In 'SPIE Conference on Advanced Focal Plane Arrays and Electronic Cameras' AFPAEC'96, Berlin, October 96

Design issues on CMOS space-variant image sensors

F. Pardo, J.A. Boluda, J.J. Pérez
Instituto de Robótica-Univ. de Valencia, C/ Hugo de Moncada 4-E, 46010 Valencia, SPAIN
Fernando.Pardo@uv.es

B. Dierickx, D. Scheffer
IMEC, Kapeldreef 75, 3001 Leuven, BELGIUM
Dierickx@imec.be

ABSTRACT

A new image sensor, using CMOS technology, has been designed and fabricated. The pixel distribution of this sensor follows a log-polar mapping, thus the pixel concentration is maximum at the center reducing the number of pixels toward the periphery, having a resolution of 56 rings with 128 pixels per ring. The design of this kind of sensors has special issues regarding the space-variant nature of the pixel distribution. The main topic is the different pixel size that requires scaling mechanisms to achieve the same output independently of the pixel size. This paper present some study results on the scaling mechanisms of this kind of sensors. A mechanism for current scaling is presented. This mechanism has been studied along with the logarithmic response of these special kind of sensing cells. The chip has been fabricated using standard 0.7 μm CMOS technology.

1 INTRODUCTION

Looking at any image system, it is easy to discover that the sensing element usually has not the cartesian structure that is mainly employed in artificial image processing. The natural common structure of the sensory part of any biological system is space-variant, meaning that image resolution decreases toward the periphery being maximum at the center of the sensor. Million years of evolution have given as result a space-variant structure for the main part of almost any biological visual system. It is easy to demonstrate that this space-variant representation is not capricious since it has, apart from other special mathematical advantages, a very interesting property for image processing, and it is the selective data reduction that allows the system to reduce the data to be processed.[1] Selective data reduction means that image information is reduced for those parts of less interest for the environment perception (usually the periphery of the attention focus), and increased in the interesting parts (usually the center of attention).

Thus it seems very interesting to design a sensor with a space-variant structure. There are many possibilities for a space-variant structure, but it is probably the log-polar mapping the most interesting for image processing,[2] and one of the closest to any biological visual system. Pixels in the image plane, also called retinal plane for its similarities with the human retina, are distributed in concentric rings around the center of the sensor. The distance of each ring, or radius, increases exponentially with the ring number. This image plane is transformed to the computation or cortical plane through the log-polar mapping. This plane is called computation plane since it is where the image processing takes place in; it is also called cortical for its similarities with the representation of a retinal image in the human cortex. The figure 1 shows the log-polar mapping and the equation (1) and (2) shows the mathematical definition of this transformation also called log-spiral mapping.[3]

$$\begin{cases} \xi = \begin{cases} r & r \leq r_o \quad fovea \\ \log r & r > r_o \quad retina \end{cases} \\ \gamma = \theta \end{cases} \tag{1}$$

where

$$\begin{cases} r = \sqrt{x^2 + y^2} \\ \theta = \arctan\left(\frac{y}{x}\right) \end{cases} \tag{2}$$

and r_o is the radius of the fovea circle.

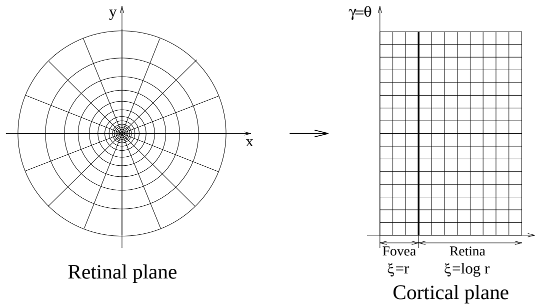

Figure 1: *Graphic representation of the log-polar transform*

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. This high quality is good enough for humans, but one could wander if it is really necessary such a good quality for any artificial vision system. For a robot, quality has to do with the possibility or not of performing a specific task; if the robot can complete the task, the image has good quality, otherwise, the image is not good. From this point of view, the CCD process has not to be necessarily the best technology for robots, industry, or image systems since they are less demanding on image quality. The standard CMOS process, usually employed in ASIC design, has proven its capability to be used also to design image sensors, with good quality even for human observers,[4] although it is a technology that cannot easily reach the quality of conventional CCD cameras.

