A NEW FOVEATED SPACE-VARIANT CAMERA FOR ROBOTIC APPLICATIONS

J. A. Boluda[1] *F. Pardo*[1] *T. Kayser*[2] *J. J. Pérez*[1] *J. Pelechano*[1]

[1]Institut de Robòtica. Universitat de València, València, Spain
[2]Fachhochschule Mannheim, Mannheim, Germany

ABSTRACT

A foveated camera has been designed and fabricated. The camera is implemented using a new foveated CMOS sensor which incorporates a log-polar transformation. This transformation has specially interesting properties for image processing including the information selective reduction and some invariances. The structure of this sensor permits an individual access to each pixel, exactly in the same way that the access to a RAM. This is a very interesting property that makes a big difference between a CCD based camera and a CMOS based camera. In the other hand, the CMOS nature of the sensor implies a very important fixed pattern noise due to the mismatch of the cell transistors. A fixed pattern noise correction circuitry has been included in the camera, in order to achieve a good image quality. The solutions adopted for achieving a low noise/signal ratio are also presented. Finally, some guidelines for including this camera in an autonomous navigation system are shown. The easy way for calculating the time-to-impact through this camera suggests the utilization of this sensor system for real-time applications.

1. THE LOG-POLAR TRANSFORMATION

In the human visual system the receptors of the retina are distributed in space with increasing density toward the center of the visual field (the fovea) and decreasing density from the fovea toward the periphery. The log-polar transformation is described as a conformal mapping of the points on the polar (retinal) plane onto a cartesian (log-polar) plane following the equations:

$$\begin{cases} \xi = \log r \\ \gamma = \theta \end{cases} \tag{1}$$

The resulting log-polar projection is invariant, under certain conditions, to linear scalings and rotations of the retinal image. These complex transformations are reduced to simple translations along the coordinate axes of the log-polar image. This property is valid if, the scene and/or the sensor moves along (scaling) or around (rotation) the optical axis. The same properties hold in the case of a simple polar mapping of the image, but a linear dilation around the fovea is transformed into a linear shift along the radial coordinate in the (ξ,γ) plane, while the log-polar transformation produces a constant shift along the radial coordinate of the log-polar projection [1].

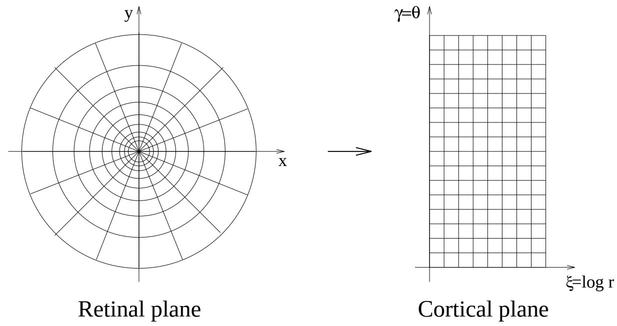

Figure 1. The log-polar mapping

This structure has good resolution for image processing in the central parts of the scene, keeping enough resolution in the periphery allowing a wide view field. The result is a reduction of the image data to be processed, making these information representation specially interesting for real-time image systems and robotic navigation [2] [3]. We can see the figure 2, as an example of how is possible to reduce the quantity of data, without reducing the subjective image quality. In this way, an important role of space-variant image processing is to keep centered and focused the interesting parts of the image. In this case it is possible to use this camera for solve problems with a high computational cost.

2. THE CMOS SENSOR

The core of this camera is the new log-polar CMOS sensor [4]. There are, basically, two kind of CMOS sensing cells. The most used method is to integrate the incident light in a photodiode that is charged during an integration time. This exactly the same way used for the CCD sensors for capturing light. There exits a retinal sensor implemented in standard CCD technology [5]. The second way for light capture avoids the light integration and the photocell elements transform the light into current, and then in voltage without charge integration. This is the photocell used for the foveated CMOS sensor.

The sensor has as input the address of the pixel to be read an as output an analog level, that afterwards, is amplified and adjusted to convert it in a digital grey value. All this operations are performed out of the sensor. This sensor uses the log-polar distribution, consequence of this special structure is the different size of the sensing elements, achieving more resolution in the center and less in the periphery. Moreover, the shape of the photodiodes follows exactly the log polar transformation through non-orthogonal shapes. Finally, like a direct consequence of this different size, a scaling mechanism is needed.

