CMOS realisation of analogue processor for early vision processing

W. JENDERNALIK*, J. JAKUSZ, G. BLAKIEWICZ, R. PIOTROWSKI, and S. SZCZEPAŃSKI

Department of Microelectronic Systems, Gdańsk University of Technology, 11/12 Narutowicza St., 80-952 Gdańsk, Poland

Abstract. The architecture concept of a high-speed low-power analogue vision chip, which performs low-level real-time image algorithms is presented. The proof-of-concept prototype vision chip containing 32×32 photosensor array and 32 analogue processors is fabricated using a 0.35 μm CMOS technology. The prototype can be configured to register and process images with very high speed, reaching 2000 frames per second, or achieve very low power consumption, several μW. Finally, the experimental results are presented and discussed.

Key words: CMOS imager, analogue processor array, smart sensor, vision chip.

1. Introduction

Image processing is frequently used in systems for monitoring and controlling of objects to help in effective management of their resources and safety, and robots controlling [1]. The practical systems for monitoring big objects, like for example complex of buildings, comprise many vision sensors recording images that have to be transmitted to, and processed in the central unit. One of the most challenging problem in such cases is effective transmission and processing of huge amount of image data. To avoid overloading of transmission channels and a central unit some early vision algorithms are frequently performed at the sensors by an integrated low-level image processor. As a result the rough image data, generated by the sensor, can be compressed or replaced by useful information extracted from images. This approach significantly improves the overall efficiency and the cost of a system implementation by relaxing requirements on throughput of transmission channels and demands on processing speed of the central unit. A complete vision chip consisting of a photo-detector array and accompanying low-level vision processors can effectively be implemented in contemporary CMOS technologies on a single substrate, which becomes a very good platform for realisations of modern and cheap image sensors.

The very early solutions of integrated vision chips were mostly dedicated to a specific image algorithm and could not be reconfigured. The next generations of programmable vision chips, designed in multiple instruction multiple data (MIMD) or single instruction multiple data (SIMD) architectures, were able to perform several image algorithms [2–5].

The newest chips have fully programmed architectures with a parallel analogue data processing, which significantly reduces the time necessary for an image processing [6–9]. Although, the recent development in vision chip implementations is great, there are still some areas for improvements. Most of the reported vision chips realise the convolution algorithms based on a reduced kernel, where only four neighbouring pixels (the top, bottom, left, and right) from the full 3×3 kernel are taken into calculations. The consequence of

the omitting the diagonal pixels is degradation of the resulting image. Some of the reported implementations [8, 9] have serious speed limitation resulting from the sequential manner of instruction processing, which means that a typical convolution algorithm requires several clock cycles to calculate single data. There is also much to do with reduction of power consumption.

The presented in this paper CMOS implementation of a vision chip addresses some of the mentioned weaknesses of the known solutions. The key advantage of the proposed chip is very small power consumption, which can be reduced to several μW depending on the speed of image recording and processing. The circuit can also be reconfigured to have very high speed of image processing reaching 2000 frames per second. The architecture of the low-level analogue processors enable the convolution calculation using the full 3×3 kernel, where the kernel coefficients can be reprogrammed on the fly, even during the image processing. This property can be exploited in improving the efficiency of the image processing or image compression. The convolution calculation for a single pixel is completed in a single clock cycle for the full 3×3 kernel.

Section 2 of the paper presents the architecture of the proposed implementation. Section 3 discusses details of the photo-pixel, and low-level analogue processing element (APE). Section 4 presents a test set-up, and results of the chip measurements. The last section contains the final discussion and conclusions.

2. Architecture of a vision chip

The sensor architecture was designed to improve performance of the vision chip in comparison to reported known solutions. A special attention was paid to minimisation of the fabrication cost and achieving a good trade-off between speed of the image processing and power consumption. The possible architectures and circuit solutions can be greatly simplified because the primary application of the sensor is monitoring and controlling of objects, which can be done based on low

