Pyramidal Architecture for CMOS Image Sensor

Fayçal Saffih[1], Richard Hornsey[2]

[1] Electrical and Computer Engineering Department, University of Waterloo, Ontario, Canada

[2] Centre for Vision Research, Department of Computer Science, York University, Toronto, Ontario, Canada

***Abstract*—A new architecture for a CMOS image sensor is presented along with its capabilities and applications. Named pyramidal architecture after its shape of image sampling, this novel architecture is primarily suggested for smart robotic imaging, videophones and wireless image transmission. By replacing the rows of conventional image sensors with rings, the dimensionality of processing itself is increased from 1 dimension (1D) to two dimensions (2D). The pyramidal architecture is also an attempt to mimic the foveated biological sampling architecture found in the human retina. In the present example, a scanning technique is demonstrated to allow the capture of images with two different integration times, thereby facilitating dynamic range enhancement through image fusion. The ring architecture leads to a sensor in which dynamic range enhancement is greatest towards the central rings in a fovea-like arrangement.**

***Index Terms*— CMOS image sensors, raster scanning, ring level processing, two dimensional processing, sampling, pyramidal architecture, Fovea.**

I. Introduction

THE demand for multimedia applications has increased tremendously in recent years especially after the advent of portable computers and other electronic devices, as well as decreasing communication costs. CMOS image sensors are particularly attractive because of their low cost and integrated processing capabilities[1]. Other advantages of CMOS imagers include low power consumption and standard fabrication process which has made it the primary choice for applications such as videophones and Web cameras[2]. On the other hand CMOS image sensors have traditionally suffered from higher noise such as fixed-pattern noise (FPN) and photo-response non-uniformity (PRNU) compared to their CCD counter part. Although, many successful attempts has been suggested to minimize FPN and PRNU noise by implementing some smart vision processing [3], comparatively little effort has been spent to sampling architectures and their applications in smart imaging and robotic vision. Our pyramidal architecture is suggested as an attempt to fill this gap.

In the classical CMOS image sensor architecture, raster scanning was adopted to sample the charge-integrated image at the focal plane, as shown in Fig. 1. Raster scanning was originally implemented for displaying video images in Cathode Ray Tubes (CRT), and was later adopted to be the scanning scheme for early MOS imagers [4] in order to maintain compatibility between acquisition and display systems. Raster scanning in a CMOS image sensor is depicted in Fig. 1. The integrated image is sampled row by row, from the first top row till the last bottom row. Each row is selected and its output current gets converted to voltage at the bottom of the imager through bank of active resistors. The resultant voltage is sampled at the same level through a bank of MOSFET switches to a bank of storage capacitors. The same row is reset to V_{dd} and then read out in the same way to a separate bank of sample-and-hold capacitors. Therefore, for each row of pixels, there exists a double sampling: row output voltages sample that is stored in a bank of capacitors, and the row reset voltages sample stored in the other bank of capacitors.

This double sampling of a row is used to reduce the noise (mostly FPN) of each row of pixels, and is conventionally known as Correlated Double Sampling (CDS). The subtraction is either performed on-chip or off-chip.

Finally, a column select logic is used to select a column sampled voltage capacitor at a time to buffer it out. This process reiterates indefinitely from the first top row to the last row at the bottom of the image sensor for each frame. More complex read-out schemes can be implemented, for example to allow integration time and frame rate to be controlled independently.

The imager-sensing unit, the pixel, is constructed based on the standard active pixel sensor comprising a reset transistor, reverse biased N^+/P-substrate diode, source follower transistor and a select transistor. All the transistors involved in the active pixel sensor are N-type MOSFETs, as shown in Fig. 2. The reverse biased photodiode first get a pulse of charge through the reset transistor after which the photodiode is left discharging by the incident light generated electron-hole pairs. The rate of discharge is ideally proportional to the light intensity.

