An Insect Vision-Based Motion Detection Chip

Alireza Moini, Abdesselam Bouzerdoum, Kamran Eshraghian, Andre Yakovleff, Xuan Thong Nguyen, Andrew Blanksby, Richard Beare, Derek Abbott, and Robert Eugene Bogner

***Abstract*— The architectural and circuit design aspects of a mixed analog/digital very large scale integration (VLSI) motion detection chip based on models of the insect visual system are described. The chip comprises two one-dimensional 64-cell arrays as well as front-end analog circuitry for early visual processing and digital control circuits. Each analog processing cell comprises a photodetector, circuits for spatial averaging and multiplicative noise cancellation, differentiation, and thresholding. The operation and configuration of the analog cells is controlled by digital circuits, thus implementing a reconfigurable architecture which facilitates the evaluation of several newly designed analog circuits. The chip has been designed and fabricated in a 1.2-μm CMOS process and occupies an area of 2 $\times$ 2 mm^2.**

***Index Terms*— Analog CMOS, analog VLSI, insect vision, machine vision, motion sensors, smart vision sensors, template model, vision chips.**

I. Introduction

CONVENTIONAL vision systems based on mathematical algorithms tend to be very complex, and their hardware implementation requires powerful computers. Biological models of the insect visual system, however, suggest that simpler solutions might exist for restricted tasks of motion detection [1], [2]. A mixed analog/digital implementation of the *template model* of the insect visual system [3] has been successfully accomplished in a proof-of-concept demonstration chip [4].

However, several problems affect the operation of continuous time very large scale integration (VLSI) motion detection chips. Two dominant problems are the presence of a strong 100 Hz (or 120 Hz in some other countries) ac component in artificial lighting, and the difficulty in detecting slow motion, which requires very large time constants for temporal differentiation. These issues have been addressed in the design of this chip. The chip also utilizes a fully testable architecture and reconfigurable circuits.

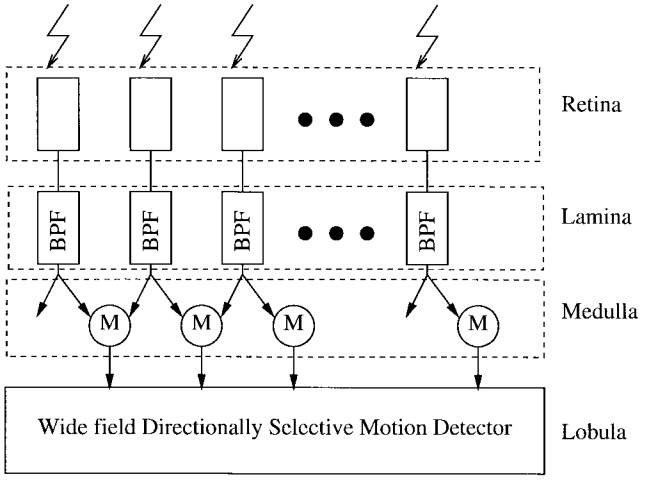

Fig. 1. The simplified model of the insect visual system. The **M** blocks are small-field motion detection elements.

This paper is organized as follows. In Section II the organization of the primary insect visual system and the template model for motion detection are described. Section III introduces the chip architecture. Section IV presents some of the newly designed circuits, including the multiplicative noise cancellation (MNC) circuit and the channel length modulation-based operational transconductance amplifier (CLM-OTA) and the role of these circuits in the system. Test results of the main parts of the circuit are presented in Section V, followed by some concluding remarks in Section VI.

