Smart Image Sensing System

Jie Yang, Cong Shi, Zhongxiang Cao, Ye Han, Liyuan Liu, and Nanjian Wu
State Key Laboratory of Superlattices and Microstructures,
Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, P. R. China
Phone/Fax: +86-10-8230-4754 E-mail: nanjian@red.semi.ac.cn

***Abstract*— Vision chips have achieved excellent performance in machine vision applications, but it is insufficient to most industrial applications due to low resolution image sensors and poor image quality. To solve the problems, a smart image sensing system which integrates image sensor and vision chip architecture image processor is proposed. The CMOS image sensor consists of an 800×600 pixel array and two column-parallel ADCs. It is capable of capturing images at 1000 fps. The image processor contains a 64×64 processing elements array, 64 row processors, and dual-core RISC. It can exploit data-level parallelism, therefore massively accelerate both low- and middle-level image processing. The proposed image sensing system is successfully applied to various applications like edge detection, motion detection, target tracking at a processing rate of 1000 fps.**

I. Introduction

Machine vision can be widely used in industry automation applications such as assembly line control, automatic classification, and defect detection [1], [2]. A lot of effort has been dedicated to develop high speed, low cost, low power, smart machine vision systems. The vision chip [3] which integrates CMOS image sensor and image processor in a single silicon chip is a promising candidate for such kind of system. Firstly, the vision chip has lower power consumption and costs compared to conventional monitor plus PC system [4]. Secondly, vision chip adopts massively parallel [5], [6] single-instruction-multiple-data (SIMD) architecture, which greatly accelerates image processing speed and then achieves high speed image processing. However reported vision chips [3], [4], [6]–[9] have low image resolution and poor quality pixels which limits the usage of vision chips.

In this paper, a 800×600 resolution CMOS image sensor with dedicated imaging pixel for high speed was developed to meet the image quality requirements for industrial machine vision applications. Furthermore, an image processor adopting the vision chip architecture was also developed. The proposed CMOS image sensor and the image processor are interface compatible. They were further integrated into a compact system to overcomes the image quality issue of vision chips and maintaining the high speed, low power and low cost features at the same time. The rest of this paper is organized as follows. In section II, the image sensor design is introduced. In section III, the architecture of the image processor is represented. In section IV, experimental results are shown. Finally, section V concludes this paper and indicates future directions.

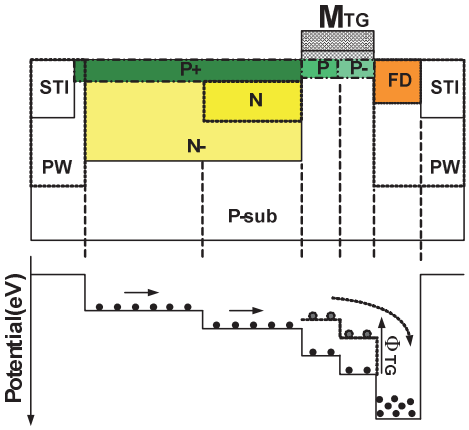

Fig. 1: Cross-section view and the signal electron transfer of the proposed pixel

II. High speed pixel and sensor design

For high speed image sensor, the exposure time is limited, hence large-area photodiode (PD) [10], [11] is required to increase the sensor sensitivity. However, a large PD results in low lateral velocity for the signal charges. Besides, image lag problem is also a design challenge for high speed image sensor. In order to enhance the lateral electric field and reduce the image lag problem, a new kind of pixel for high speed image sensors is adopted [12]. As shown in Fig. 1, additional ion implantation layers N and P are inserted in PD and TG channel respectively. Compared with conventional pixel devices, this type of pixel enhances the lateral electric field and speeds up the signal charge transfer from the PD to the transfer gate.

Imaging sensor pixels is implemented with the above techniques. The CMOS image sensor was fabricated using the $0.18\mu m$ 1P4M CIS technology, it consists of a 800×600 pixel array, two column-parallel ADCs and other necessary digital circuits. The microphotograph of the image sensor is shown in Fig. 2. The image sensor chip size is 10.0 mm × 15.0 mm, it operates with 1.8 V supply voltage.