For space-variant sensing, as it will be shown next, the best technology is not the CCD but the CMOS, since this technology can deal better with space-variant sensing structures and it is specially suited for specific applications, in which a space-variant sensor has the largest field of application. With CMOS is not only possible to design a space-variant structure but it also provides the possibility of an easy interface to a simple processor for embedded applications.

2 SPACE-VARIANT DESIGN

As a first approach, a space-variant sensor has nothing to do with a cartesian constant resolution image sensor. Many problems arise changing the pixel size, pixel distribution, total sensor topology and information read-out. Three important problems will be pointed out in this section. This problems have to do with the polar distribution of pixels, pixel response scaling, and topology of the center part of the sensor.

2.1 Polar layout

The structure shown in the figure 1 is polar and the polygons that define every pixel have not an orthogonal disposition of their lines. Almost any integrated circuit fabrication process expect designs snapped to a cartesian grid and polygon lines in manhattan disposition; there are even some foundries that only allow lines rotated 90 degrees. Thus, the first approach for a space variant sensor can be a sensor with log-polar distribution of pixels but having each pixel a cartesian structure. This approach is not good since the aperture ratio is very low and it present read-out and design problems difficult to solve.

The best solution is then to design polygons with lines following any orientation. This solution has a problem and it is that the polygons still have to be snapped to a fixed cartesian grid that can cause distortion and non uniformities among pixels. This snapping problem is shown in figure 2(a) and 2(b), while the figure 2(c) shows the solution adopted for the retinal CMOS sensor presented in this paper.

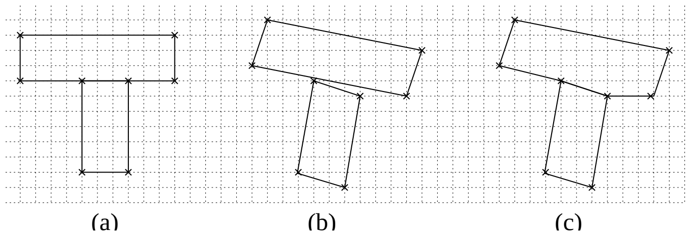

Figure 2: *Polygons adjustment: (a) cartesian (b) rotated (c) adopted solution*

CAD tools for integrated circuit design are usually oriented to cartesian designs since most of ASIC circuits are digital and are based on standard cells that are cartesian polygon structures. These CAD tools allow some automation on the design process adding cell replication, autoplacement, autorouting, etc; all these features do not work when the design is not orthogonal. It means that a different approach has to be taken for the layout drawing and design. The easiest way to solve this problem is the use of a high level description language for the layout generation. Unfortunately it seems that there is no specific language for high level layout description, therefore, a normal language as C or SKILL (language embedded in Cadence CAD tools) has to be employed for the automatic layout generation.

2.2 Signal scaling

Another problem, present in any space-variant image sensor, is the signal scaling. Outer cells yield larger response than inner ones. In the case of integration sensing cells (as CCD for example), where the light produced charges are integrated during a fixed time, the response is directly proportional to the photodiode area. For any sensor following a log-polar mapping, it is easy to calculate the scaling factor, or ratio between the sensing area of the cells in the last ring and those of the first ring, with the expression:

$$S_{max} = \left(1 + \frac{2\pi}{N_c}\right)^{2(N_r-1)} \tag{3}$$

where N_r is the number of pixels per circumference and N_c is the number of circumferences or rings in the sensor.

If the number of circumferences is large, this factor can reach values around 200 and 300, which means that a scaling mechanism is necessary to correct the response differences between inner and outer cells. This scaling mechanism is not easy to implement since the amplification factors, one per circumference, have a big variation (say from 1 to 200). Even providing a scaling mechanism there will be a factor non uniformity among rings that will produce different response slopes depending on the circumference.

