Real-Time Image Edge Extraction with Integrated CMOS Sensors

Denis Dubois
Denis Poussart

Laboratoire de Vision et Systèmes Numériques
Département de Génie Electrique
Université Laval,
Québec, Canada, G1K 7P4

Abstract

This paper describes two models of a real-time edge extracting system implemented in VLSI CMOS technology. Each unit is complete and include processing circuitry as well as photosensitive devices which perform light acquisition directly from a scene. The first model (LVRET) uses an algorithm which is based on principles of lateral inhibition and the second one (LVHEX) takes advantage of the behavior of a large resistive network as a smoothing operator. The circuits are built upon an hexagonal tessellation and their output is derived from a scanning operation performed on a 28 x 36 planar array in the LVRET chip and 27 x 35 in the LVHEX one. The paper outlines some basic concepts which arise from biological and neuronal architecture and describes how the relevant concepts mesh with the intrinsic properties of VLSI.

Resumé

Cet article présente deux modèles de systèmes d'extraction d'arêtes en temps réel realisés sur circuits intégrés à très grande échelle. Les unités ne nécessitent pas de caméra puisque les éléments de calcul ainsi que les photosenseurs qui recoivent l'information lumineuse directement de la scène sont inclus sur le même substrat de silicium. Le premier modèle (LVRET) utilise un algorithme décisionnel derivé du principe d'inhibition latérale tandis que le second (LVHEX) exploite l'effet d'adoucissement de signal d'un grand réseau resistif. Une tessellation hexagonale est utilisée pour couvrir tout le plan image et la sortie est le résultat du balayage d'une matrice de 28 x 36 pixels pour LVRET et de 27 x 35 pour LVHEX. L'article expose les principes architecturaux inspirés des systèmes biologiques et illustre comment ceux-ci s'intègrent naturellement et efficament à la technologie microélectronique.

Keywords: Integrated sensing, analog VLSI implementation, edge detection, parallel computing, photosensitive silicon devices.

Introduction

With the recent developments in the area of semiconductors, it has become increasingly apparent that very complex, special purpose systems can be implemented on a single chip and that different types of devices can be integrated on the same silicon wafer. In order to obtain the very high levels of density and high speed necessary for real-time implementation of computer vision algorithms, one must take full advantage of the intrinsic properties of conventional silicon devices. This translates into two observations: *i)* typically, analog circuitry which directly implements computations to be performed have significant advantages of processing speed over digital approaches and *ii)* global communications are very costly in terms of utilization of the silicon real estate, and therefore architectures with dense local connectivity must be priviledged. Biological systems follow similar principles, and vision in particular provides a very pertinent model of organization [12]. In the same spirit, it is logical to attempt an integration of the transducting and of the processing components in order to achieve high performance, "intelligent" sensing. The devices described in this paper combine photosensitive devices and analog elements on the same silicon chip [16], thus allowing direct data connections and providing a computation which proceeds in a massively parallel manner.

Neural networks models are good examples of the kinds of architectures which are directly inspired from studies of biological systems. These networks are complex, highly organized and fairly regular structures, and are composed of a large number of simple processors connected to one another, as in nervous tissue [9, 8, 1, 4]. The results presented here are very much influenced by neural network principles.

Parallel Architecture

Typically, conventional, sequential computing is not capable of providing the level of performance required for real-time image processing. Although the speed of single digital devices is impressive, and far superior to that of biological elements, only a small fraction of the available ressources is utilized at any particular time and intensive tasks such as early-processing and recognition are overwhelming.

Vision processing is a field where a strategy of parallel architecture is most indicated. Indeed, *i)* the very nature of light is such that the incoming optical information impinges upon the sensor in a parallel manner and *ii)* the early processing operations to be performed involved a high degree of local support and connectivity. Therefore a sensing architecture which integrates light transduction within an array of local processors which have efficient communication with their surrounding neighbors has the required attributes and the potential of high performance. This is a path which Nature has largely developped.

