A CMOS Focal-Plane Motion Sensor With BJT-Based Retinal Smoothing Network and Modified Correlation-Based Algorithm

Chung-Yu Wu, *Fellow, IEEE,* and Kuan-Hsun Huang, *Student Member, IEEE*

***Abstract*—This work presents and implements a CMOS real-time focal-plane motion sensor intended to detect the global motion, using the bipolar junction transistor (BJT)-based retinal smoothing network and the modified correlation-based algorithm. In the proposed design, the BJT-based retinal photoreceptor and smoothing network are adopted to acquire images and enhance the contrast of an image while the modified correlation-based algorithm is used in signal processing to determine the velocity and direction of the incident image. The deviations of the calculated velocity and direction for different image patterns are greatly reduced by averaging the correlated output over 16 frame-sampling periods. The proposed motion sensor includes a 32×32 pixel array with a pixel size of 100×100 μm^2. The fill factor is 11.6% and the total chip area is 4200×4000 μm^2. The dc power consumption is 120 mW at 5 V in the dark. Experimental results have successfully confirmed that the proposed motion sensor can work with different incident images and detect a velocity between 1 pixel/s and 140,000 pixels/s via controlling the frame-sampling period. The minimum detectable displacement in a frame-sampling period is 5 μm. Consequently, the proposed high-performance new motion sensor can be applied to many real-time motion detection systems.**

***Index Terms*—Direction sensor, focal-plane motion sensor, motion sensor, retinal processing circuit, velocity sensor, vision chip.**

I. Introduction

REAL-TIME motion estimation or detection is an important research topic in the field of image processing, as used in many applications. Much effort has been devoted to realizing detection functions through hardware or software implementation. Typically, hardware implementation is more feasible than software implementation in achieving real-time processing under a critical speed requirement. Parallel and pixel-level computation is required to detect motion in real time [1]. Besides, appropriate image acquisition and preprocessing is needed to accurately locate the edges of a moving object in the presence of noise and under varying illumination, to ensure the robustness of the motion detector.

Many hardware-implemented motion detectors have recently been elucidated [2]–[15]. Their adopted algorithms for motion detection can be separated into three categories—gradient-based [3], energy-based [4], [5], and correlation-based algorithms [6]–[15]. The correlation-based algorithm, which is inspired by a biological model [16], has the benefits of robustness and compactness, and it is thus the most practical algorithm for very large scale integration (VLSI) implementation. However, in the previous designs [7], [11], [14], the delay element is realized by RC circuits. The resistor is typically implemented by a MOSFET resistor or a transconductance amplifier to control the time constant. In such designs, the delay time cannot be accurately controlled. Moreover, the tunable range of the delay time is also somewhat limited. Furthermore, in conventional correlation-based computation of motion, either the temporal or the spatial edge detector is adopted for computing motion [6], [8], [10], [11], [14], [15]. However, both spatial and temporal edge detectors may occupy a large area of silicon, so that the pixel area is large.

The detected motion is usually related with the incident image pattern. Thus, detecting motion involves significant deviations among various incident image patterns. The deviations make the sensor's handling of arbitrary incident images very difficult. One attempt to solve this problem is described in [15]. However, a specially designed pattern is required.

The required function of image acquisition and preprocessing in a motion sensor can be performed by using the retinal processing circuit. The retinal processing circuit mimics parts of functions of the cells in the outer plexiform layer of the real retina. As in the real retina, the retinal processing circuit has similar advantages of high dynamic range, edge enhancement, and noise immunity. Of the proposed structures for retinal function [12]–[14], [17]–[19], the BJT-based retinal structure [14], [18], [19] is very compact and suitable for VLSI implementation.

In this work, the adopted motion computation method is based on the correlation-based algorithm, with some modifications. It should be noted that the proposed motion sensor is intended to calculate the global velocity and direction and the terms "velocity" and "direction" hereafter mean the global velocity and the global direction. The motion computation method is firstly to sample two image frames at a frame sampling period of 100 clock cycles. Subsequently, the previously sampled image frame is shifted to four positions along the $\pm x$ and $\pm y$ directions. Then, correlation between the shifted previous frame and the current frame is computed. The correlation results are averaged over 16 frame-sampling periods. The averaged correlation results are used to calculate displacement and direction. Notably, the motion computation method is digital, and thus robust.

Manuscript received February 15, 2002; revised October 16, 2002. This work was supported by the National Science Council, R.O.C., under Contract NSC91-2215-E-009-071. The associate editor coordinating the review of this paper and approving it for publication was Dr. Ralph Etienne-Cummings.

The authors are with National Chiao Tung University, Hsinchu 300, Taiwan, R.O.C.

Digital Object Identifier 10.1109/JSEN.2002.807774

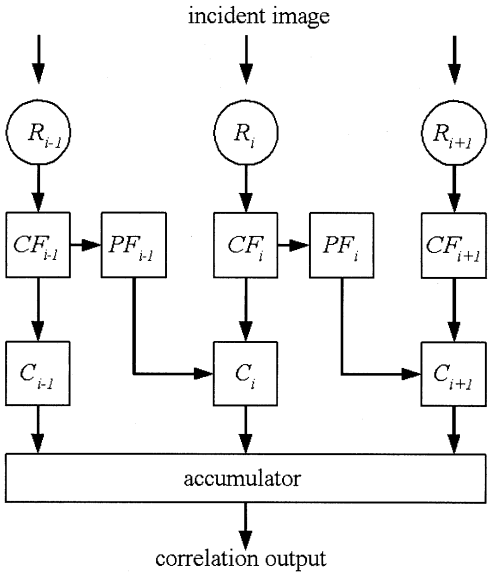


Fig. 1. Conceptual structure of the motion computation method. R is the retinal processing circuit; *CF* and *PF* are registers used to store the current and previous sampled outputs of R, and C is the correlator. All the outputs of C are summed by the accumulator to obtain the correlation output.

The delay time of the proposed motion sensor is governed by the sampling period. Consequently, the delay time is accurately controlled via adjusting the clock rate. The tunable range of the clock rate is thus high. The proposed design greatly reduces the pattern-related deviations in calculating the displacement and direction by averaging the correlation results. Averaging over longer time generally corresponds to more accurately calculated velocity and direction.

In the proposed motion sensor, the BJT-based retinal smoothing network is used to obtain the smooth signal [18], which is subtracted from the photoreceptor signal. The difference after subtraction is amplified by the current-input Schmitt trigger [20], which enhances noise immunity and eliminates ambiguity [14]. The outputs are further amplified to digital levels by two inverters. The image contrast is thereby enhanced. This structure is very suitable for motion detection applications.

The experimental chip of the proposed motion sensor is designed and fabricated by the 0.5-μm double-ploy-triple-metal (2P3M) CMOS process. Experimental results confirmed the correct functioning of this chip. This two-dimensional (2-D) motion sensor is suitable for the applications of image stabilization, ego motion detection, and optical mouse.

Section II discusses the motion computation method. Section III describes the architecture and circuit. Section IV gives experimental results. A concluding section ends the paper.