II. Template Model for Insect Vision System

Insect vision has been the subject of research for several decades [5], and it is now clear that in many insects, the motion detection mechanisms play a predominant role. A simplified architecture of the insect visual system is illustrated in Fig. 1. Three main layers compose the primary insect visual system: the *lamina*, the *medulla*, and the *lobula* or *lobula complex* [6], [7]. The lamina contains band-pass or high-pass filters for contrast detection and enhancement. The medulla provides many complicated functions and in particular, small-spatial-field motion detection. Finally, the primary wide-field motion computation is located in the lobula complex. *Lobula plate* (the posterior part of the lobula complex) is also characterized by large directionally sensitive motion detection (DSMD) neurons [8]. From retina to medulla, the information is transferred in parallel channels with a small lateral interaction. Therefore, it is suitable for implementation in VLSI.

Manuscript received December 4, 1995, revised October 2, 1996. This work was supported in part by Britax Rainsfords, the Australian Research Council (ARC), and Department of Education, Employment and Training (DEET).

A. Moini, A. Bouzerdoum, A. Blanksby, R. Beare, D. Abbott, and R. E. Bogner are with the Department of Electrical and Electronic Engineering and the Centre for GaAs VLSI Technology, the University of Adelaide, SA 5005, Australia.

K. Eshraghian was with is the Department of Electrical and Electronic Engineering and the Centre for GaAs VLSI Technology, the University of Adelaide, SA 5005, Australia. He is now with Edith Cowan University, Perth, WA 6027, Western Australia.

A. Yakovleff was with the Department of Electrical and Electronic Engineering and the Centre for GaAs VLSI Technology, the University of Adelaide, SA 5005, Australia. He is now with the Defence Science and Technology Organization (DSTO), Information Technology Division, Salisbury, SA 5108, Australia.

X. T. Nguyen was with the Department of Electrical and Electronic Engineering and the Centre for GaAs VLSI Technology, the University of Adelaide, SA 5005, Australia. He is now with the Air Operation Division, Defence Science and Technology Organization (DSTO), Salisbury, SA 5108, Australia.

Publisher Item Identifier S 0018-9200(97)01118-9.

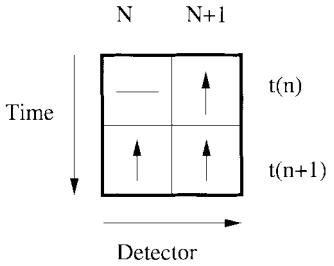

Fig. 2. Forming a template from sampled spatiotemporal contrasts.

The template model, proposed by Horridge [9], models the function of the small-field motion detection neurons in the medulla. In the template model, the temporal contrast is quantized to three levels: increase $(\uparrow)$, decrease $(\downarrow)$ and no-change $(-)$. Therefore, the signals from two neighboring detectors can have one of nine combinations: $(-,-)$, $(-,\downarrow)$, $(-,\uparrow)$, $(\downarrow,-)$, $(\downarrow,\downarrow)$, $(\downarrow,\uparrow)$, $(\uparrow,-)$, $(\uparrow,\downarrow)$, $(\uparrow,\uparrow)$. The combination of the outputs of two detectors at two sampling times gives rise to 81 possible states, each called a template (see Fig. 2).

In the VLSI implementation of the template model, two basic analog building blocks are required; a temporal differentiator and a thresholding circuit. The rest of the model relies on the digital processing of the outputs of the thresholding stage.

III. Chip Architecture

The template model was implemented in the first generation of our insect vision chips, called Bugeye I [10]. The architecture of that chip was a close replica of the simplified model of insect vision. The retina was implemented by parasitic p-well junction diodes and logarithmic current-to-voltage converters for providing a wide dynamic range [10]. The band-pass filter in the lamina was implemented by using a subthreshold differentiator. Template formation and encoding were implemented digitally, after thresholding the outputs of the band-pass filter [4]. Finally, a higher level processing layer, containing six parallel wide-field motion detection engines, was included as the hardware counterpart of the lobula.