It will be shown next that it is possible to use a sensing cell with logarithmic response that minimizes the factor variation and converts it to an additive factor being easier to perform a complete compensation.

2.3 Central zone

Like a biological visual system, the retina sensors have the maximum resolution in the center. This part of the sensor is called fovea since its functionality is similar to the biological eyes. The fovea in a log-polar sensor is always a problem since the logarithmic transformation cannot be accomplished in the center of the sensor; there is a singularity when the radius becomes zero meaning that the pixel concentration per area unit becomes infinitum.

There are some solutions for this fovea depending on the technology process possibilities. A simple solution consist on just putting a conventional cartesian grid in the center of the sensor. This is the solution adopted for the first retinas designed, but it is not a good solution. With this solution there is no continuity between the retina and fovea and there are some blind areas in the transition that are not covered by pixels. The lost of continuity between the retina and fovea makes very difficult the image processing in the center of the sensor, different topologies means different algorithms being difficult to process objects divided in two different spaces. The blind parts are also a problem since the fovea is supposed to have the highest resolution in the sensor.

A proposed solution, adopted for the retinal sensor presented in this paper, consist on keeping the polar pixel distribution modifying the exponential growing law as shown in the equation (1) and figure 1. In this way, there is a continuity between fovea and retina, the image processing takes place in the same space, and there is no blind parts between the retina and the fovea keeping the highest possible resolution that the technological process allows.

3 FIRST SPACE-VARIANT IMAGE SENSORS

The log-polar CMOS sensor presented in this paper is not the first intent to make a space-variant image sensor, although it is probably the first one that offers enough performance and quality to be employed in real applications. There were two sensors before, one was made using CCD technology, and the second employed standard CMOS technology. The problems of these sensors will be discussed.

3.1 The CCD retinal sensor

The first log-polar sensor ever designed was done using CCD technology.[5,6] The design was carried out by IMEC, in Belgium, as part of a collaborative project of many universities and research institutes. The figure 3 shows the structure of this CCD retinal sensor. From this figure it is possible to see one of the main problems of this sensor, the *blind* sector on the image plane. This blind sector, that occupies around 2.5 radial pixels, is directly one consequence of the use of the CCD technology.

The main characteristic of CCD sensors is the readout through CCD devices instead of wires; the signal is formed by charges instead of current. CCD devices are larger than simple wires, and they are silicon structures so they cannot be placed over any other semiconductor structure in the sensor as wires can. The polar structure of the sensor forces a read-out from the center of the sensor, thus there is a need of a CCD radial structure to take out the charges from the inner cells in the sensor. This CCD read-out structure produces the blind sector in the CCD retinal sensor, that cannot be minimized or avoided using other read-out schemes, at least, using standard CCD technology. Therefore, a blind sector is an inherent problem of CCD space-variant sensors.[7]

The CCD sensor has a retinal resolution of 30x64 (30 rings and 64 pixels per ring). Regarding the signal scaling problem, and making use of the equation (3), it is easy to calculate that the outer pixels yield a response 230 times bigger than the inner cells. A scaling mechanism is provided in this sensor to scale the charges, but it is not easy to make such a circuit and it did not work very well, obtaining large slope non uniformities among different rings.

