal of the four DS units, whereas PFQB from the neighboring pixel that can provide identical directional inhibition is selected as the inhibitory signal in each DS unit. If the stimulus is moving along +Y direction with the final excitation at the previous frame, PFQB_d is 1 and PFQ is 1. Thus, $d(+Y)$ is 1 and others are 0 to indicate the motion direction is +Y. In this way, the structure can detect the correct motion direction at the back end of the stimulus. In Fig. 3, the level signals instead of transient signals are taken as excitation and inhibition signals for reliable operation. Thus, the inhibition signals are suitably selected to realize the correct function of the DSGC model.

By combining the four outputs $d(+X)$, $d(-X)$, $d(+Y)$, and $d(-Y)$ along with $\overline{\text{on}}$ and $\overline{\text{off}}$ signals, as shown in Fig. 3, a local motion vector (LMV) and a local edge direction (LED) in terms of vector representation can be obtained. LMV and LED can be expressed as (See equation at bottom of page) where (i, j) represent the position indexes of the pixel. According to (1) and (2),

if the four outputs $d(+X)$, $d(-X)$, $d(+Y)$, and $d(-Y)$ are 1, 0, 1, and 0, the LMV is (1, 1) indicating an up-right motion. If the four outputs are 1, 1, 1, and 1, the LMV is (0, 0) indicating no motion. There exist nine possible LMVs. They are (0, 0), (0, 1), (0, -1), (1, 0), (1, 1), (1, -1), (-1, 0), (-1, 1), and (-1, -1) representing stillness, up, down, right, up-right, down-right, left, up-left, and down-left motion, respectively. LED is used to extract the direction of a local edge defined by the perpendicular direction of the edge. The values of LED are used to enhance the accuracy of the motion direction computation. DS units are shared for both extraction of LMV and LED to save hardware. The major difference between LMV and LED is that LED gives static information regardless of the “on” or “off” status of pixel, whereas LMV depends on the pixel moving status. According to the proposed structure, if the pixel is not altered or turned off ($\overline{\text{on}} = 1$), LMV/LED are extracted from PFQ/PFQB. If the pixel is turned on, $\overline{\text{on}} = 0$, only LMV is extracted and LED is not extracted. Thus, LED is not double counted.

A 1-D example showing how the proposed structure works is given in Fig. 4. The figure shows two successive frames with a bar pattern moving one pixel pitch to the right. The positions of transient pixels are indicated by “on” and “off” signals. Two DS units with output signals $d(+X)$ and $d(-X)$ in Fig. 3 are required for 1-D direction extraction. The turn-on pixel denoted as A is examined first. At the pixel A, $\overline{\text{on}} = 0$. According to the structure shown in Fig. 3, CFQB = 1 from the left pixel inhibits the excitation CFQ and $d(-X) = 0$, whereas CFQB = 0 in the right pixel makes $d(+X) = 1$. This is consistent with the previously explained DSGC behavior. According to (1), the values

$$\text{LMV}(i, j) = \begin{cases} (d(+X) - d(-X), d(+Y) - d(-Y)) & \text{if } \overline{\text{on}} = 0 \text{ or } \overline{\text{off}} = 0 \\ (0, 0) & \text{otherwise} \end{cases} \quad (1)$$

$$\text{LED}(i, j) = \begin{cases} (d(+X) - d(-X), d(+Y) - d(-Y)) & \text{if } \overline{\text{on}} = 1 \\ (0, 0) & \text{otherwise} \end{cases} \quad (2)$$

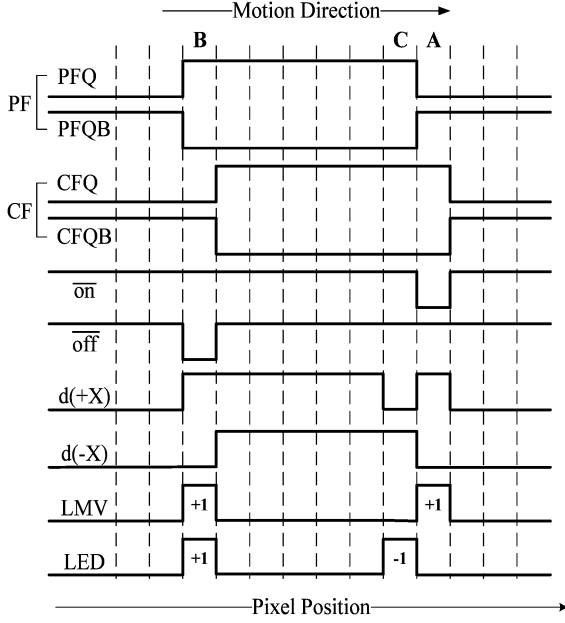


Fig. 4. 1-D example of the direction computation structure.

of $d(+X)$ and $d(-X)$ make $LMV = +1$, which indicates right motion.

At the turn-off pixel B where $\overline{off} = 0$, PFQ(PFQB) is chosen as the excitation(inhibition) signal. Thus, $LMV = +1$ can be obtained. If CFQ and CFQB are chosen as those of pixel A, incorrect results of $d(+X)$ and $d(-X)$ would be obtained at B.

The purpose of choosing different signals at pixels A and B through the multiplexer is to utilize static excitation and inhibition signals rather than transient signals and thus simplify the hardware. Therefore, only static signals plus \overline{on} and \overline{off} signals are used to realize the exhibition-inhibition operation in the proposed structure. As the motion pixels are located in the static edges, performing the excitation-inhibition operation with static signals is sufficient to extract both local motion vector (LMV) and local edge direction (LED) as long as “on” and “off” conditions are known. Another advantage of the proposed structure using static signals and the multiplexers is that the local edge direction LED can also be extracted without any extra hardware overhead. At the pixel B, $LED = +1$.

As for the no-motion pixels, $d(+X)$ and $d(-X)$ are 1 except at pixel C where $d(+X) = 0$ and $d(-X) = 1$. Because $\overline{on} = 1$ and $\overline{off} = 1$ at pixel C, the value of LMV is still 0 according to (1). However, LED of this pixel is -1 because this pixel is indeed located at the edge.