Retinal Physiology

The retina is not a simple sensor which gathers intensity and color information to be forwarded to the visual cortex [6]. Indeed, the retina consists of a complex and highly organized structure, with each cell acting as a small analog processor which processes information in a cooperative, parallel, hierarchical manner before a relevant representation is finally made available through the optic nerve. This early segmentation operation is a key step toward the goal of successful recognition of objects and symbolic representation of visual scenes. The various processing operations can be viewed, in part, as a form of spatial and temporal filtering, dealing for instances with noise or redundancy [18]. Among these early visual processes, edge detection is considered to plays a most significant role; therefore it has been selected as a preferred target for an articial implementation in VLSI technology.

The input layer of the retina is composed of a large number of receptive cells that are connected to one another in a center-surround antagonistic manner. Figure 1 shows a highly simplified, one-dimensional representation of the type of connectivity which each receptor cell supports with its neighbours. Cells located in the processing layer(s) receive a positive (excitatory) contribution from the photosensitive cell which is located immediately above and negative (inhibitory) contributions from other cells in the immediate surround. The spatial extent of connections defines the size and geometry of the field associated with a specific channel. This receptive field may include a large collection of cells and its characteristics, as a spatial processor, are determined by the dependence of connection weights as a function of distance and orientation from the central element. The processing behavior of this architecture is called lateral inhibition and corresponds to a particular configuration of a neural network where weighted connections are given fixed values and have a high degree of local regularity.

Looking at one particular neuron in the processing level we note that it acts as a simple adder and subtractor followed by a comparator [18]. This arrangement is shown in Figure 2 and represents the basic node of a class of neural network, with the exception that the weights of the input connections are fixed.

From Lateral Inhibition to Edge Detection

One form of edge detection with strong links to lateral inhibition is based upon the application of the Laplacian operator on an image which has been convolved with a gaussian kernel [5, 11] . For every pixel in an image, the application of the the so-called "Mexican Hat" $\nabla^2 G$ function, shown in Figure 3(b), computes a weighted intensity difference between the central and neighbouring regions. This operator is insensitive to absolute levels of intensity with large spatial extend and emphasizes local variations which may be conceptually associated with edges [18]. As shown in Figure 3, the location of the detected edge is determined by the position of the zero-crossing of the processed image. A proper selection of connection weights in a structure of lateral inhibition can yield a suitable mean-square approximation of the desired $\nabla^2 G$ function.

Simplification of the Problem

A critical aspect of a 2D VLSI implementation of the approach described above, which would integrate photosentive cells *and* processing on the same surface, relates to the interplay between sufficient spatial resolution of the pixel array and silicon area available for supporting a functional algorithm for edge detection. In other words, there is a trade-off to be exercised between the respective fractions of the overall area devoted to photo transduction and to actual processing, since each pixel of the receptive plane must be accompanied by its own small processor on the silicon die. Therefore the lateral inhibition principle must be simplified and approximated in order to reduce the area used by the physical devices necessary to implement the algorithm.

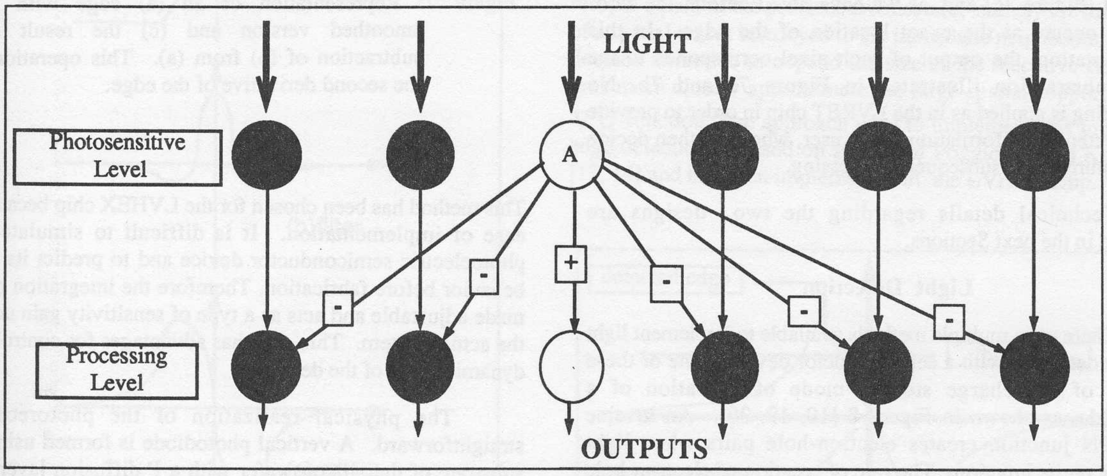

Figure 1: This diagram is a representation of the lateral inhibition in the retina. Only a one-dimensional array is shown here. In reality the cells are also connected to other cells that are over and under the plane of this sheet. This view is a cut through the retina plane. The incoming light falling on cell A produces a positive reaction in the corresponding cell of the processing level and a negative reaction in the surrounding cells. The negative contribution decreases when we get farther from the center. Many connections are not shown for clarity of the diagram.