This second generation of the insect vision chip, called Bugeye II, only implements the retina and lamina. Template model decoder and wide-field motion detection units, which were implemented in Bugeye I, have not been included in the Bugeye II, which has principally been designed as a test bed for evaluating several different algorithms and new circuits. For this purpose, testability and flexibility have been paid considerable attention. Special switches have been used to route the signals from each individual circuit in the array to the outside of the chip. Signals from outside the chip can also be applied directly to the circuits, resulting in a fully testable structure.

IV. Analog Building Blocks

A. The CMSA and MNC Circuits

Spatial smoothing of the input signal is one of the operations required in many front-end processing stages for reducing the input noise. Voltage mode circuits which use resistive networks have been used elsewhere [11]. However, these circuits occupy large areas. Current mode smoothing networks, which have already been implemented in several vision chips [12], are area-efficient and can provide a variable smoothing window as a function of biasing voltages.

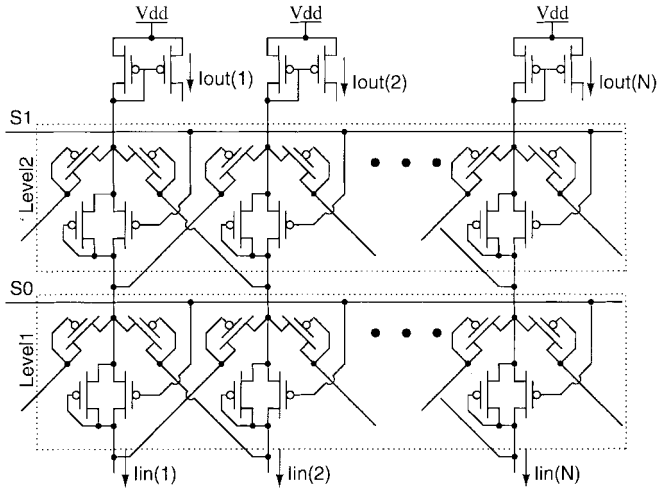

Fig. 3. The implemented CMSA circuit.

Another important reason for making use of spatial averaging in the chip is MNC. It should be mentioned that the MNC operation is not a part of the template model, and it can be considered as a signal conditioning stage before the differentiation stage. Our previous experiments indicated that the 100 Hz frequency component of light sources, operating with mains power, severely affect the detection of motion because of the temporal differentiation stage. In this stage, noise is sometimes amplified to a point where it dominates the signals induced by the motion of objects. As this noise is usually common among several neighboring photodetectors and has a multiplicative nature, spatial averaging is ineffective, but dividing the signal in one channel by the spatial average of the input signals should effectively reduce the effect of the noise [13]. Therefore, the MNC consists of spatial averaging and division-by-average stages. Although in our chip the main reason for using the MNC is to cancel the 100 Hz noise, somewhat similar principles have been exploited in other vision chips, such as the neuromorphic chips mimicking the early visual processing in the retina of various animals [11], [12], [14], [15].

In order to realize the spatial smoothing, we have designed a current mode spatial averaging circuit (CMSA) which features a digitally adjustable smoothing window. In comparison, the circuit described in [12] has a continuously adjustable smoothing window and requires a few less transistors.

The transistor level circuit of the implemented CMSA is shown in Fig. 3. The circuit has two levels of averaging circuitry and comprises current mirrors at the output. The averaging window of this circuit can be changed simply by switching on or off those transistors that are controlled by $S1$ and $S0$.

The division operation required in the implementation of the MNC is performed by a translinear divider which operates in