A 2-D example of the excitation-inhibition operation is shown in Fig. 5 where a pattern moves along $+X$ direction. The arrows show the LMVs produced by DS units. It can be observed that the LMVs of all the diagonal edges are $(1, 1)$ or $(1, -1)$, i.e., up-right or down-right motion, whereas those of vertical edges are $(1, 0)$. They are the same as the local edge directions (LEDs) defined as the perpendicular direction to edges. This phenomenon is caused by the aperture phenomenon. To reduce the error caused by the aperture problem in motion direction computation, a normalization method is proposed in the computation.

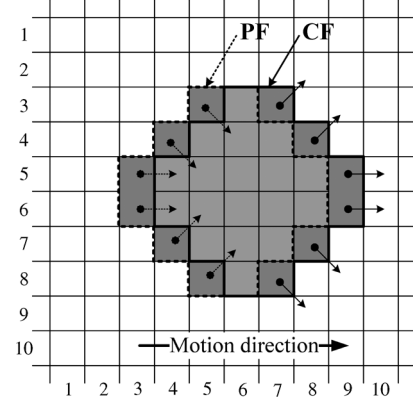


Fig. 5. 2-D example of the motion computation method when the symmetrical pattern moving right along the x-axis is shown with LMVs.

Local motion numbers (LMNs) comprising LMN_X , LMN_Y , LMN_{D1} , and LMN_{D2} are defined by the values of LMV as

$$LMN_X(i, j) = \begin{cases} 1 & \text{for } LMV(i, j) = (1, 0) \\ -1 & \text{for } LMV(i, j) = (-1, 0) \\ 0 & \text{Otherwise} \end{cases} \quad (3)$$

$$LMN_Y(i, j) = \begin{cases} 1 & \text{for } LMV(i, j) = (0, 1) \\ -1 & \text{for } LMV(i, j) = (0, -1) \\ 0 & \text{Otherwise} \end{cases} \quad (4)$$

$$LMN_{D1}(i, j) = \begin{cases} 1 & \text{for } LMV(i, j) = (1, 1) \\ -1 & \text{for } LMV(i, j) = (-1, -1) \\ 0 & \text{Otherwise} \end{cases} \quad (5)$$

$$LMN_{D2}(i, j) = \begin{cases} 1 & \text{for } LMV(i, j) = (1, -1) \\ -1 & \text{for } LMV(i, j) = (-1, 1) \\ 0 & \text{Otherwise.} \end{cases} \quad (6)$$

where the subscripts X, Y, D1, and D2 represent the components of LMV projected on X, Y, 45° diagonal, and -45° diagonal axes, respectively. A LMN can be 1, -1 , or 0 showing positive direction, negative direction, or zero motion on the axes. For example, $(LMN_X, LMN_Y, LMN_{D1}, LMN_{D2}) = (0, 1, 0, 0)$ indicates that the pixel's LMV is $(0, 1)$ showing an up motion. LMNs equal to $(0, -1, 0, 0)$ represent a down motion.

On the other hand, local edge numbers (LENs) comprising LEN_X , LEN_Y , LEN_{D1} , LEN_{D2} are defined by the value of LED as

$$LEN_X(i, j) = \begin{cases} 1 & \text{for } LED(i, j) = (1, 0) \text{ or } (-1, 0) \\ 0 & \text{Otherwise} \end{cases} \quad (7)$$

$$LEN_Y(i, j) = \begin{cases} 1 & \text{for } LED(i, j) = (0, 1) \text{ or } (0, -1) \\ 0 & \text{Otherwise} \end{cases} \quad (8)$$

$$LEN_{D1}(i, j) = \begin{cases} 1 & \text{for } LED(i, j) = (1, 1) \text{ or } (-1, -1) \\ 0 & \text{Otherwise} \end{cases} \quad (9)$$

$$LEN_{D2}(i, j) = \begin{cases} 1 & \text{for } LED(i, j) = (1, -1) \text{ or } (-1, 1) \\ 0 & \text{Otherwise.} \end{cases} \quad (10)$$

LENs are used to indicate the direction of an edge at a pixel. They can be either 0 or 1 showing that there is no edge or an edge at a particular pixel. For example, $(LEN_X, LEN_Y, LEN_{D1}, LEN_{D2})$ of a pixel located in a diagonal edge whose direction is the down-right with $LED = (1, -1)$ is equal to $(0, 0, 0, 1)$. As edge direction is of static information, down-right and up-left directions are classified as the same type. Therefore, $LENs = (0, 0, 0, 1)$ represent either a down-right or up-left edge.

The summations of all LEMs and LMNs of pixels in the whole $N \times N$ array except those at the boundary are calculated to obtain global motion numbers (GMNs) and global edge numbers (GENs), respectively. GMNs and GEND2 can be expressed as

$$GMN_S|_{S \in \{X, Y, D1, D2\}} = \sum_{i=2}^{N-1} \sum_{j=2}^{N-1} LMN_S|_{S \in \{X, Y, D1, D2\}}(i, j) \quad (11)$$

$$GEN_S|_{S \in \{X, Y, D1, D2\}} = \sum_{i=2}^{N-1} \sum_{j=2}^{N-1} LEN_S|_{S \in \{X, Y, D1, D2\}}(i, j). \quad (12)$$

As the boundary pixels have incomplete inhibitory inputs, their LMVs and LEMs are not considered in (11) and (12) to avoid accuracy degradation.

The motion velocity vector components along the X-axis and Y-axis, denoted as v_x and v_y , are calculated in terms of GENs and GMNs as

$$v_x = W_X \cdot \frac{GMN_X}{GEN_X} + \frac{GMN_{D1}}{GEN_{D1}} + \frac{GMN_{D2}}{GEN_{D2}} \quad (13)$$

$$v_y = W_Y \cdot \frac{GMN_Y}{GEN_Y} + \frac{GMN_{D1}}{GEN_{D1}} - \frac{GMN_{D2}}{GEN_{D2}}. \quad (14)$$

In (13) and (14), GMNs are normalized to GENs and the normalized terms are then summed together with weighting factors W_X and W_Y . The reason for adopting the weighting factors W_X and W_Y in (13) and (14), respectively, is to make the projection of diagonal vectors D1 and D2 onto the X-axis and Y-axis with suitable weighting. In (13), the projection of the two diagonal axis terms GMN_{D1}/GEN_{D1} and GMN_{D2}/GEN_{D2} onto the X-axis should be multiplied by a factor of $1/\sqrt{2}$ so that each term has the same contribution as GMN_X/GEN_X . Thus physical values of W_X and W_Y are close to $2\sqrt{2} \approx 2.8$. The weighting factors W_X and W_Y are also used as semi-empirical factors to minimize the errors. Therefore, the weighting factors are chosen as 3 to minimize the calculation errors of motion directional angles for different patterns according to system simulation results.