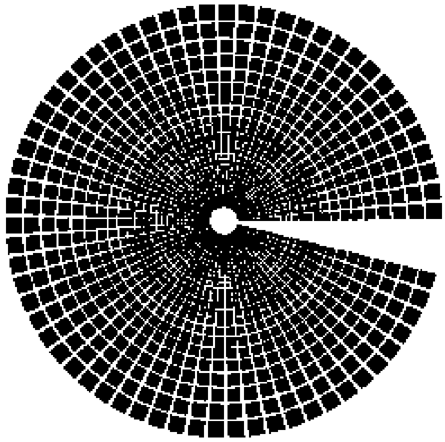

Figure 3: *Retinal structure of the log-polar CCD sensor*

The layout transformation was done by vertex position recalculation from a cartesian grid using a Pascal program. The problem with this approach is that straight lines in the cartesian plane are transformed into straight lines also in the retinal plane, and not curves as it should be. This problem can be solved using multiple vertices in straight lines in the cartesian representation. The problem, that the designers of the CCD sensor had to deal with, was that the employed CAD tool did not allow multiple vertices in straight lines. The solution they adopted was to define each squared polygon by multiple squared polygons. Taking into account that the polygons have to be snapped to a squared grid after the transformation, they overlapped each other to avoid lost of continuity in the transformation (see figure 2(b)).

Regarding the fovea problem, using the CCD technology it is very difficult to maintain a polar pixel distribution, specially for the read-out CCD. The solution adopted by the designers of this CCD sensor was a cartesian grid like any other CCD sensor. This lost of continuity between retina and fovea makes this sensor useless for image processing in the fovea and surroundings since the concept of pixel *neighbor* is lost, apart from the blind areas found between the retina and fovea. The problem anyway was even worst since the fovea did not work properly.

3.2 The McGill University CMOS sensor

A complete different approach was taken for the CMOS foveated sensor designed by Wodnicki at the McGill University.[8] They used CMOS technology instead of the CCD. The objective seems not to design a retinal sensor but to design a normal cartesian sensor with a retinal surrounding. From this objective, they developed a retina with only 16 rings, being the fovea the largest part of the sensor.

The schematic structure of the sensor is shown in the figure 4 where it is possible to see the retinal and the fovea simplified structures. The retina is formed, in the real sensor, by 16 rings containing 64 pixels each. The fovea is a cartesian array with 40x52 pixels. All cells in the sensor, including the retinal cells, are based on light integration; light is captured in a photodiode during a fixed time and then is read-out.[9] For these integration sensors, the response is proportional to cell area, therefore the expression (3) is valid and can be used to calculate the response difference between the inner and outer cells. In the case of this CMOS sensor there is no so much scaling problems since there are only 16 rings, and the maximum scaling factor, calculated using the equation (3), is just 16.6 that does not necessarily require internal scaling. In fact, the sensor does not provide any scaling mechanism, and the area growing compensation is done outside the chip.

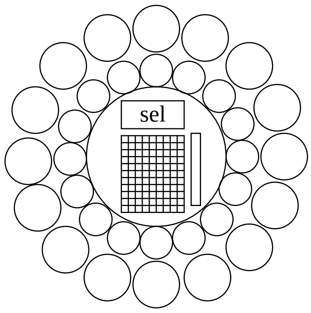

Figure 4: *Retinal structure of the McGill CMOS sensor*

The layout problem in this sensor was solved by the designers using a simple and modular cell for the retina. There is only two components in each cell, in one hand there is a photodiode for the light capturing, and in the other, the selection transistor. Two wires are necessary in each cell, one for the analog read-out and the other for the digital selection. The designers employed the C language to generate the retina. A circular shape for the photodiode was chosen to facilitate the retina generation. Since the structure is very simple, no other considerations were taken in the layout.

With this sensor, the problems of the CCD are not solved. The scaling is not solved since the resolution is so low that these effects are not very large. The retina and fovea continuity does not exist and it is even worst since there is a large separation between the fovea and the retina to allocate the selection circuits of the fovea. Topologically the fovea and the retina are completely different, therefore it is not possible a continuous application of the image processing algorithms.

4 THE NEW RETINAL CMOS SENSOR

A new retinal sensor has been designed in order to solve most of the problems that these kind of space-variant sensor present.[10] Some solutions are presented for the fovea, scaling and layout. The result is a sensor with higher resolution, fovea and retina continuity and minimized scaling effects.

