A Low-power Data Acquisition System for Image Contrast Detection

L. Gasparini[1], D. Macii[2], M. Gottardi[1], D. Fontanelli[2]

[1]*Smart Optical Sensors and Interfaces – Fondazione Bruno Kessler (FBK)*
Via Sommarive 18, 38123, Trento, Italy
E-mails: gasparini@fbk.eu, gottardi@fbk.eu

[2]*DISI – University of Trento*
Via Sommarive 14, 38123, Trento, Italy
E-mails: macii@disi.unitn.it, fontanelli@disi.unitn.it

Abstract- This paper deals with a data acquisition system based on a low-power digital sensor for image contrast detection. This system is suitable for geometric pattern recognition as well as for event monitoring. Possible applications are: surveillance, people access monitoring, and road lines recognition in car-like vehicles. The adopted imager was developed by the *Fondazione Bruno Kessler* (FBK), Trento, Italy, and it is particularly useful whenever the available energy budget and the channel communication bandwidth are too limited to use common video cameras. Since neither the standard testing procedures for Analog-to-Digital Converters (ADCs) nor those for cameras are suitable to assess the performance of the proposed system, one possible ad-hoc testing methodology is also described and some experimental results are reported.

I. Introduction

Digital cameras generally consist of Complementary Metal-Oxide Semiconductor (CMOS) or Charge-Coupled Devices (CCDs) pixel matrices, which directly return the digitally encoded light values associated with each pixel, through multiple ad-hoc analog to digital converters (ADCs) [1]-[2]. Due to the increasing diffusion of vision-based measurement systems, e.g., for robotic as well as for mechanical monitoring and biomedical applications [3]-[5], accurate modeling and testing procedures of image acquisition systems are increasingly important. Unfortunately, the metrological characterization of vision-based instruments is still a quite open and challenging research field in spite of some recent advances [6]. In fact, neither the standard IEEE 1241 for ADC testing, nor the photography-oriented camera testing procedures are suitable when atypical image sensors are used [7]-[9]. This is the case of low-power, low-resolution, grayscale cameras for event monitoring applications, whose main goal is to recognize specific patterns or to detect possible moving objects, regardless of the details of a traditional photograph. In fact, if just some limited and well-defined amount of information has to be detected from a given image, acquiring a full-color, high-resolution picture is an unnecessary waste of energy and communication bandwidth, which might not be even bearable when battery-operated wireless devices are employed. In order to reduce acquisition time, amount of transferred data and power consumption, several image contrast sensors have been developed in the last years [10]-[12]. The main advantage of such sensors is their ability to detect only the light contrast patterns from the scene in view, possibly at a higher frame rate than usual cameras and with a reduced energy dissipation. In [10], Ruedi et al. propose a 128x128 pixels image sensor extracting intensity and direction of spatial contrast from four neighboring pixels. However, pixels look fairly complex and large, requiring more than 50 transistors and the total sensor power consumption is quite high, i.e. about 300 mW at 3.3 V. In [12], Kim et al. present a 3.3 V imager with 128x128 pixels, 11 transistors per pixel and 40% fill factor. In this case, power consumption is much lower (i.e. in the order of 1.2 mW when a bright contrast contour is detected), but a further reduction is possible. This is shown in Section II, where a more efficient image contrast sensor is described, along with the corresponding data acquisition system. In Section III, a possible testing procedure is presented. Finally, in Section IV some experimental results are reported.

II. The Image Acquisition System

A. Image Sensor Description

The present imager is a custom, 128×64 pixels, 35-μm CMOS sensor with image processing capabilities on the focal plane. The imager detects the light contrast of the observed scene and it performs on-chip frame differencing [13]. Each output frame results from a three-stage process: i.e. (1) image acquisition, (2) frame differencing and (3) address-based read-out.