The normalization proposed in (13) and (14), called edge-number normalization, is important to enhance the calculation accuracy for arbitrary patterns. Two examples are demonstrated to show the accuracy of the proposed calculation method. In Fig. 5 where a symmetrical octagon pattern moving along the X-axis is shown, LMVs are all perpendicular to the contour of the octagon because of the aperture phenomenon. The LMVs of the diagonal edges are obviously wrong because the pattern is moving to the right along the X-axis. The calculated GENs and GMNs from (11) and (12) are listed in Table I. All nonzero GMNs are the same because of the symmetry of the pattern.

For the purpose of comparison, a set of equations for the calculation of motion velocity vector components without normalization are given as

$$v'_x = W_X \cdot GMN_X + GMN_{D1} + GMN_{D2} \quad (15)$$

$$v'_y = W_Y \cdot GMN_Y + GMN_{D1} - GMN_{D2} \quad (16)$$

TABLE I
CALCULATED GENs AND GMNs FOR THE OCTAGON AND TRIANGLE PATTERNS

GMNs/GENs	Octagon	Triangle
GMN_X	4	9
GMN_Y	0	0
GMN_{D1}	4	4
GMN_{D2}	4	1
GEN_X	4	9
GEN_Y	4	0
GEN_{D1}	2	4
GEN_{D2}	2	1

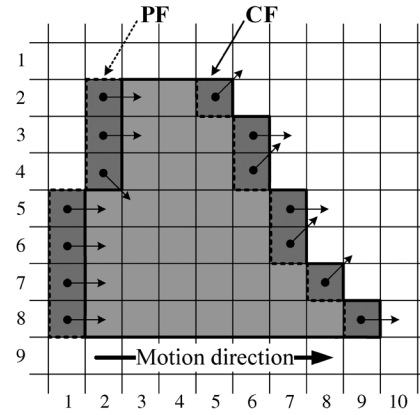


Fig. 6. 2-D example with asymmetric pattern moving right along the x-axis to show the effectiveness of edge-number normalization.

Using the values in Table I, one has $(v_x, v_y) = (7, 0)$ and $(v'_x, v'_y) = (12, 0)$. In this case, the calculated directions with or without normalization are the same, i.e., right movement along the X-axis. The two terms GMN_{D1} and GMN_{D2} in (16) are the same because of the symmetry of the test pattern.

Fig. 6 shows the second example where an asymmetrical triangle pattern moving along the X-axis is tested. The results are also listed in Table I. It can be seen that the values of diagonal GMN_{D1} and GMN_{D2} are not the same because of the pattern asymmetry. The calculated $(v_x, v_y) = (3, 0)$ and $(v'_x, v'_y) = (14, 3)$. The direction with normalization is correct and the other one is not. The result with normalization is correct because the normalization of GMN_{D1} and GMN_{D2} to GEN_{D1} and GEN_{D2} makes them cancel each other. Without normalization, GMN_{D1} and GMN_{D2} in (15) and (16) cannot be cancelled.

The binary imager of the retinal processing circuit is used to digitize the image to 1-bit. This may change the pattern shapes. Since edge-number normalization is adopted in the directional angle calculation, its accuracy is less pattern shape dependent. Thus the use of a binary imager does not severely degrade accuracy.

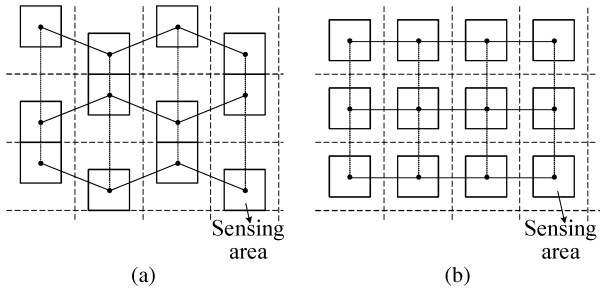


Fig. 7. (a) Proposed pseudorandom tessellation structure of pixels. (b) Conventional rectangular tessellation structure.

The directional angle denoted as θ ranging from 0 degree to 360 degrees is calculated using

$$\theta = \begin{cases} 90^\circ & \text{for } v_x = 0 \text{ and } v_y > 0 \\ 270^\circ & \text{for } v_x = 0 \text{ and } v_y < 0 \\ \tan^{-1}\left(\frac{v_y}{v_x}\right) & \text{for } v_x > 0 \text{ and } v_y > 0 \\ \tan^{-1}\left(\frac{v_y}{v_x}\right) + 180 & \text{for } v_x < 0 \\ \tan^{-1}\left(\frac{v_y}{v_x}\right) + 360 & \text{for } v_x > 0 \text{ and } v_y < 0. \end{cases} \quad (17)$$

C. Pseudo-Random Tessellation(PRT) Structure

As mentioned in subsection B, the effect of the aperture phenomenon on asymmetrical patterns degrades the accuracy of the motion direction computation. Hence, another technique called the *pseudo-random tessellation* (PRT) structure is proposed to improve the accuracy. The idea comes from the natural distribution of retinal cells of human or other primates that are largely random in all areas of the retina [18]. The random tessellation structure increases the degree of the irregularity of the perceived patterns; thus, the directional distribution of the edges is equalized to some degree. The concept of random tessellation structure can be applied to the motion direction sensor array to improve the computational accuracy. Nevertheless, a truly random tessellation structure is not realizable in silicon chips because of the layout complexity. Therefore, the PRT structure as shown in Fig. 7(a) is proposed to mimic the characteristic of the random distribution of retina cells while the regular structure is still retained to keep the layout simple. As compared with the conventional rectangular tessellation structure in Fig. 7(b), the pixels of PRT in a column are positioned in a saw shape. The next column is the mirrored image of the previous one shown in Fig. 7(a) where the interconnections among neighboring pixels are also sketched. The simulation results of rectangular tessellation, PRT, and the truly random tessellation structures will be compared in the following subsection.

