MVA '98 IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

11—2

A Light Adaptive 4000 Pixel Analog Silicon Retina for Edge Extraction and Motion Detection

Gilles Sicard, Gérard Bouvier, Vincent Fristot, *
Laboratoire des Images et des Signaux
Institut National Polytechnique de Grenoble

Alan Lelah †
Centre National d'Etudes des Télécommunications
Grenoble

Abstract

An analog vision chip has been designed on the basis of retina models. Biological retinas perform a massively parallel computation well suited to low level visual processing applications such as edge extraction, motion detection and light adaptation. An analog electrical model has been developed that is suitable for designing analog neural network circuits with low power consumption. This 4000 pixel neuromorphic circuit extracts edges or detects motion over a large dynamic range of luminosity and velocity. Experimental results are presented.

1 Introduction

Numerous digital visual algorithms are used in a wide range of applications : edge extraction, texture analysis and motion sensing and analysis. In such systems, the amount of information delivered by the source is substantial. The required investment in hardware and computation time is often very costly for just low level visual processing.

On the other hand, bio-inspired retina models represent a massively parallel spatiotemporal filter that is effective for low level visual processing applications such as edge extraction and motion detection [1] [7]. Our model contains a two-layer filter structure representing the first functional retina layer. It provides an easy way to implement analog continuous time filtering in the focal plane array [6] [8]. The design of low power smart sensors is facilitated by the simple structure of the network and by use of CMOS transistors in weak inversion mode. Each pixel of the array includes a light adaptive bio-inspired system to increase sensitivity by better information encoding [3]. Also, this system reduces noise effects by adaptive filtering [5].

Our chips have been fabricated using a 3.3V, $0.5\mu m$ CMOS technology. We present a circuit including an array of 64x64 pixels with both analog and digital outputs. This chip can extract edges of static and moving objects or detects motions. These two functions and filtering features may be selected by bias voltage modulation of the network parameters. Experimental results demonstrate both operating modes.

2 Biological inspiration

The retina is the first neural structure to be used in visual data processing (Fig. 1). It filters input signals and extracts only the relevant information used in visual cortex for high level visual processing tasks.

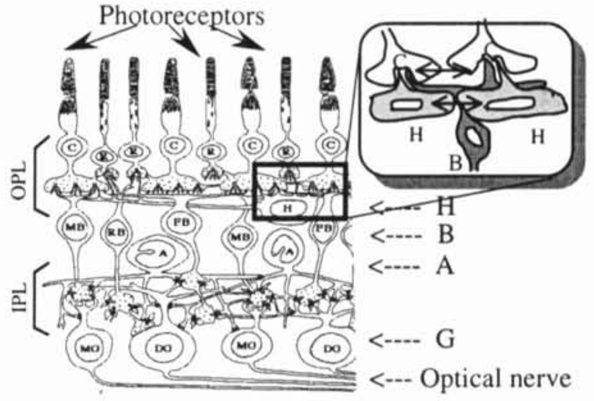

Figure 1: Schematic cross section of the Macaque retina showing the five basic neural cells. The synaptic triad (the basic neural set of our model), which connects the three OPL cells, is enlarged.

The photoreceptors convert incident light into electrical signals and also compute a first low pass filtering. Each horizontal cell (H) spreads information over neighboring horizontal cells and thus performs

*Address : LIS, 46, Av. Félix Viallet, 38031 Grenoble cedex, France. E-mail :[sicard or bouvier]@tirf.inpg.fr
†Address : CNET, DTM/CET, 28, Chemin du vieux chêne, BP 98, 38243 Meylan cedex, France. E-mail :alan.lelah@cnet.francetelecom.fr

spatiotemporal low-pass filtering on the photoreceptor outputs. Bipolar cells (B) deliver the difference between the receptor and horizontal cell responses. This band-pass filtering result forms a first functional layer: The Outer Plexiform Layer (OPL). The second functional layer, the Inner Plexiform Layer (IPL), is composed of amacrine cells (A) and ganglion cells (G). It uses OPL outputs for among others detecting motion and encoding the information before transmitting it through the optical nerve.