*e-mail: waldi@ue.eti.pg.gda.pl

W. Jendernalik, J. Jakusz, G. Blakiewicz, R. Piotrowski, and S. Szczepański

or medium image resolutions and dynamics. Typically, accuracy of image signal processing equivalent to 5 or 6 bits resolution is sufficient [10] for satisfactory results. The relatively low accuracy of such systems allows application of analogue signal processing, which contrary to digital one, enables greater reduction of power consumption and a chip area. The considered low-level image processing algorithms like for example: convolution filtering, smoothing, edge detection or segmentation, require only four basic operations: addition, four-quadrant multiplication or four-quadrant division, level discrimination, and storage of signal samples. From the point of view of power consumption and chip area reduction, the multipliers or dividers are the most challenging functional blocks which need special attention. The key difficulty in implementation of low-level image processing is its numerical intensity. For example, even a low-resolution 128×128 pixel array working with 25 Hz frame frequency requires about 3.7 million multiplications and 3.3 million additions per second to calculate 3×3 kernel convolution filtering.
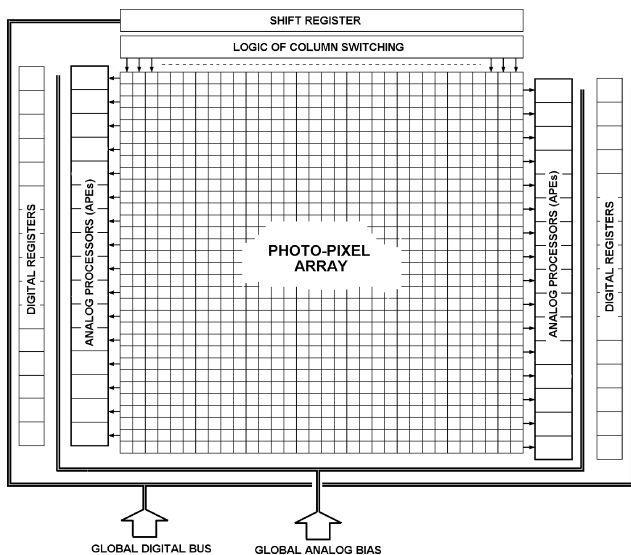

Fig. 1. Architecture of the proof-of-concept prototype chip

To achieve a good trade-off between the factors discussed above, a SIMD architecture using analogue processors containing multipliers and summers is selected. The general architecture of the prototype chip is shown in Fig. 1 [11]. The photo-pixel array in the proof-of-concept prototype is of 32×32 pixels and is located in the centre of the chip. Two sets of APE are arranged in columns placed on the right and left sides of the pixel array. APE assures parallel and very fast processing of image signals coming from each row of the array. This kind of circuit arrangement enables relatively simple and short signal paths, having small stray capacitances, which help to achieve fast signal transmission and reduction of power dissipated on switching. By making the photo-pixel sensor as simple as possible, containing only the necessary circuits for signal acquisition, it is possible to significantly reduce the array dimensions, and reduce the length of signal paths. The logic circuits control the order of processed image samples.

The presented in Fig. 1 architecture guarantees, that in a single master clock cycle a complete result for a pixel is calculated, regardless of complexity of the realised image algorithm. This is achieved due to direct processing of a complete set of nine signals coming from all the neighbouring photo-pixels, as shown in Fig. 2.
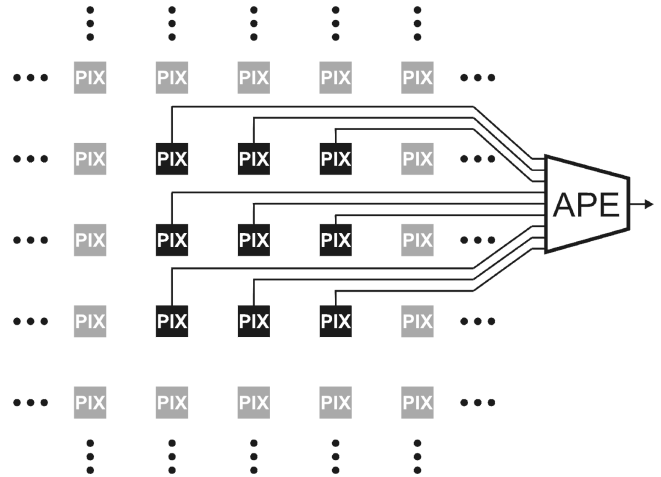

Fig. 2. Mettalic signal paths arrangement for a single APE

The time needed for finishing calculations for a single image frame is equal to product of number of columns and the clock period. Because of very fast response of APE, the time can be made very small.

3. Vision chip implementation

3.1. Photo-pixel. The photo-electrical conversion in the sensor module is performed by the high frame rate CMOS device. All pixels in the array process images at the same time slot due to using a global electronic shutter, that eliminates image smearing caused by fast movement of object. The shutter time can be varied in order to adjust photosensitivity to different illumination conditions. The important feature of this CMOS sensor is complete decoupling of the shutter time from the readout pixel clock. This feature enables application of the sensor to optical measurements, like for example object speed evaluation, which is especially useful in monitoring systems. The schematic of an active pixel sensor (APS) is shown in Fig. 3. The pixel is composed of five functional circuits: photodiode formed of p- and n-well diffusions, the reset transistor M_1, the source-follower M_2 biased by the current sink M_3, the shutter switch M_4, the storage capacitor C_{MEM}, and the buffer for non-destructive readout and driving a row line. For the purpose of reduction of power consumption, the drain current of M_3 is reduced to zero after closing the shutter switch. Furthermore, the signal "Enable" activates the output buffer only during the read time of the selected column. As a result, each functional circuit is only activated while needed for a short time interval.