D. Simulation Results

System simulation with MATLAB is performed to verify the proposed method for motion direction computation. The test patterns shown in Fig. 8(a), 8(b), and 8(c) are down sampled to the resolution of 32×32 in rectangular tessellation, PRT, and truly random tessellation structures. The sampled image is then converted to the binary one with a spatial template that

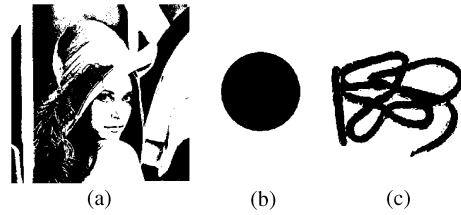


Fig. 8. Patterns tested in the system simulation. (a) Lena, (b) disc, and (c) wire.

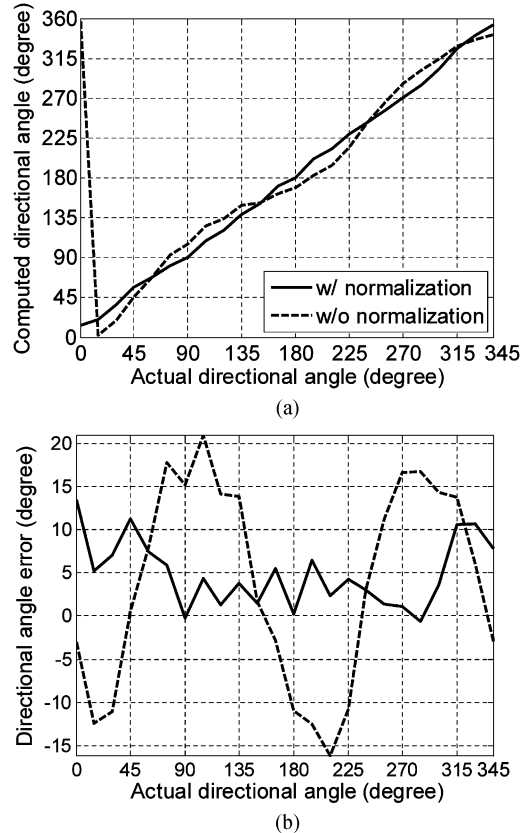


Fig. 9. Simulation results of directional angles with and without edge-number normalization: (a) computed directional angle versus actual ones; (b) directional angle error versus actual ones.

models the behavior of the retinal-processing circuit. The sampled binary image is processed with the proposed computation method to extract global motion direction. The comparison of calculated directional angle θ versus actual ones with the Lena pattern between the methods with and without normalization for rectangular tessellation structure is shown in Fig. 9(a). It can be seen that the curve with normalization is more linear than that without normalization. The errors are also smaller for the method with normalization as shown in Fig. 9(b) where the errors for most actual directional angles are less than 10 degrees for the Lena pattern.

Table II lists the results of the simulation for different tessellation structures. Fig. 10 shows the truly random tessellation structure used in the simulation. The computational angle is calculated by summing the GENs and GMNs over 11 consecutive frames when the patterns are moving 0.2 pixel pitch per frame time for each actual directional angle. The maximum direction angle error is chosen as the index of evaluation. According to the

TABLE II
SIMULATION RESULTS OF MATLAB SYSTEM SIMULATION ON
MAXIMUM DIRECTIONAL ANGLE ERRORS OF THREE DIFFERENT
TESSELLATION STRUCTURES

Edge-number normalization	Tessellation structure	Lena (degrees)	Disc (degrees)	Wire (degrees)
No	Rectangular	21	5	20
Yes	Rectangular	13	3	10
Yes	Truly random	11	5	14
Yes	PRT	10	5	8

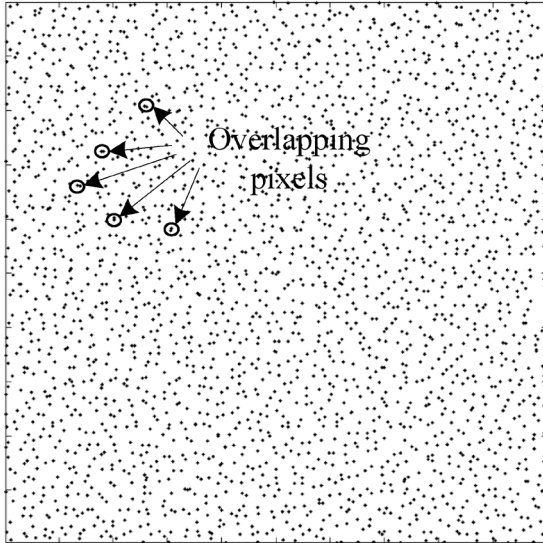


Fig. 10. Truly random tessellation used in the simulation. The dots indicate the center of the sensing region.

results in Table II, the normalization always decreases the error for all the patterns, even in the rectangular structures. The PRT reduces the error for complex patterns, but slightly increases the error for the simple disc pattern. The errors of the truly random tessellation structure are similar to those of the PRT structure with edge-number normalization except for the wire pattern. This is caused by the inevitable overlapping of pixels, which occurred in the truly random tessellation structure, as shown in Fig. 10.

The patterns are also simulated at the velocity of 1 pixel pitch per frame time. The maximum directional angle errors are 10, 4, and 8 degrees for Lena, disc, and wire patterns, respectively. The results of moving at 0.2 and 1 pixel pitch per frame time are almost the same. This shows that the computation error is less relevant to the velocity of the patterns. To conclude system simulations, the method with edge-number normalization in the pseudo-random tessellation structure is the most accurate and suitable one in computing motion direction angles.