II. Motion Computation Method

The modified correlation-based algorithm is adopted in the proposed motion sensor. Fig. 1 depicts the conceptual structure of the adopted motion computation method. The incident image is acquired and processed by the retinal processing circuit R. The retinal processing circuit includes the retinal smoothing network, photoreceptor, current-input Schmitt trigger, and two inverters. The incident image is processed by R to yield a contrast-enhanced bilevel black-and-white image. The outputs

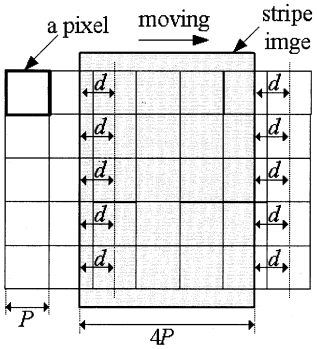


Fig. 2. Pattern of a stripe with a width of $4P$, moving to the right. d is the distance from an edge to the center of the nearest pixel in the direction of motion. P is the distance between two adjacent photo sensors in two adjacent pixels, or the length of each pixel.

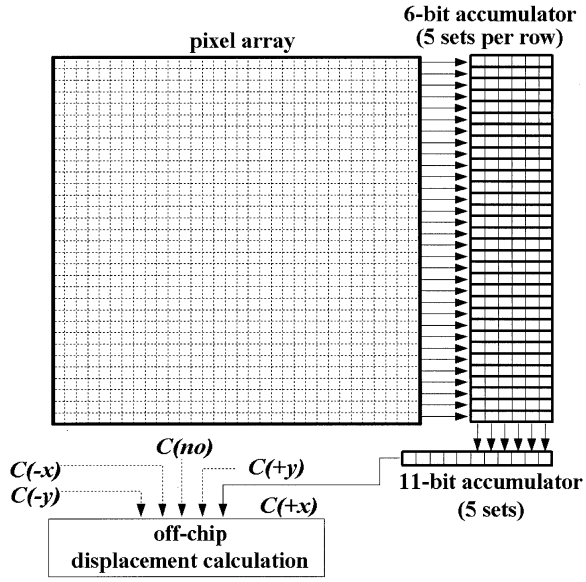


Fig. 3. Architecture of the proposed focal-plane motion sensor.

of R are binary signals which are sampled and stored in two registers *CF* and *PF*, which store the current and previous sampled outputs of R, respectively. The output of *PF* is then shifted to the nearest neighbors, along the preferred direction, to correlate with the output of *CF*. All the outputs of the correlator C along the preferred direction are accumulated by the accumulator throughout the array to determine the correlation output. The correlation output is used to calculate the displacement in a sampling period. The term "displacement" hereafter refers to the displacement in a sampling period, except where specified. In the proposed motion sensor, the output of *PF* is shifted to the four nearest neighbors in the $+x$, $-x$, $+y$, and $-y$ directions so that the displacement along these four directions can be determined. The corresponding correlation outputs are $C(+x)$, $C(-x)$, $C(+y)$, and $C(-y)$, respectively. The correlation output between the output of *PF* without shift and the output of *CF* as defined by $C(no)$ is also determined and is used to calculate the displacement.

++The five kinds of correlation outputs are averaged over 16 sampling periods and are defined as $C_a(+x)$, $C_a(-x)$, $C_a(+y)$, $C_a(-y)$, and $C_a(no)$, corresponding to $C(+x)$,

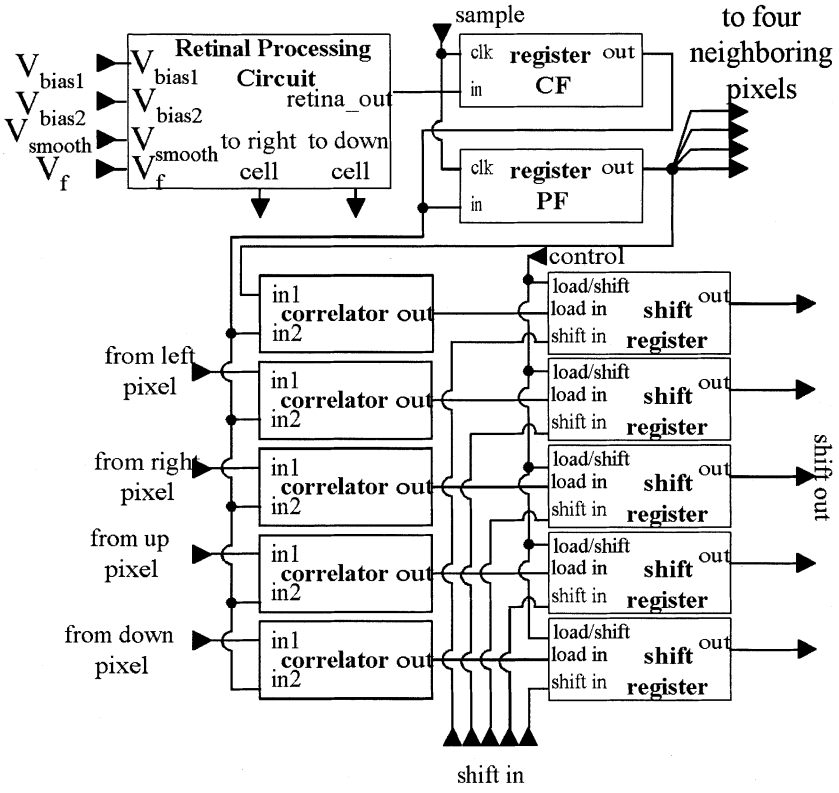


Fig. 4. Structure of a single pixel.

$C(-x)$, $C(+y)$, $C(-y)$, and $C(no)$, respectively. The calculated displacement Δx in the $+x$ or $-x$ direction normalized to the distance between two adjacent photosensors P can be expressed by the averaged correlation outputs as

$$\frac{\Delta x}{P} = \frac{A - B}{2*[C_a(no) - B]}$$
$$A \equiv \max\{C_a(+x), C_a(-x)\}$$
$$B \equiv \min\{C_a(+x), C_a(-x)\}. \quad (1)$$

The introduction of A and B is used to obtain the absolute value of the difference between $C_a(+x)$ and $C_a(-x)$ and to decide whether the image is moving along $+x$ or $-x$. Similarly, the displacement, Δy, in the $+y$ or $-y$ direction can be expressed as

$$\frac{\Delta y}{P} = \frac{G - H}{2*[C_a(no) - H]}$$
$$G \equiv \max\{C_a(+y), C_a(-y)\}$$
$$H \equiv \min\{C_a(+y), C_a(-y)\}. \quad (2)$$

