Design of Integrated Circuits and Systems, DCIS'96, Sitges, Spain, November 1996

Design of a foveated log-polar image sensor in standard CMOS technology

F. Pardo J.A. Boluda J.J. Pérez S. Felici
B. Dierickx†

Instituto de Robótica - Universidad de Valencia
C/ Hugo de Moncada, 4 Entlo. 46010-Valencia, Spain
Email: Fernando.Pardo@uv.es

† IMEC
Kapeldreef 75, B-3001 Leuven, Belgium
Email: Dierickx@imec.be

Abstract

This paper deals with the design and implementation of a CMOS foveated, or retinal, image sensor. The sensor performs a spatial image transformation from cartesian to log-polar coordinates. The image sensor geometry is very similar to the human eye, where the receptive fields are concentrated in the center of the sensor, and the resolution decreases toward the periphery. Though the layout of the sensor and the human retina are very similar, the aim is not to design biological sensors, but to design a sensor topology with interesting features for mobile robotics and machine vision applications. The sensor has been made using standard 0.7 μm CMOS process. The non-squared *shape of the elements in the chip, makes the design approach different from other ASICs using the same standard CMOS process.*

1 Introduction

Information organization plays an important role in environment image perception. One of these organizations of the information is the space-variant representation, that has interesting properties [1]: image data is reduced in a selective way, maintaining accurate resolution in the foveal area, or center, while globally minimizing the amount of information necessary to describe the image. One example of this space-variant representation can be found in the eye of many animals including human beings. In these eyes there is a central zone, called fovea, where the concentration of photoreceptive fields is higher; around the fovea there is another area, called retina, where the concentration of pixels decreases toward the periphery. With this special representation there is high resolution in the interesting parts of the scene, usually in the center, while still keeping a wide view field, and reducing the amount of data for the image.

Though the concept of space-variant vision is not new, only recently with the realization of retinal, or foveated, sensors, are vision techniques based on this representation becoming feasible for practical applications [2]. Chip sensors speed-up the image processing, performing a log-polar transformation in real-time, on the chip itself, in the very instant of acquisition.

The main technological process for image sensors is the CCD process (Charge Coupled Device). The reason is the high image quality obtained from these kind of sensors. Quality for humans is a subjective concept while for systems or robots, quality is more objective and it is possible to say that, for a robot, a good image quality is that one that allows the system to correctly perform the tasks programmed. From this point of view, the CCD process has not to be necessarily the best technology for robots, industry, or systems [3]. The standard CMOS process, usually employed in ASIC design, has proved its capability to design also image sensors, with good quality even for human observers [4].

Having in mind the interesting properties of the space-variant sensing (specially performing the log-polar transform) and the better characteristics of the CMOS process for this specific application, a new foveated sensor based on CMOS technology has been designed and fabricated.

2 Sensor structure

There are two precedents to the CMOS sensor presented here. The first approach to a foveated retinal sensor was a retinal CCD [5]. Though CCD is the best technology for image sensing, it has been proved from that CCD chip [6], that the CCD process is not at all the best process for space-variant resolution cameras. Problems arise from charge scaling, and specially, from the read out of charges that is made through CCD elements instead of simple metal wires. The second approach is a recently developed CMOS foveated sensor designed by Wodnicki [7] at the McGill University in Canada. There are many differences between the Wodnicki's CMOS sensor and the CMOS sensor presented in this paper. The main differences are the resolution (lower in the precedent sensor), fovea structure, and cell type.

Following the same structure of these kind of sensors, there are two areas in the focal plane. One zone is the fovea or center of the sensor where the resolution is higher, and the other zone is the retina, where the pixel distribution follows the log-polar transformation. The equation (1) defines the log-polar transformation performed in the chip, and the figure 1 shows the retinal or image plane and the transformation to the cortical or computation plane[1].

$$\begin{cases} \xi = \log r = \log\left(\sqrt{x^2+y^2}\right) \\ \gamma = \theta = \tan^{-1}\left(\frac{y}{x}\right) \end{cases} \tag{1}$$

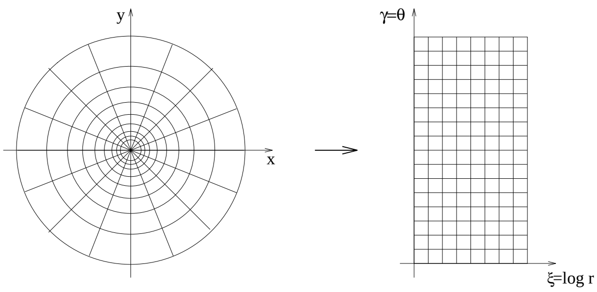

Figure 1: *Graphic representation of the log-polar transformation*

The retina of the sensor is formed by 56 circles everyone containing 128 pixels, giving a retinal resolution of 56x128 pixels. The retinas designed to date had less resolution, they were 30x64 for the CCD sensor, and 16x64 for the other foveated CMOS sensor. The final structure of the retina can be seen in the figure 2, that is a microphotograph of the fabricated image sensor.