III. ARCHITECTURE AND CIRCUIT DESIGN

The chip architecture that realizes the proposed computation structure is shown in Fig. 11. The architecture consists of a pixel array with a pixel structure of Fig. 3, row selector, column multiplexer, LENS/LMNs decoders, summation unit, accumulator, controller, address counter, and output drivers. There are $32 \times$

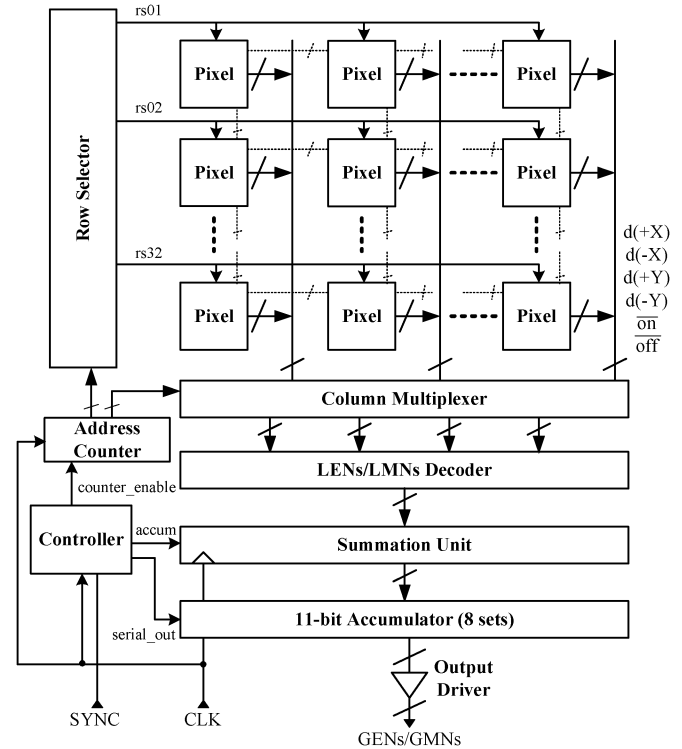


Fig. 11. Whole chip architecture.

32 pixels in the array. Firstly, the controller sends control signals to sample the CF and the PF image frames. Each pixel produces $d(+X)$, $d(-X)$, $d(+Y)$, $d(-Y)$, $\overline{\text{on}}$, and $\overline{\text{off}}$ output signals. The row selector and the column multiplexer select four pixels' outputs simultaneously. These outputs are further decoded to LENS and LMNs by LEN/LMN decoders. The summation unit calculates the partial summation of LENS and LMEs. The partial summation of the four pixels is then accumulated by accumulators to calculate GENs and GMEs. The accumulation repeats until all of the pixels are counted. It takes 256 clock cycles to complete all the accumulation operations. Finally, the accumulators send out the calculated GENs and GMNs for off-chip computation of the motion direction angles.

The circuit schematic of the retinal-processing circuit in the pixel structure of Fig. 3 is shown in Fig. 12. A photodiode rather than a phototransistor is adopted to improve the transient response. The photodiode senses the incident light and converts it to photocurrent. The amplifier along with MN0 clamps the voltage at the negative terminal of the photodiode to V_x to ensure appropriate bias. The clamp circuit and the associated MN0 also isolate the large capacitance of the photodiode and improve the speed of the current mirror. The photocurrent is mirrored and amplified to reach MN3–MN4 whose current is denoted as I_{ph} , by cascading two cascode current mirrors. The current of MN3 flows in or out through transistors MR1 and MR2 that are used as voltage controlled resistors and connected to the upper and the right neighboring pixels, respectively. The voltage-controlled resistors connected to the lower and the left neighboring pixels are not shown in Fig. 12. The four transistors form a resistive smoothing network that computes the weighted local average of the original image [19]–[22]. The smoothed current denoted as

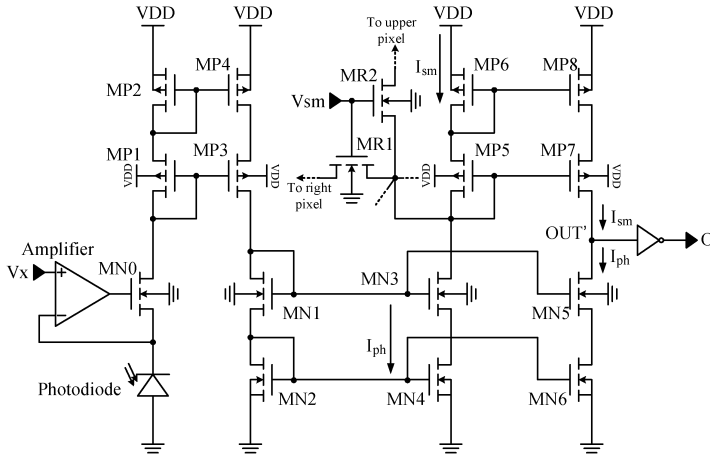


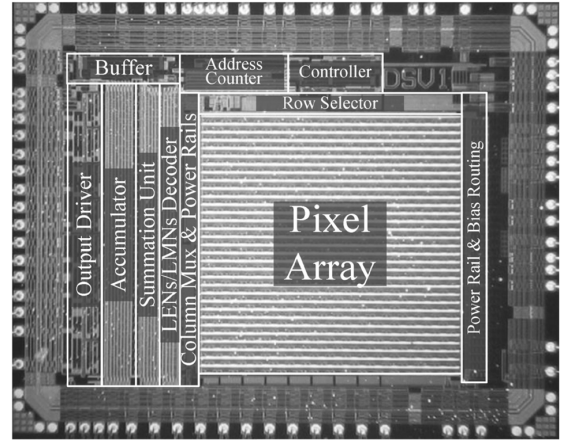
Fig. 12. Schematic of the retinal-processing circuit.

I_{sm} is then mirrored to MP7 and MP8 through MP5 and MP6 for further processing. The subtraction of current I_{ph} and I_{sm} is converted to voltage at the node OUT' buffered by an inverter. The subtraction of unsmoothed and smoothed image performs edge enhancement [23]. The final output carrying edge-enhanced information is the binary conversion of the original image. The pixel structure is similar to [14] and [17] except that an inverter instead of a Schmitt trigger is used.

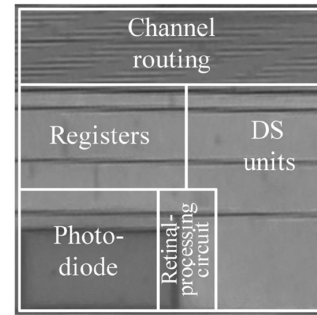
IV. EXPERIMENTAL RESULTS

A test chip is designed and fabricated with the TSMC 0.35- μm double-poly-quadruple-metal standard CMOS technology. The photographs of the whole chip, a single pixel, four pixels as a subset, and part of the array are shown in Fig. 13(a), (b), (c), and (d), respectively. The whole chip area including ESD protecting pads is 3.3 mm \times 4.2 mm. The pixel pitch is 63 μm and the photosensing area is 17.4 μm \times 29.1 μm with a fill factor of 12.8%. The rest of the region in the pixel is covered by the top metal layer (not shown) to shield the incident light and prevent the light from interfering with the operation of other transistors. Top metals also serve as power lines for the pixel array. Fig. 13(b) and 13(c) show the chip photographs of a single pixel and a subset of the array, respectively. It contains four pixels as a basic cell to form the whole array. Every single pixel in the subset is the mirror image of any neighboring pixels. The layout area required for the PRT structure is the same as that of the regular rectangle structure. Part of the array is shown in Fig. 13(d) where the interpixel connections are marked to highlight the PRT structure.