3 Electrical model

The first analog retina model was proposed by C. Mead et al. [6]. This model implements the biological structure which connects the three OPL cells, i.e. the synaptic triad. It constitutes the basic neural set of our model. In order to develop a biologically and structurally consistent retina model, a more complex OPL model was developed in our Laboratory [1]. This model incorporates a resistive layer which implements photoreceptor coupling (Fig. 2). It produces in $C_k(t)$ a spatial low pass filter that attenuates the photoreceptor's dispersion and reduces the noise. $E_k(t)$ is the output voltage of the light adaptive phototransduction stage. It provides an incident light-electrical signal conversion in accordance with the illumination condition. This light adaptive system control the resistor r_c.

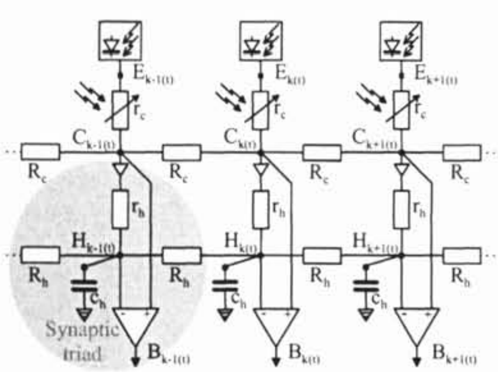

Figure 2: Electrical schematic of the OPL model.

The second resistive layer performs in $H_k(t)$ a low pass filtering, which is equivalent to deliver a spatiotemporal averaged value of the photoreceptor outputs. The differential output in $B_k(t)$ provides a spatial band-pass response and a temporal high pass response for a range of a low spatio-temporal frequencies. The transfer function is given by:

$$G(f_x, f_t) = \frac{B(f_x, f_t)}{E(f_x, f_t)} = F_h(f_x, f_t)[1 - F_c(f_x, f_t)] \quad (1)$$

with

$$F_c(f_x, f_t) = \frac{C(f_x, f_t)}{E(f_x, f_t)} \quad (2)$$

$$F_c(f_x, f_t) = \frac{1}{1 + 2\alpha_c(1 - cos(2\pi f_x))} \quad (3)$$

and

$$F_h(f_x, f_t) = \frac{H(f_x, f_t)}{C(f_x, f_t)} \quad (4)$$

$$F_h(f_x, f_t) = \frac{1}{1 + 2\alpha_h(1 - cos(2\pi f_x)) + 2\pi j f_t \tau_h} \quad (5)$$

Two types of constants modulate the outputs of each layer :

- The space-constants $\alpha_c = \frac{r_c}{R_c}$ and $\alpha_h = \frac{r_h}{R_h}$ control the spatial filtering.
- The time-constant $\tau_h = r_h C_h$ controls the temporal filtering.

These constants are modulated by external biases which control resistor values.

This network possesses very attractive properties for image processing by the modulation of the α_h parameter:

- It inherently does edge extraction in an image (Fig. 8.b and c).
- It can detect only the edges of moving objects (Fig. 8.d).

The motion detection emerges from the second term of the network transfer function (1). If the space-constant α_h (which controls spatial behavior of the second layer of the filter) tends towards zero (with horizontal resistors R_h much higher than vertical resistors r_h) then the network transfer function is transformed in such that spatial filtering disappears for low temporal frequencies. Edges of fixed objects are therefore strongly attenuated as α_h tends towards zero. On the other hand, edges of moving objects are continuously extracted and enhanced.