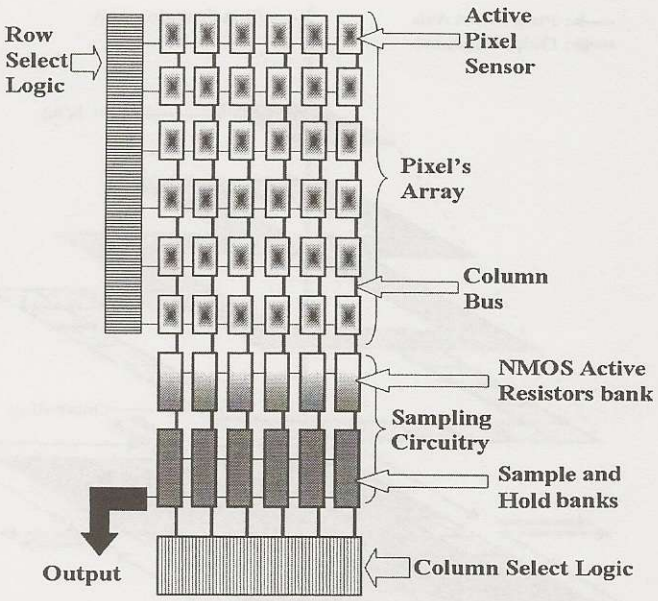

Fig. 1. Standard CMOS image sensor architecture.

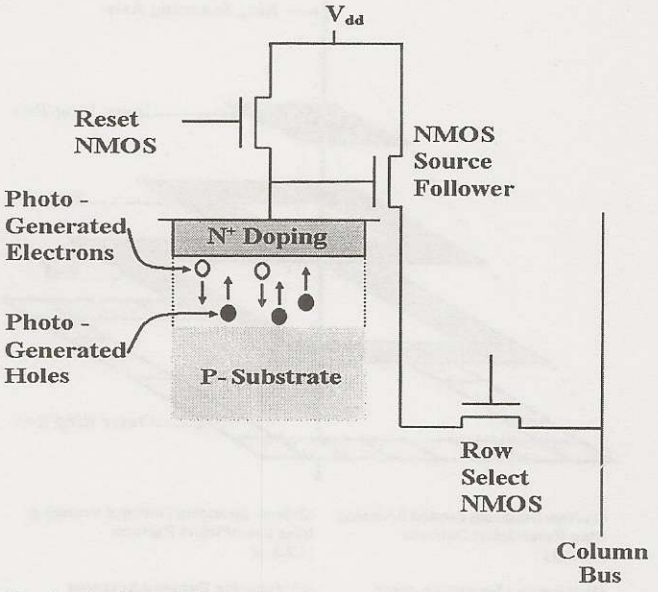

Fig. 2. Active pixel sensor structure

At the readout phase the row select transistor is on and the source follower drives a current proportional to the voltage of photodiode discharge. The output current will be then be converted to voltage through a biased NMOS transistor at the sample and hold circuitry at which the output voltage is sampled and stored.

From the above description of raster scanning, we easily can see that its horizontal sampling of the sensed image is faster than the vertical sampling. How will this fact affect the output image, especially if the sensed image is of a moving object? Is a display scanning scheme suitable for scanning acquired/sampled image? In the following sections, we will try to answer these questions while presenting our new architecture for scanning the sensed image.

In section 2, the pyramidal architecture is described through its biological motivation and then its structural and scanning timing is presented and analyzed. Section 2 is then concluded with the timing assessment of the suggested bouncing scanning scheme and the resulting foveated intrascene dynamic range enhancement. The following section will be dedicated to the experimental implementation of the pyramidal architecture along with a preliminary result. We conclude with a discussion of the advantages and disadvantages of the suggested architecture.

II. Pyramidal Architecture

A. Biological Motivation

An observation of the retina anatomy reveals that the human visual acquisition system is composed of two categories of cells, the rods and the cones. The distribution of the visual cells is however not uniform, the cones being highly dense on the central region of the retina, named the fovea, and the rods being spread all over around in an almost circular symmetry. Fig. 3 sketches the distribution of the human visual cells on the retina[5].