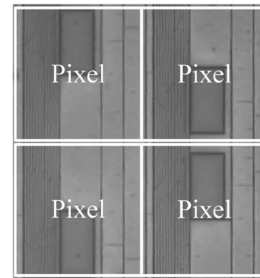
The experimental setup is shown in Fig. 14 where a laser is used as the light source. The laser beam goes through the patterns and projects the image on the chip plane. The pattern is fixed on a linear motor controlled by a PC. The PCB carrying the chip is attached to a goniometer that can rotate an object and measure the slope angle simultaneously. The goniometer rotates the chip to the desired direction on each measurement. Because the goniometer only covers rotation angles of -20 degree to 20 degree, the PCB is placed at different angles so that all the desired motion direction angles can be realized. The power supply and bias voltage are provided by programmable



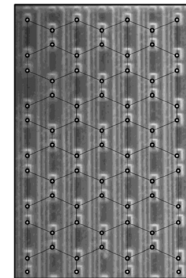
(a)



(b)



(c)



(d)

Fig. 13. Photographs of (a) the whole chip, (b) a single pixel, (c) four pixels as a subset, and (d) part of the array showing PRT structure.

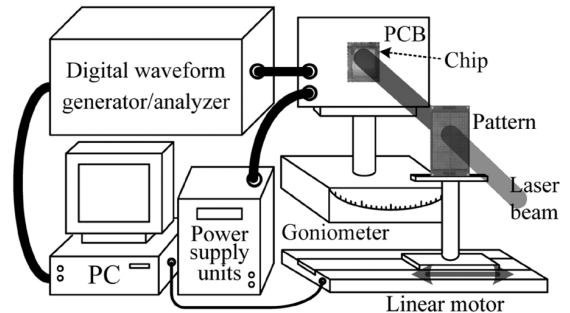


Fig. 14. Experimental setup.

power supply units. Clock and synchronization signals are generated by a PC-controlled digital waveform generator/analyzer that also receives the output signals of the chip and stores them on the PC for further direction angle computation and analysis.

The disc and twisted wire patterns have been tested. The disc pattern is produced by an opaque plastic board with small via

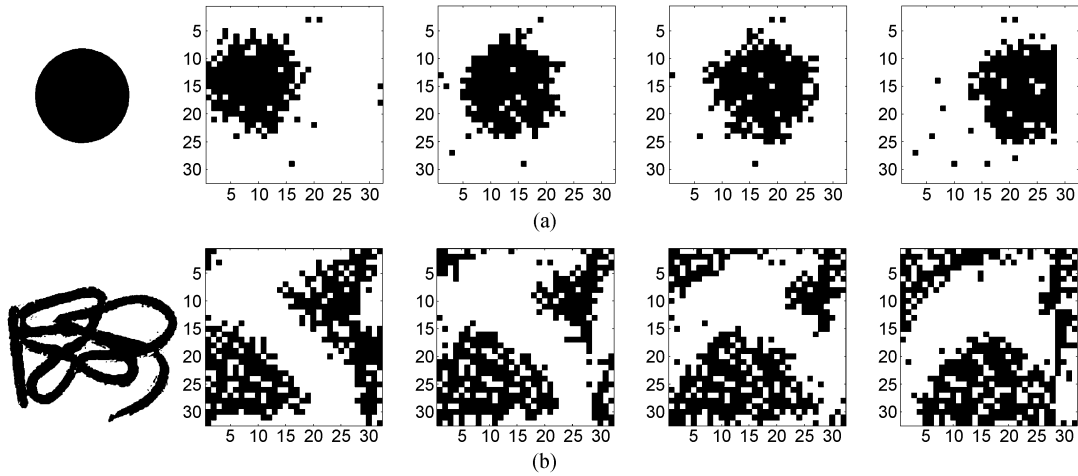


Fig. 15. Test patterns used in the experiment and four consecutive output frames of the retinal-processing circuit. (a) Disc. (b) Wire.

through holes around 0.9 mm in diameter on it. The twisted wire pattern is produced by a thin metal wire twisted arbitrarily. The wire width is about 0.25 mm. Because the two patterns are produced by nontransparent objects, they produce extremely high contrast images on the chip, which is larger than the fixed pattern noise (FPN). The captured images of four consecutive frames produced by the two patterns when they are moving are shown in Fig. 15(a) and 15(b). The directional angle is calculated from GENs and GMNs of three consecutive frames. Five calculated angles with a pattern moving in a specific direction are averaged to yield the final computed angle. The computed directional angles, their errors, and simulation results are plotted in Fig. 16(a) and 16(b). The speed of the movement is fixed to 250 mm/s, equivalent to one pixel pitch per frame time sampled at 3.6 kS/s with the clock of 1 MHz. It can be seen from Fig. 16(a) and 16(b) that for all tested motion directions, the largest errors are 9 and 11 degrees, whereas the rms errors are 5 and 8 degrees for disc and wire pattern, respectively. The experimental results are close to the simulation results. Linearity is also calculated with linear regression to evaluate the overall figure of merit. The average linearity is 98.6%.

The measurement results of the disc pattern moving at different speeds and at 0 and 45 degrees are shown in Fig. 17. Since the range for the inhibition operation is one pixel, the maximum speed is set to one pixel pitch per frame time. It can be seen from Fig. 17 that the detectable velocity range with angle error less than 10 degrees is from 1 to 0.8 of the maximum speed. In Fig. 17, the angle error increases with decreasing speed. As the speed decreases, the number of extracted GMNs per frame is smaller because the number of edges crossing the pixel boundaries is fewer. Smaller GMNs are more susceptible to fixed pattern and temporal noises. Therefore, angle measurements are degraded with decreasing speed in the results of Fig. 17. Because noises are not included in system simulations, the speed of disc pattern does not affect the calculated angle errors in simulation.