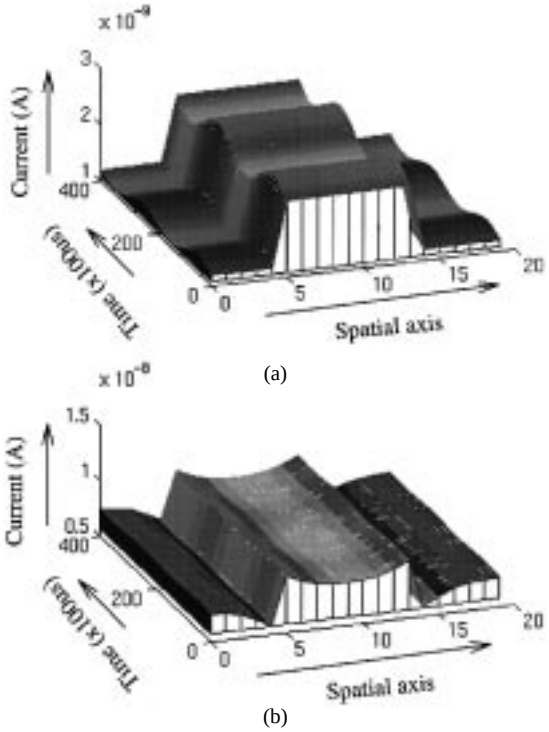

Fig. 4. Simulation output for input signals containing a 50% ac noise in an array of implemented MNC circuits: (a) without activating the MNC circuit (b) with the MNC circuit activated.

the subthreshold region [12]. The input and output variables of this divider are all currents, and the divider is thus suited to our CMSA circuit.

An array of 20 averaging cells and division circuits has been simulated. The input to the array is a spatial square function with current levels of 1 nA and 10 nA. Multiplicative noise with a frequency of 50 Hz and a magnitude of 50% of the signal level in each channel has been added. Fig. 4 shows the input and output signals of the circuit. The relative noise in the output signal has been significantly reduced from the initial 50% to less than 5% of the dc level of the signal. Different simulations with different models and current levels yield the same improved result.

B. Channel Length Modulation OTA (CLM-OTA) for Temporal Differentiation

One principal operation of a motion detection algorithm is temporal differentiation. In the implementation of a time differentiator, the bandwidth of the signals (in our case the photocurrents) to be differentiated is of major concern. Differentiator circuits implemented in VLSI have minimum and maximum frequencies for proper operation. These frequencies are determined by the gain of the differentiator, the parasitic capacitance and impedances present in the circuit, and the dynamic range and saturation voltage of the circuits used in the realization of the differentiator. The OTA-based circuit, shown in Fig. 5, was used in the previous design [10]. The OTA in this circuit requires a biasing current of

$$I_{\text{bias}} = \frac{kT}{nqR_{\text{eq}}} \tag{1}$$

to yield a specified equivalent resistance R_{eq} where n is the

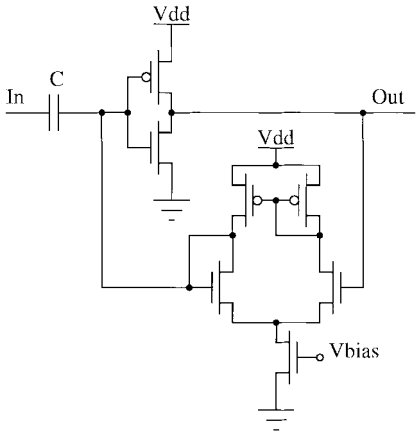

Fig. 5. The OTA-based differentiator implemented in the chip.

subthreshold ideality factor. Therefore, very small currents in the deep subthreshold region are needed to yield the desired equivalent resistances (in the order of 1 GΩ).

To obtain high resistance values while maintaining feasible biasing currents, we have designed a circuit based on the channel length modulation in MOS transistors. The simplest expression characterizing the drain-source current I_{ds} as a function of the channel length modulation coefficient λ is given by

$$I_{ds} = f(V_{gs}, V_T)(1 + \lambda V_{ds}). \tag{2}$$

Fig. 6 shows a simple circuit which achieves this purpose. I_{bias} is the bias current. Transistors M2 and M1 form a source follower stage and V_b follows the input voltage V_{in}. Therefore, the current through this branch, I_1, is modulated by the input voltage. Transistors M6 and M5 provide a reference current, I_2, which is subtracted from the modulated current using the current mirror formed by M3–M4. By changing the state of M7, the circuit can operate either in the channel length modulation mode, when M7 is turned off, or in the simple OTA mode, when M7 is turned on.