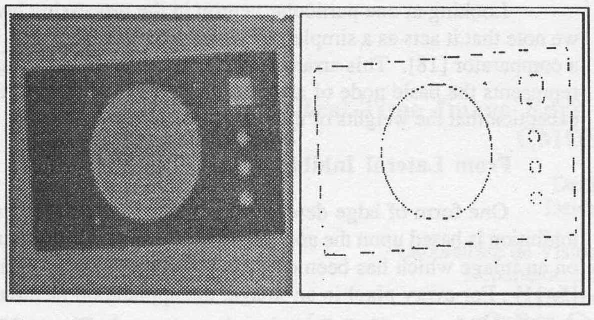

Figure 5: Computer simulation of the decisional algorithm based on the lateral inhibition principle. The left image is the input and the right one is the output.

authors have proposed an analog model of computation in electrical networks for vision problems that can be characterized as mathematically ill-posed problems. Such problems can be solved by applying variational principles and Poggio and Koch have shown that there exists a direct and natural mapping between these and analog networks. In the case of the early vision process of edge detection the strategy of a large resistive network is able to solve the problem [7]. With a distributed resistive network covering the entire image, all pixels become functionally interconnected with a specific relationship between attenuation and distance. It is thus possible to effectively implement receptive fields with suitable on-center and off-surround responses while maintaining the area devoted to parallel connections to a small fraction of the available silicon.

The principle applied in the LVHEX chip is the following: when a discontinuity is applied to the hexagonal, two dimensional, resistive network of Figure 6, the signal is spatially smoothed as in Figure 7b, where a one dimensional case is shown. The subtraction of the two signals, direct and smoothed, yields an approximation of the second derivative of the input (Figure 7c) and, as we have noted before, the zero-crossing occurs at the exact location of the edge. In this implementation, the output of each pixel corresponds to the direct substraction illustrated in Figure 7a and 7b. No thresholding is applied as in the LVRET chip in order to provide a more extensive information to the user, who may then decide on the nature of the subsequent processing.

Technical details regarding the two designs are presented in the next Sections.

Light Detection

There are a multiple methods available to implement light intensity detection with a semiconductor device. One of these consists of the charge storage mode of operation of a photodiode, as shown in Figure 8 [10, 19, 20]. An inverse biased PN junction creates electron-hole pairs when light impinges on its junction. The rate of creation of electron-hole pairs is a function of light intensity and generates a weak photocurrent. This signal is difficult to detect because of its small size, and is best measured through charge accumulation in a capacitor during a given, well chosen, integration interval.

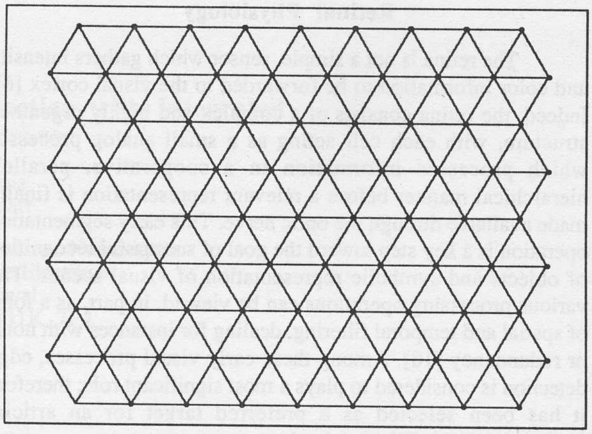

Figure 6: Resistive network with an hexagonal tessellation, the LVHEX chip.