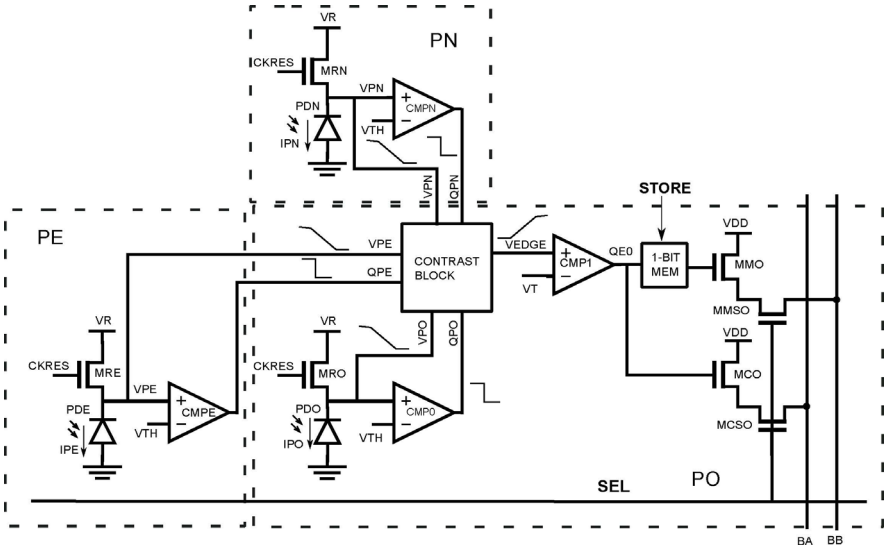


Fig. 1 — *Kernel* schematic for 1-pixel contrast extraction.

In the first stage, the voltage proportional to the luminance sensed by each photodiode of the pixel matrix is integrated as soon as the voltage across the photodiode exceeds a reference threshold V_{th}. The user can set the integration time, which ranges between 1 ms and 100 ms. Afterwards, each value of contrast is estimated from triples of adjacent pixels (referred to as PO, PN and PE), as shown in Fig. 1. In the following we will refer to each triple of pixels as a *kernel*. A *Contrast Block* (CB) generates a voltage value V_{EDGE} proportional to the ratio between the maximum luminance difference detected by the photodiodes of the kernel pixels and the corresponding minimum luminance within the kernel. V_{EDGE} is then compared with a threshold voltage level and the resulting quantized value can be either directly connected to several bit-lines (one per matrix column) for immediate read-out, or it can be stored into a 1-bit memory cell. Bit storage occurs anytime the *current frame* has to be used as a *reference frame* for subsequent frame differencing.
In the second stage, the *output image* is obtained from the pixel-by-pixel difference between the *current frame* and the *reference frame*. Thus, each pixel in the output image can just take the values -1, 0, or +1. Notice that if the reference frame is empty, then the output values can just be either 0 or 1.
In the third stage, two read-out modes are supported: *active mode* and *idle mode*. In the *active mode* the sensor scans the output image, row by row, and it returns both the column address and the sign of non-zero pixels only. Data read-out is asynchronous: a flag signal is generated anytime new valid data are placed on the output pins of the chip. Since just those pixels that are different in the *current* and *reference* frames are actually transferred, a considerable amount of energy and bandwidth is saved when data are transmitted to the PC. Further savings can be achieved in the *idle mode*, because in this case only the number of non-zero pixels is returned by the sensor. This is particularly useful for motion detection. In fact, if each frame is saved as a temporary reference, then some moving object can be detected by simply monitoring the variations in the number of active pixels over time. In Table I some basic features of the image sensor are shortly reported. Each pixel consists of 45 transistors and it is approximately square in size, i.e. 26x26.5 μm, with a fill factor of 20%. The average power consumption of the sensor ranges between 30 μW and 90 μW in *idle mode* and in *active mode*, respectively. Such values are much smaller than those of other similar sensors, such as those described in [10]-[12]. The power consumption of the various sections of the sensor was estimated by measuring the average current drawn from a 3.3 V stable power supply under stationary operating conditions, i.e. with an integration time equal to 10

PARAMETERS	VALUE
Technology	CMOS 0.35 μm 4M
Array size	128 x 64 pixels
Pixel size	26 μm x 26.5 μm
Fill factor; N of transistors	20% ; N = 45 T
Power consumption @ 50 fps and 25% activity	
Active Mode	90 μW
Imager	20 μW
Digital (core + pads)	70 μW
Idle Mode	30 μW
Imager	20 μW
Digital (core + pads)	10 μW

Table I – Some basic features of the CMOS light contrast image sensor.