The operation of the CLM-OTA can be affected by device mismatch, which stems from two main sources: the mismatch between the λ factors of the two transistors M1 and M5, and the mismatch between the current mirror transistors M3 and M4. In an ideal circuit, the output current would be zero if the input voltage was equal to the reference voltage V_{ref}. However, an offset voltage may be caused by mismatches. For analysis, it is convenient to assign all the fluctuations of β (the transconductance of transistors) and V_T (the threshold voltage) to the current mirror, which results in an offset of bias current as a shift of the $I_{ds} - V_{ds}$ curve in Fig. 7(a), and the fluctuations of V_A (the Early voltage) to M1 and M5, as a change of the slope of the $I_{ds} - V_{ds}$ curve indicated in Fig. 7(b). The total offset voltage ΔV_{offset} can be expressed as

$$\begin{aligned}\Delta V_{\text{offset}} &= \sqrt{(\Delta V_1)^2 + (\Delta V_2)^2} \\ &= \sqrt{\left(\frac{\Delta I_{\text{bias}}}{I_{\text{bias}}} V_A\right)^2 + \left(\frac{\Delta V_A}{V_A} V_{\text{ref}}\right)^2}\end{aligned} \tag{3}$$

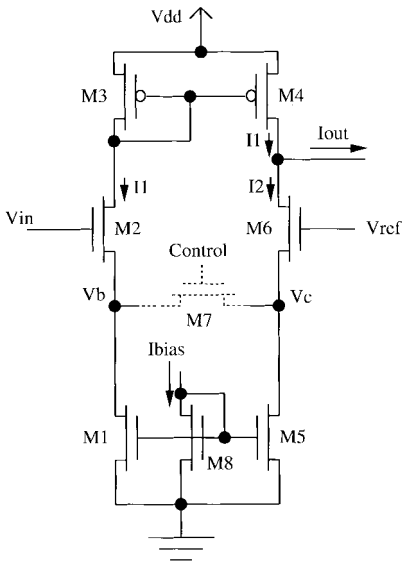

Fig. 6. The reconfigurable transconductance amplifier, which can switch between the simple OTA-mode and the CLM-OTA-mode.

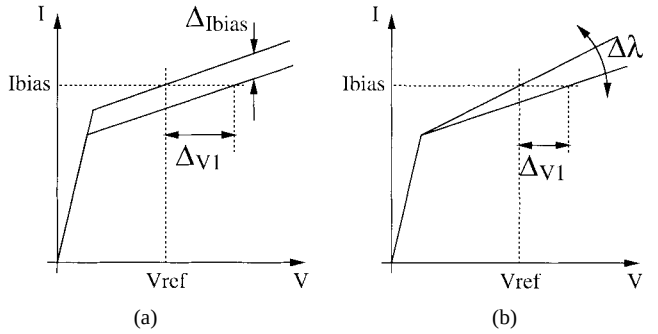

Fig. 7. Offset voltage introduced by mismatch between transistors in current mirrors, and biasing transistors. (a) Bias current shift and (b) a change in the slope of the I-V curve due to Early voltage mismatches.

where I_{bias} is the bias current and V_{ref} is the reference voltage. $\Delta I_{\text{bias}}/I_{\text{bias}}$ and $\Delta V_A/V_A$ are the relative variances of the bias current and Early voltage due to transistor mismatch, respectively. $\Delta I_{\text{bias}}/I_{\text{bias}}$ is inversely proportional to the transistor gate length, L, V_A is directly proportional to L, and $\Delta V_A/V_A$ is relatively constant. Therefore, the equation seems to result in a constant value. However, V_A in the first term of (3) may significantly increase the offset due to bias current mismatch. Hence, V_A should be chosen as small as possible, which means that the transistor gate length should be proportionally small. Also, the width of transistors M1 and M5 should be made large enough to decrease mismatch. In our design we have used the minimum gate length for biasing transistors, which is 1.2 μm and have an Early voltage of about 5 V.