4.1 Non orthogonal layout

At the beginning of the design there was a try of using a CAD tool that could help on the design of such a space-variant structure. At that moment, most of CAD tools for integrated circuit design, were specially addressed to the digital circuit design, that are usually cartesian polygon structures. For such a space-variant design it was necessary to draw every polygon, using a high level description language, in order to have the complete layout. The design tool employed was Cadence, and the high level language to automatically generate the sensor was SKILL. This language is not specially addressed to full custom polygon drawing, but it can be used for this purpose with not many difficulties. The result has been a program, written in SKILL, that automatically generates the retina, the fovea, and the external selection circuits, from a given retinal resolution.

One important issue in the design of any non-cartesian circuit is the grid snapping. Every vertex has to be snapped to a cartesian grid which size depends on the technological process and foundry in which the chip is fabricated. This snapping is not specially critical in cartesian designs, but it has a large influence if the circuit is slightly rotated from its original position. A rectangle usually loose its shape when it is rotated and then snapped to a cartesian grid, and this is specially important in the transistor channel where the constancy of the channel length is important for good property matching among transistors in the circuit. Another problem, derived from the grid snapping, is the layout rule violation. In a cartesian design it is very easy to check the layout rule violations, but when that design is rotated and then snapped to a grid, it looses the shape and the distances among polygons change in a very unpredictable way. To solve this problem, most of polygon distances have been increased a 10%.

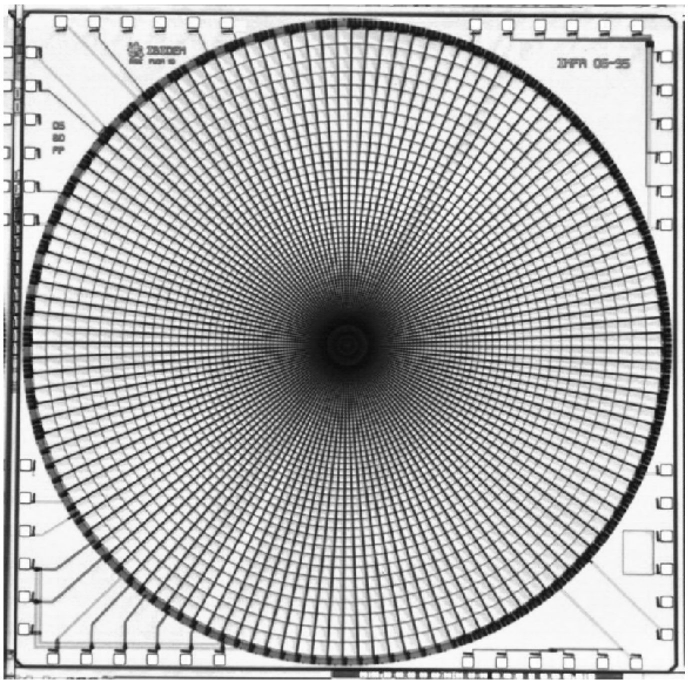

Figure 5: *Microphotograph of the sensor layout*

The last problem, coming from the grid snapping, is the abutment connection of polygons. Every cell in the sensor is connected by abutment, including the selection and read-out wires. This abutment connection can be lost after rotation as shown in the figure 2. If the vertices of one polygon match with the vertices of the connecting polygon there is no problem after rotation, but if the vertices are connected to the side of the polygon, some problems arise, as shown in the figure. In the CCD retinal sensor they have the same problem and they solved it overlapping the polygons to ensure that after rotation the connection still remains. This is not a very elegant solution. The approach taken in the new CMOS retinal design has been the addition, to the side of the polygon to be connected, of the necessary vertices to match the vertices of the original polygon; in cartesian coordinates the shape is exactly the same, but after rotation the connection still remains as shown in the figure 2(c).

The result of these consideration is the sensor shown in the figure 5 where the log-polar structure of the retina is clearly identified. The retina has a resolution of 56 circumferences with 128 pixels per circumference, and the fovea (black hole in the middle of the figure) has 20 rings with decreasing number of pixels per ring, finishing in one single pixel in the middle of the sensor. The sensor has been designed at IMEC, Belgium.