The image sensor shows a linear response to illumination and a good noise performance [12]. The photodiode area is 131 μm^2, which yields a sensor fill factor of 8%. To max-

imise pixel sensitivity no metal wires pass over the photodiode region. With 500–1000 lux illumination the shutter time is within the range of 10–20 μs.

A single cycle of the circuit operation for each image frame starts with resetting, which is achieved by applying low value of the signal "Reset" (see Fig. 3) and charging the photodiode capacitance C_D to V_{DD}. At the same time the signal "Shutter" is set to V_{DD} and C_{MEM} is charged to $V_{REF} = V_{DD} - V_{GS1} \approx 2.1$ V. When the signal "Reset" goes to V_{DD}, the photodiode current integration starts. The discharging speed of C_D and C_{MEM} is proportional to the energy of incident light, which guaranties a linear photo-electrical conversion. At the end of the integration time, the shutter switch is opened, and the final voltage is stored on the capacitor C_{MEM}. The readout of the voltage stored on C_{MEM} is started by activating the "Enable" signal of the pixel output buffer.
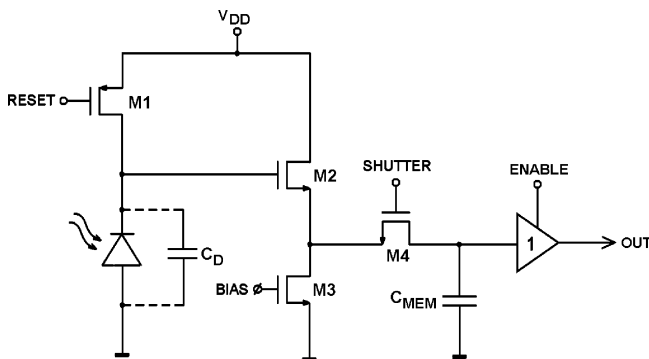

Fig. 3. The simplified scheme of the active pixel sensor (APS)
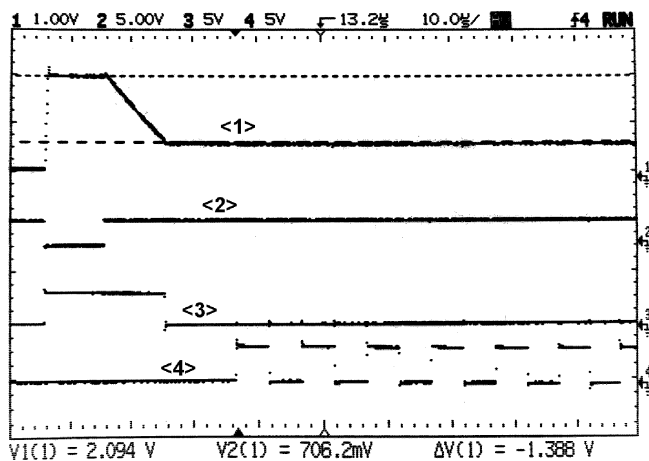

Fig. 4. The measured pixel signal waveforms: <1> voltage on C_{MEM}, <2> reset, <3> shutter, <4> column readout clock

The exemplary waveforms, illustrating operation of the circuit, are shown in Fig. 4. The waveform <1> shows voltage on the memory capacitor, which represents image data. The logic signals <2>, <3>, and <4> drive the reset, the shutter, and process of array columns reading, respectively. To improve clarity of Fig. 4, the low level of the reset signal is set to relatively large width of 20 μs, which in reality is about 500 ns. Since the reset reaches high level, the integration

starts, and the voltage <1> decreases linearly until the shutter <3> drops to low. At this moment voltage on C_{MEM} capacitor represents actual image data, and can be further processed. The sequential reading of consecutive array columns can be observed based on signal <4>, which is activated and becomes the master clock synchronising the reading.