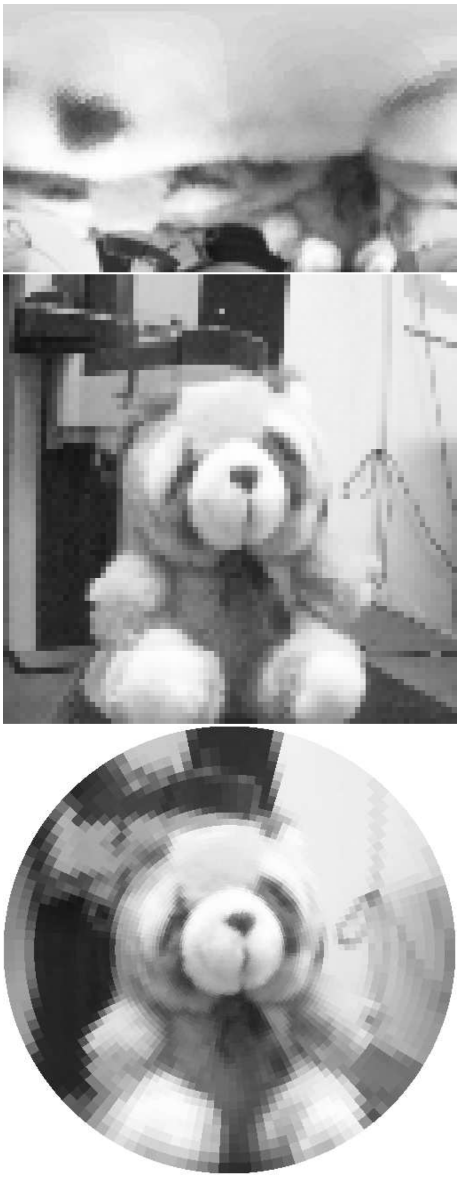

Figure 2. The lab pet in the cortical plane, in a CCD image and in a retinal image

There are two different parts in the sensor, the retina or periphery and the fovea or central part. The retina follows the log-polar transformation explained before. This is not possible for the fovea part due that following the log-polar equations should implies a very small transistors and, like consequence, rule violations in the layout design. This limitation is a big problem for the sensor symmetry structure, and then, for the image processing algorithms. Due to that the solution adopted is a fovea with 20 rings. The ten outer rings have 64 pixels, the following inner five have 32 pixels, the next two 16 pixels and finally a ring with 8 pixels, another one with 4 pixels and the last one with one pixel, the central one. The resolution of the retina is 56x128. That means 56 circumferences with 128 pixels each. This resolution is the biggest never used in a foveated sensor to date. The sensor itself has been designed at IMEC, Belgium, and the physical layout follows the log-polar transformation.

The 56 rings in the retina area follow an exponential growing of the radius. That means that the pixel area grows also exponentially. The outer cells are about 250 times bigger than the inner cells, then the response of the outer cells must be attenuated by the adequate factor in order to obtain the same response than the inner cells. The sensor has been implemented using a cell where the output is proportional to the logarithm of the incident light and photodiode size. This growing factor becomes an additive term instead of a multiplicative term, in other words, the difference from one cell to another is just an offset.

Moreover, there are more factors than include a fixed pattern noise inside the sensor. The mismatching properties of MOS transistors are also a problem for analog design [6]. This is an unavoidable problem that is also present in other CMOS visual sensors.The idea is to have exactly the same response of two cells receiving the same quantity of light, but this is impossible.

There are several sources for the transistor mismatching. Some of these sources are basically the randomly variation of some technological parameters, like the doping concentration and the oxide thickness. This variation depends of the wafer area where the transistors are located. There are also geometrical parameters like the length (L) and the width (W) of the transistor channels that can be modified, but it is not easy to know the influence of each parameter in the final offset non uniformity. There has been made a detailed study of this influence in this sensor and what to do in order to minimize it [7]. Unfortunately, due to the fixed pattern noise nature, it is not possible to avoid it completely. In this way, it is necessary to take into account this offset non uniformity for acquire the log-polar images. The obvious way for doing that is to have previously measured this fixed pattern noise, and afterwards, subtract it to each sensing cell.

3. THE CAMERA STRUCTURE

The camera has a C-mount lens holder for the optical elements. The CMOS sensor is mounted in a PLCC socket of 68 pins. Afterwards, an operational amplifier connected to the sensor output. This OPAMP is necessary since the output is not amplified and its amplitude is minimal, near of 10 mv per light octave. Moreover, the signal is inverted, then is necessary a negative amplification. This OPAMP is also used for change the gain and the offset in a very accurate way. This characteristic permits to change the contrast and adjust the signal to the range of the A/D converter. This OPAMP also filters the great quantity of high frequency noise that is in the sensor output avoiding introduce delays in the analog signal, due to that a high frequency OPAMP has been used. The use of most populars OPAMPs like the 741 introduces a delay bigger that the delay produced by the sensor itself. With this, the delay obtained is around 500 ns, but in order to ensure a right stabilized value a time of 1 (s is used.