The physical interpretation is discussed later. The direction of motion can be determined by comparing the amplitude of $C_a(+x)$ with that of $C_a(-x)$ as well as $C_a(+y)$ with that of $C_a(-y)$. For example, if $C_a(+x)$ exceeds $C_a(-x)$ and $C_a(+y)$ exceeds $C_a(-y)$, then the image is moving in the $(+x, +y)$ direction. The directional angle θ is given by

$$\tan\theta = \frac{\Delta y}{\Delta x} = \frac{G - H}{A - B}. \quad (3)$$

If the displacement and direction of motion is calculated without using the averaged correlation outputs, the pattern-related deviations are significant. A variable d is defined as the distance from the edge of the incident image patterns to the center of the front pixel, in the direction of motion, as shown in Fig. 2. The value of d is between 0 and P. Assume that the edges must move over the center of a pixel to be detected by that pixel. The influence of the pattern on the deviation of the calculated displacement is illustrated by the example presented in Fig. 2, in which the image pattern of a vertical stripe with a width of $4P$ is used. If the displacement within a sampling period equals P, then the detected image will move to the right by exactly one pixel. If the displacement of motion is calculated without using the averaged correlation outputs, the calculated displacement will become P, according to (1) with $C_a(+x)$, $C_a(-x)$, and $C_a(no)$ replaced by $C(+x)$, $C(-x)$, and $C(no)$, respectively. If the displacement within a sampling period is less than P, d will gradually decline over successive sampling periods since the edge is approaching the center of the front pixel in the direction of motion. The calculated displacement is 0 in these sampling periods. Finally, the edge will cross the center of the front pixel in the direction of motion and the calculated displacement will become P in that sampling period. In the example of Fig. 2, it is found that the calculated displacement is either 0 or P. Neither value, however, is the actual displacement. As can be found in the example of Fig. 2, the calculated displacement depends on the number pixels crossed by the edge within a sampling period. The number of pixels crossed by the edge within a sampling period is determined by the distribution of the value of d for all

edges, which is a property related to the pattern of the incident image. The phenomenon discussed above indicates that the deviation in the displacement and direction calculation is large and dependent on input patterns. This problem can be solved by using the averaged correlation outputs in the proposed design.

To illustrate why it is advantageous to average the correlation outputs, an example is analyzed below. Assume that the displacement is $0.2P$ and d is P, so the calculated displacements, without using the average correlation outputs, in the five subsequent sampling periods are $\{0, 0, 0, 0, P\}$ according to (1), with $C_a(+x)$, $C_a(-x)$, and $C_a(no)$ replaced by $C(+x)$, $C(-x)$ and $C(no)$, respectively, in each sampling period. An average of correlation outputs is taken to calculate the displacement so that the deviation between calculated and actual displacement can be eliminated. Using (1) with an average over five sampling periods yields a calculated displacement of 0.2, which is exactly the displacement normalized to P. A more general analysis on how accuracy is improved by averaging the correlation results is discussed below.

Assume that a rectangular pattern moves in the $+x$ direction. $C(+x) - C(-x)$ equals the number of pixels crossed by the edges as they move to the right. $C(no) - C(-x)$ equals half the number of pixels at the vertical edges. For example, in Fig. 2, when the edges do not cross any pixels, then $C(+x) - C(-x)$ is zero. When the edges cross the pixels, $C(+x) - C(-x)$ is ten, which is exactly the number of pixels crossed by the edges. $C(no) - C(-x)$ is fixed at five in every sampling period, which is half the number of the pixels at the vertical edges. The ratio of $C(+x) - C(-x)$ to twice $C(no) - C(-x)$ is equal to the number of pixels crossed by edges divided by the total number of pixels at the vertical edges, which is the displacement normalized to P in the $+x$ direction [15].

Assume that $C_a(+x)$, $C_a(-x)$, $C_a(+y)$, $C_a(-y)$, and $C_a(no)$ are the average correlation outputs over N sampling periods and E_t is defined as the total number of the pixels at the vertical edges, which is equal to $2 * [C(no) - C(-x)]$. If the boundary condition is ignored, then (1) can be rewritten as

$$\begin{aligned}\frac{\Delta x}{P} &= \frac{C_a(+x) - C_a(-x)}{2*[C_a(no) - C_a(-x)]} \\ &= \frac{\frac{1}{N}\sum_{i=1}^{N} C(+x)_i - \frac{1}{N}\sum_{i=1}^{N} C(-x)_i}{2*\left[\frac{1}{N}\sum_{i=1}^{N} C(no)_i - \frac{1}{N}\sum_{i=1}^{N} C(-x)_i\right]} \\ &= \frac{\sum_{i=1}^{N} [C(+x)_i - C(-x)_i]}{N*E_t} \end{aligned} \tag{4}$$

where $C(+x)_i$, $C(-x)_i$, and $C(no)_i$ are the correlation outputs at the ith sampling period.

If the velocity is P/m, the number of pixels that are crossed by the edges in every m sampling periods will be equal to E_t. This is because the vertical edges move to the right by exactly one pixel in every m sampling periods so that the number of pixels at the vertical edges will equal the total number of pixels that are crossed by the edges in every m sampling periods. Thus,

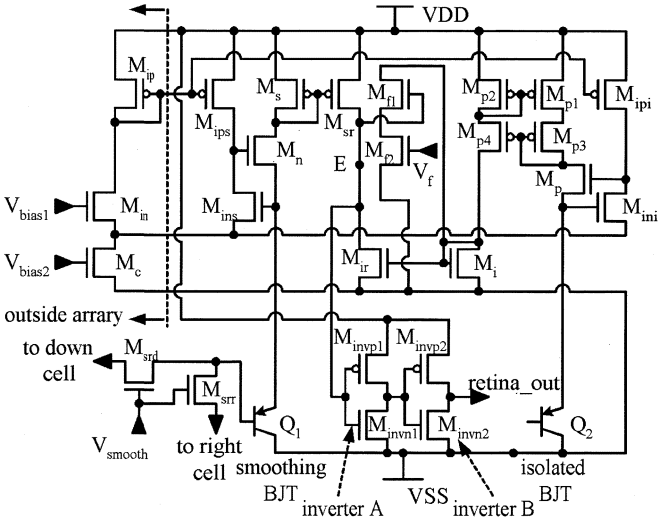


Fig. 5. Structure of the BJT-based retinal processing circuit.

the number of pixels that are crossed by the edges within N sampling periods can be expressed by

$$\begin{aligned}\sum_{i=1}^{N} [C(+x)_i - C(-x)_i] &= \sum_{i=1}^{qm} [C(+x)_i - C(-x)_i] \\ &\quad + \sum_{i=qm+1}^{N} [C(+x)_i - C(-x)_i] \\ &= q*E_t + \sum_{i=qm+1}^{N} [C(+x)_i - C(-x)_i] \\ N &= qm + r \end{aligned} \tag{5}$$

where q is the quotient of N/m and r is the residue of N/m. Substituting (5) into (4) yields

$$\begin{aligned}\frac{\Delta x}{P} &= \frac{q*E_t + \sum_{i=qm+1}^{N} [C(+x)_i - C(-x)_i]}{N*E_t} \\ &= \frac{1}{m} - \frac{r}{m*N} + \frac{\sum_{i=qm+1}^{N} [C(+x)_i - C(-x)_i]}{N*E_t} \end{aligned} \tag{6}$$

$1/m$ is the displacement normalized to P, and the other terms in (6) are deviations. From the above derivation, it is shown that the accuracy is improved by using the averaged the correlation results and a larger N corresponds to a more accurately calculated displacement. The above analysis is more complex for arbitrary patterns. However, $C(+x) - C(-x)$ remains a function of the number of pixels crossed by the edges and $C(no) - C(-x)$ remains a function of half the number of pixels at the vertical edges. Hence, the conclusion that the deviation can be reduced by averaging the correlation outputs still holds.