In order to make possible simultaneous connection of three adjacent pixels in a line, three multiplexed analogue buses are applied for every row, as it is shown in Fig. 5. Such an arrangement of signal buses allows simultaneous connection of each APE to all necessary nine signals coming from adjacent pixels. The multiplexing is done by digital circuits synchronised by the "Column Readout Clock" signal (see the signal <4> in Fig. 4).
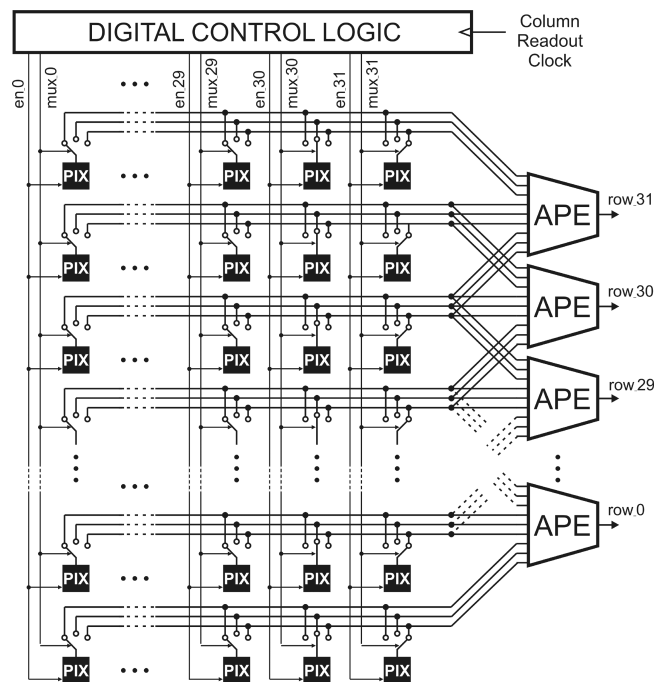

Fig. 5. Pixels readout circuit architecture

3.2. Analogue processor. Most of low-level early vision algorithms operate on a 3×3 pixel kernel to calculate result for a single central pixel. The majority of tasks can be done using a kernel of coefficients in the range of -1 to $+1$ with a 3-bit resolution. A 9-input signed, 3-bit resolution multiplier can complete all the necessary calculations from a circuit point of view. A general architecture is presented in Fig. 6. The APE consists of nine voltage to current converters (V-I), nine passive current scaling ladders (R-2R), and a single output buffer (I-V) adding up all the scaled input signals.

As shown in Fig. 6, the first functional block at each input of APE is a voltage to current converter. The converter separates a photo-pixel from the rest of the circuits and additionally improves overall linearity of the signal processing path. All the converters are identical and based on a linearised MOSFET differential pair as shown in Fig. 7. The transistors M_1 to M_4 form a switching circuit, and are used for changing signal polarity. The differential pair is composed of M_5 and M_6. The parallel connected transistors M_7 and M_8 together with high

W. Jendernalik, J. Jakusz, G. Blakiewicz, R. Piotrowski, and S. Szczepański

resistive polysilicon resistors R_1 and R_2 improve linearity of voltage to current conversion [13]. The parameters of those elements were selected to achieve good linearity at relatively small biasing currents. The voltage to current converter can alternatively be implemented as a linearised transconductance amplifier [14], but this solution increases power consumption.
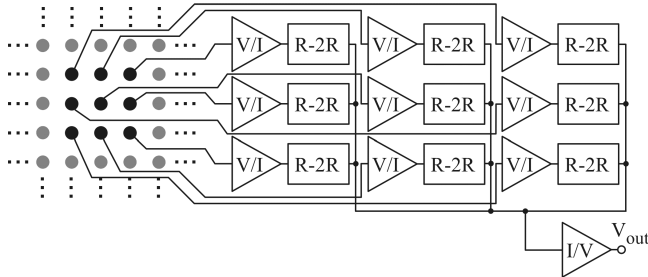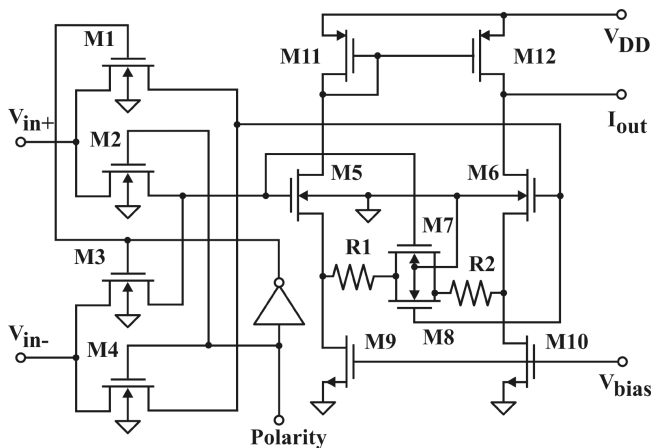

Fig. 6. Architecture of a nine-input APE

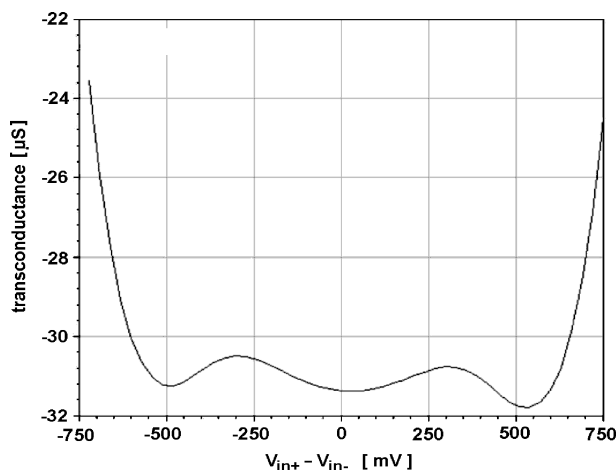
Fig. 7. Voltage to current converter

Fig. 8. Transconductance of the voltage to current converter as a function of a differential input voltage