The fovea of this sensor is one of the main differences respect the other two foveated sensors [8]. While the structure of other foveas were a simple square matrix inside the foveal circle, in the new sensor, the fovea follows a similar structure like the retina. The fovea structure is formed by circles with decreasing number of pixels approaching the center. There are a total 20 rings in the fovea, the outer 10 rings has 64 pixels each, the following 5 have 32 pixels, the following 2 have 16, and the last three have 8, 4 and 1 respectivelly. This makes a total of 845 pixels for the fovea, that jointly with the pixels in the retina makes a total of 8,013 pixels for the whole sensor. The radius of the rings in the fovea decreases linearly instead of exponentially as in the retina. This structure is shown in the figure 3. One of the problems that other foveas had was the image processing in the edge between the retina and fovea, since it is difficult to match a square distribution of pixels with a circular one, specially trying to find pixel neighbors. Another problem is the space between a squared fovea and a circular retina not covered with pixels. With the new fovea, there is a continuous transition between the fovea and retina, and there is no area not covered by pixels. It makes easier the image processing since there is no problem to determine the neighbors of a pixel in a pseudo-cortical plane that includes both, the retina and the fovea.

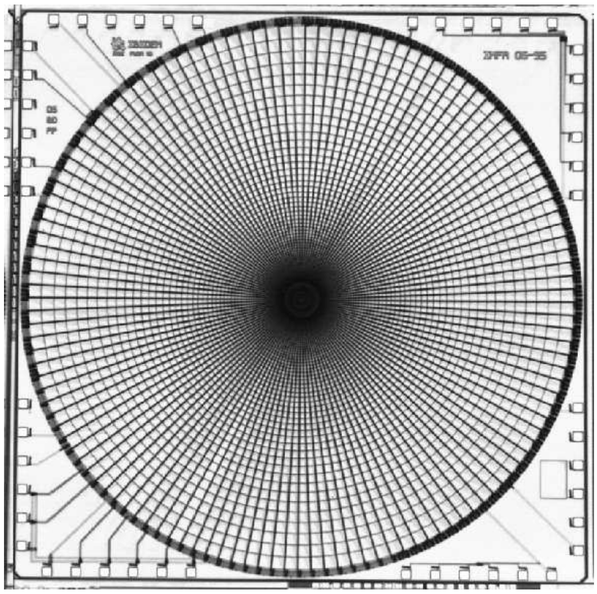

Figure 2: *Sensor retina*

The retinal diameter is about 8.5 mm. while the fovea is around 285 μm. The chip was designed at IMEC, and has been fabricated using the Mietec 0.7 μm CMOS process.

3 Basic sensing cell

There are many possibilities, using the CMOS process, of converting light in electrical information. As

[1]The words *retinal* and *cortical* come from the observation that the log-polar transformation is very similar to the retinal image transformation in the human eye cortex.

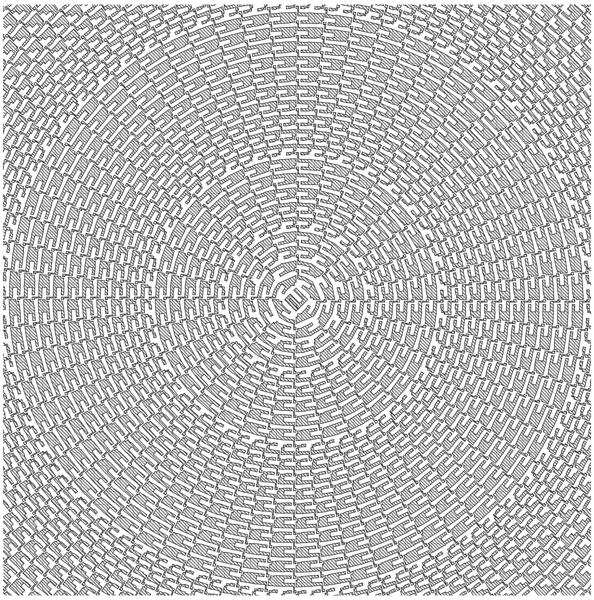

Figure 3: *Sensor fovea*

a first approach, it is possible to make two groups of sensors. The first group is formed by those sensors [9] where the small light generated current in the inversely polarized photodiode, charges or discharges a capacitor during a fixed time. The voltage of this capacitor after the integration time is proportional to the incident light intensity and the fixed integration time. This method is the most appropriate for CMOS image sensors, and it offers good image quality, specially due to the amplification factor of the integration time.