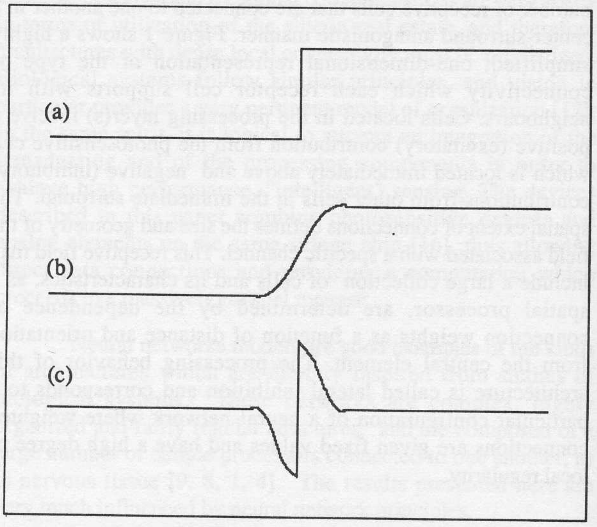

Figure 7: Representation of an (a) edge with (b) its smoothed version and (c) the result of the subtraction of (b) from (a). This operation gives the second derivative of the edge.

This method has been chosen for the LVHEX chip because of its ease of implementation. It is difficult to simulate such a photoelectric semiconductor device and to predict its detailed behavior before fabrication. Therefore the integration period is made adjustable and acts as a type of sensitivity gain control in the actual system. This also has advantages for controlling the dynamic range of the device.

The physical realization of the photoreceptor is straightforward. A vertical photodiode is formed using the N substrate of the silicon wafer with a P diffusion layer on top, two basic materials of the CMOS process used. Figure 9 shows this configuration. The substrate is biased at a voltage of 5 volts in order to polarize the diode in the backward direction and thus produce the small backward photocurrent as an output.

With such constraints in sight, the choice of analog circuitry becomes fairly obvious since a digital implementation of the system and the digital coding of information would involve much larger silicon real estate. Furthermore, it is possible to achieve analog devices with wide bandwidth and therefore realize the level of performance corresponding to real-time applications.

In these early experiments on integrated sensors, we have exploited two simplifications of the general principle of lateral inhibition which lead to a simple implementation. These have been incorporated into the LVRET chip.

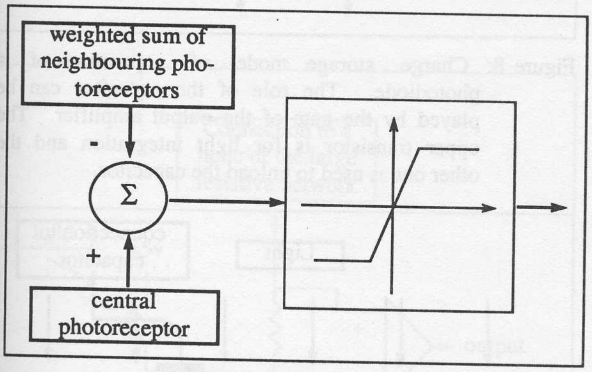

Figure 2: Representation of a single neuron of a simplified retina.

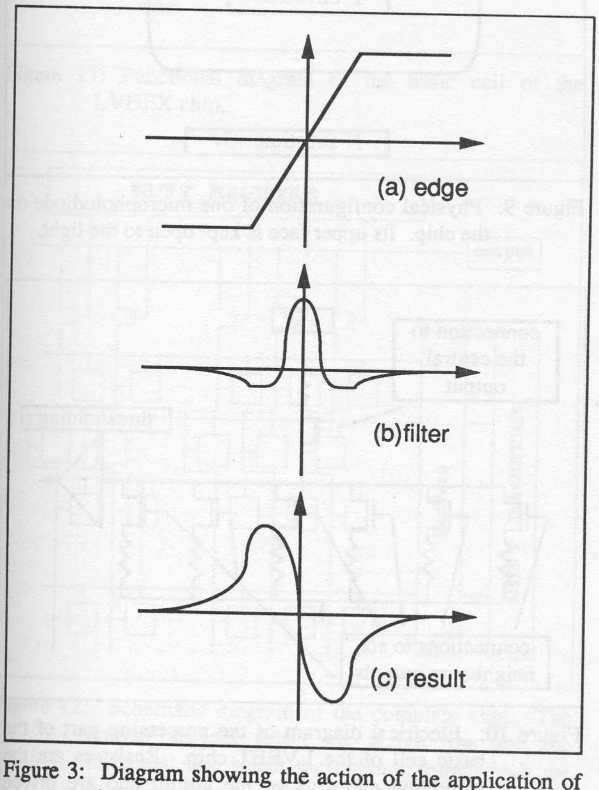

Figure 3: Diagram showing the action of the application of the "Mexican Hat" to a conventional edge.