V. Physical Implementation and Test Results

Two arrays of analog circuitry, each containing 64 rows, have been implemented in the chip. Each row comprises a photodetector, CMSA cell, division circuit, differentiator, thresholding circuit, and analog switching circuits for testability purposes. The chip has been designed in a 1.2-μm double-poly double-metal p-well CMOS process, and occupies an area of 2×2 mm^2. *Metal-2* layer has been used for protecting the analog circuits from light exposure.

The CMSA and MNC stage were tested by projecting a stationary light bar onto the chip (Fig. 8). The sharp fluctuations observed on the input and output signals are caused by mismatches between transistors used in the CMSA circuit and in the divider. The original photocurrent profile obtained directly from the output of photodetectors is smooth (not shown). Although the magnitude of the fluctuations is relatively high compared with the input signal, it appears only as a stationary spatial background. In the implementation of the template model, which does not rely on the spatial information in the early stages of analog processing, this spatial background does not affect the operation of the circuit. However, in other motion detection models, which depend on spatial information, significant performance degradation can occur.

To test the CLM-OTA, the reference input was fixed at 2.5 V, while the other input was varied from the ground voltage to the supply voltage range. The output current of the CLM-OTA is very linear over a range of 1 V to 5 V, i.e., from about the threshold voltage of the transistor to the supply voltage.

The I-V characteristics of all the available CLM-OTA's were measured in all 40 chips received from the manufacturer. At a biasing current of 100 nA, a mean offset voltage of 0.37 V with a standard deviation of 0.55 V is obtained. The maximum offset is about 1.2 V. As the circuit is used in a feedback loop, and the dc offset of the circuit does not affect the operation of the next processing stages, large offset values do not affect the function of the circuit.

For testing at the system level, the chip was exposed to moving objects. To obtain a reasonable output, the biasing currents needed to be tuned for the particular lighting conditions. Fig. 9 shows the coded output of the chip. The straight pattern is from a person moving 3 to 4 m away from the chip, and the curved pattern is from a pen waved in front of the chip at approximately 30 cm. Each hexadecimal number represents a particular template which has been encoded to single characters for display purposes.

VI. Conclusion

The architecture and building blocks of a motion detection chip, inspired by the insect visual system and based on the template model for insect vision, were described. The model lends itself to VLSI implementation, as the processing is done in parallel and no spatial processing in the analog front-end is required. Spatial information is only used after thresholding the time differentiated photocurrents.

A new current mode spatial averaging circuit has been implemented for use in the multiplicative noise cancellation process to filter out the effect of the 100 Hz ac noise.

The channel length modulation OTA shows a significant transconductance reduction compared with a simple five-transistor OTA. Transconductance values as low as 10^{-11} S can be achieved with a practical bias current of 1 nA. Therefore, time constants in the order of 0.1 s are achievable in a very small area.