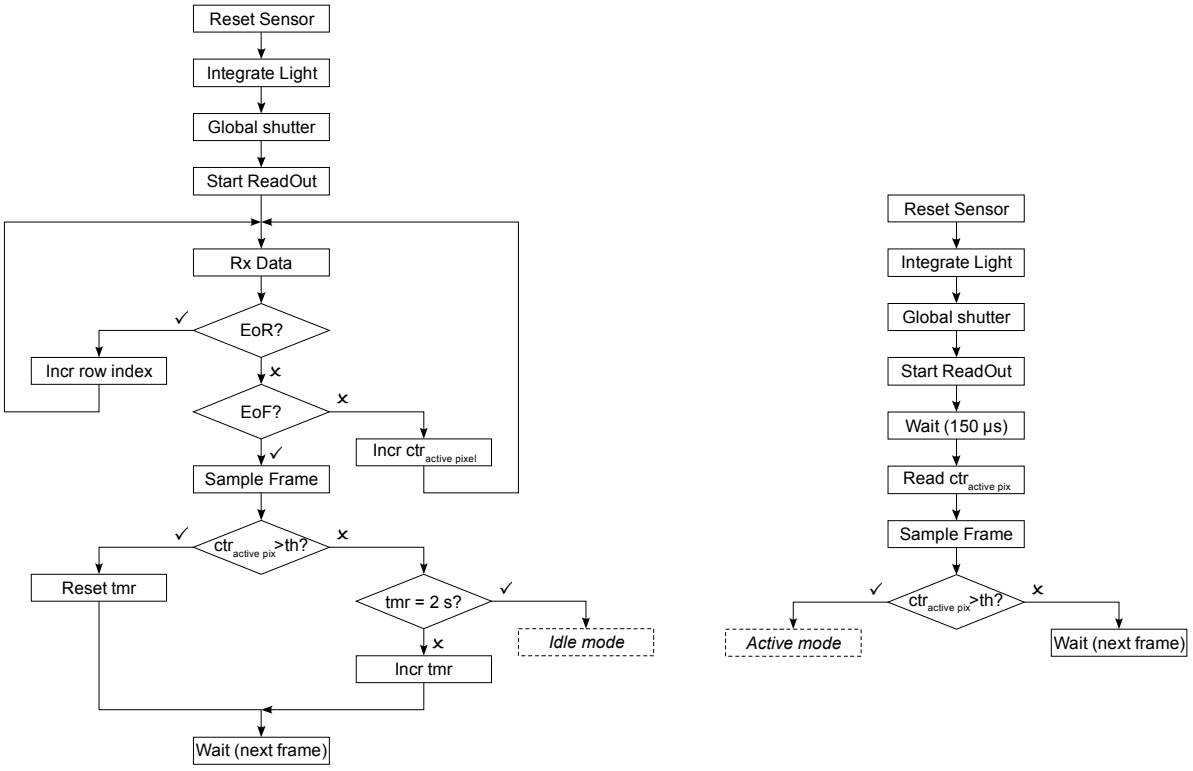


Fig. 2 – Data acquisition procedures in *active mode* (a) and in *idle mode* (b).