The second method, less common, is to directly measure the current produced by light in the photodiode [10]. This current is usually converted to voltage through a MOS transistor. This transformation follows a logarithmic function, thus the response of these kind of cells, though noisier due to the low current generated by light, is logarithmic with still enough quality for most applications.

For the retinal sensor the second kind of sensing cell was preferred [11]. The reasons were the better acknowledge of these kind of sensors and specially, the logarithmic response of the basic cell, that gives to the system a more "zoological" appearance, apart from the better signal scaling of the logarithmic response. The basic sensing cell is shown in the figure 4. This cell works as follows [12]: the current generated by light in the diode flows through the transistor M1. This current is very small (between pico and nano amperes) biasing the transistor in its weak inversion region. The second transistor M2 is just a source follower. In the weak inversion region, the simplified expression for the voltage between gate and source is:

$$V_{gs} = \frac{kT}{q} ln(\frac{L}{W}\frac{I_d}{I_{do}}) \qquad (2)$$

where V_{gs} is the gate-source voltage, I_d is the drain current, W and L are the width and length of the transistor channel, T is the temperature, and k, q and I_{do} are constants. From this expression it is very simple to extract conclusions about the cell response. The most important conclusion is the logarithmic dependency of the voltage with the current. It directly means that the response (voltage) is logarithmic with the incident light intensity (current). The second conclusion is that the slope or gain of this logarithmic response, is a constant ($\frac{kT}{q}$) totally independent of the technology, cell size, design, etc. it only depends on the temperature, but this is a small problem since the temperature is usually common for the sensor chip and most of image processing algorithms compute relative instead of absolute values.

The only dependency with cell size is the ($\frac{L}{W}$) factor inside of the logarithmic term. This factor is specially important in space-variant sensors since it can be utilized to compensate the different pixel size in every circle. The current generated by light is directly proportional to the photodiode area. To compensate the increasing area there is only one parameter to change, and it is the width of the transistor since the length is kept minimum and cannot be reduced. The area increases quadratically (two-dimensional) while the width increases linearly (one-dimensional), it means that it is not possible to perform a complete compensation but just half.

The result, of this lack of compensation, is a different offset for every circle, offset that is not very high taking into account that it is produced inside of the logarithmic expression. This different response of the cells needs an offset compensation circuit outside the chip. Depending on the contrast of the image, this compensation can be performed by software, just subtracting the fixed image when the original image has low contrast, or by hardware [13] doing this subtraction with an operational amplifier for high contrast images. For the hardware solution, a memory is necessary to store the fixed image to be subtracted, and a digital to analog converter to feed the OA.

4 Layout

Most of CAD tools for ASIC design are able to handle cell libraries, automatic place and route, orthogonal designs, and many other things specially

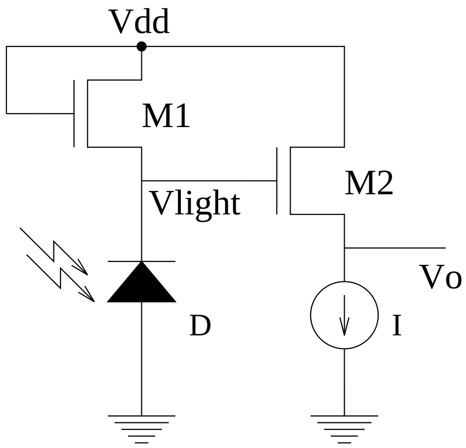

Figure 4: *Basic sensing cell of the sensor*

addressed to digital and generally rectangular designs. When the layout involves non orthogonal shapes, the design becomes very complicated due to the lack of automation of current CAD tools [14]. Almost every layer of the design has to be made by hand. Nevertheless, there are some tools that have an internal specification language where different shapes can be defined, giving the possibility of designing a complete layout from a pseudo description language.