The phototransduction stage and the first resistive layer implement a non-linear photoreceptor model with lateral interactions which are deduced from the biological cone adaptation system [1] [2]. Several authors have shown interest in light adaptive systems. Usually, this property was used to increase the sensitivity of photoreceptors [3] or to improve regularization [5]. Our model offers both these possibilities. It increases the sensitivity in comparison with a classical logarithmic compressor circuit [7]. The new compression law is logarithmic only for illumination levels close to the average value (Fig. 3). The curve shifts according to the mean illumination. For an input range ΔI (Fig. 3), a higher ouput range ΔV_2 is obtained with the light adaptive phototransduction law than the logarithmic compressor law (ΔV_1). For convenience, phototransduction curves are plotted using the phototransistor photonic current and not the illumination value. In figure 4, experimental results are presented. The edges are extracted with an 1D network using logarithmic compressor (a.1 and b.1) and with a light adaptive network (a.2 and b.2). The network parameters are the same for both 1 D retinas except the differentiator gain. Thus, results obtained are equal for a maximum contrast (a.1 and a.2). On the other hand, a pattern with a less stronger contrast is better extracted with the adaptive network (b.2). The sensitivity to weak contrasts increases through better information encoding.

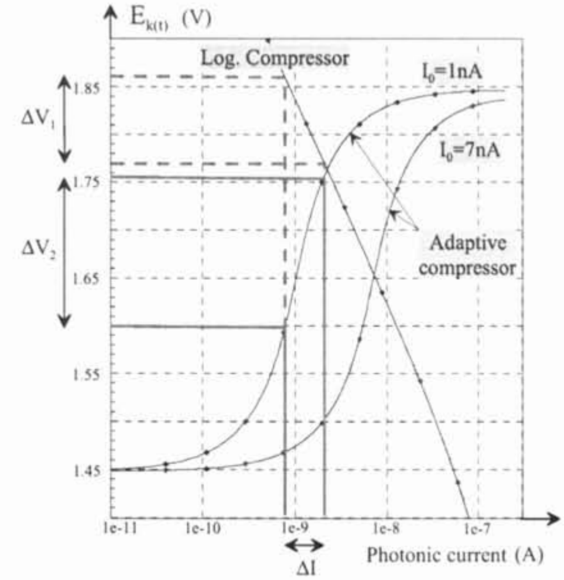

Figure 3: Comparison between a classical logarithmic compression law and light adaptive phototransduction law.

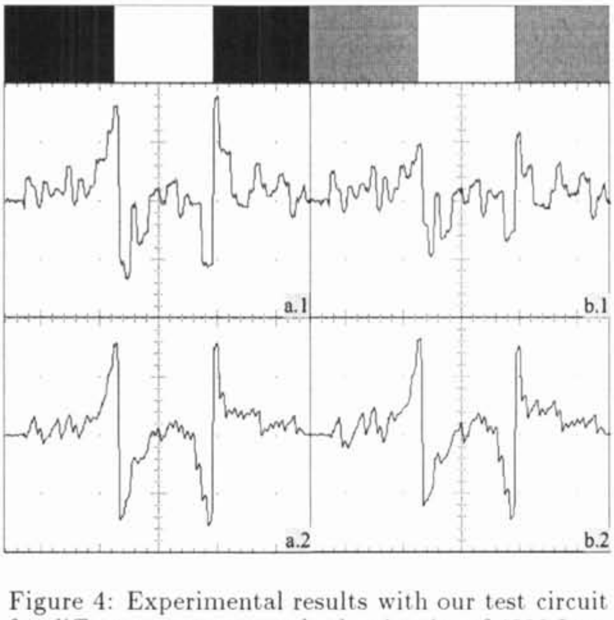

Figure 4: Experimental results with our test circuit for different contrasts and a luminosity of 4000 Lux. There are two 1D 48 cells retinas on this circuit. One uses logarithmic compressor (a.1 and b.1). The second one is a light adaptive network (a.2 and b.2).

Furthermore, the resistor r_c is controlled by the incident light and the mean illumination. The high cut-off spatial frequency of the network is therefore modulated by the resistor variation. The figure 5 shows experimental and simulation transfer function curves obtained with our OPL model [9] [4]. The bandwidth and the gain decrease with the illumination condition. This enables an adapted regularization by automatic signal-to-noise ratio optimization [5]. Note that the cut-off frequency and the illumination conditions are different for simulation results and experimental results.