First, we must select the size of the receptive field and the number of neighbouring cells that contribute to the response of a given cell. We have opted for an hexagonal tessellation with lateral connections to the six immediate neighbours. This choice was motivated by two major reasons. The six nearest neighbours, on an hexagonal grid, are all at the same distance from the central cell and hence all connection weights are equal. This fact reduces the complexity of each cell and provides for efficient modularity, key qualities are in dense VLSI designs. Furthermore, the need for (long) connections from more distant cells is eliminated, thereby saving significant area on the chip.

The second simplification aims at reducing the complexity of the processing cell. The detection of a pure zero-crossing is not easy to achieve in a microelectronic analog circuit. To avoid this difficulty, the circuit is designed instead to perform a simple thresholding operation on the result of the $\nabla^2 G$ function, as shown in Figure 4. This approximation does reduce the efficiency of the algorithm but, at this early stage, has been considered sufficient for our purpose, at least for images with good signal to noise ratios.

With these design choices, the processing algorithm may be described as follows:

-IF more than a given number of neighbours receive an intensity level higher than a given level;

-THEN the middle cell yields a NON-EDGE output;

-ELSE the middle cell yields an EDGE output.

This algorithm has been tested and a typical computer simulation is shown in Figure 5.

From Resistive Networks to Edge Detection

The motivation to implement a second model of integrated sensing arises from the unsatisfactory size of the receptive field of the previous architecture, namely the limitation of a spatial extent restricyed to six immediate neighbours. It was deemed attractive to expand the size of the receptive field and thus more closely approximate the large spatial support of the $\nabla^2 G$ function. The approach used to implement this improved edge detection is based on studies of Poggio, T. and Koch, C. [13, 14] and has been implemented in the LVHEX chip. These

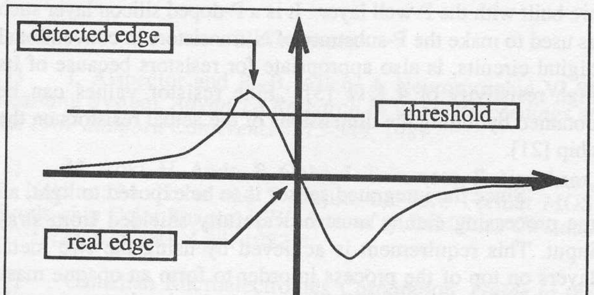

Figure 4: Approximation of the edge using a thresholding effect instead of the zero-crossing detection.

Cell Design and Implementation

Based on the previous simplifications and concepts, the chips can now be designed in order to be implemented on a silicon wafer. For the two chips, a central objective is to synthesize a unique cell that will be duplicated in as many incarnations as required to create the large pixel matrix on the single chip. The two main cells are designed with a rectangular shape and with a precise side to side ratio of 1.155 in order to realize the hexagonal tessellation. Furthermore, all connections to the neighbours are brought to the edges of the cells, thus allowing direct connections by simple placement of cells, side by side. For this purpose, the cells must include the following elements: one microphotosensor open to the light, the analog processing circuit, connecting paths to neighbours, global communication and power.

In the first chip, LVRET, the processing subsystem of the cell is an analog implementation of the simple decisional algorithm discussed before. The contribution from each neighbour is provided through a resistor which weights the input, and the resulting analog voltage is sensed via a thresholding comparator. Figure 10 shows that neighbouring cells do not have to drive resistors directly. Instead they drive them through transistor gates [4]. This configuration is very useful since without it, the photodiodes would have to be able to drive resistors. This would necessitate the addition of a few buffers in the cell in order to get larger fan-out, and would waste considerable area on the chip.

In the second chip, LVHEX, each node of the hexagonal resistive network is connected to a circuit which subtracts the smoothed input from the raw signal. Figure 11 shows a functional diagram of this circuit where phototransistors have been used instead of photodiodes. Advantages include a faster response to light excitation and an economy of silicon area by eliminating control for the photodiodes.