Figure 8 shows a plot of the measured transconductance of the converter as a function of a differential input voltage. The transconductance characteristic is relatively flat within the range of −0.5 V to 0.5 V, which means good linearity of input voltage to current conversion. The total conversion error is

about 3% for the full input voltage range. The MOSFET-only R-2R ladder [15] is presented in Fig. 9. The circuit has excellent linearity and can be easily programmed, by switching the transistors M_{2a}, M_{2b}, M_{4a}, M_{4b}, which play twofold role of switches and resistors. The binary signals b_{N-1}, b_{N-2},... are used to program the value of the multiplier coefficients.
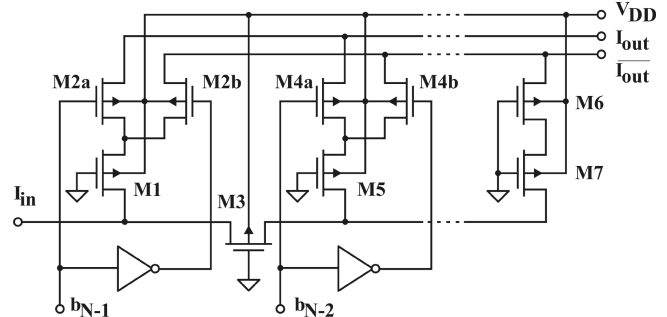

Fig. 9. MOSFET-only R-2R ladder

To achieve good linearity and high accuracy of current division, the output signals I_{out} and $\overline{I_{out}}$ should be connected to low-impedance nodes. The circuit is passive and does not consume any static power. The total processing error is below 0.05% for transistors having dimensions 3 μm×3 μm. The output buffer, the last circuit in a signal processing path, adds together all the currents flowing from each R-2R ladder. In order to preserve high precision of the ladders, the buffer was designed to have very low input resistance. The flipped follower configuration [16] was chosen as the input stage of the buffer, because it can have input resistance as small as 10–20 Ω due to an internal negative feedback. The output buffer comprising two identical followers (transistors $M_{1a(b)}$–$M_{3a(b)}$) and an additional output stage (M_{1c}–M_{3c}) is shown in Fig. 10.
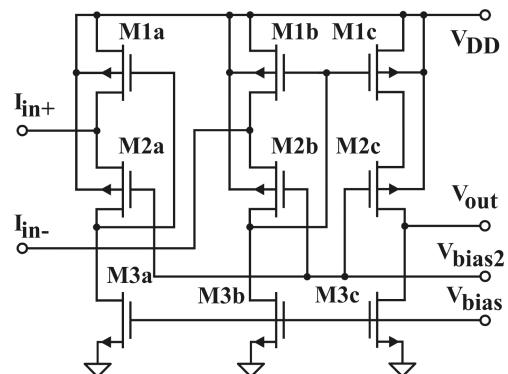

Fig. 10. Low-resistance output buffer

The input resistance of the circuit is practically independent of biasing currents and mainly depends on relations of the transistors aspect ratios. The input resistance is $R_{in} \cong 2/(g_{m1}g_{m2}r_{o1}) = A((W/L)_1(W/L)_2)^{-0.5}$, where g_{m1} and g_{m2} are transconductances of transistors $M_{1a(b)}$ and $M_{2a(b)}$, respectively, r_{o2} is the output resistance of $M_{2a(b)}$, $(W/L)_1$ and $(W/L)_2$ are the aspect ratios (width to length of a channel) of transistors $M_{1a(b)}$ and $M_{2a(b)}$, A is a constant, dependent on a selected technology. The dimensions of transistors $M_{1a(b,c)}$

and $M_{2a(b,c)}$ were chosen to achieve sufficiently low input resistance, which is assumed to be a hundred times smaller than an equivalent output resistance of nine R-2R MOSFET ladders connected together. The biasing currents, generated by transistors $M_{3a(b,c)}$ were adjusted to satisfy the assumed dynamic range $\pm 200~\mu$A. The circuit consumes 600 μA static current, and assures signal processing accuracy better than 0.1%. A complete processor occupies 98 μm $\times$ 220 μm area on a chip including registers for the array coefficients programming, and all biasing circuits.