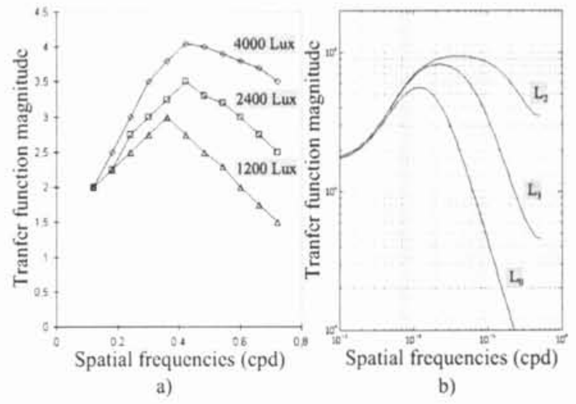

Figure 5: Adaptive filtering: a) Experimental transfer function curves versus spatial frequency (in cycles per degree) for different mean illuminations. b) Simulation results.

4 A 64x64 pixel silicon retina

Our first 2D retina (an array of 24x24 pixels, $13mm^2$, $250\mu W$, CMOS $0.5\mu m$) [8] showed the feasibility of analog neuromorphic circuits with good filtering features.

Thus, we designed a more complex circuit in order to demonstrate its uses in different applications. This one contains an array of 64x64 adaptive pixels. The pixel size is $110\mu m$ x $110\mu m$ and it contains 55 transistors. The phototransistor is a vertical PNP with an area of $900\mu m^2$. The chip is designed in a 3.3V CMOS $0.5\mu m$ technology (Fig. 6).

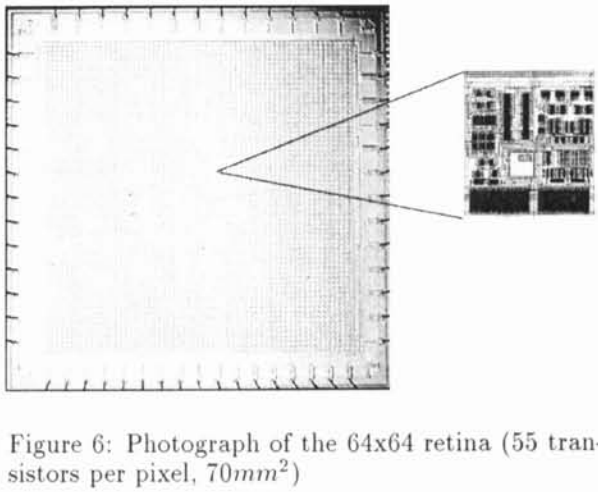

Figure 6: Photograph of the 64x64 retina (55 transistors per pixel, $70mm^2$)

The circuit includes an analog output (for video, 25 images per second) and a digital output (6-bit AD flash converter) with the possibility of random addressing (Fig. 7).

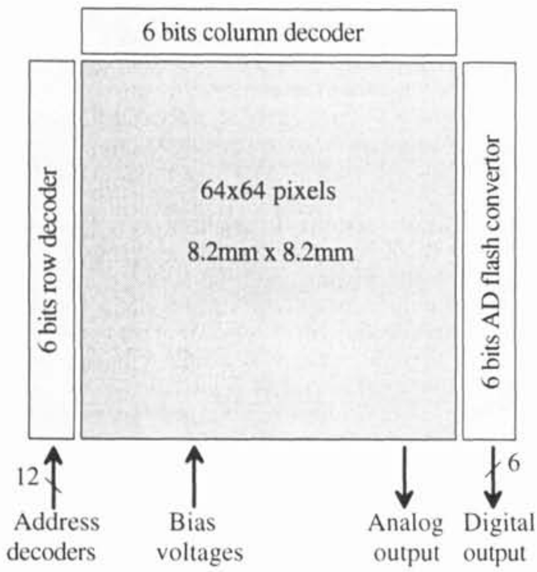

Figure 7: The chip's functional blocks.