For a design like the retinal sensor, where there is no orthogonal shapes in a manhattan disposition, it was necessary to use one of these languages to define the complete layout of the sensor. The design was made using the Cadence Design Framework. Cadence has a command language, called SKILL, that has not been thought specifically for layout design although it can successfully be utilized for this purpose. This language offers hundreds of functions giving a very large programming flexibility, although only a few of them have been employed on the design, basically those for program flow control as loops, and others to just write polygons that have been the main "brick" to build the complete layout. Some mathematical functions were employed to calculate the layout for the log-polar transformation. The result has been a program that generates the layout of a retinal sensor as specified by the log-polar transformation.

Design rule checking is always desirable in any full custom design, nevertheless, it is not easy to find DRC tools able to handle non orthogonal designs. Usually, common DRC tools are used to check orthogonal designs, or even with 45^o slopes, but most of them definitely fail when different slopes are present in the layout. As an example, the retina layout produces thousands of errors when a normal design rule checker is run, while measuring it by hand it is possible to see that everything is correct. In any case, it is always very risky to visually check the layout, although it has been the method employed in this special design. At least for this chip, this method worked out since the resulting sensor seems to work as expected and it has not been noticed any layout problem.

5 Experimental results

The sensor, after fabrication, showed the expected behavior. The obtained image has enough quality for most image processing applications. Only one problem has to be solved, and it is the Fixed Patter Noise (FPN) that is very high in this sensor. This FPN can be eliminated subtracting it from the image.

This FPN is mainly due to the mismatching properties of MOS transistors and there are several sources for this non uniformity [15]. The first source are the technological parameters, specially the thin oxide and substrate doping. In this design the geometrical parameters are also specially relevant since the length of the weak inversion transistor has been kept minimum. The length non uniformity, apart from the expected technological variation, is due to the non orthogonal shape of the channel. Since the polygon vertices have to be snapped to a squared grid it becomes a source of extra non uniformity.

There are other sources for the FPN in the sensor. The variation among radii is due to the offset non uniformity of the different output amplifiers (one per each radio). The non uniformity among rings is due to the half scaling of the signal as explained in the basic sensing cell section. This variation among rings is partially compensated by the narrow channel effect that has opposite sign though it only compensates part of the ring non uniformities.

A normal image obtained from the sensor after subtracting the FPN can be see in the figure 5.

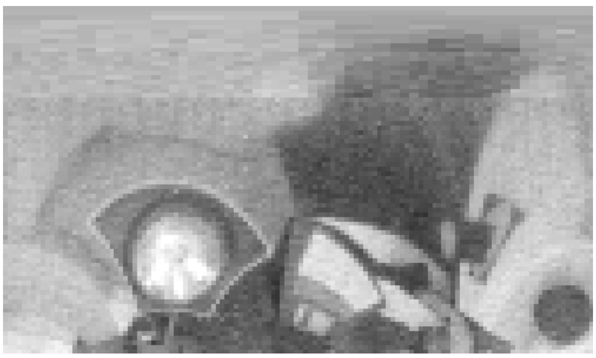

Figure 5: *Cortical image obtained from the camera*

This picture corresponds to the cortical or computation representation of the image. For the computer, this representation is very interesting for the mathematical properties it has, but for the human eye it is useless since it is a complex transformation of the reality. It is possible to transform this image obtaining the original scene. This retransformation can be seen in the figure 6.

Figure 6: *Original scene after transformation*

The acquisition rate can ba as fast as 100 or 200 images per second, depending on the image quality. The pixel access is not fast (between 500 and 1,000 nanoseconds), but since one image is only around 8,000 pixels, it is possible to read complete images at very high rates. This rates can be even increased since the access is random, thus there is no need to read complete images but only part of them.

The response is exactly logarithmic as expected. The slope is also very constant for most of the rings, although there is a difference between the inner and outer rings; the slope is really constant for rings above the fourth starting to decrease for rings lower than the fourth. The maximum difference is around a five percent. The reason for this difference is the influence of the leakage current of the photodiode that for the outer rings is negligible but for the inner rings changes a little the response slope, that anyway is always logarithmic.

6 Conclusion

A retinal visual sensor has been designed. The structure of the retina and fovea, and the resolution of the sensor, are better than the predecessors.

The FPN (Fixed Pattern Noise) in the response is a problem inherent to these space-variant sensors, even in the CCD version there were non uniformities that had to be fixed doing post normalization of data. In the current CMOS design there are three types of non uniformities. One is a random non uniformity along the chip due to CMOS transistors matching properties, non uniformities of the oxide thickness, doping concentration and also geometry non uniformities in the channel width and length. The second is a qualitatively predictable non uniformity due to the narrow channel effect [16], specially for the first rows of the retina and fovea where the width is very narrow. The last non uniformity, which value is very easy to calculate, is due to the different scaling of the area of the pixel and the MOSFET channel width. These last two non uniformities arise only in space-variant sensors, like the retina, where the pixel size is not constant. For devices where all pixels are exactly the same (say a normal CMOS camera sensor) only the first non uniformity is present. Other non uniformity is due to the use of individual amplifiers for each radio than can be solved using a single output amplifier.