III. Architecture of the image processor

Fig. 3 illustrates the architecture of the proposed image processor. It mainly consists of a N×N processing element (PE) array, N row processors (RP), and a dual-core RISC. An additional sensor interface (SI) is also implemented to

978-1-4673-4642-9/13/$31.00 ©2013 IEEE

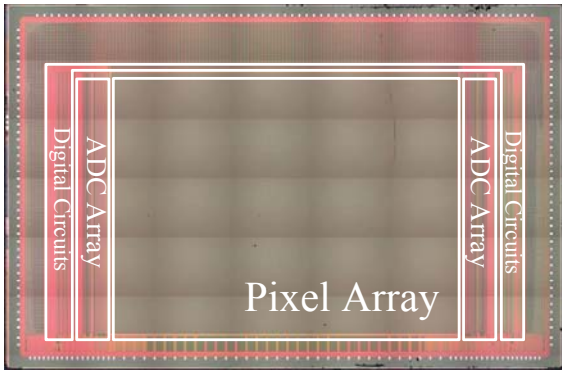

Fig. 2: Microphotograph of the image sensor

serve as the interface between the image sensor and the processor. This heterogeneous architecture gives consideration to both flexibility and parallelism in image processing. The PE array and the RPs have an *O(N × N)* parallelism and *O(N)* parallelism respectively [9]. The PE array has maximum parallelism and data throughput, but the simplest control path and the least flexibility. RPs have moderate parallelism and flexibility. RISC has least parallelism but the highest flexibility. Low-, mid- and high-level image processing algorithms [13] are tightly coupled with PE array processor, RPs and dual-core RISC respectively.

Both the PE array and the RPs are controlled by dedicated SIMD instructions. Each PE contains a 1-bit arithmetic unit and data memory. PE can also access data memory from its four neighbor PEs through the arithmetic unit. The 1-bit arithmetic unit finishes *addition, or, inversion, and* operations. Multiple bits operation can be finished by repeating 1-bit operations [9]. PE array is suitable for carrying out near field spatial filtering algorithms such as linear smoothing, and line detection. A simple example of PE array accomplishing 3×3 smoothing is shown in Fig. 4. Using intra-PE data transfer, near neighbor data can be moved to the memory of the center PE within the filter mask. Multiplication and division is accomplished by shifting the data left or right certain bits respectively. Since all PEs are controlled by the same instruction, the PE array finishes the smoothing for the entire image plane as long as one PE finishes the smoothing.

RP supports sub-set of RISC instruction set. It can access one row of PE memory. RP can perform arithmetic operations, data communication with adjacent RPs and carry out index addressing. With the help of RPs, statistical operations can be finished efficiently. RPs are used to generate vectors that represent the image feature, and they are the interface between the PE array and the RISC. The image processor chip microphotograph and specifications are shown in Fig. 5 and Table I respectively.

IV. System Implementation and Experimental Results

The proposed smart image sensing system is illustrated in Fig. 6. The major hardware components are the proposed