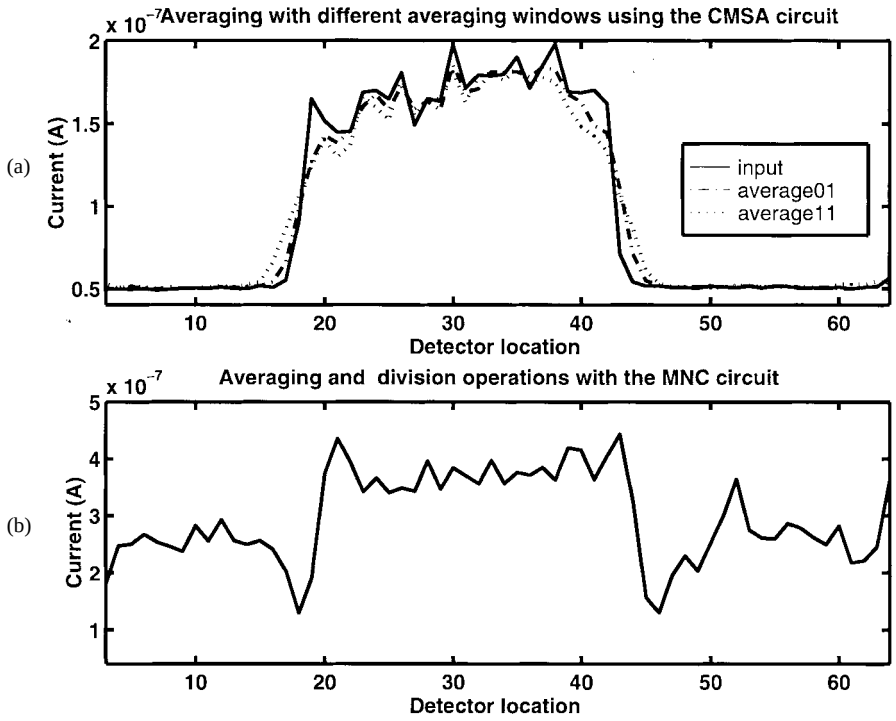

Fig. 8. Test results of the CMSA and MNC circuits with an input light bar. (a) The input and the spatially averaged output of the CMSA circuit with two different averaging windows. "average01" is the output of the CMSA when $S1 = 0$ and $S0 = 1$ and "average 11" is the output when $S1 = S0 = 1$. (b) The output of the MNC circuit in response to the light bar.

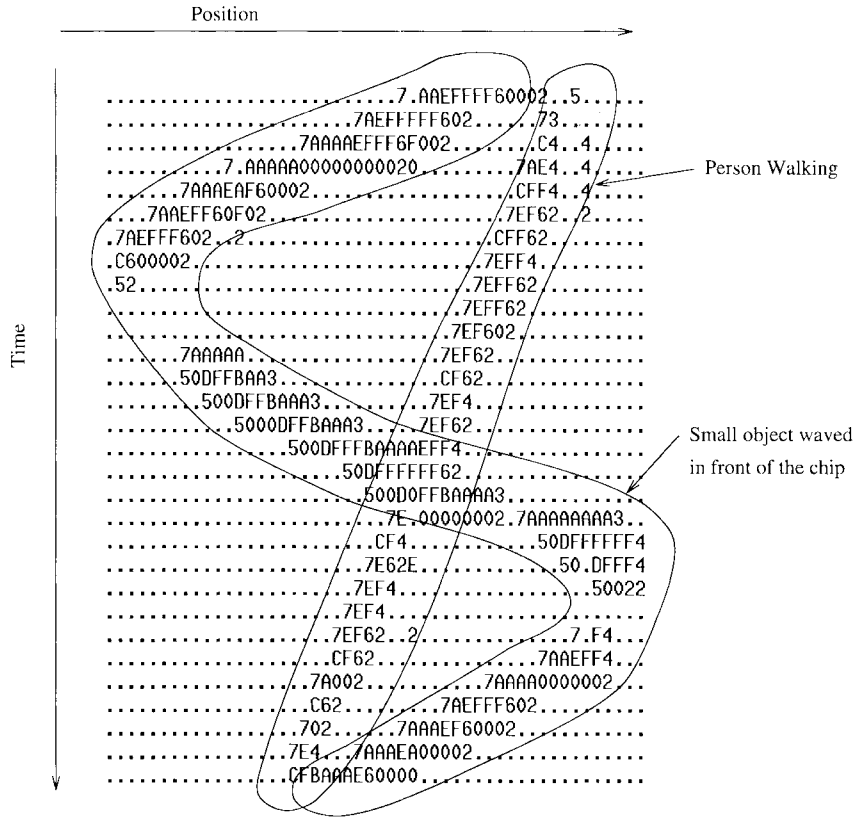

Fig. 9. Output of the chip when looking at two different objects at different distances from the chip. The "·" represents no-motion.