Another pattern tested is a Lena pattern printed on a transparency by a laser printer. Because the contrast of the pattern is not good enough, serious FPN is observed in the measurement when a transparency pattern is used. The image is distorted and

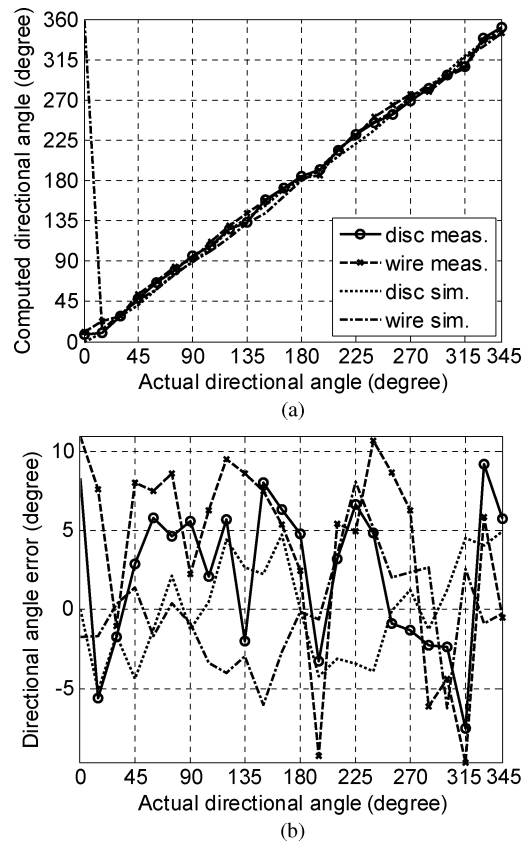


Fig. 16. Measurement and simulation results. (a) Computed directional angles and (b) the directional angle errors.

the computational accuracy is degraded by the added FPN offsets to GENs. It also masks the moving patterns so that the calculated GMNs are fewer than the values they should be. Thus, the measured error with the Lena transparency is as high as 31 degrees. The FPN and the temporal noise that may degrade the accuracy of computation can be reduced by a Schmitt trigger for binary conversion. With less noise, the accuracy of computation can be enhanced even when a lower contrast pattern is used.

The fabricated chip consumes 17 mW and 21 mW when it operates at 1 MHz and 10 MHz, respectively, with a 3.3-V power

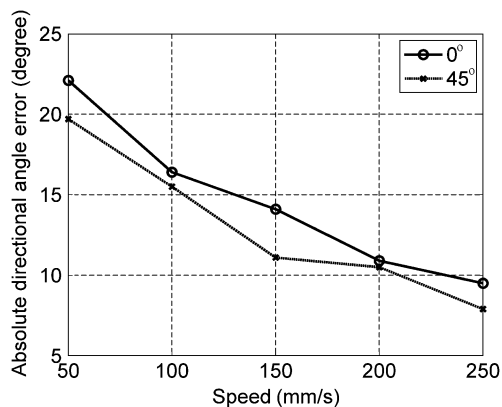


Fig. 17. Measured angle error versus speed of disc pattern moving at 0 and 45 degree.

TABLE III
THE MEASURED PERFORMANCE OF THE FABRICATED CHIP

Process	TSMC 0.35- μm 2P4M
Power Supply	3.3 V
Resolution	32 \times 32
Photosensing Area	500 μm^2 (17.4 μm \times 29.1 μm)
Pixel Pitch	63 μm
Fill Factor	12.8%
Chip Size	3.3 mm \times 4.2 mm
Maximum Frame Rate	36 kS/sec
Maximum Clock Rate	10 MHz
Power Consumption	9.9 mW @ dark & standby 16 mW @ standby 17 mW @ 1 MHz 21 mW @ 10 MHz

supply. The standby power consumptions with and without laser illumination are 16 mW and 9.9 mW, respectively. The standby power consumption of the fabricated chip is 5 times smaller than that in [14]. The measured performance of the fabricated chip is given in Table III.

V. CONCLUSION

A direction computation structure and method based on the DSGC model in the retina are proposed to calculate the motion direction of scenes. The method extracts local and global motion and edge numbers to calculate the final directional angle. Edge-number normalization in computation equations and pseudo-random tessellation structure in pixel layout are proposed to enhance the computational accuracy. Simulations have been performed to verify the proposed method. A CMOS test chip based on the proposed method is designed and measured. A maximum error of 11 degrees and a linearity of 98.6% is measured. The total power consumption is 9.9 mW at dark standby and 21 mW at the maximum clock rate of 10 MHz under 3.3-V power supply. The proposed structure is suitable for motion direction sensor applications that require accurate real-time motion direction extraction with low-power consumption and com-

compact chip size. Further research on applications such as target tracking systems, wireless optical mice, pointing devices, optical remote controllers, and digital image stabilization systems will be conducted.

ACKNOWLEDGMENT

The authors thank the Chip Implementation Center (CIC), Taiwan, for the fabrication of the testing chip.