Notably, determining the correlation is complicated by the boundary condition because of the shifting of the previous image frame. This boundary is fixed at zero for simplicity. Therefore, the correlation at the boundary yields errors in calculating $C(+x)$, $C(-x)$, $C(+y)$, and $C(-y)$. However, the

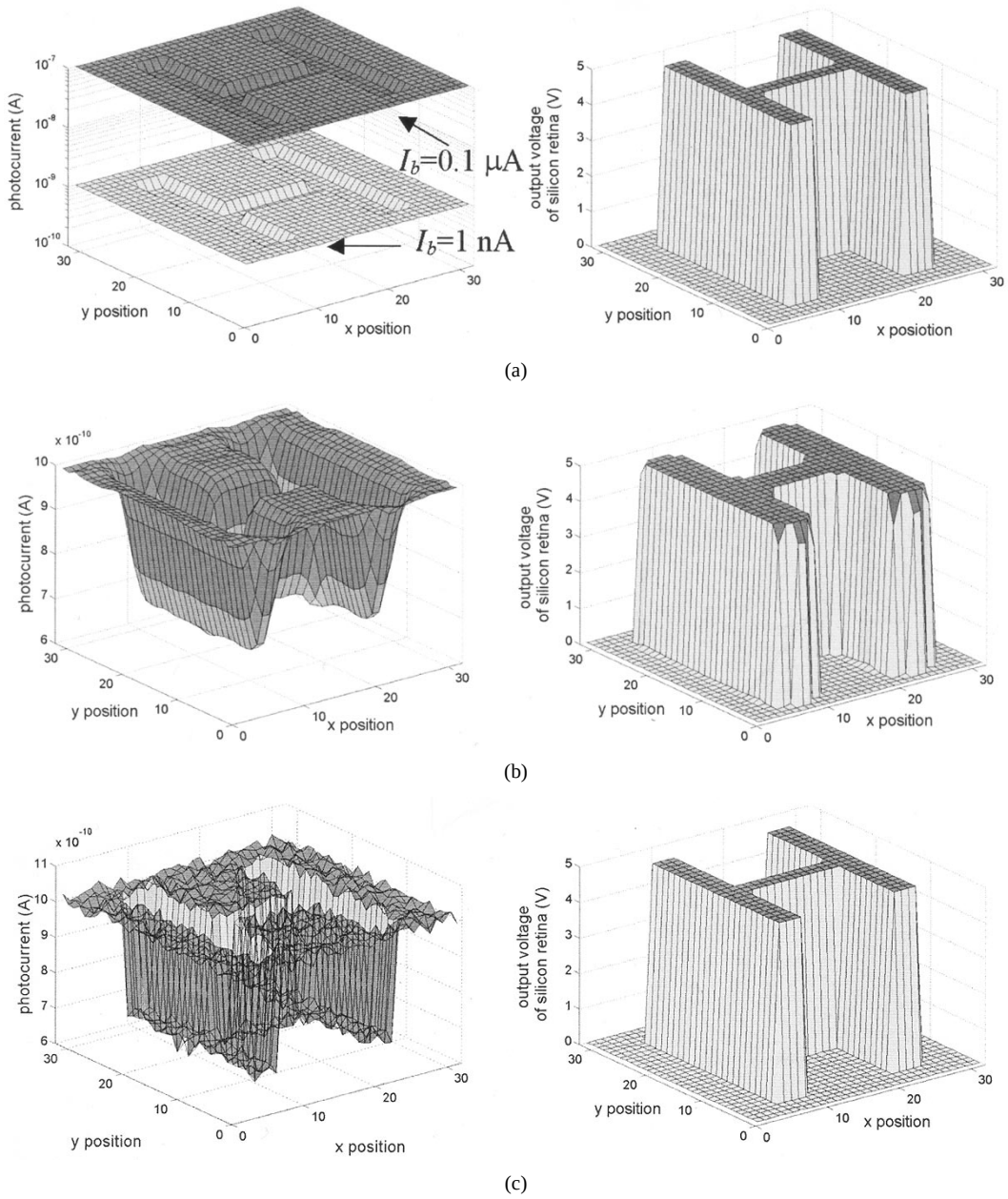


Fig. 6. HSPICE simulation results, illustrating the output of the retinal processing circuit (a) with the background photocurrent's varying from 1 nA to 0.1 μA, (b) with the image blurred, and (c) with normally distributed noise, with a standard deviation of 0.04 nA.

boundary condition does not influence the calculation of $C(no)$ because the previous image frame is not shifted when $C(no)$ is calculated. Two boundary factors are added in (1) and (2) to calibrate the effect due to boundary conditions and thus solve this problem in calculating displacement. Equations (1) and (2) are modified to

$$\frac{\Delta x}{P} = \frac{A - B + b_1}{2*[C_a(no) - B - b_2]} \tag{7}$$

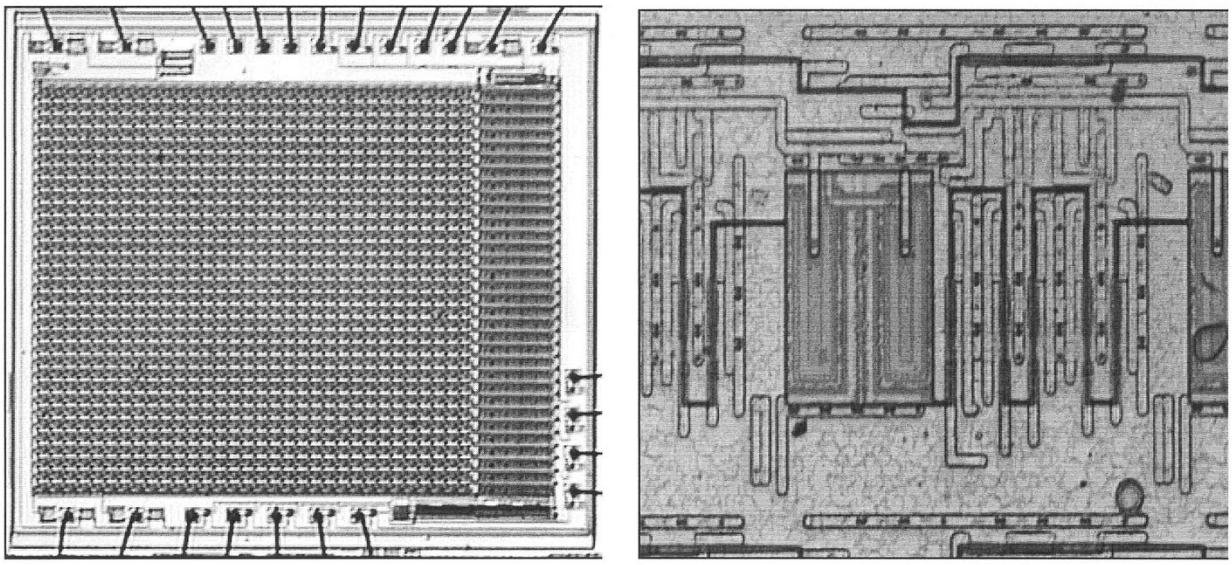
$$\frac{\Delta y}{P} = \frac{G - H + b_1}{2*[C_a(no) - H - b_2]} \tag{8}$$

where b_1 and b_2 are the calibrating factors. The value of b_1 is chosen to be half the number of pixels at one side of the boundary while b_2 is half of b_1. Since the size of the pixel array is 32×32, b_1 and b_2 are chosen to be 16 and 8, respectively.

The maximum detectable displacement is P because the output of *PF* is shifted only to the nearest neighbors. The minimum detectable displacement depends on the accuracy of the calculation. The sampling period can be adjusted to fit the desired range of velocity, according to the detectable range of displacement. In practice, the sampling period may be fixed or automatically adjusted according to application.