The complete multiplier is very fast, because the internal parasitic capacitances are relatively small and the signal nodes are of low resistance. The small signal -3 dB bandwidth of the circuit is 90 MHz, whereas the rise and fall times are $t_r = 10.5$ ns and $t_f = 11$ ns, respectively. The time response for the input test signal is presented in Fig. 11. During that test, the input signal where sequentially changing from $+0\nabla000$ to $1\nabla000$. The simulations show that for the full range of the array coefficient values, the total error is within the acceptable limit for 5-bit resolution, which is 3.13%. The broadband frequency and fast time response make the circuit well suited for time division and sequential processing of data flowing from a photoreceptor array.
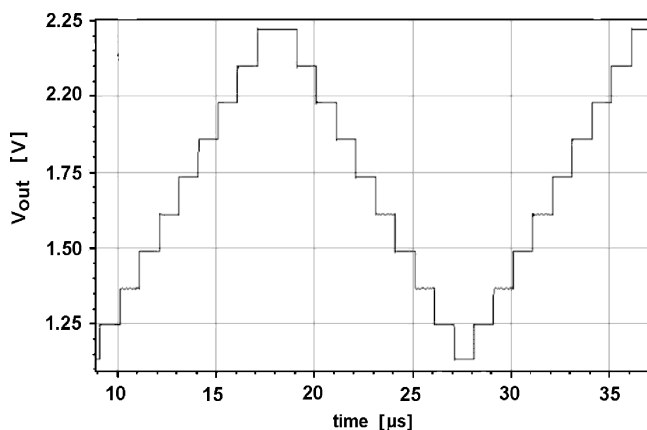

Fig. 11. APE time response for a test sequence

4. Prototype testing and measurements

The experimental image sensor is fabricated in a 0.35 μm CMOS process from AMS for 3.3 V power supply. Its functional testing and characterization were performed using a hardware platform. The hardware part of the imaging system contains a Virtex-4 FPGA. An interface acquisition circuit includes four 8-channel ADC MAX158, high speed TS464 amplifiers and others elements such as lens. Figure 12a shows the prototype chip with a removed top cover to place the lens. Figure 12b presents a test board with peripheral circuits and lens installed on top of the prototype chip. The output signals from the chip, after digitisation, are stored in FPGA, where complete image frames are collected for next stages of processing. All the necessary logic signals controlling operation of the sensor ($<2>$–$<4>$ from Fig. 4) are generated using FPGA platform.
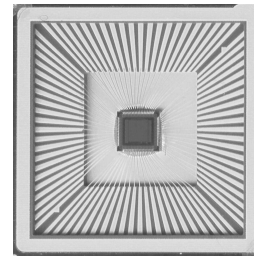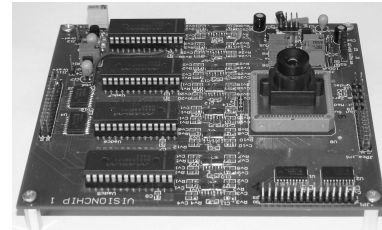

Fig. 12. Prototype vision-chip: a) chip photograph, b) test board

4.1. DC offset and fixed pattern noise. The image sensors always suffer from technology related nonidealities that can limit the performances of the vision system. In the presented visionchip two main sources of nonidealities are: the current offsets generated by the low-resistance output buffers in APEs and a fixed pattern noise (FPN) produced by the pixels. The APE offsets are caused by mismatch between individual biasing currents of the buffers. For a uniform illumination it results in a specific/characteristic image are shown in Fig. 13a. However, DC offset can be easily removed in a data acquisition process after image read. The result of DC offset removing is shown in Fig. 13b where only FPN is present. Fixed pattern noise is the variation in output pixel signals, under uniform illumination, due to a device and an interconnect mismatch across the image sensor. FPN can be reduced by implementing a correlated double sampling (CDS). To implement CDS, each pixel output needs to be read twice, after reset and at the end of integration time. The correct pixel signal is obtained by subtracting those two values. In the presented chip, CDS is realised externally in FPGA. For this purpose, the pixel value just after the reset signal and the value at the end of integration are transferred to the FPGA, which produces the difference. For the tested chips the difference is about 200 μV RMS, and is mainly caused by random variations of the offset voltages in the pixel-level analogue structures. Figure 13c shows an image after CDS correction, where FPN is reduced to $\sim$10 μV RMS.
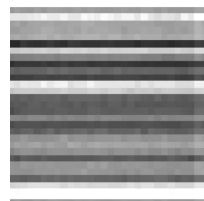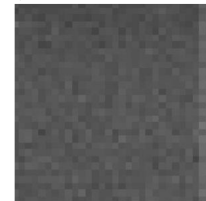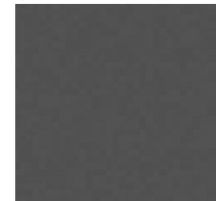

Fig. 13. Images achieved from the prototype: a) raw image, b) after removing the processors offset, c) after CDS