All transistor dimensions and resistor values have been adjusted with the help of SPICE [17], a software for semiconductor circuits simulation. This made possible a precise modelling of the circuit, using a complete list of parameters defining the properties and electrical behavior of the CMOS technology used, in this case, the CMOS3-DLM [3] process of Northern Telecom. This foundry provides a 3 microns process with two metal levels and one polysilicon level.

The resistors of the analog processors in the basic cells are built with the P-well layer. It is a P-doped silicon layer such as used to make the P-substrate of N-transistors in conventional digital circuits, is also appropriate for resistors because of its high resistivity of 4 KΩ [3]. Fine resistor values can be obtained by tuning the dimensions of the actual resistors on the chip [21].

Since the integrated sensor is to be exposed to light, all the processing cicuits must be carefully shielded from stray input. This requirement is achieved by using the two metal layers on top of the process in order to form an opaque mask over the entire chip. Metallic communication lines are made wide enough to hide all underlying circuitry. The entire design was laid out using the ELECTRIC Layout Tool [15].

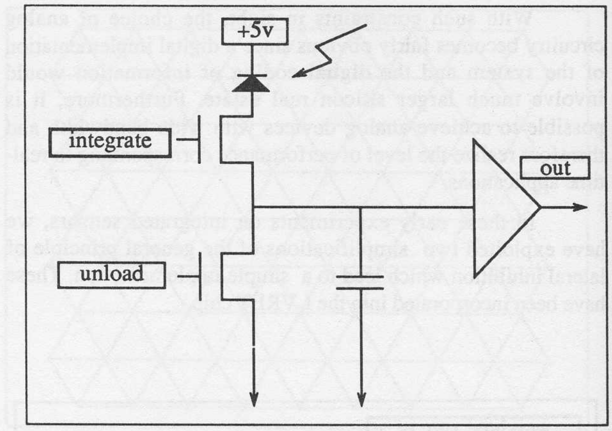

Figure 8: Charge storage mode of operation of a photodiode. The role of the capacitor can be played by the gate of the output amplifier. The upper transistor is for light integration and the other one is used to unload the capacitor.

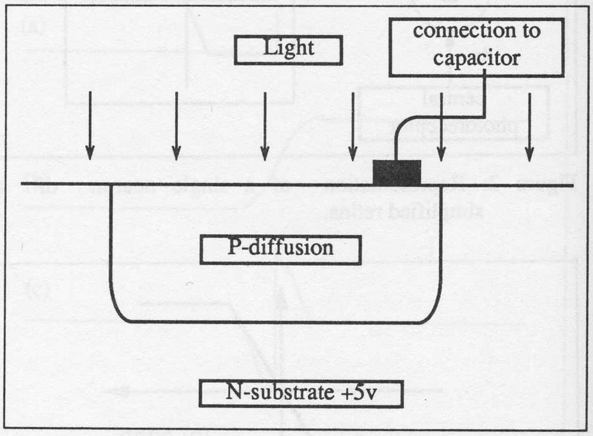

Figure 9: Physical configuration of one microphotodiode on the chip. Its upper face is kept open to the light.

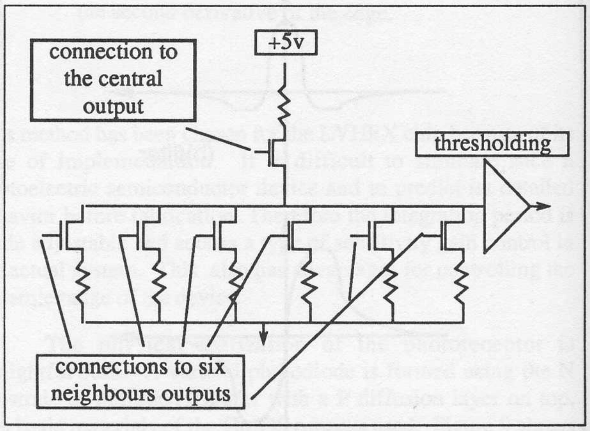

Figure 10: Electrical diagram of the processing part of the basic cell of the LVRET chip. Resistors are the weighting elements of the circuit and are driven by the neighbouring cells. The output amplifier executes the thresholding function.