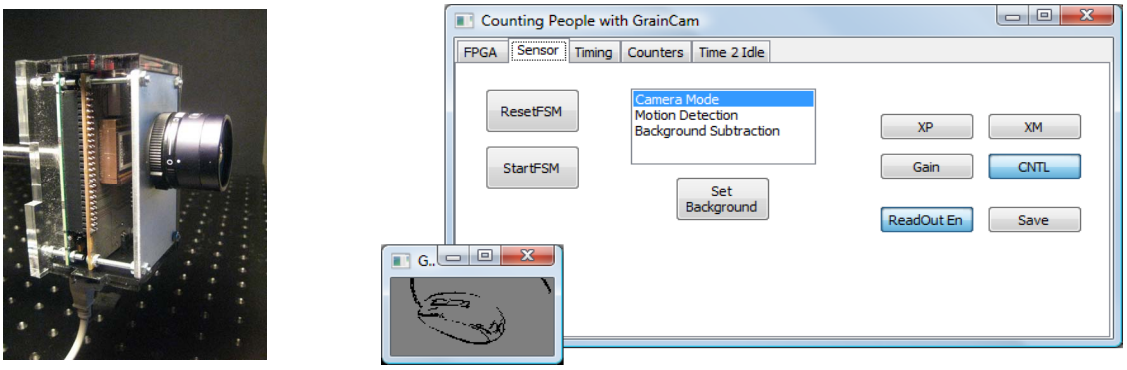
ms, an acquisition rate of 50 frame/s, and constant artificial light conditions. During the test about 25% of pixels were kept active just using a simple black-and-white line pattern. The average current drain was measured by a digital multimeter Agilent 34411A, with 6 and ½ digits, remotely controlled by NI Labview, according to a procedure similar to that described in [14].

B. Data Acquisition

The read-out and data acquisition circuitry connected to the sensor consists of an FPGA-based board linked to a host PC via a USB port. The FPGA-based board is an Opal Kelly XEM 3001v2 equipped with a Xilinx Spartan-3 XC3S400. A picture of the whole system is shown in Fig. 3(a). The control unit implemented in the FPGA generates the signals to drive the imager into the *active* or *idle* mode, respectively according to the flow charts shown in Fig. 2. When the number of non-zero pixels falls below a user-defined threshold for a given amount of time, the sensor switches from *active* to *idle*, whereas it returns to *active mode* as soon as the number of non-zero pixels exceeds the chosen threshold again. In *active mode*, data are temporarily stored into a FIFO buffer before being transferred to the host PC. The maximum achievable data acquisition rate, in the order of 100 frame/s, is lower than the best possible frame rate of the sensor, as it is limited by the operating system latencies, the USB endpoint access delays and the execution time of the image reconstruction algorithms running on the PC. However, parameters such as the frame rate, the pixel integration time and the threshold value for active-to-idle mode switching can be set through a Python Graphic User Interface (GUI). A screenshot of this application is shown in Fig. 3(b). The same GUI can be also used to set the *reference frame*. This frame can be empty (in this case the visualized output image is just the contrast pattern), it can be fixed by the user to have a static background to build a differential image, or it can be updated at a constant rate to detect motion. The GUI also displays the stream of frames and enables users to save data on the PC for further off-line processing.

III. Performance Parameter and Estimation Procedure

Unlike typical color or grayscale cameras, the proposed system just detects the maximum luminance contrast ratio between triples of adjacent pixels. Therefore, the image acquisition system returns a three-color image, which reduces to a binary black-and-white one when the *reference frame* is empty. Accordingly, some typical camera-oriented performance parameters such as the exposure index, the ISO film speed ratings and the dynamic



(a) (b)

Fig. 3 – Acquisition system picture (a) and screenshot of the GUI for image acquisition control (b).

range do not apply in our case. Other parameters, such as the *sensitivity* and the *noise*, can not be estimated using the procedures described in standard documents such as the ISO 15739:2003 or the ISO 14524:2009, and they need to be defined in a different way. Also, no univocal definition of *contrast* exists. The most widely accepted ones are the *Weber contrast* and the *Michelson contrast*, which are defined respectively as:

$$C_W = \frac{L_x - L_b}{L_b} \qquad \text{and} \qquad C_M = \frac{L_{MAX} - L_{MIN}}{L_{MAX} + L_{MIN}} \tag{1}$$

where L_x is the luminance of a given feature in the image, L_b is the luminance of the background (usually dark), whereas L_{MAX} and L_{MIN} are the maximum and minimum luminance in the scene. Notice that, in both cases the contrast is defined as an adimensional quantity. Also, if L_x=L_{MAX} and L_b= L_{MIN}, the two contrast metrics are related, i.e. $C_W = \frac{2C_M}{1-C_M}$.