4.2. Examples of image processing. The presented prototype is a small-size, proof-of-concept device used mainly for testing quality of early vision algorithms implementation. The selected examples of early image processing (smoothing, edge detection and sharpening) are presented in Fig. 14. In this figure the upper images refer to the results after application of the algorithms implemented in the prototype chip, whereas the bottom images are achieved with "ideal", infinite-precision numerical processing. The prototype low-resolution image-array does not allow processing images of practical importance, but the used simple images allow evaluation of image processing quality.
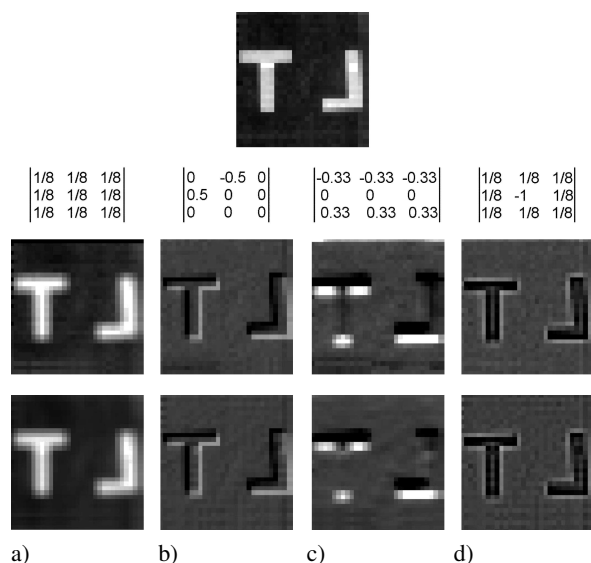

Fig. 14. Image processing examples: a) lowpass filter, b) Roberts cross edge detection, c) Prewit edge filter (horizontal edge detection), d) highpass filter. Top: acquired image, middle: results of focal-plane processing on the visionchip, bottom: results of "ideal" (numerical) image processing

As analogue operations can only be completed with finite accuracy, it is interesting to compare the experimental results with results obtained using "ideal" numerical computations. For the images in Fig. 14 the rms difference between "ideal" and experimental results (ignoring the border effects) are equal to 1.9%, 2.1%, 2.9% and 1.9%, respectively. Even though the analogue computations are limited in accuracy, the final result should be satisfactory for many computer vision applications. All the images in Fig. 14 are obtained without CDS.

It has also to be noted that some additional errors can be caused by discharging of the analogue signals stored in the pixels memory capacitors as a result of leakage currents, particularly since the chip is exposed to light. At 125 lux illumination the value stored on the C_{MEM} decreases by approximately 25 mV (i.e., 0.16% of maximum data value) per millisecond.

4.3. Performance. As explained in Sec. 3, the time needed for image acquisition can be made relatively small (below 20 μs, typically 15 μs), whereas the time necessary for data transfer from a photo-pixel to APE is extremely short, shorter than 100 ns. The speed of image recording of the presented test circuit is mainly limited by the external ADCs and throughput of a connection between the test board and PC, nevertheless moving object can be effectively observed by the prototype vision-chip. Exemplary result of observing a fast moving object, represented by a black and white wheel, is presented in Fig. 15. Fig. 15a shows a raw image of 10 cm diameter wheel rotating at 2000 rpm. The image, achieved in real time by application of edge detection algorithm, is shown in Fig. 15b.
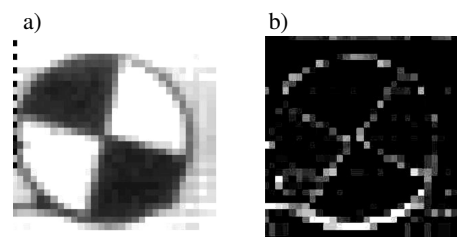

Fig. 15. Pictures of a black and white wheel rotating 2000 rpm: a) raw image, b) edge detection

A single photo-pixel consumes 0.4 μA of the supply current while recording an images and needs 10 μA during readout period. The shutter is opened for the entire array of photo-pixels, which means that the array consumes 0.4 μA$\times$32$\times$32 $\approx$ 410 μA during a time interval of 15 μs. The image data transfer and processing requires activation of only three columns and all APE at the same time. For typical working conditions of 25 fr/s, the time for processing of a complete row of image is 3.2 μs, which makes an average power consumption of 7.2 μW, assuming 780 μA for a single APE. The complete chip, including biasing circuits and digital logic, consumes 21 μW in average. The same circuit without the supply save mode consumes 82 mW. The summary of the vision-chip main parameters are given in Table 1.