ACKNOWLEDGMENT

The authors are also indebted to Emeritus Professor G. A. Horridge for his inspiration and encouragement during the course of this research.

REFERENCES

[1] G. A. Horridge, "Ratios of template responses as the basis of semivision," *Phil. Trans. Royal Soc. London, B*, vol. 331, pp. 189–197, 1991.
[2] J. M. Pichon, C. Blanes, and N. Franceschini, "Visual guidance of a mobile robot equipped with a network of self-motion sensors," *SPIE Proc. Mobile Robots IV*, vol. 1195, pp. 44–53, 1989.
[3] G. A. Horridge and P. Sobey, "An artificial seeing system copying insect vision system," *Int. J. Optoelectron.*, vol. 6, no. 1/2, pp. 177–193, 1991.
[4] A. Yakovleff, A. Moini, A. Bouzerdoum, X. T. Nguyen, R. E. Bogner, and K. Eshraghian, "A micro-sensor based on insect vision," in *Workshop on Computer Architectures for Machine Perception*, Dec. 1993, pp. 137–146.
[5] G. A. Horridge, *The Compound Eye and Vision of Insects*. Oxford: Clarendon, 1975.
[6] N. J. Strausfeld, "Beneath the compound eye: neuroanatomical analysis and physiological correlates in the study of insect vision," in *Facets of Vision*, D. G. Stavenga and R. C. Hardie, Eds. Berlin: Springer-Verlag, pp. 317–359, 1989.
[7] S. R. Shaw, "The retina-lamina pathway in insects, particularly diptera, viewed from an evolutionary perspective," in *Facets of Vision*, D. G. Stavenga and R. C. Hardie, Eds. Berlin: Springer-Verlag, pp. 186–212, 1989.
[8] A. Bouzerdoum, "The elementary movement detection mechanism in insect vision," *Phil. Trans. R. Soc. London*, vol. B-339, pp. 375–384, 1993.
[9] G. A. Horridge, "A template theory to relate visual processing to digital circuitry," *Proc. Royal Soc. London*, vol. B-239, pp. 17–33, 1990.
[10] A. Moini, A. Bouzerdoum, A. Yakovleff, D. Abbott, O. Kim, K. Eshraghian, and R. E. Bogner, "An analog implementation of early visual processing in insects," in *Proc. 1993 Int. Symp. VLSI Technology, Systems, and Applications*, Taipei, Taiwan, May 1993, pp. 283–287.
[11] C. Mead, *Analog VLSI and Neural Systems*. Reading, MA: Addison-Wesley, 1989.
[12] A. G. Andreaou, K. A. Boahen, P. O. Pouliquen, A. Pavasovic, R. E. Jenkins, and K. Strohbehn, "Current mode subthreshold MOS circuits for analog VLSI neural systems," *IEEE Trans. Neural Networks*, vol. 2, pp. 202–213, Mar. 1991.
[13] A. Moini, A. Blanksby, R. Beare, A. Bouzerdoum, and K. Eshraghian, "Multiplicative noise cancellation (MNC) in analog VLSI vision sensors," in *Electronics Technology Directions for the Year 2000 (ETD2000)*, Adelaide, Australia, 1995, pp. 253–257.
[14] E. A. Vittoz and X. Arreguit, "Linear networks based on transistors," *Electron. Lett.*, vol. 29, no. 3, pp. 297–299, Feb. 1993.
[15] W. Bair and C. Koch, "Real-time motion detection using an analog VLSI zero-crossing chip," *SPIE Proc. Visual Information Processing: From Neurons to Chips*, vol. 1473, pp. 59–65, 1991.