In the following we will regard the whole system as a special *contrast* ADC, in which the K quantization thresholds, instead of being related to K distinct amplitude levels of the input signal, refer to the same level of Weber contrast in K distinct points of the image plane. Ideally, if no reference frame is used and if we assume to apply the same levels of luminance L_1 and L_2 (with L_1>L_2) to each pair of adjacent pixels through a proper checker-board patterned test image, the sensor output results from:

$$\boldsymbol{V}_k = Q_{\boldsymbol{T}_k}\left[C + \boldsymbol{n}_k\right] \qquad k=0,\ldots,K\text{-}1 \tag{2}$$

where $C = \frac{L_1 - L_2}{L_2}$ is the ideal Weber contrast ratio measured by each kernel, $\boldsymbol{n}_k$ is a random variable modeling the noise affecting the contrast detection and $Q_{\boldsymbol{T}_k}[\cdot]$ is the 1-bit quantization operator related to the contrast threshold $\boldsymbol{T}_k$ of the k-th pixel. Notice that the values of $\boldsymbol{T}_k$ for k=0,…,K-1 slightly differ from one another due to the circuitry tolerances, i.e. $\boldsymbol{T}_k$=T+$\boldsymbol{\varepsilon}_k$, where the sequence of random variables $\boldsymbol{\varepsilon}_k$ can be assumed to be normally distributed with zero-mean and standard deviation σ_T. Similarly, in a first approximation, $\boldsymbol{n}_k$ can be simply modeled as a sequence of normally distributed random variables with zero-mean and standard deviation σ_n over the whole pixel matrix. Thus, the activation probability of pixel k is:

$$\boldsymbol{p}_k = \Pr\{C > \boldsymbol{T}_k\} = 1 - \Phi\left(\frac{\boldsymbol{T}_k - C}{\sigma_n}\right) = 1 - \Phi\left(\frac{T - C + \boldsymbol{\varepsilon}_k}{\sigma_n}\right) \qquad k=0,\ldots,K\text{-}1 \tag{3}$$

where $\Phi(\cdot)$ is the normal cumulative distribution function. If the second-order Taylor series expansion of (3) is computed around 0, after a few steps, it can be easily proved that the mean value of the pixel activation probability is:

$$p = E[\boldsymbol{p}_k] \approx 1 - \Phi\left(\frac{T - C}{\sigma_n}\right) + \frac{T - C}{2\sigma_n^3\sqrt{2\pi}} e^{-\frac{(T-C)^2}{2\sigma_n^2}} \sigma_T^2 \tag{4}$$

However, in practice, the actual value of p is smaller than what reported in (4) either because some pixels could be faulty or because exciting each kernel exactly with the same luminance difference is impossible due to optics distortions, pixel to checker-board square misaligment and pixel size differences. Therefore, if we assume in a

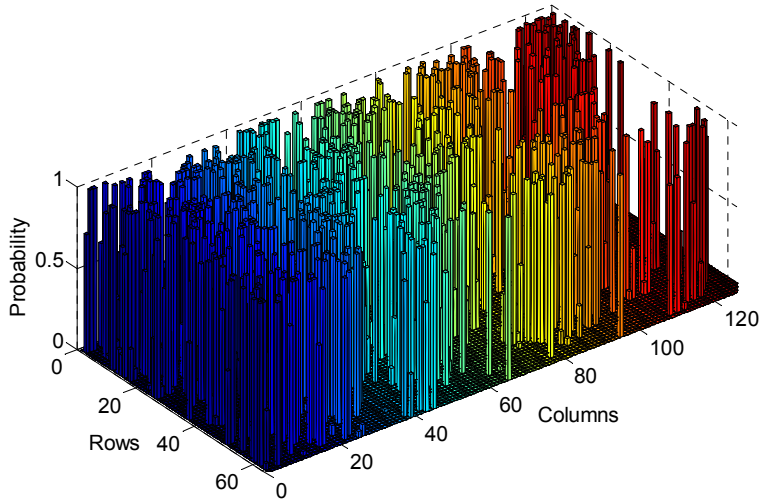


Fig. 4 – Pixel-by-pixel activation probability histogram for C_m≈15, with integration time equal to 10 ms.