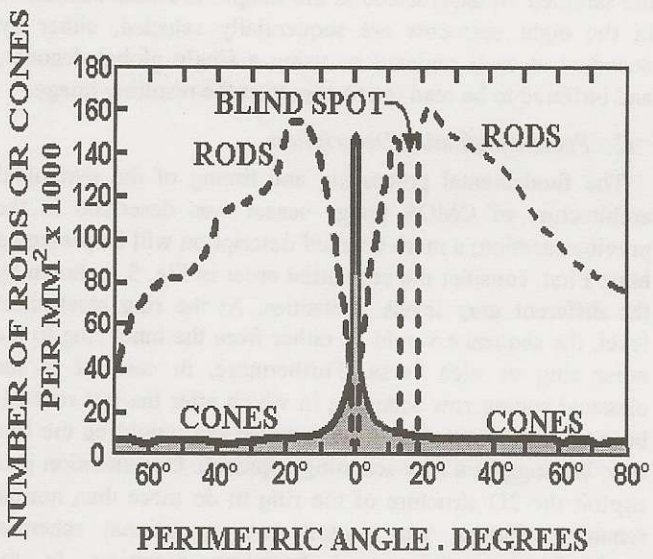

Fig. 3. The distribution of rods and cones photoreceptors in human eye, after [5].

The circularly symmetrical distribution of rods and cones is very interesting besides the large order of magnitude of light sensitivity of the rods over the cones [5]. This certainly presumes that both cells have different temporal response to sample the incident light. Both of these observations were the motivation of suggesting a new scanning/sampling architecture for CMOS image sensor based on a concentric 2D rings instead of the 1D row sampling found in raster scanning.

The addressability of CMOS image sensors was the key feature that allows us to implement such geometric scanning.

B. Hardware Description

Our suggested architecture is proposed primarily as an alternative for the classical raster scanning architecture widely used in conventional image sensors to sample the sensed image. A sketch of this architecture is shown in Fig. 4. The figure is divided into two sections; the reset and select concentric rings and the readout scanning of our pyramidal CMOS imager. The gradual gray color is used to show the sequences of the readout without giving any directional preference. For example, reset and select sequence can be from outer ring towards the inner ring or vice versa. Also the buffering of the pixels' output from the sampling capacitors can be from the pyramid diagonals towards the middle of the rings, where a buffer should be connected, or vice versa. Each ring is built of pixels that share the same reset and select signals. Each ring bus is activated by a decoder dedicated for resetting or selecting any given ring of pixels. After a ring is selected, its sampled voltages are dumped through the pyramid diagonal busses (equivalent to columns in a conventional array) into the eight sample-and-hold capacitor banks at the base of the image sensor pyramid. This distribution of ring readings into eight banks of capacitors will result to dividing the read image into eight clusters. Finally, the sampled voltages stored at the sample-and-hold capacitors in the eight segments are sequentially selected, either by decoders at each segment or using a single global decoder, and buffered to be read out to construct the resulting image.

C. Processing/timing Description

The fundamental processing and timing of the pyramidal architecture of CMOS image sensor was described in the previous section; a more detailed description will be presented here. First, consider the sequential order in Fig. 5 is shown by the different gray levels intensities. At the ring reset/select level, the sequence would be either from the inner ring to the outer ring or vice versa. Furthermore, in contrast to the classical rolling row scanning in which after the last row has been read the next row to be sampled out would be the first row, we suggest a new scanning approach. Our intention is to exploit the 2D structure of the ring to do more than normal scanning. Hence, we suggest two operational schemes, conventional scanning and bouncing scanning. In the bouncing scheme, the scan direction reverses after all concentric rings have been read out, leading to alternating outward and inward scanning directions. Hence, for example, the outermost ring would be followed by the next closest ring, rather than by the innermost ring.

The conventional scanning scheme is similar to raster scanning in which the sequence order of the rows to be scanned does not change from frame to frame. To analyze the usefulness of the bouncing scanning scheme we first recall some useful parameters of CMOS image sensor, namely the dynamic range related to light range detectability and the integration time related to imager frame rate.