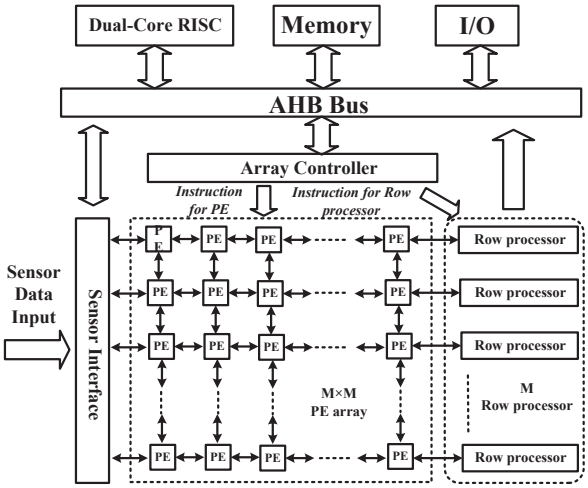

Fig. 3: Architecture of image processor

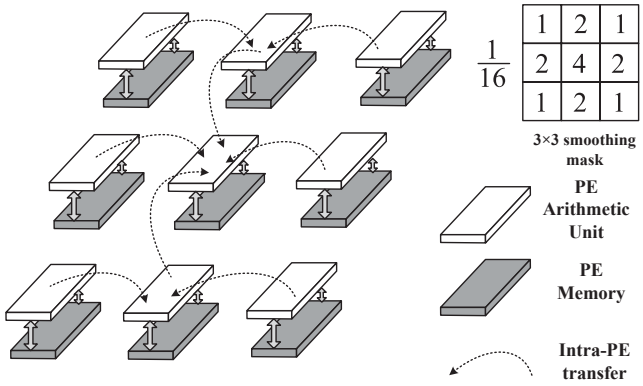

Fig. 4: 3×3 smoothing filter mask and its implementation on PE array

high speed image sensor and the image processor. A FPGA board is used to configure the image sensor parameters at system startup and it is also responsible for loading binary codes to the image processor. An on board DDRII can store 1 second duration of raw image data from the image sensor. The PC-end Debug software tool is used to code the PE, RP, and RISC for certain applications. Image processing kernel functions were implemented in the debug tool for user-friendly operation. Programming language and a compiler as in [9] were implemented. When application development and debugging were finished, the compiler generate binary codes for the image processor, and then loads codes into the program memory.

A. Motion Detection

Motion detection is a basic need for industry automation and event monitoring. We adopt the background subtraction algorithm [8], [14], [15] to implement motion detection in the proposed image sensing system. Since the image processor resolution is much larger than the image processor, thus the

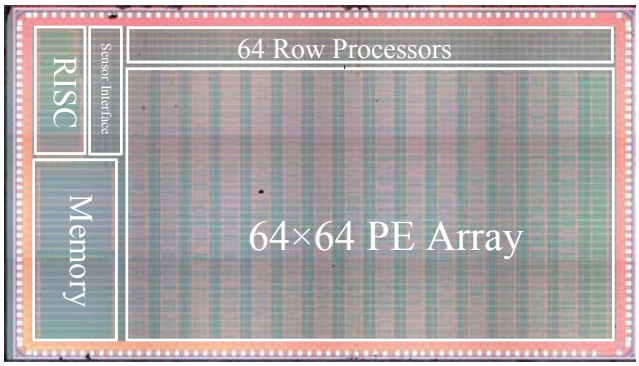

Fig. 5: Image processor microphotograph

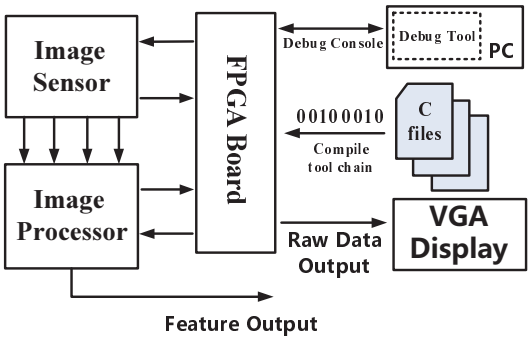

Fig. 6: Architecture of the proposed smart sensing system

TABLE I
PERFORMANCE SUMMARY OF THE IMAGE PROCESSOR

Parameter	Value
Technology	$0.18\mu m$ 1P4M standard process
Total area	5 mm × 8 mm
Array size	64×64 PE array, 64 RPs
PE memory	64 bits
RP memory	32 Bytes
Supply voltage	1.8 V
Clock frequency	50 MHz
Power consumption	420 mW

TABLE II
SYSTEM PERFORMANCE SUMMARY

Algorithm	Cycles number
3×3 weighted average	30
3×3 nonweighted average	39
Sobel edge detection	31
3×3 binary erosion	4
3×3 binary dilation	4
Global thresholding	9
Background reduction	35
Sensor interface reconfiguration	20
Harris corner detector	4.3k
Susan corner detector	7.5k
Motion detection	4.1k
Target Tracking	12.4k
0th moment	4.1k
1st moment	64k

on chip configurable sensor interface subsamples the image data to fit the PE array size. The PE array continuously carries out background subtraction, and global thresholding in every captured frame. RPs make statistical calculations on the "moving" pixels, finally the RISC decides whether a motion is detected in scene or not. Cycle numbers for PE and RP operation are shown in Table II.

B. Target Tracking

Chips like [8], [9], [14], [16] can finish target tracking algorithms. However they neither have adequate resolution nor optimized pixels for image sensors, thus their target tracking precision and speed are limited. With the help of a carefully designed high speed, high resolution image sensor and the high performance image processor, our image sensing system can finish target tracking with better precision at high speed.