The FPN can be easily solved subtracting it from the image, so the sensor can successfully be employed in almost any image processing task where a log-polar transformation, or also selective image data reduction, are convenient.

7 Acknowledgements

The fabrication of the chip has been funded by the IBIDEM project of the European Community TIDE program. The design of the sensor has been funded in part by the EC Human Capital and Mobility program.

References

[1] G.D. Hager. Using resource-bounded sensing in telerobotics. In *Fifth Int. Conference on Advanced Robotics*, Pisa, Italy, June 1991.

[2] M. Tistarelli and G. Sandini. Dynamic aspects in active vision. *CVGIP: Image Understanding*, 56 No.1:108–129, 1992.

[3] F. Pardo, F. Vegara, J.A. Boluda, and S. Felici. Sensores CMOS para robótica e industria: Sensor retínico espacio variante y visión activa. In *Cuarto Congreso AER-ATP*, Zaragoza, Spain, October 1995.

[4] D. Rensaw, P.B. Denyer, G. Wang, and M. Lu. ASIC vision. In *IEEE Custom Integrated Circuits Conference*, pages 7.3.1–7.3.4, 1990.

[5] I. Debusschere, E. Bronckaers, C. Claeys, G. Kreider, J. Van der Spiegel, G. Sandini, P. Dario, F. Fantini, B. Bellutti, and

G. Soncini. A retinal CCD sensor for fast 2D shape recognition and tracking. *Sensors and Actuators*, pages 456–460, 1990.

[6] F. Pardo and E. Martinuzzi. Hardware environment for a retinal CCD visual sensor. In *EU-HCM SMART Workshop: Semi-autonomous Monitoring and Robotics Technologies*, Ispra, Italy, April 1994.

[7] Robert Wodnicki, Gordon W. Roberts, and Martin D. Levine. A foveated image sensor in standard CMOS tecnology. In *Custom Integrated Circuits Conference*, Santa Clara, California, May 1995.

[8] D. Scheffer, B. Dierickx, J. Vlummens, and F. Pardo. Log-polar image sensor in CMOS technology. In *European Symposium on Lasers, Optics, and Vision for Productivity in Manufacturing*, Besancon, France, June 1996.

[9] Christer Jansson, Per Ingelhag, Christer Svensson, and Robert Forchheimer. An addressable 256x256 photodiode image sensor array with an 8-bit digital output. *Analog Integrated Circuits and Signal Proccessing*, pages 37–49, 1993.

[10] Nico Ricquier and Bart Dierickx. Addressable imager with a logarithmic response for machine vision. In *ISIIROM*. Brussels, 1993.

[11] Fernando Pardo. Development of a retinal image sensor based on CMOS technology. Technical Report LIRA-TR 6/94, University of Genoa, Dept. of Computing, LIRA Lab, October 1994. http://carpanta.eleinf.uv.es.

[12] Nico Ricquier and Bart Dierickx. Pixel structure with logarithmic response for intelligent and flexible imager architectures. *Microelectronic Engineering*, 19:631–634, 1992.

[13] J.A. Boluda, F. Pardo, T. Kayser, J.J. Pérez, and J. Pelechano. A new foveated space-variant camera for robotic applications. In *IEEE, International Conference on Electronics Circuits And Systems, ICECS'96*, Rodos, Greece, October 1996.

[14] F. Pardo, J.A. Boluda, B. Dierickx, and D. Scheffer. Design issues on CMOS space-variant image sensors. In *European Symposium on Advanced Imaging and Network Technologies, AFPAEC'96*, Berlin, Germany, October 1996.

[15] F. Pardo, J.A. Boluda, J.J. Pérez, S. Felici, B. Dierickx, and D. Scheffer. Response properties of a foveated space-variant CMOS image sensor. In *IEEE, International Symposium on Circuits And Systems, ISCAS'96*, Atlanta, USA, May 1996.

[16] Kelvin Kuey-Lung Hsueh, Julian J. Sanchez, Thomas A. Demassa, and Lex A. Akers. Inverse-narrow-width effects and small-geometry MOSFET threshold voltage model. *IEEE Transactions on Electron Devices*, 35(3):325–338, March 1988.