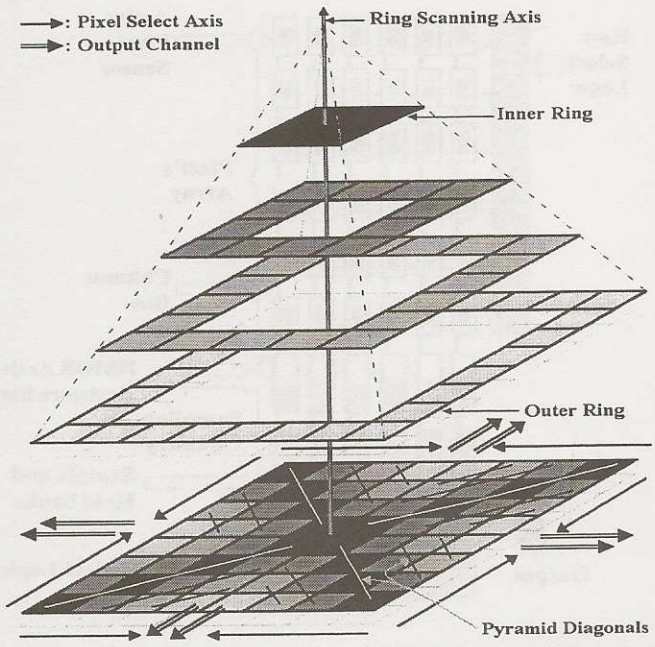

Fig. 4. Sketch of the pyramid architecture of CMOS image sensors.

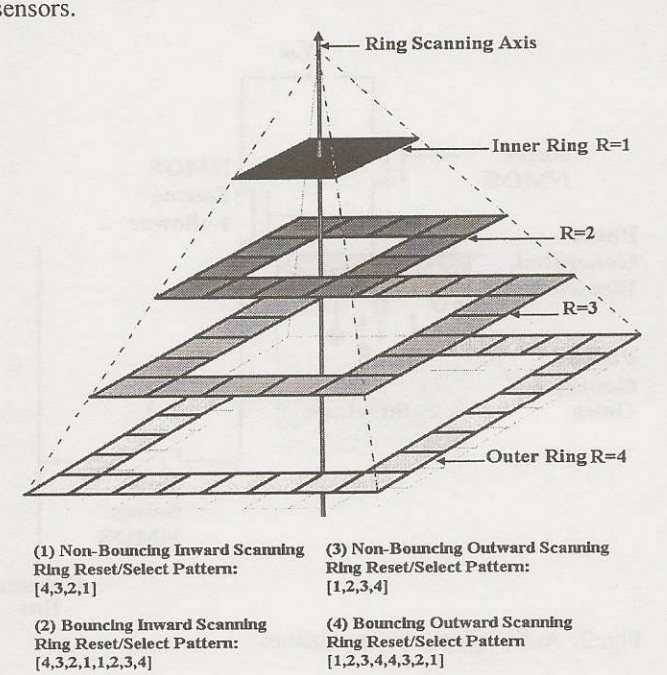

Fig. 5. The different ring scanning pattern of pyramidal CMOS image sensor.

Integration time is the time during which the light is being converted to charge and accumulated in a pixel. In other words, integration time is the period of time between the consecutive reset and read events. Dynamic range is the ratio of highest non-saturating optical flux to the smallest detectable flux for a sensor. So, for short integration time, the image sensor resolves nicely only high light intensities and the

rest of the imager stays dark (very low pixel discharge). On the other hand, for long integration times, the imager is better able to resolve low light intensities, while it saturates everywhere that the light intensities are relatively high. Thus, if images from a given imager are captured using two integration times, the image sensor dynamic range will be extended by fusing the two images. This improvement is given by the logarithm of ratio of the two periods of integration time, $20\log(T_{\text{int}1}/T_{\text{int}2})$ where T_{int1} and T_{int2} are the longer and shorter integration times, respectively. This technique of using dual integration time periods to extend the dynamic range has been suggested by [6] while other techniques of fusing many acquired images of different integration times have been suggested by recently by[7].