first approximation, that this reduction in pixel activation probability affects uniformly the whole pixel matrix, (4) can be rewritten as

$$p' = E[\boldsymbol{p}_k] \approx A\left(1-\Phi\left(\frac{T-C}{\sigma_n}\right)+\frac{T-C}{2\sigma_n^3\sqrt{2\pi}}e^{-\frac{(T-C)^2}{2\sigma_n^2}}\sigma_T^2\right) \tag{5}$$

where the coefficient A<1 implies that, even when the pixel-by-pixel contrast level is maximum, the overall activation probability can never reach 1. Thus, if M average pixel activation probabilities $\bar{p}'_m$ corresponding to different levels of contrast C_m with m=1,..., M are measured with negligible uncertainty using large records of N collected images, the four model parameters T, σ_T, σ_n and A in (5) can be estimated through some nonlinear (e.g. least mean squares) fitting procedure. Also, we can determine the *contrast noise power* σ_n^2, the *q-level sensor sensitivity* S_Q (i.e., the contrast level for which the pixel activation probability is equal to Q) and the Signal-to-noise Ratio (SNR), i.e. $SNR = 10\,\boldsymbol{log}_{10}\frac{C_m^2}{\sigma_n^2}$. Notice that, according to the proposed definition of S_Q, $S_{0.5}$=T.

IV. Experimental Results

In order to evaluate the performance of the system according to the procedure described in Section III, some experiments were conducted in a laboratory equipped with an optical bench. M=10 checker-board images were used for testing. This kind of pictures is indeed also commonly used for estimating the ANSI contrast ratio in digital displays. Each printed image consists of 128x64 squares with a different Weber contrast level, in the range 10%-50%. Out of this range the image detected by the sensor does not change significantly, i.e. pixels are either mostly off or on. The level of contrast was set by properly adjusting the level of grayscale of the dark squares of each test pattern. The illuminator used for testing is a Thorlabs high-intensity fiber-coupled light source. The maximum illumination generated by the lamp is about 430 klx. The illuminator was placed about 20 cm in front of the test images in order to keep the luminance as much uniform as possible over the area observed by the sensor. The focal distance was adjusted accordingly. Since two pixel columns of the sensor are faulty and no perfect checker-board-to-pixel-matrix alignment is possible, no more than 82% of pixels can be active even under strong contrast conditions. For each test image, approximately N=500 frames were collected. The corresponding individual activation probabilities were estimated from the ratio between the number of times in which each pixel is active and N. Fig. 4 displays an example of pixel-by-pixel activation probability histogram for C_m=15%. Fig. 5 instead shows the result of the least squares nonlinear fitting procedure described in Section III. The model parameter values defined in (5) are: T=0.19, σ_T=1, σ_n=0.06 and A=0.81. Accordingly, the 0.75-level contrast sensitivity is $S_{0.75}$=0.28 and the noise power is σ_n^2=3.6·10^{-3}. Quite unexpectedly, we did not observe any significant dependence of such performance parameters on the integration time of the sensor.

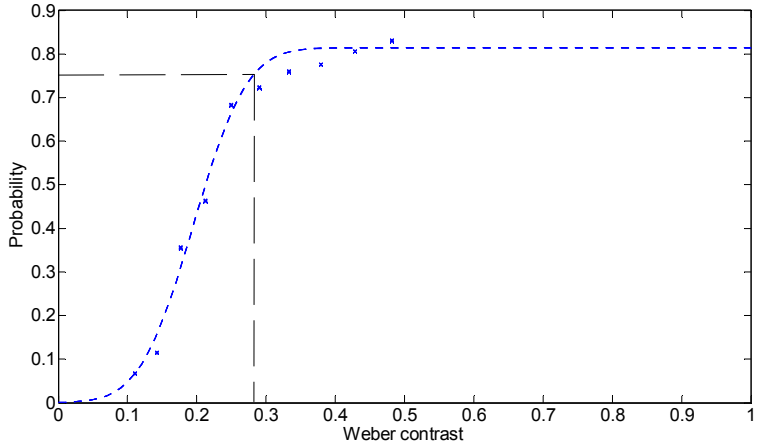


Fig. 5 – Estimated activation pixel probability (dotted line) as a function of the image light contrast. The cross-shaped dots refer to the experimental data.