This circuit, which is currently used in a camera, can work in different modes (edges extraction or motion detection). This is easily selected via only one external modulation (bias voltage) of a network parameter (α_h). The figure 8 b and c shows results in edge extraction mode. Edges of moving objects (the car in c) are more enhanced than edges extracted with an immobile car (in b). The image d present the motion detection mode result, where only the moving car appears.

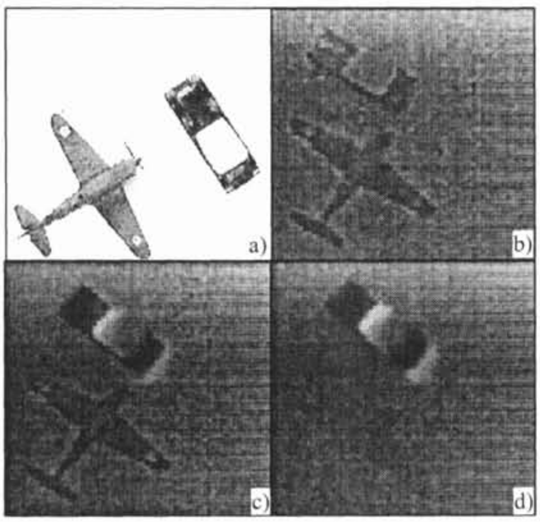

Figure 8: Video sequence with the 4000 pixel retina. a) Original scene contains a fixed plane and a car. b) Edge extraction mode with a fixed car. c) Edge extraction mode with a slowly moving car. d) Motion detection mode.

The velocity range detected with nominal parameters is suitable for the detection of all mobile objects in a street scene.

5 Conclusions

We designed a 4000 pixel analog neuromorphic vision circuit. It contains 250000 transistors, an analog and a digital outputs with the possibility of random addressing. This silicon retina is currently used in a camera.

It can work in edge extraction mode or motion detection mode by easily external modulation of the network parameters. The motion detection is effective over a wide range of speeds. The light adaptive system offers high sensitivity to weak contrasts with an automatic signal to noise ratio optimization.

Results obtained with this intelligent sensor are suitable for real time applications in robotics, detection and machine vision.

6 Acknowledgments

This work benefits from the support of the CNET (market title N^o 95 3B).

References

[1] W.H.A. Beaudot, "Le traitement neuronal de l'information dans la rétine des vertébrés", M.S. thesis, LTIRF-INPG, Grenoble France, December 1994.

[2] W.H.A. Beaudot, "Sensory coding in the vertebrate retina: towards an adaptive control of visual sensitivity", Computation in neural systems, N^o7, pp. 317-323, February 1996.

[3] T. Delbrück, C. Mead, "Time derivative adaptive silicon photoreceptor array", Proc. SPIE, Infrared sensors : Detectors, Electronics and Signal Processing, 1541 : 92-99, 1991.

[4] R.L. De Valois, K.K De Valois, "Spatial vision", Oxford Psychology series N^o14, Oxford Science Publications, 1990.

[5] H. Kobayashi, T. Matsumoto, K. Tanaga, T. Yagi, "Light adaptive architectures for regularization vision chips", Neural Networks, vol. 8, no. 1, pp 87-101, 1995.

[6] C. Mead and M.A. Mahowald, "A silicon model of early vision processing", Neural Networks, vol.1, pp. 91-97, 1988.

[7] C. Mead, "Analog VLSI and neural systems", Addison-Wesley, Reading, Massachusetts, 1989.

[8] A. Mhani, G. Sicard, G. Bouvier, "Analog vision chip for sensing edges, contrasts and motion", ISCAS'97, pp. 2769-2772, June 1997.

[9] A. Torralba, "Traitements spatio-temporels non-linéaires dans un modèle de vision d'inspiration biologique", Technical report, LTIRF-INPG, Grenoble, France, 1996.