Prior to the timing assessment, it can be thought that the bouncing scheme will result in a non linear image sampling transfer due to different integration time periods for each ring. Note that the ring that is the sampling unit/level in our pyramidal CMOS image sensor is a two-dimensional (2D) structure of pixels in contrast to the 1 dimensional (1D) raster scanning row structure. This is a very important difference between the two scanning methods. In the classical raster scanning scheme, the integration time has to be equal for every single pixel in the image sensor in order to have a consistent output image. Hence, if the bouncing scanning scheme is applied to the raster scanning architecture, a 1D variation of response from one side of the array to the other will result. In the pyramidal architecture however, this is not a problem because of the fact that the sampling structure is a 2D structure similar to the projected image dimension. Hence, the consistency of the output image can be reduced to the ring level instead of the whole frame. In other words, the output image will be pyramidal in its detection of light intensities integration. High light intensities are more sensed at the bouncing scanning scheme edges, especially at the inner rings, while low light intensities are more sensed near the other edges. This way of scanning would result an extra dimension for light sampling giving the shape of pyramid light intensities sampling. Back to our new suggested scanning approach we find that the dynamic range is gradually enhanced from ring to ring. This will be more examined in the following formulation of the integration time periods for the bouncing scanning schemes.

D. Integration Time Assessment for Bouncing Scanning

Firstly, the parameters for assessing the integration time of sampled rings in bouncing scanning schemes are discussed. The readout can be divided into two steps; the sampling step that takes a period of time, T_{spl} ,from 1μs to 10μs, and the scanning step to buffer out the row pixels that takes a period, T_s per pixel, in the range of 0.1μs to 1μs. In the pyramidal architecture, the rings contain different numbers of pixels. We take the inner ring to be the reference ring, $r = 1$ and the subsequent ring is r = 2, and so on. Secondly, we assume the output image is built from the eight output channels of the pyramid imager. Hence, the analysis of the rings' integration time can be reduced to the analysis of the pyramid cluster, instead of the entire sensor. In other words, since the pyramidal imager has eight output channels in parallel, frame scanning is therefore based on the cluster timing. Cluster ring 1 contains one pixel, ring 2 has two pixels…etc. Hence ring r contains r pixels, as shown by Fig. 6. Finally, the number of rings is $R = N/2$ where N is the resolution of the imager edge, assumed to be square. Similar analysis can be applied in the case of a rectangular image sensor. Therefore, the time required to read a pyramid cluster ring is:

$$Tring = rT_s + T_{spl} \quad \ldots \qquad (1)$$

Taking the following example of four-ring pyramid as shown in Fig. 6, we will show the integration time periods for every ring.

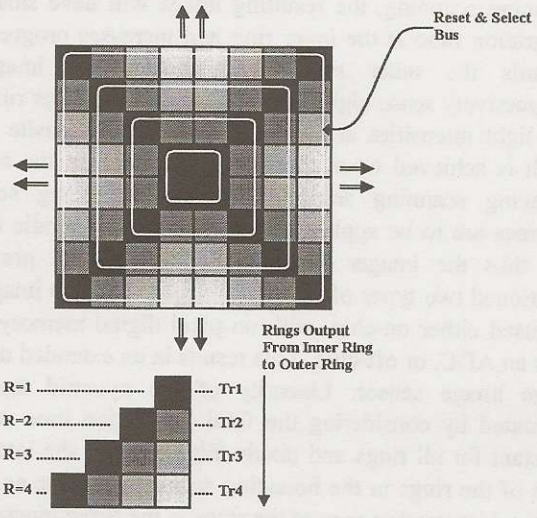

Fig. 6. Schematic of a Pyramid Cluster Scanning

Suppose that we have started scanning the pyramid imager from the outer ring towards the inner ring, from which we bounce back the scanning towards the outer ring, at which ring scanning bounce back again. Fig. 7.a sketches the timing of the bouncing at the inner ring and Fig. 7.b sketches the timing of the bouncing at the outer ring. From both diagrams, integration times are deduced.