4.2 Sensor fovea

For this new retinal sensor, a different approach for the fovea has been taken. Instead of just putting a squared matrix in the middle of the sensor, as any other existing foveated sensor did, we preferred to follow the structure of the retina, keeping the circular distribution of pixels, while linearly reducing the radius of the circumferences, instead of the exponential reduction that is not possible to be implemented due to the design rule limits. The result is a fovea with decreasing number of pixels per circumference as shown in figure 6; there is one ring with one pixel (the central pixel), one with 4 pixels, another one with 8 pixels, two with 16 pixels, five with 32, and ten with 64 pixels.

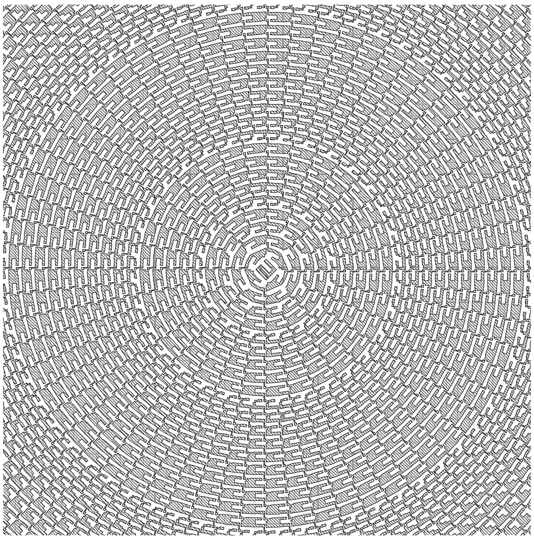

Figure 6: *Sensor fovea*

4.3 Signal scaling

The foveated sensor, presented here, has 56 rings in the retinal part following a exponential growing of the radius. This exponential growing of the radius means a exponential growing also for the pixel area. The growing factor depends on the shape of the cell, etc. Supposing a squared shape for the pixel, and 56 rings, the outer cells are about 250 times bigger than the inner cells using the equation (3). This number is really a large difference, meaning that the response of the outer cells should be attenuated by a factor 250 to obtain the same response than the inner cells. Nevertheless the problem is even worst. This factor has been calculated supposing that the photodiode occupies all the area of the pixel, and this is not true for the retinal sensor. For the outer cells the photodiode covers most of the pixel (around 95%), but for the inner cells, the photodiode is a small part (around 20%) since other cell elements and transistors cover most of the pixel. Therefore, the photodiode size (and not the pixel size) of the outer cells is around 900 times bigger than the inner cell size. Supposing integration cells, as CCD or some CMOS, thus with linear response, a mechanism for charge scaling is necessary. This mechanism should handle differences of 900 from cell to cell and it is not simple at all, specially since the scaling is referred to the slope of the response.

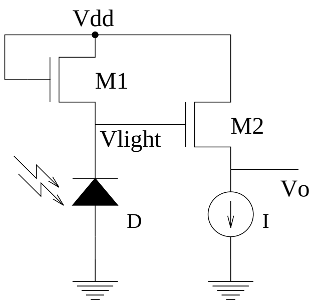

Figure 7: *Basic sensing cell of the foveated sensor*

To solve this problem a logarithmic response cell has been employed.[11] The basic sensing cell is shown in the figure 7. The transistor M1 is biased in the weak inversion region by the small current generated by light in the photodiode. The transistor M2 is just a source follower and the transistor M3 is the selection of the cell for read-out. The current source is common for every transistor in a ring and it is located outside the sensing array. Since the transistor M1 is working in its weak inversion region, the output voltage is proportional to the logarithmic of the current thus proportional to the logarithmic of light intensity as shown in the following equation:

$$V_{gs} = \frac{kT}{q} ln \left(\frac{L}{W} \frac{I_d}{I_{do}} \right) \tag{4}$$

where V_{gs} is the gate-source voltage, I_d is the drain current, W and L are the width and length of the channel, T is the temperature, and k, q and I_{do} are constants. From this expression it is simple to see the constancy of the slope of V_{gs} against $ln(I_d)$, since the term $\frac{kT}{q}$ is independent of the technology, shape, position, etc. This slope only depends on the temperature, but T is suppose to be constant for the whole chip.