III. Architecture and Circuit

Fig. 3 illustrates the architecture of the proposed focal-plane motion sensor, which includes the 32×32 pixel array and the peripheral circuits, including five sets of 6-b accumulators for each row, and five sets of 11-b accumulators. The data in the five 11-b accumulators are read out as $C(+x)$, $C(-x)$, $C(+y)$,



(a) (b)

Fig. 7. Photographs of the proposed motion sensor for (a) the whole fabricated chip and (b) a single pixel.

$C(-y)$, and $C(no)$, respectively. The average, displacement, and direction are calculated off-chip by software.

Each pixel of the 32×32 array includes the BJT-based retinal processing circuit, two registers, five correlators, and five shift registers, as depicted in Fig. 4.

In the implementation of the BJT-based retinal processing circuit, as shown in Fig. 5, an isolated photo-BJT is used as the photoreceptor, a smoothing photo-BJT with adjustable N-channel MOSFET resistors is used to form the retinal smoothing network, and a current-input CMOS Schmitt trigger [14], [20] and two inverters are included. The base of the smoothing photo-BJT Q_1 is connected to the photo-BJT Q_1's four nearest neighbors, via the N-channel MOSFET resistors M_{srd} and M_{srr}, whose resistance is controlled by the gate voltage $V_{\rm smooth}$, forming the smoothing network. The transistors $M_{\rm ins}$, $M_{\rm ips}$, M_n and $M_{\rm ipi}$, $M_{\rm ini}$, M_p are used to virtually bias the emitters of Q_1 and Q_2 at $V_{\rm bias1}$. $M_{\rm ip}$, $M_{\rm in}$, and M_c are common to all pixels. The current-input CMOS Schmitt trigger comprises M_s, M_{sr}, M_i, M_{ir}, M_{f1}, and M_{f2} transistors. The voltage V_f is used to adjust the threshold level. The transistors M_{p1}, M_{p2}, M_{p3}, and M_{p4} are used to mirror the emitter current of Q_2 to the current-input CMOS Schmitt trigger. The inverters A and B amplify the output of the current-input CMOS Schmitt trigger to VDD or VSS so that the signal is converted from analog to binary.

The output of the retinal processing circuit is then sent to *CF*, which memorizes the current frame, as depicted in Fig. 4. The *PF* in Fig. 4 stores the previous frame and sends its output to the four nearest neighboring pixels in the $+x$, $-x$, $-y$, and $+y$ directions. The sample signal in Fig. 4 causes *CF* and *PF* to sample their inputs every 100 clock cycles. The output of *CF* is then correlated with the output of *PF* and the outputs of *PF* from its neighbors. The outputs of the correlators are stored in five shift registers and then shifted to the 6-b accumulators in each row. The motion computation method precisely controls the sampling period by controlling the clock rate and the velocity can be calculated as the displacement divided by the frame-sampling period. Moreover, the tunable range of the clock rate exceeds that of the RC delay circuits, so the detectable range of the velocity is greatly increased.

NAND rather than XOR is used in the correlator because the former is less likely to have errors. If a noise exists in the signal path, then the inputs of the correlator have a probability g of error. The probability of error in the output of the correlator is $2 * (g - g^2)$ for XOR and $g - g^2$ for NAND. The probability of error for XOR is clearly double that for NAND.

The outputs of the correlator stored in five shift registers are shifted to the 6-b accumulators in each row to be accumulated. This chip includes five sets of 11-b accumulators. The outputs of the 6-b accumulators are shifted to the 11-b accumulators, which accumulate the outputs of the 6-b accumulators. The data in the 11-b accumulators are sent out of chip serially as the correlation outputs $C(+x)$, $C(-x)$, $C(+y)$, $C(-y)$, and $C(no)$. It takes 100 clock cycles in total to sample the image, shift the outputs of the correlators to the 6-b accumulators, shift the data in the 6-b accumulators to 11-b accumulators, send the data out of chip, and wait for the next sampling period. The software calculates the average of the correlation outputs, displacement, and direction.

Fig. 6 presents the results of the HSPICE simulation. As shown in Fig. 6, a pattern "H" is incident on the motion chip under various conditions to confirm the advantages of high dynamic range, contrast enhancement, and noise immunity. In Fig. 6(a), the photocurrents of the dark and bright regions are $0.66I_b$ and I_b, respectively, where I_b is defined as the photocurrent generated by the light source. Assuming that the illumination of the light source varies, the simulation result of HSPICE indicates that the outputs of retinal processing circuit are constant as I_b varies from 0.1 μA to 1 nA. As can be calculated from Fig. 6(a), the minimum acceptable contrast of input image is about 20%. Fig. 6(b) offers another example of contrast enhancement. The pattern is the same

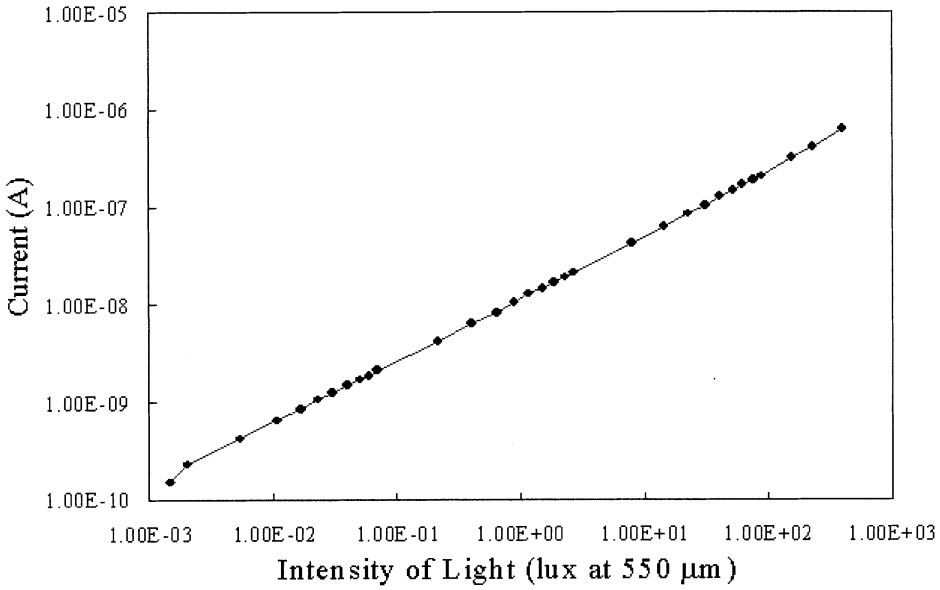


Fig. 8. Photo response of the photo-BJT.