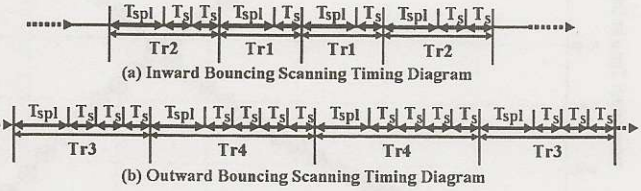

Fig. 7. Timing diagram of inward and outward bouncing scanning

Note also that sampling period T_{spl} contains three consecutive phase: ring output sampling, ring reset and then ring reset voltage sampling. Thus, the integration time of a

ring r for inward bouncing scanning is Tr_{in} and estimated to be:

$$Tr_{in} = 2\left[\sum_{\substack{i=R \\ i\to i-1}}^{r+1} iT_s + (R-r)T_{spl}\right] + rT_s \ldots (2)$$

The integration for outward bouncing scanning is similarly calculated to be:

$$Tr_{out} = 2\left[\sum_{\substack{i=1 \\ i\to i+1}}^{r-1} iT_s + (r-1)T_{spl}\right] + rT_s \ldots (3)$$

The pyramidal CMOS image sensor when using bouncing scanning scheme will output two versions of the sensed image corresponding to these two integration times. At the inward bouncing scanning, the resulting image will have small ring integration time at the inner ring and increases progressively towards the outer ring. This enables the imager to progressively sense high light intensities at the outer rings and low light intensities at the outer rings. The opposite of this result is achieved when the imager becomes on the outward bouncing scanning scheme. The two bouncing scanning schemes are to be applied continuously in a periodic manner and thus the imager continuously delivers the previously mentioned two types of acquired images. The two images can be fused either on-chip, with on-pixel digital memory joined with an ADC, or off-chip. This results in an extended dynamic range image sensor. Linearity of the pyramid imager is estimated by considering the final integration time which is constant for all rings and pixels. Fig. 8 shows the integration time of the rings in the bouncing scanning scheme as well as the total integration time of the rings in the fused image.

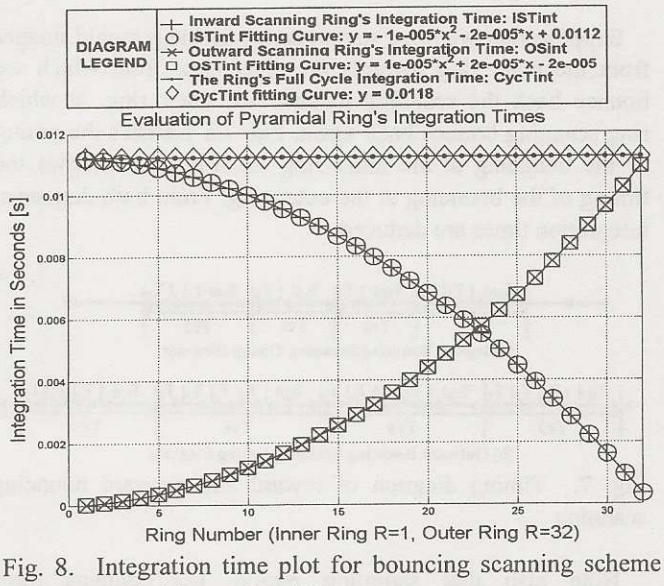

Fig. 8. Integration time plot for bouncing scanning scheme with Ts=Tspl=10μs and R=32.

After fusing the two resulting images from inner scanning and outer scanning as described earlier [6], the enhancement in the intrascene dynamic range for each ring (although it is an imager characteristic parameter) is plotted in Fig. 9.