Table 1
The main parameters of the visionchip

Technology	0.35 μm CMOS
Total chip area	2.7 mm^2
No. of transistors	60815
Pixel array size	32$\times$32
Pixel array area	1.8 mm^2
Total power consumption	
at 25 frames/s	21 μW
at 2000 frames/s	589 μW
Pixel dimensions	40 μm $\times$ 35 μm
Photodiode fill factor	8%
No. of transistors per pixel	22
APE dimensions (including digital registers)	220 μm $\times$ 98 μm
No. of APEs	32
Power per APE	
at 25 frames/s	0.2 μW
at 2000 frames/s	18 μW
Max. frame rate	2000 frames/s
Accuracy (image filtering)	$\approx$ 2.5%

5. Conclusions

A prototype chip that allows real-time focal-plane processing of grey-scale images has been fabricated, and is fully functional. The experimental results prove correctness of the assumed array architecture and circuit solutions. The proposed approach can also be used for a high-resolution vision chip realisation. The applied analogue continuous-time technique allows for a compact, low-cost, high speed and low-power implementation of a visionchip. The main parameters, summarized in Table 1, show that the prototype visionchip assures a good trade-off between power consumption and accuracy of image processing.

The proposed architecture is scalable and even quite large arrays can be integrated on a single silicon die using present-day CMOS technologies. Based on the presented design it is estimated, that 128×128 array fabricated in a 0.18 μm technology would measure 6 mm^2 while dissipating 100 μW of power.

REFERENCES

[1] P. Dutkiewicz, M. Kiełczewski, K. Kozłowski, and D. Pazderski, "Vision localization system for mobile robot with velocities and acceleration estimator", *Bull. Pol. Ac.: Tech.* 58 (1), 29–41 (2010).

[2] D.A. Martin, H.S. Lee, and I. Masaki, "A mixed-signal array processor with early vision applications", *IEEE J. Solid-State Circuits* 33 (3), 497–502 (1998).

[3] R. Etienne-Cummings, Z.K. Kalayjian, D. Cai, "A programmable focal plane MIMD image processor chip", *IEEE J. Solid-State Circuits* 36 (1), 64–73 (2001).

[4] J. Schemmel, K. Meier, and M. Loose, "A scalable switched capacitor realization of the resistive fuse network", *Analog Integrated Circuits and Signal Processing* 32 (2), 135–148 (2002).

[5] A. Dupret, J. O. Klein, and A. Nshare, "A DSP-like analog processing unit for smart image sensors", *Int. J. Circuit Theory and Application* 30, 595–609 (2002).

[6] P. Dudek and P.J. Hicks, "An analogue SIMD Focal Plane Processor Array", *IEEE Int. Symp. Circuits and Systems* 1, 490–493 (2001).

[7] P. Dudek and P.J. Hicks, "A general-purpose processor-per-pixel analog SIMD vision chip", *IEEE Trans. Circuits and Systems–I* 52 (1), 13–20 (2005).

[8] D. Hillier and P. Dudek, "Implementing the grayscale wave metric on a cellular array processor chip", *11th Int. Workshop on Cellular Neural Networks CNNA'2008* 1, 120–124 (2008).

[9] P. Dudek, A. Lopich, and V. Gruev, "A pixel-parallel cellular processor array in a stacked three-layer 3D silicon-on-insulator technology", *Proc. Eur. Conf. Circ. Theory and Design ECCTD'2009* 1, 193–196 (2009).

[10] P. Dudek and P.J. Hicks, "A CMOS general-purpose sampled-data analog processing element", *IEEE Trans. on Circuits and Systems–II* 47 (5), 467–473 (2000).

[11] W. Jendernalik, J. Jakusz, G. Blakiewicz, and R. Piotrowski, "CMOS realisation of the specialized analogue processor aiding early vision processing", *IX State Electronics Conf. KKE'2010* 1, CD-ROM (2010), (in Polish).

[12] A. El Gamal and H. Eltoukhy, "Cmos image sensors", *IEEE Circuits & Devices Magazine* 21 (3), 6–20 (2005).

[13] F. Krummenacher and N. Joehl, "A 4-MHz CMOS continuous-time filter with on-chip automatic tuning," *IEEE J. Solid-State Circuits* 23 (3), 750–758 (1988).

[14] S. Szczepański and S. Kozieł, "Phase compensation scheme for feedforward linearized CMOS operational transconductance amplifier", *Bull. Pol. Ac.: Tech.* 52 (2), 141–148 (2004).

[15] K. Bult and G.J.G.M. Geelen, "An inherently linear and compact MOST-only current division technique", *IEEE J. Solid-State Circuits* 27 (12), 1730–1735 (1992).

[16] R.G. Carvajal, J. Ramirez-Angulo, A.J. Lopez-Martin, A. Torralba, J.A.G. Galan, A. Carlosena, and F.M. Chavero, "The flipped voltage follower: a useful cell for low-voltage low-power circuit design", *IEEE Trans. Circuits and Systems–I* 52 (7), 1276–1291 (2005).