Another bloc included is the automatic circuitry for the fixed pattern noise compensation we can see in the figure 3. This compensation is implemented through a current injection in the OPAMP input which is equivalent to subtract an offset to the sensor output. The offset that must be subtracted is pixel-dependent. Initially, before of the normal camera utilization, this offset must be measured and calibrated. This operation is made positioning an ideal totally uniform grey level color in the front of the camera and recording the sensing cell values in an SRAM. Ideally the normal pixel value should be uniform, but due to the offset non-uniformity there is a difference between the real and the theoretical value. When this fixed pattern noise is in the SRAM, it is possible subtract it converting this digital offset value to an analogue value through a D/A converter.

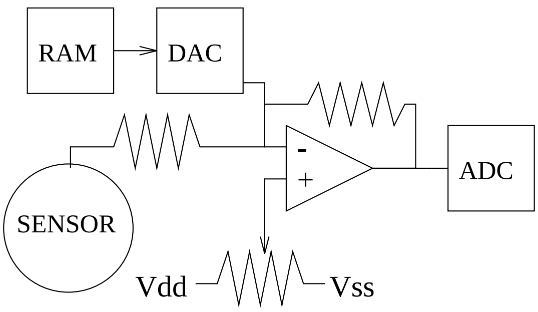

Figure 3. The fixed pattern noise correction circuitry

The camera is connected through a plane cable and an IDC26 connector to a PC board. This board is basically an Intel 8255 chip that has 3 I/O ports. The A port is configured for write the pixel address. The B port has the different camera control signals, like y-select for the address multiplexer latch and the clk-adc for the A/D converter capture. The C port is used for read or write the binary value of the retinical pixel selected. In this way, it is very simple to capture an image. Firstly the camera should be calibrated as has been described before, and afterwards, if we want a complete image, all the addresses must be generated and all the pixel values must be read, exactly in the same way that using a ROM. The bottle-neck is the speed for writing the address and for reading the value when the idea is to obtain a full image. When the image processing algorithm do not use the whole image the advantages are clear. The out/in PC operations are slower than the memory access. If we take into account the software time needed for generate all the pixel addresses and for reading all the values the ratio obtained is about 15 frames per second.

The solution to this time constrain is very simple.

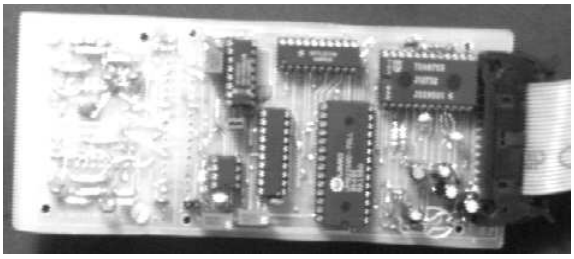

Figure 4. Camera PCB

A frame grabber should be implemented. In fact, a frame grabber for the log-polar camera is now being designed. This board will control the access to the camera and will have a memory buffer that will have the last image acquired. This memory will be mapped directly to the PC memory, achieving a fast frame acquisition. The card will have several configuration registers for select a window in the image or choose the acquiring speed. The initial simulations give a frame ratio close of 100 frames per second.

4. EXPERIMENTAL RESULTS

The software for calibrating and for acquiring images has been also developed. This software is very simply and there are basically only 3 routines for access to each pixel. The first routine for write the high address or circumference (the addresses are multiplexed), the second one to write the low address or pixel position in the circumference and the third, and last one, to read the pixel value. Many measures has been made with the retinal camera, specially taking into account the problem on the offset non uniformity. The retinal sensor yields a good response but the fixed pattern noise needs a correction that is performed by the circuitry described before. In the figure 5 it is shown an image before of correction through the subtraction of the fixed pattern noise and after this correction.

The response time measured of the camera developed is 500 ns, that means than the stabilization time in the sensor output when the pixel select signal has been activated is 500 ns. This speed will give us, when the frame grabber is implemented, a ratio of 100 images per second, in any case enough for robotics applications. The final image quality is not very good for video transmission but is not bad for image processing.