To implement target tracking in our system, the described motion detection is first applied to detect the approximate location of the moving object. The RISC reconfigures the sensor interface so that it can sample or sub-sample a small image region that contains only the moving object. ROI extraction reduces the data transfer and data storage of the PE array, thus it reduces computation needed for the tracking algorithm.

Self-window target tracking algorithm [16] is implemented in our image sensing system. Self-window algorithm searches only inside the object region and its one-pixel neighbor. The algorithm utilize the property of high speed vision that the change between two successive frames is very small. By taking advantage of high-speed imaging, this algorithm is much less computationally expensive than tracking algorithms like mean-shift and kalman filtering, and it is easy to implement on vision chip architecture. After the self-window capturing, RPs calculate the 0th and the 1st moments of the captured target. The results of tracking a free-fall ping-pong ball are shown in Fig. 7. The ball position was given by the image processor, and the captured image was temporarily stored in the DDRII on FPGA board.

To accomplish multi-target tracking in the proposed system using self-window algorithm, all of the tracking images need to be stored in the PE memory. Aside from the memory needed for necessary calculation and background storage, 30 tracking images can be stored in the PE memory, which means 30 objects can be tracked simultaneously. RISC records the relative positions and the traces of objects, thus collision and separation can be predicted and monitored.

V. CONCLUSION

This paper proposes a compact image sensing system which integrates a high speed image sensor and an image processor. The image sensor pixel uses additional ion implementation layer to enhance the lateral electric field thus speeding up

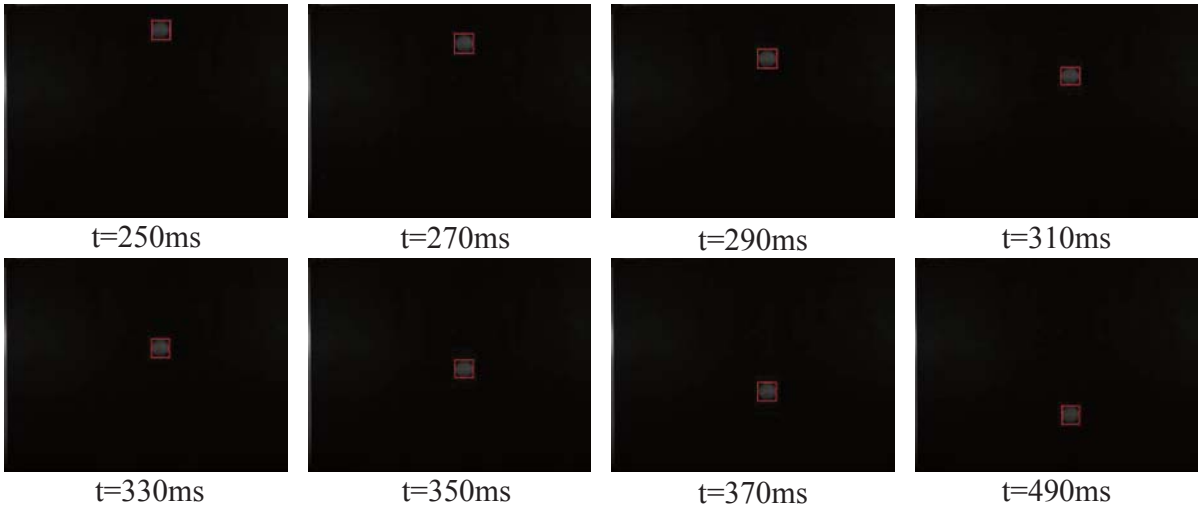

Fig. 7: Target tracking experiment result (first row from left to right, frames captured at 250, 270 290, 310ms respectively, second row from left to right frames captured at 330, 350, 370, 390ms respectively)

charge transfer and reducing image lag problems. The image quality and resolution of the image sensor exceeds recently reported vision chips. The image processor is designed based on the massively parallel SIMD architecture. Multiple levels of parallel processors makes it suitable for various vision applications. The proposed image processor inherits the high speed processing performance of the vision chips, which is tightly coupled with the image sensor. The image sensor and the image processor were designed and fabricated in $0.18\mu m$ standard CMOS technology and CIS technology respectively. The smart image sensing system was built with the two chips, FPGA board and peripheral software tool. The proposed system has been utilized in applications like edge detection, target tracking at 1000 fps.