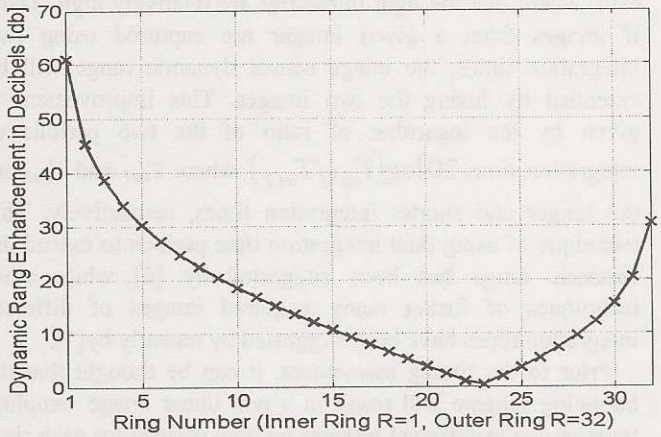

Fig. 9. Ring intrascene dynamic range enhancement after bouncing scanning resulting images fusion.

Image fusion can be achieved on chip or off chip linearly using bit concatenation or non-linearly using addition. It is clear from Fig. 9 that dynamic range enhancement is higher in the inner rings than in the outer rings. It also get minimizes to zero in ring R=23 where the inward and outward scanning integration times are equal. This is an interesting result especially when considering the biological motivation behind our pyramidal CMOS image sensor discussed in II.A.

III. EXPERIMENTAL IMPLEMENTATION

We have designed our pyramidal image sensor using standard 0.18μm CMOS technology. The floorplanning of the pyramidal imager containing 64x64 pixels (32 rings) of 3.5x3.5 mm^2 die area is shown in Fig.10.

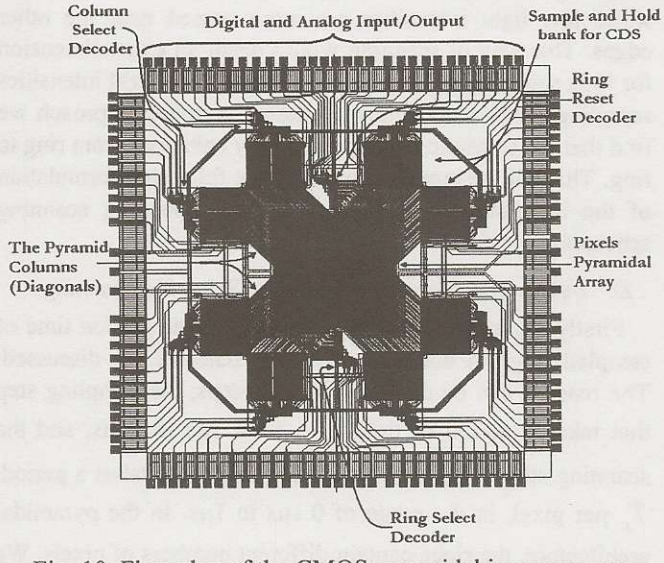

Fig. 10. Floorplan of the CMOS pyramidal image sensor.

The inner ring with the pixel physical layout description is

shown in Fig. 11.

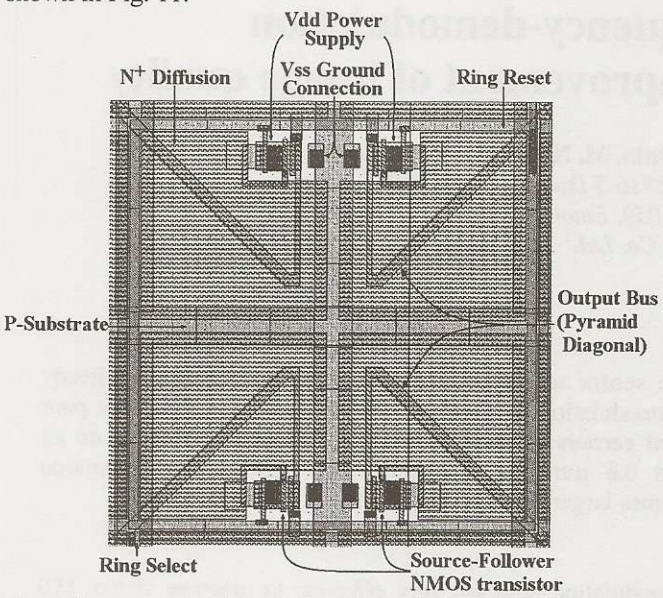

Fig. 11. Layout of the inner ring of the pyramidal image sensor along with pixel description.