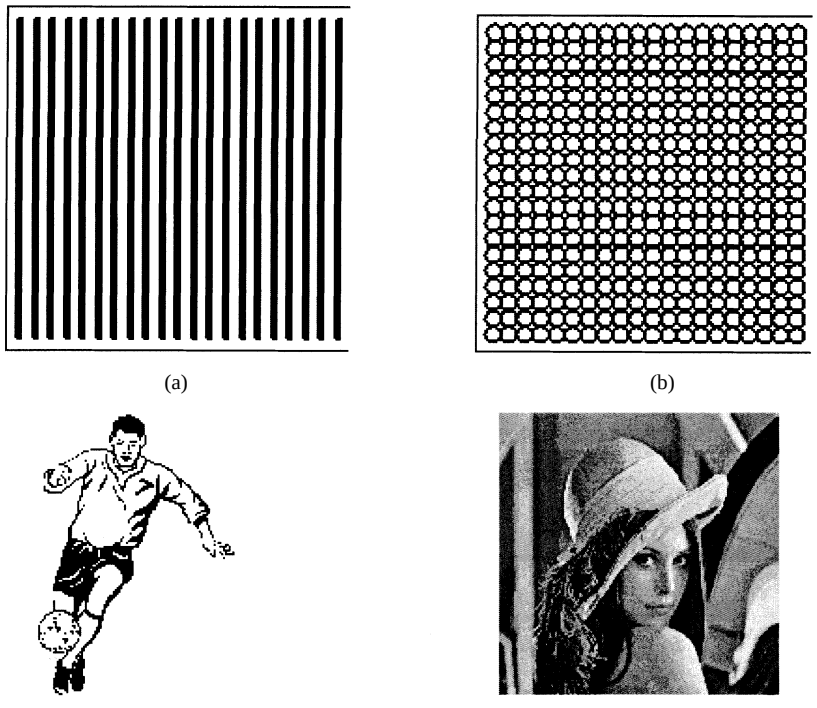
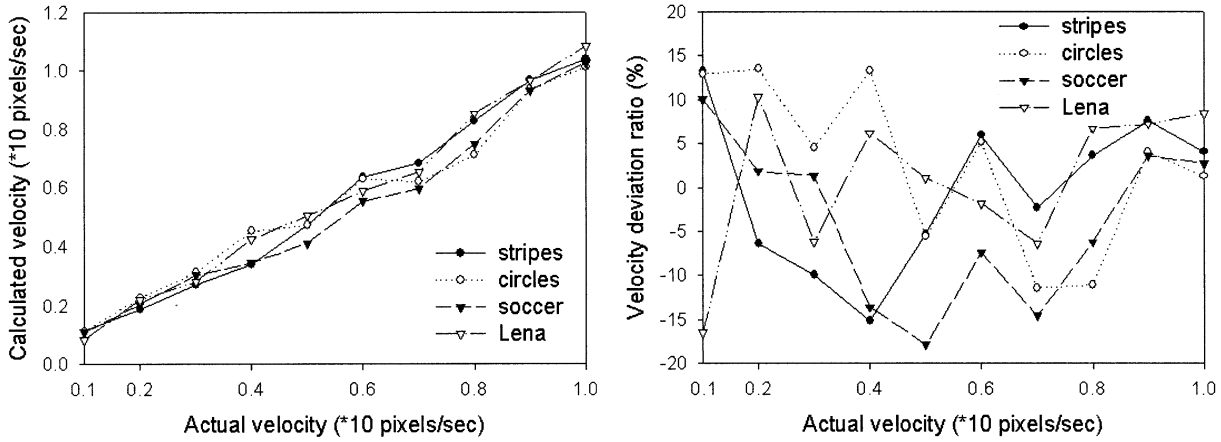


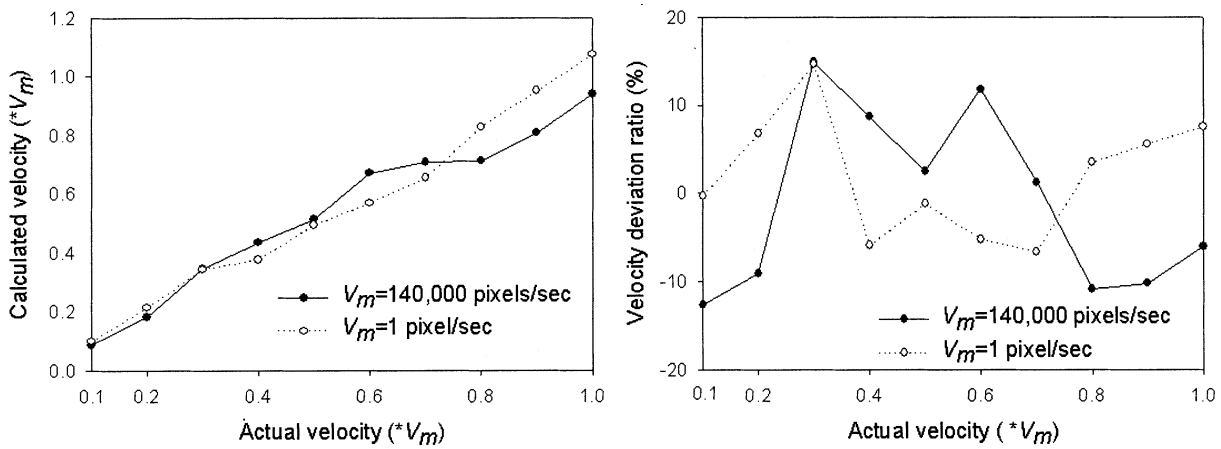
Fig. 9. Four test patterns: (a) stripes, (b) circles, (c) soccer, and (d) Lena.

but the image is blurred. The simulation result shows that the contrast can be enhanced. The location of the edge is governed by the resistance of the MOSFETs in the smoothing network. Fig. 6(c) depicts a situation in which the pattern is disturbed by noise. The noise immunity is enhanced by the current-input CMOS Schmitt trigger, which is controlled by V_f. A normally distributed noise is added to the original image with a standard deviation of 0.04 nA and V_f is set to 0.4 V. The simulation result indicates that the outputs of the retinal processing circuit are not affected by this noise level. The retinal processing circuit can work with the standard deviation of noise greater than 0.04 nA, if V_f is greater than 0.4 V. The simulation results confirm the advantages of the retinal processing circuit. First, the output of the retinal processing circuit can adapt to the background illumination, so that a high dynamic range can be obtained. Second, the error in the computation of motion can be reduced since the contrast of image is enhanced with the minimum acceptable contrast of 20%. Finally, the retinal processing circuit is immune to noise and the noise immunity is enhanced by the current-input CMOS Schmitt trigger.



(a) (b)

Fig. 10. Calculated velocity for the four kinds of patterns: (a) calculated velocity and (b) velocity deviation ratio. The velocity of the moving image varies from 1 to 10 pixels/s in the $-x$ direction; the sampling rate is 10 Hz.



(a) (b)

Fig. 11. Calculated velocity at different sampling rates: (a) calculated velocity and (b) velocity deviation ratio. The sampling rate is 140 kHz for $V_m = 140{,}000$ pixels/s and 1 Hz for $V_m = 1$ pixel/sec. Motion is in the $-y$ direction.

IV. Experimental Results

An experimental focal-plane motion sensor chip is designed and fabricated by the 0.5-μm single-poly-triple-metal CMOS process. Fig. 7(a) presents a photograph of the fabricated chip and Fig. 7(b) displays the photograph of a single pixel. The area of this chip is 4200×4200 μm^2 and that of a single pixel is 100×100 μm^2 with a fill factor of 11.6%. The dc power dissipation is 120 mW at 5 V in the dark.