Data Read-Out and Scanning

The chips implement "large" arrays of basic cells, 28 x 36 for the LVRET chip and 27 x 35 for the LVHEX chip. Each array is read out using an horizontal and a vertical scanner [2]. These scanners consist of two shift registers in which a single bit circulate to activate rows and columns sequentially, with the entire array being scanned during a complete cycle. As shown in Figure 12, the shift registers select the pixel that is read out at any given time. A complete cycle of the LVRET system includes all of the following steps. First, all the capacitors of the microphotodiodes are unloaded simultaneously with a single bit command. Next the target image is recorded by enabling all the photodiodes during a period of time that is adjusted depending on the average intensity. This variable integration interval is helpful as a sensitivity gain control. The adjustment is performed by an external circuitry, which also generates timing and synchronization information for the sensor. The final step of the cycle consists of scanning the entire array as explained earlier and feeding the information to the subsequent system. Following this sequence, the cycle restarts and is repeated as needed. In the case of the LVHEX chip, the initialization step is not present since the phototransistors give an virtually instantaneous output. Scanning simply proceeds continously.

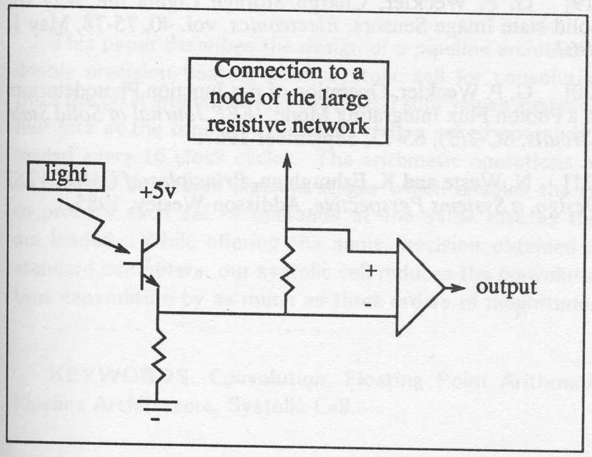

Figure 11: Functional diagram of the basic cell of the LVHEX chip.

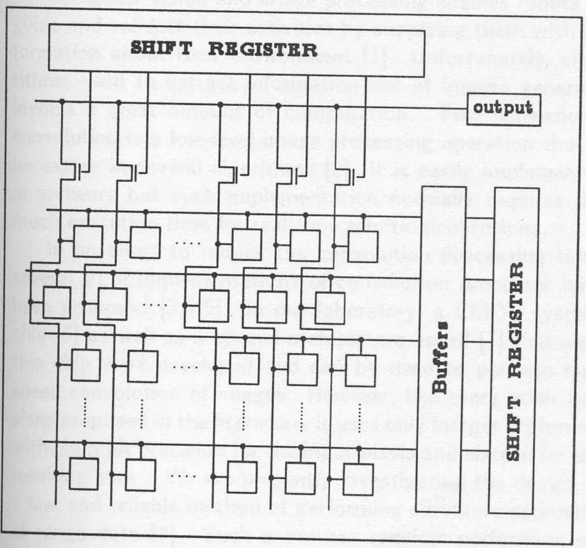

Figure 12: Schematic diagram of the complete chip. The vertical scanner select the active row and the horizontal one select the column from which the output is taken. The diagram also shows the way we deal with the hexagonal tessellation to scan the entire plane.

Conclusions

The integrated optical sensors described in this paper include analog, digital and photosensitive devices implemented on the same silicon substrate. The overall design philosophy is inspired by biological architectures and exploit local connectivity and parallel processing. The CMOS3-DLM process used is neither specialized in analog nor in photosensitive devices. It has been selected because of its availability and has provided encouraging initial results. To be utilized, these sensors have to be mounted within an an optical system and suitable electronic interface must be provided for control, timing, and display of the processed image. Early prototypes are being evaluated for special purpose computer vision cameras.

Although the resolution of the order of 30x30 pixels may appear somewhat limited, it has provided an interesting test bench toward the development of optical sensors with built-in processing capabilities. Silicon foudries now available are capable of increasing this density quite significantly. Although this paper has been limited to the support of (some) edge detection algorithms, there exist numerous other applications, such as sub-pixel interpolation, region connectivity, or motion detection. Further work is currently in progress in our laboratory with the objective of integrating sensing and early processing in various applications of 1D, 2D, and 3D imaging.