The chip has been fabricated through Canadian Microelectronics Corporation (CMC) at TSMC and has been recently received. Testing is currently ongoing, but a preliminary image captured with the pyramidal sensor is shown in Fig. 12. The image in Fig. 12 is captured using a rolling ring scanning in which scanning starts from the outer ring towards the inner ring then the scanning is back again at the outer ring for the next frame.

IV. Conclusion

The pyramidal CMOS image sensor was primarily suggested to inherit the parallel 2D acquisition of the biological vision such as the fovea. Due to the dimensionality of the sampling ring structure, many applications that required fast imaging can be targeted. This is argued by the fact that for a pixel out buffered from a cluster of the imager the imager output also 7 other pixel from the other cluster in parallel. This comes at the cost of off chip memory requirement to build the final image. Furthermore, The ring selection ability is very important feature which allows us to fixate the acquisition to just the central part, using a minimum number of rings, thus reducing the data to be sampled, and hence increasing speed of processing and acquisition. This might be very useful in videophones that need to minimize video data transmission. This feature, which is very similar to the function of the iris in the human eye, is unachievable with the raster scanning CMOS image sensors due again to the additional dimensionality of the sampling structure. Finally, our pyramidal sampling architecture is an attempt to pave the way for more smart acquisition systems to be easily handled by machines or robots.

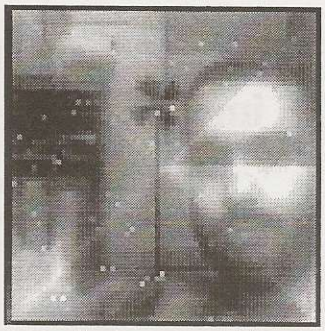

Fig. 12. Image captured using pyramidal architecture CMOS image sensor.

Acknowledgment

The authors are grateful for fabrication support provided by the Canadian Microelectronics Corporation. Funding from NSERC Canada and Betacom Inc. are also gratefully acknowledged. Fruitful discussions with members of the Integrated Camera Group are greatly appreciated.

References

[1] E. R. Fossum, "Active pixel sensors: are CCD's dinosaurs?," Charge-Coupled Devices and solid-state Optical Sensors III, Proc. SPIE 1900, pp. 2-14, 1993.

[2] P.B. Denyer, L. Mingying, D. Renshaw, W. Guoyu, "CMOS image sensors for multimedia applications," in Proc. of the IEEE Custom Integrated Circuits Conference, 9-12 May 1993, pp. 11.5.1-11.5.4.

[3] Alireza Moini, "Vision Chips," Kluwer Academic Publishers, 1999

[4] G. P.Weckler, "Operation of p-n junction photodetectors in a photon flux integration mode," *IEEE J. Solid-State Circuits*, vol. SC-2, pp. 65–73, 1967

[5] B. A. Wandell, "Foundations of Vision," Sinauer Associates, Inc. Publishers. Sunderland, Massachusetts, 1995

[6] O. Yadid-Pecht, E. R. Fossum, "Wide Intrascene Dynamic Range CMOS APS Using Dual Sampling," IEEE Trans. Electron Devices, vol. 44, no 10, pp. 1721-1723, October 1997.

[7] B. Wandell, P. Catrysse, J. DiCarlo, D. Yang and A. El Gamal , "Multiple Capture Single Image Architecture with a CMOS Sensor," In *Proc. of the Intern. Symp. on Multispectral Imaging and Color Reproduction for Digital Archives*, pp. 11-17, Chiba, Japan, October 21-22 1999.