The emitter and the collector were biased at 3.5 and 0 V, respectively, and the base was floating to investigate the photo response of a photo-BJT. The intensity of the incident light on the photo-BJT was then varied and the collector current, corresponding to various illuminations, was recorded. The generated photo current is expressed as $I_{\rm ph} = I_c/(\beta + 1)$, where β is the current gain of BJT and I_c is the collector current. The current gain is measured as 48 and the photocurrent is evaluated as in Fig. 8. The illumination for all the following experiments is set at 36 lux at wavelength of 550 μm.

Fig. 9 shows four patterns used to test the motion sensor to show that the proposed design can work on the different patterns. The four patterns are chosen for their different characteristics. Fig. 9(a) shows a striped pattern. The stripes are equally spaced so that d is a constant, as in the example in Fig. 2. Fig. 9(b) depicts the pattern formed by the circles. This pattern has a more uniform distribution of d than does the stripes because the edge of the circles is round. Both the stripes and the circles are spatially periodic. The third pattern presented in Fig. 9(c) is "soccer." The third pattern involves a rather random distribution of d and is spatially aperiodic, unlike the stripes and the circles. The fourth pattern is Lena and is used to test the motion sensor to determine whether the displacement can be calculated for a grayscale image with a randomly distributed d. The contrast is 80% for all test patterns except for Lena. Only a part of the image pattern is incident on the chip since the size of the four patterns exceeds that of the motion sensor chip.

Fig. 10 shows the calculated velocity and the deviation ratio. The displacement is calculated from (7) and (8) and the velocity is the ratio of the displacement to the frame sampling period. The deviation ratio is defined as

$$\text{deviation ratio } (\%) = \frac{(V_e - V_a)}{V_a} * 100 \qquad (9)$$

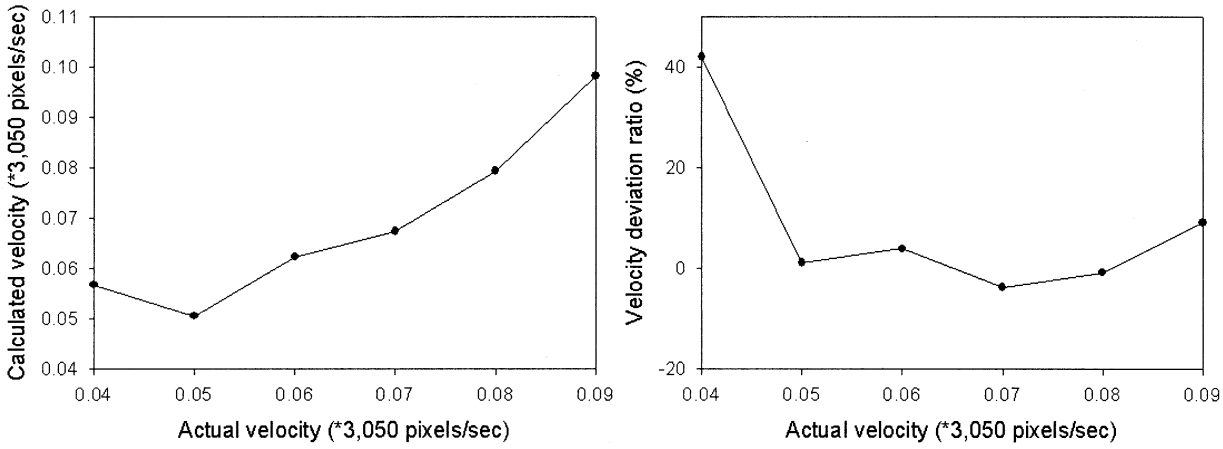


Fig. 12. Minimum detectable velocity: (a) calculated velocity and (b) velocity deviation ratio. The velocity of the moving image varies from 274.5 to 122 pixels/s in the $-x$ direction; the sampling rate is 3.05 kHz.

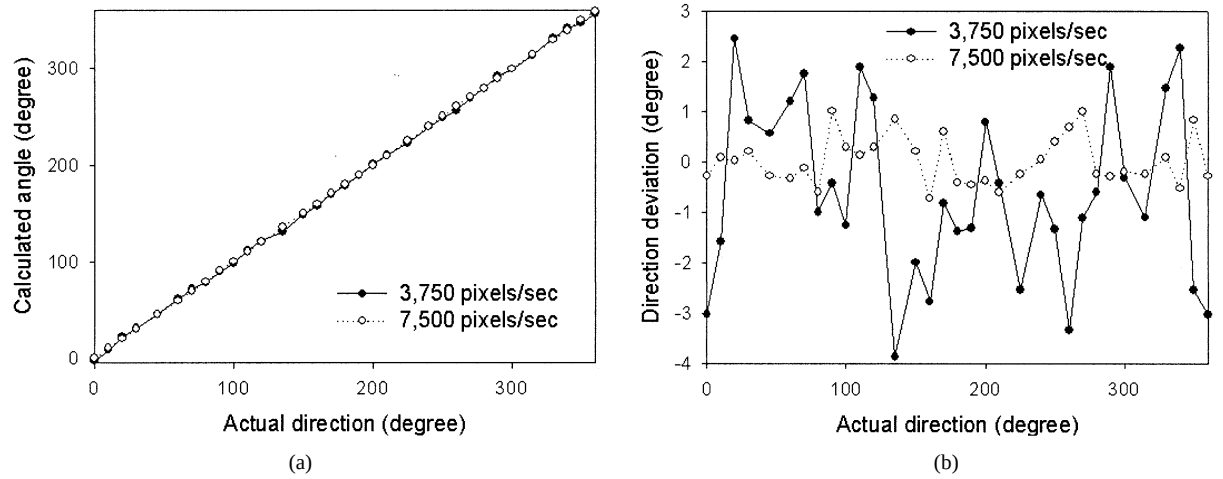


Fig. 13. Calculated direction as the image moves at two different velocities: (a) calculated direction and (b) direction deviation. The velocities of the moving image are 3750 and 7500 pixels/s, respectively. The sampling rate is 75 kHz for both curves.

where V_e is the calculated velocity and V_a is the actual velocity. The sampling rate is 10 Hz. The correlation results are averaged over 16 frame-sampling periods in all experiments. The image moves in the $-x$ direction with the velocity varying from 1 to 10 pixels/s; only the calculated result for the x direction is shown, since the calculated displacement in the y direction is zero. According to Fig. 10, the four curves rather closely resemble each other and the deviation is within $\pm 20\%$. Experimental results confirm that the proposed design can work for different patterns. Only the striped pattern is used in the following experiment because the choice of pattern will not affect accuracy.

Fig. 11 shows the detectable range of the velocity. The velocity of the moving image decreases one decade from 140 000 pixels/s and 1 pixel/s along the $-y$ direction, and the sampling rate is 140 kHz and 1 Hz, respectively. As shown in Fig. 11, the proposed design could work with a sampling rate that varied by six decades and a velocity deviation within $\pm 20\%$.

The minimum detectable velocity depends on the number of frame-sampling periods over which the correlation results are averaged. The correlation outputs are averaged over 16 sampling periods in the proposed motion sensor. In commercial optical mouse systems, the maximum detectable velocity is around 12 in/s (304.8 mm/s). If the effect of the lens is ignored, then the on-chip velocity will be 304.8 mm/s. According to this velocity, the sampling rate is set to 3050 Hz. As shown in Fig. 12, the proposed design can calculate a velocity of 152.5 pixels/s, that is, a displacement of 5 μm within a sampling period.