ACKNOWLEDGMENT

The authors would like to thank Qi Qin and Runjiang Dou for their work on PCB board and DDRII interface.

REFERENCES

[1] M. Sonka, V. Hlavac, and R. Boyle, *Image processing, analysis, and machine vision*, 2nd ed. Thomson-Engineering, 2007.

[2] J. Ohta, *Smart CMOS image sensors and applications*. CRC Press, 2010, vol. 129.

[3] M. Ishikawa, K. Ogawa, T. Komuro, and I. Ishii, "A CMOS vision chip with SIMD processing element array for 1 ms image processing," in *Dig. Tech. Papers IEEE Int. Solid-State Circuit Conf*, 1999, pp. 206–207.

[4] A. Lopich and P. Dudek, "A SIMD cellular processor array vision chip with asynchronous processing capabilities," *IEEE Trans. Circuits Syst. I*, vol. 58, no. 10, pp. 2420–2431, 2011.

[5] K. Yamaguchi, Y. Watanabe, T. Komuro, and M. Ishikawa, "Design of a massively parallel vision processor based on multi-SIMD architecture," in *IEEE Int. Symp. Circuits and Syst*, 2007, pp. 3498–3501.

[6] J. Dubois, D. Ginhac, M. Paindavoine, and B. Heyrman, "A 10 000 fps cmos sensor with massively parallel image processing," *IEEE J. Solid-State Circuits*, vol. 43, no. 3, pp. 706–717, 2008.

[7] T. Komuro, S. Kagami, and M. Ishikawa, "A dynamically reconfigurable SIMD processor for a vision chip," *IEEE J. Solid-State Circuits*, vol. 39, no. 1, pp. 265–268, 2004.

[8] W. Miao, Q. Lin, W. Zhang, and N.-J. Wu, "A programmable SIMD vision chip for real-time vision applications," *IEEE J. Solid-State Circuits*, vol. 43, no. 6, pp. 1470–1479, 2008.

[9] W. Zhang, Q. Fu, and N.-J. Wu, "A programmable vision chip based on multiple levels of parallel processors," *IEEE J. Solid-State Circuits*, vol. 46, no. 9, pp. 2132–2147, 2011.

[10] M. Furuta, Y. Nishikawa, T. Inoue, and S. Kawahito, "A high-speed, high-sensitivity digital CMOS image sensor with a global shutter and 12-bit column-parallel cyclic A/D converters," *IEEE J. Solid-State Circuits*, vol. 42, no. 4, pp. 766–774, 2007.

[11] B. Cremers, M. Agarwal, T. Walschap, R. Singh, and T. Geurts, "A high speed pipelined snapshot CMOS image sensor with 6.4 Gpixel/s data rate," in *Int. Image Sensor Workshop*, 2009, pp. 70–72.

[12] Z. Cao, Y. Zhou, Q. Li, Q. Qin, and N.-J. Wu, "Design of pixel for high speed image sensors," *Int. Image Sensor Workshop*, pp. 229–232, 2013.

[13] S. Kyo, S. Okazaki, and T. Arai, "An integrated memory array processor architecture for embedded image recognition systems," in *IEEE Int. Symp. Comput Archit*, 2005, pp. 134–145.

[14] B. Zhao, X. Zhang, S. Chen, K.-S. Low, and H. Zhuang, "A 64 $\times$ 64 CMOS image sensor with on-chip moving object detection and localization," *IEEE Trans. Circuits Syst. Video Technol*, vol. 22, no. 4, pp. 581–588, 2012.

[15] N. Cottini, M. Gottardi, N. Massari, R. Passerone, and Z. Smilansky, "A 33 μw 64 $\times$ 64 pixel vision sensor embedding robust dynamic background subtraction for event detection and scene interpretation," *IEEE J. Solid-State Circuits*, vol. 48, no. 3, pp. 850–863, 2013.

[16] T. Komuro, I. Ishii, M. Ishikawa, and A. Yoshida, "A digital vision chip specialized for high-speed target tracking," *IEEE Trans. Electron Devices*, vol. 50, no. 1, pp. 191–199, 2003.