All other parameters of expression (4) are included in the logarithm as product terms. It has a straight forward meaning: any change of any parameter as the length, substrate doping, etc., produces a change in the offset of the curve that defines the dependency between light intensity and output voltage. This is specially interesting since it is easier to perform offset corrections than slope corrections that usually also implies an offset correction. These interesting characteristics are always true if transistors are big enough to discard second order effects, that for small size transistor are specially important.

With the cell used in the foveated sensor presented here, the scaling problem is not so big. Since the output is proportional to the logarithm of the incident light and photodiode size, the growing area has less effect in the output since it is included inside of the logarithmic term. In the other hand, the growing factor becomes an additive term instead of a multiplicative term, so the difference from one cell to another is an offset while the slope of the curve remains constant. The maximum difference between inner and outer cells was a factor of 900; this proportion, after the logarithmic expression, is converted to an offset, or additive term, of only 6.8, that is less traumatic than the 900 multiplication factor.

It is simple to perform a half compensation in the design increasing the transistor channel width with the pixel area.[12] In the design the area increases squarely while the width increases linearly, this is the reason it is only possible a half compensation since the 2 exponent comes out the logarithmic expression like a multiplicative 1/2 in the additive term. Thus the maximum additive term 6.8 is divided by two simply increasing the weak inversion transistor channel width.

5 CONCLUSIONS

A new space-variant CMOS image sensor has been designed and fabricated. The sensor performs the log-polar transformation in the focal plane. This strategy for image acquisition has interesting properties for image processing, robotics navigation, image transmission, etc. There are three important issues in any space variant sensor. The first is the polar structure of the sensor layout, the second is the necessary scaling of the cell response, and the last is the topology for the center part, or fovea, of the sensor. Three solutions has been presented for these problems and tested with the new retinal sensor. The layout has been generated automatically with a high level language to define the vertex positions. The scaling, main problem of these sensors, has been minimized using a logarithmic response cell that converted the scaling difference in just additive terms that can be easily subtracted with a simple external circuit (D/A converter and RAM memory to store the terms). The scaling problem has been also reduced designing an increasing channel width for the weak inversion transistor for the basic cell. An last, the fovea problem has been solved using a retinal similar topology for the fovea, keeping the polar structure and making a small change in the radius transformation.

The sensor fabricated has enough resolution and performance to be employed in real image processing applications. The main problem that it presents in the high FPN (Fixed Pattern Noise) that has as contributors the transistor mismatching and the scaling half compensation described in the paper. Anyway, this FPN can be easily canceled with the external circuit commented before and it does not suppose any problem for the normal image acquisition. The speed of the sensor can be as fast as 100 images per second, allowing real-time application with not very complicated processing hardware since one image is just 8 Kbytes.

6 ACKNOWLEDGEMENTS

This sensor has been funded with the IBIDEM project of the European Community TIDE program. F. Pardo has been supported in part with an European Community grant within the Human Capital & Mobility program. The authors would like to thank many participants of the IBIDEM project for their valuable comments and suggestions on the chip structure.

7 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 R.M. Iñigo, C.Q. Xu, B.C. Arrúe, and E.S. McVey. Hardware-implementable neural network for rotation-scaling invariant pattern classification. *SPIE Journal of Electronic Imaging*, 1(3):293–312, July 1992.

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 J. Van der Spiegel, G. Kreider, C. Claeys, I. Debusschere, G. Sandini, P. Dario, F. Fantini, P. Belluti, and G. Soncini. *A Foveated Retina-Like Sensor using CCD Technology.* Kluwer Academic, Boston, 1989.

6 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.

7 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.

8 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.

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

10 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.

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

12 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.