Fig. 13 illustrates the accuracy of the calculation of direction. The sampling rate is 75 kHz. Two velocities are tested. The selected test velocity is fixed and the direction of motion is varied from 0 to 360 degrees. Since the striped pattern is used, the stripes are rotated to be perpendicular to the direction of motion so that the aperture problem is avoided. Experimental results indicate that the proposed design can determine the direction of motion with an angular deviation less than 4 degrees while the velocity varies.

V. Conclusion

This work presents a real-time CMOS focal-plane motion sensor that uses the BJT-based retinal processing circuit and a modified correlation-based algorithm. The correlation-based algorithm is modified and applied to calculate the velocity and direction and the BJT-based retinal processing circuit is

used to acquire images and enhance contrast. The presented motion sensor greatly reduces the deviation of the calculated displacement and direction for different image pattern by averaging correlation results over 16 frame-sampling periods. Consequently, the proposed design can work for different incident images. Furthermore, the velocity ranging from 1 to 140 000 pixels/s can be calculated by the proposed design via adjusting the clock signal without reducing accuracy. The smallest displacement that can be calculated within a sampling period is 5 μm. The direction can be calculated correctly between 0 and 360 degrees with an angular deviation less than 4 degrees. The proposed motion sensor is comprised of a 32×32 pixel array and peripheral circuits. The area of a pixel is 100×100 μm^2 with a fill factor of 11.6%, and the total chip area is 4200×4000 μm^2. The supply voltage is 5 V and the dc power consumption is 120 mW. The average, velocity, and direction are calculated off-chip by software. In the future, the functionality of the software will be further integrated into the chip and the sampling rate will be adaptively adjusted according to desired applications.

ACKNOWLEDGMENT

The authors acknowledge the LITE-ON Semiconductor Corporation for their support in chip fabrication.

REFERENCES

[1] A. Moini, *Vision Chips*. Norwell, MA: Kluwer, 1999, ch. 4.
[2] A. Moini, A. Bouzerdoum, K. Eshraghian, A. Yakovleff, X. T. Nguyen, A. Blanksby, R. Beare, D. Abbott, and R. E. Bogner, "An insect vision-based detection chip," *IEEE J. Solid-State Circuits*, vol. 32, pp. 279–284, Feb. 1997.
[3] J. Tanner and C. Mead, "An integrated analog optical motion sensor," in *VLSI Signal Processing II*. New York: IEEE Press, 1986, pp. 59–76.
[4] A. B. Torralba and J. Hérault, "An efficient neuromorphic analog network for motion estimation," *IEEE Trans. Circuits Syst. I*, vol. 46, pp. 269–280, Feb. 1999.
[5] R. Etienne-Cummings, J. Van der Spiegel, and P. Mueller, "Hardware implementation of a visual-motion pixel using oriented spatiotemporal neural filters," *IEEE Trans. Circuits Syst. II*, vol. 46, pp. 1121–1136, Sept. 1999.
[6] C. M. Higgins, R. A. Deutschmann, and C. Koch, "Pulse-based 2-D motion sensors," *IEEE Trans. Circuits Syst. II*, vol. 46, pp. 677–687, June 1999.
[7] T. Delbruck, "Silicon retina with correlation-based velocity-tuned pixels," *IEEE Trans. Neural Networks*, vol. 4, pp. 529–541, May 1993.
[8] R. Etienne-Cummings, J. Van der Spiegel, and P. Mueller, "A focal plane visual motion measurement sensor," *IEEE Trans. Circuits Syst. I*, vol. 44, pp. 55–66, Jan. 1997.
[9] J. Tanner and C. A. Mead, "Correlation Optical Motion Detector," U.S. Patent 4 631 400, Dec. 23, 1986.
[10] J. Kramer, R. Sarpeshkar, and C. Koch, "Pulse-based analog VLSI velocity sensors," *IEEE Trans. Circuits Syst. II*, vol. 44, pp. 86–101, Feb. 1997.
[11] S. C. Liu, "A neuromorphic aVLSI model of global motion processing in the fly," *IEEE Trans. Circuits Syst. II*, vol. 47, pp. 1458–1467, Dec. 2000.
[12] R. C. Meitzler, K. Strohbehn, and A. G. Andreou, "A silicon retina for 2-D position and motion computation," in *Proc. 1995 IEEE Int. Symp. Circuits and Systems*, vol. 3, 1995, pp. 2096–2099.
[13] A. G. Andreou, K. Strohbehn, and R. E. Jenkins, "Silicon retina for motion computation," in *Proc. 1991 IEEE Int. Symp. Circuits and Systems*, Singapore, June 1991, pp. 1373–1376.
[14] 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, pp. 241–247, Feb. 1999.
[15] X. Arreguit, F. A. van Schaik, F. V. Bauduin, M. Bidiville, and E. Raeber, "A CMOS motion detector system for pointing devices," *IEEE J. Solid-State Circuits*, vol. 31, pp. 1916–1921, Dec. 1996.
[16] B. Hassenstein and W. Reichardt, "Systemtheoretische analyze der zeit-, reihenfolgen- und vorzeichenauswerttung bei der bewegungsperzeption des russelkafers chlorophanus," *Z. Naturforsch*, vol. 11b, pp. 513–524, 1956.
[17] C. Mead, *Analog VLSI and Neural System*. Reading, MA: Addison-Wesley, 1989.
[18] C. Y. Wu and C. F. Chiu, "A new structure of 2-D silicon retina," *IEEE J. Solid-State Circuits*, vol. 30, pp. 890–897, Aug. 1995.
[19] C. Y. Wu and H. C. Jiang, "An improved BJT-based silicon retina with tunable image smoothing capability," *IEEE Trans. VLSI Syst.*, vol. 7, pp. 241–248, June 1999.
[20] Z. Wang and W. Guggenbuhl, "Novel CMOS current schmitt trigger," *Electron Lett.*, vol. 24, no. 24, pp. 1514–1516, Nov. 1988.

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

In addition, he conducted post-doctoral research at the University of California at Berkeley in the summer of 2002. Since 1980, he has served as 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 at National Chiao-Tung University. During 1984 to 1986, he was a Visiting Associate Professor in the Department of Electrical Engineering, Portland State University, Portland, OR. Since 1987, he has been a Professor at NCTU. From 1991 to 1995, he was rotated to serve as the Director of the Division of Engineering and Applied Science on the National Science Council, Taiwan. From 1996 to 1998, he was honored as the Centennial Honorary Chair Professor at NCTU. He has published more than 250 technical papers in international journals and conferences. He also has 19 patents including nine U.S. patents. His research interests are nanoelectronics and VLSI including circuits and systems in low-power/low-voltage mixed-signal design, and systems, biochips, neural vision sensors, RF circuits, and CAD analysis.

Dr. Wu is a member of Eta Kappa Nu and Phi Tau Phi Honorary Scholastic Societies. He was a recipient of the IEEE Fellow Award in 1998 and the Third Millennium Medal in 2000. In Taiwan, he received numerous research awards from the Ministry of Education, the National Science Council, and professional foundations.

Kuan-Hsun Huang (S'00) was born in Taiwan, R.O.C., in 1975. He received the B.S. degree in electronics engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1998. He is currently working toward the Ph.D. degree at the Institute of Electronics in the same university.

His main research interests have been in analog integrated circuits, vision chips, and neural networks.