Acknowledgments

The authors would like to thank Mr. C. Scaffidi for reviewing the manuscript and Dr. D. Laurendeau for his helpful advice. We also wish to extend special thanks the Canadian Microelectronics Corporation and Northern Telecom who have provided the ressources required for VLSI implementation. This work was supported, in part by NSERC grant A1544, FCAR grant EQ2830 and by an NSERC fellowship to DB.

References

[1] J. Alspector and R. B. Allen, A Neuromorphic VLSI Learning System, *Advanced Research in VLSI, Proceedings of the 1987 Stanford Conference*, 313-349, 1987.

[2] M. Aoki, H. Ando, S. Ohba, I. Takemoto, S. Nagahara, T. Nakano, M. Kubo and T. Fujita, 2/3-inch Format MOS Single-chip Color Imager, *IEEE Journal of Solid State Circuits, SC-17(2)*, 375-380, April 1982.

[3] Canadian Microelectronics Corporation, *Guide to the Integrated Circuit Implementation Services of the CMC*, version 3.0, January 1987.

[4] H. P. Graf and P. deVegvar, A CMOS Implementation of a Neural Network Model, *Advanced Research in VLSI, Proceedings of the 1987 Stanford Conference*, 351-367, 1987.

[5] E. C. Hildreth, Edge Detection, MIT AI Memo No. 858 Sept. 1985.

[6] A. Kaneko, Physiology of the Retina, *Ann. Rev. Neurosci.*, 2:169-191, 1979.

[7] C. Koch, J. Luo and C. Mead, Computing Motion Using Analog and Binary Resistive Networks, *IEEE Computer*, 52-63, March 1988.

[8] C. Koch, J. Marroquin and A. Yuille, Analog "Neuronal" Networks in Early Vision, *Proc. Natl. Acad. Sci. USA*, Vol. 83, 4263-4267, June 1986.

[9] R. P. Lippmann, An Introduction to Computing with Neural Nets, *IEEE ASSP Magazine*, 4-22, April 1987.

[10] W. F. List, Solid-State Imaging Methods of Approach, *IEEE Trans. Electron Devices, ED-15(4)*, :256-261, April 1968.

[11] D. Marr and E. C. Hildreth, Theory of Edge Detection, *Proc. R. Soc. London B*, 207:187-217, 1980.

[12] C. Mead, Silicon Models of Neural Computation, *IEEE First International Conference on Neural Networks, Vol 1*, 91-106, June 1987.

[13] T. Poggio and C. Koch, An Analog Model of Computation for the Ill-posed Problems of Early Vision, A.I. Lab. Memo No. 783, MIT, Cambridge, May 1984.

[14] T. Poggio and C. Koch, Ill-posed Problems in Early Vision: from Computational Theory to Analogue Networks, *Proc. R. Soc. London B 226*, 303-323, 1985.

[15] S. M. Rubin, *ELECTRIC 3.40A User's Guide*, Schlumberger Palo Alto Research, October 1987.

[16] M. Sivilotti, M. Mahowald and C. Mead, Real-time Visual Computations Using Analog CMOS Processing Arrays, *Advanced Research in VLSI, Proceedings of the 1987 Stanford Conference*, 295-312, 1987.

[17] SPICE Version 2G User's Guide, Univ. of California, Berkeley, 1981.

[18] M. V. Srinivasan, S. B. Laughlin and A. Dubs, Predictive Coding: a Fresh View of Inhibition in the Retina, *Proc. R. Soc. London B*, 216:427-459, 1982.

[19] G. P. Weckler, Charge Storage Lights the Way for Solid-state Image Sensors, *Electronics*, vol. 40, 75-78, May 1, 1967.

[20] G. P. Weckler, Operation of p-n Junction Photodetectors in a Photon Flux Integrating Mode, *IEEE Journal of Solid State Circuits, SC-2(3)*, 65-73, September 1967.

[21] N. Weste and K. Eshraghian, *Principles of CMOS VLSI Design, a Systems Perspective*, Addisson-Wesley, 1985.