V. Conclusions

Data acquisition systems based on ultra-low power image sensors rely on specific quantization schemes, which can be hardly classified as either typical ADCs or digital cameras. For this reason, no standard modeling and testing procedures can be used to analyze their performance. In this paper, one of such systems is described and some custom performance parameters, based on an ad-hoc model, are reported. In spite of the specific features of the system considered, the proposed modeling and testing methodology might be extended also to other similar digital contrast imagers with moderate resolution.

References

[1] A. Simoni, G. Torelli, F. Maloberti, A. Sartori, M. Gottardi, L. Gonzo, "256×256 P ixel CMOS digital camera for computer vision with 32 algorithmic ADCs on board," Proc. *IEE Circuits Devices and Systems*, Vol. 146, No. 4, pp. 184–190, Aug. 1999.

[2] O. Elkhalili, O. M. Schrey, P. Mengel, M. Petermann, W. Brockherde, B.J. Hosticka, "A 4×64 pixel CMOS image sensor for 3-D measurement applications," *IEEE Journal of Solid-State Circuits*, Vol. 39, No.7, pp. 1208- 1212, Jul. 2004.

[3] M. Ribo, M. Brandner, "State of the art on vision-based structured light systems for 3D measurements," Proc. *Int. Workshop on Robotic Sensors: Robotic and Sensor Environments,* pp. 2-6, Ottawa, Canada, Sep.-Oct. 2005.

[4] R. Anchini, G. Di Leo, C. Liguori, A. Paolillo, "Metrological Characterization of a Vision-Based Measurement System for the Online Inspection of Automotive Rubber Profile," *IEEE Transactions on Instrumentation and Measurement,* Vol. 58, No. 1, pp. 4 - 13, Jan. 2009.

[5] N. Faramarzpour, M. El-Desouki, M. J. Deen, Qiyin Fang, S. Shirani, L. W. C. Liu, "CMOS imaging for biomedical applications," *IEEE Potentials,* Vol. 27, No.3, pp.31-36, May-June 2008.

[6] M. Brandner, "Graphical modelling of measurement uncertainties in vision-based metrology," Proc. *IEEE International Workshop on Advanced Methods for Uncertainty Estimation in Measurement,* pp. 28-33, Bucharest, Romania, Jul. 2009.

[7] IEEE Std. 1241-2000, "IEEE Standard for terminology and test methods for analog to digital converters."

[8] *Photography—Electronic Still-Picture Cameras—Methods for Measuring Opto-Electronic Conversion Functions (OECFs)*, ISO 14524:2009, Feb. 2009.

[9] *Photography—Electronic Still-Picture Imaging—Noise Measurements*, ISO 15739:2003, May 2003.

[10] P.-F. Ruedi, P. Heim, F. Kaess, E. Grenet, F. Heitger, P.-Y. Burgi, S. Gyger, P. Nussbaum, "A 128×128 pixel 120-dB dynamic-range vision-sensor chip for image contrast and orientation extraction," *IEEE Journal of Solid-State Circuits,* Vol. 38, No. 12, pp. 2325- 2333, Dec. 2003.

[11] U. Mallik, M. Clapp, E. Choi, G. Cauwenberghs, R. Etienne-Cummings, "Temporal change threshold detection imager," Proc. *IEEE International Solid-State Circuits Conference* (ISSCC), pp. 362-603, San Francisco, USA, Feb. 2005.

[12] D. Kim, E. Culurciello, "A Compact-pixel Tri-mode Vision Sensor," *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, pp. 2434 – 2437, Paris, France, May 30 – June 2, 2010.

[13] M. Gottardi, N. Massari, S. A. Jawed, "A 100 µW 128×64 Pixels Contrast-Based Asynchronous Binary Vision Sensor for Sensor Networks Applications," *IEEE Journal of Solid State Circuits*, Vol. 44, No. 5, pp.1582-1592, May 2009.

[14] D. Macii, D. Petri, "Accurate Software-Related Average Current Drain Measurements in Embedded Systems," *IEEE Trans. on Instrumentation and Measurement*, Vol. 56, No. 3, pp. 723-730, Jun. 2007.