REFERENCES

- [1] J. E. Dowling, *The Retina: An Approachable Part of the Brain*. Cambridge, MA: Harvard Univ. Press, 1987, pp. 8–11.
- [2] F. Werblin, B. Roska, D. Balya, C. Rekeczky, and T. Roska, "Implementing a retinal visual language in CNN: A neuromorphic study," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2001, vol. 2, pp. 333–336.
- [3] L.-J. Lin, C.-Y. Wu, B. Roska, F. Werblin, D. Bálya, and T. Roska, "A neuromorphic chip that imitates the ON brisk transient ganglion cell set in the retinas of rabbits," *IEEE J. Sens.*, vol. 7, pp. 1248–1261, Sep. 2007.
- [4] H. B. Barlow and W. R. Levick, "The mechanism of directionally selective units in rabbit's retina," *J. Physiol.*, vol. 178, pp. 477–504, Jun. 1965.
- [5] S. I. Fried, T. A. Münch, and F. S. Werblin, "Mechanisms and circuitry underlying directional selectivity in the retina," *Nature*, vol. 420, pp. 411–414, Nov. 2002.
- [6] 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.
- [7] J. Kramer, R. Sarpeshkar, and C. Koch, "Pulse-based analog VLSI velocity sensors," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 44, no. 2, pp. 86–101, Feb. 1997.
- [8] C. M. Higgins, R. A. Deutschmann, and C. Koch, "Pulse-based 2-D motion sensors," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 46, pp. 677–687, Jun. 1999.
- [9] S. C. Liu, "A neuromorphic aVLSI model of global motion processing in the fly," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 47, pp. 1458–1467, Dec. 2000.
- [10] C. M. Higgins, V. Pant, and R. Deutschmann, "Analog VLSI implementation of spatio-temporal frequency tuned visual motion algorithms," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, pp. 489–502, Mar. 2005.
- [11] E. Ozalevli, P. Hasler, and C. M. Higgins, "Winner-take-all-based visual motion sensors," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 53, pp. 717–721, Aug. 2006.
- [12] K. Nakayama and G. H. Silverman, "The aperture problem—I. Perception of nonrigidity and motion direction in translating sinusoidal lines," *Vision Res.*, vol. 28, pp. 739–746, 1988.
- [13] C. C. Pack and R. T. Born, "Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain," *Nature*, vol. 409, pp. 1040–1042, Feb. 2001.
- [14] C.-Y. Wu and K.-H. Huang, "A CMOS focal-plane motion sensor with BJT-based retinal smoothing network and modified correlation-based algorithm," *IEEE J. Sens.*, vol. 2, pp. 549–558, Dec. 2002.
- [15] H.-C. Jiang and C.-Y. Wu, "A 2-D velocity- and direction-selective sensor with BJT-based silicon retina and temporal zero-crossing detector," *IEEE J. Solid-State Circuits*, vol. 34, no. 2, pp. 241–247, Feb. 1999.
- [16] K.-H. H., L.-J. Lin, and C.-Y. Wu, "Analysis and design of a CMOS angular velocity- and directional-selective rotation sensor with a retinal processing circuit," *IEEE J. Sens.*, vol. 4, pp. 845–856, Dec. 2004.
- [17] C.-Y. Wu and C.-T. Chiang, "A low-photocurrent CMOS retinal local-plane sensor with pseudo-BJT smoothing network and adaptive current Schmitt trigger for scanner applications," *IEEE J. Sens.*, vol. 4, pp. 510–518, Aug. 2004.
- [18] A. Moini, *Vision Chip*. Boston, MA: Kluwer, 2000, pp. 52–53.
- [19] M. A. C. Maher, S. P. Deweerth, M. A. Mahowald, and C. A. Mead, "Implementing neural architectures using analog VLSI circuits," *IEEE Trans. Circuits Syst.*, vol. 36, no. 5, pp. 643–652, May 1989.
- [20] C.-Y. Wu and C.-F. Chiu, "A new structure of the 2-dimensional silicon retina," *IEEE J. Solid-State Circuits*, vol. 30, no. 8, pp. 890–897, Aug. 1995.
- [21] C.-Y. Wu and H.-C. Jiang, "An improved BJT-based silicon retina with tunable image smoothing capability," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 7, no. 6, pp. 241–248, Jun. 1999.

- [22] K. A. Boahen and A. Andreou, "A contrast-sensitive retina with reciprocal synapses," in *Advances in Neural Information Processing 4*, J. E. Moody, Ed. San Mateo, CA: Morgan Kaufman, 1992, vol. 4, pp. 764–772.
- [23] H. Kobayashi, J. L. White, and A. A. Abidi, "An active resistor network for Gaussian filtering of images," *IEEE J. Solid-State Circuits*, vol. 26, no. 5, pp. 738–748, May 1991.



Wen-Chia Yang (S'05) was born in Hsinchu, Taiwan, in 1979. He received the B.S. and M.S. degrees in electronics engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2002 and 2004. He is now working towards the Ph.D. degree at the Institute of Electronics, National Chiao Tung University.

His major research interests are bio-inspired vision sensor systems, retinal prosthesis, and biomedical electronics.

Li-Ju Lin, photograph and biography not available at the time of publication.



Herming Chiueh (M'90) received the B.S. degree from the Department of Electrophysics, National Chiao Tung University, Hsinchu, Taiwan, in 1992, and the M.S. and Ph.D. degrees from Department of Electrical Engineering, University of Southern California, Los Angeles, in 1994 and 2002.

From 1996 to 2002, he was with Information Sciences Institute, University of Southern California, Marina del Rey, CA. He has participated the VLSI effort on several large projects in USC/ISI and most recently participated the development of a

55-million transistor processing-in-memory (PIM) chip. He is currently an Assistant Professor, Department of Electrical Engineering, National Chiao Tung University, Hsin-Chu, Taiwan. His research interests include system-on-chip design methodology, thermal management for VLSI, low-power integrated circuits, neural interface circuits, and biomimetic systems.



Chung-Yu Wu (S'76–M'76–SM'96–F'98) was born in 1950. He received the M.S. and Ph.D. degrees in electronics engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1976 and 1980, respectively.

Since 1980, he has been a Consultant to high-tech industry and research organizations and has built up strong research collaborations with high-tech industries. From 1980 to 1983, he was an Associate Professor with National Chiao Tung University. From 1984 to 1986, he was a Visiting Associate Professor

with the Department of Electrical Engineering, Portland State University, Portland, OR. Since 1987, he has been a Professor with National Chiao Tung University. From 1991 to 1995, he served as the Director of the Division of Engineering and Applied Science, National Science Council, Taiwan. From 1996 to 1998, he was bestowed as the Centennial Honorary Chair Professor of National Chiao Tung University. From 2007 to 2011, he served as the President of National Chiao Tung University. He is currently a Chair Professor of National Chiao Tung University and the Director General of National Program on Nano Technology, Taiwan. He has authored or coauthored over 300 technical papers in international journals and conferences. He holds 30 patents, including 17 U.S. patents. His research interests are biomedical electronic devices and systems, intelligent bio-inspired vision sensor systems, nanoelectronic circuits and systems for RF/microwave communication.

Dr. Wu is a member of Eta Kappa Nu and Phi Tau Phi. He was a recipient of the 1998 IEEE Fellow Award and a 2000 Third Millennium Medal. He was also the recipient of numerous research awards presented by the Ministry of Education, National Science Council (NSC), and professional foundations in Taiwan (1999–2003).