5. IMAGE PROCESSING FOR ROBOTIC APPLICATIONS

The camera is going to be used for obstacle avoidance in autonomous navigation for compute the time to impact, for approaching objects. In the log-polar representation only the radial component of the optical flow, represented on the polar plane, depends on the time-to-impact. The polar flow representation allows the direct computation of the time-to-impact with arbitrary egomotion [8]. It is possible to generalize the logarithmic-polar complex mapping property of transforming objects dilation into a translation along the (radial coordinate to

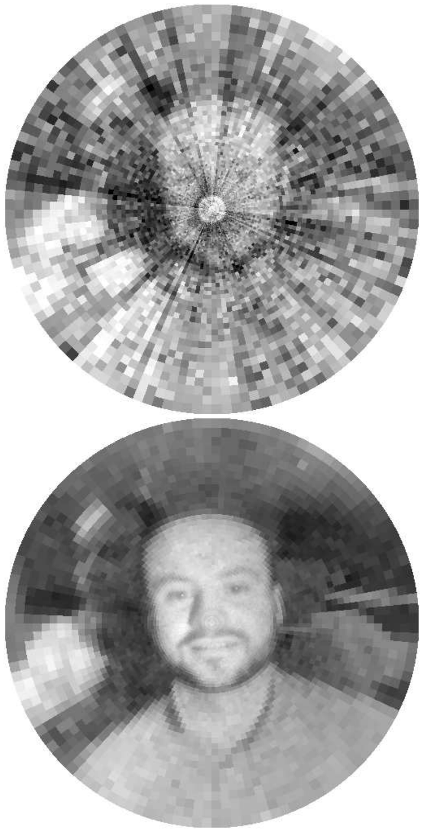

Figure 5. Top: Original sensor image. Bottom: Camera supplied image

more general and complex kind of motions. Generally, an expansion of the image of an object, due either to the motion of the camera or the object itself, will produce a radial component of velocity on the retinal plane. It has been developed a theoretical way for this estimation [8]. The log-polar representation, that has been used in the sensor of this camera, is the best suited for the computation of the scene structure mainly for three reasons.

- The problem is reduced to a small number of equations which relates the optical flow and its derivates.
- The dependence on depth is decoupled in the radial and angular component of the optical flow. Only the radial component is proportional to the time-to-impact.
- The polar mapping can be easily applied for real time applications through the hardware described in this paper.

6. CONCLUSIONS

A new camera fully oriented to robotic applications has been developed and presented in this paper. The log-polar properties have been demonstrated to be oriented to robotic autonomous navigation. Once the CMOS sensor was developed it was necessary to develop a camera for acquire log-polar images. A fixed pattern noise appears due to the log-polar representation and also due to the mismatching of MOS transistors. This problem must be taken into account in order to achieve a good image quality. The solution for this offset non-uniformity has been to have in memory this fixed pattern noise and subtract it through an OPAMP, in any pixel access. The elements included in the camera and the software employed and used for capture images are described. Images without and with the offset correction are showed like experimental result of the developed system.

Finally, the image processing topics where this camera is useful have been indicated. The next steps in the future research include the frame grabber implementation, and some hardware design for the time-to-impact estimation. In fact, now this system is being incorporated to an autonomous navigation system, in the Robotics Institute of the University of Valencia.

7. ACKNOWLEDGEMENTS

The authors wish to thank Dr. Dierickx of IMEC for his help on the retinal sensor. The sensor chip has been designed and fabricated at IMEC (Belgium). The camera design and implementation has been supported in part by the Spanish government (CICYT project TAP95-1086-C02-02).

REFERENCES

[1] G. Sandini, V. Tagliasco, "An anthropomorfic retina-like structure for scene analysis", Computer graphics Image Processing 14(3), 1980.

[2] M. Tistarelli, G. Sandini, "Dynamics aspects in active vision", Image understanding, Vol. 56, N.1 July, pp 108-129, 1992.

[3] F. Pardo, F. Vegara, J. A. Boluda, S. Felici, "Sensores CMOS para robótica e industria", IV Congreso AER-ATP. Zaragoza, Spain, October 1995.

[4] D. Scheffer, B. Dierickx, J. Vlummens F. Pardo, "Log-polar image sensor in CMOS technology", European symposium on lasers, optics and vision for productivity in manufacturing. Besancon, France, June 1996.

[5] Van der Spiegel, G. Kreider, C. Claeys, I. Debuscherre, G. Sandini, P. Dario, F. Fantini, P. Bellutini, G. Soncini, "A foveated retina-like sensor using CCD technology", Analog VLSI and Neural Network Implementation. Boston 1989.

[6] M. J. M. Pelgrom, A. C. J. Duinmaijer, A. P. G. Welbers, "Matching properties of MOS transistors", IEEE Journal of Solid-State circuits. Vol. 24. N.5, October 1989.

[7] F. Pardo, J.A. Boluda, J.J. Pérez, S. Felici, B. Dierickx, D. Scheffer, "Response properties of a foveated space-variant CMOS image sensor", International IEEE Symposium in Circuits and Systems, ISCAS'96, Atlanta 1996.

[8] M. Tistarelli, G. Sandini, "Direct estimation of time-to-impact from optical flow" , IEEE Workshop on Visual Motion. Princenton 1991.