High speed vision chips based on multiple levels of parallel processors [Invited]

Nanjian WU
State Key Laboratory for Superlattices and Microstructures,
Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, China
E-mail:nanjian@red.semi.ac.cn

Abstract This paper introduces a high speed vision chips based on multiple levels of parallel processors. It integrates an image sensor, three von Neumann-type parallel processors and a SOM neural network. The SOM network can be reconfigured from the pixel-parallel array processor and reduces the high-level image feature recognition processing time by 98%. The chip can achieve 1000fps system-level performance from image acquisition to high-level feature recognition processing.

Keywords: vision chip, parallel processor, dynamic reconfiguration, SOM neural network

1. Introduction

The vision chip is a device integrating image sensor and parallel image processor on a single chip. It overcomes serial image transmission and serial image processing bottlenecks in traditional vision systems. It can achieve an image processing rate higher than 1000 frames per second (fps) and can be widely applied in industry automation, security monitoring, robotic vision, etc. The reported chips integrated the pixel-parallel and row-parallel SIMD array processors to speed up low- and mid-level image processing.[1][2] Recently, they further embedded MPU to carry out high-level image processing.[3][4] Although excellent in low- and mid-level processing, they were poor in high-level image processing tasks due to the von Neumann bottleneck of the MPU.

This paper presents a high speed vision chips based on multiple levels of parallel processors. It integrates a high speed image sensor, von Neumann-type pixel-parallel and row-parallel array processors, and a non-von Neumann-type self-organizing map (SOM) network. The SOM network can speed up the high-level image processing remarkably in a vector-parallel fashion. Furthermore, the SOM network can be dynamically reconfigured from the pixel-parallel array processor. The chip can achieve 1000fps system-level performance even when complicated high-level recognition tasks are involved.

2. Chip architecture

Figure 1 shows the system architecture of the vision chip.[5] It mainly consists of a 256×256 4T-APS high speed image sensor, a 64×64 pixel-parallel processing element (PE) array processor, a 64 row-parallel row processor (RP) array, a thread-parallel dual-core 32b RISC MPU, and a 16×16 SOM neural network. The pixel-parallel PE array processor and the SOM network can be dynamically reconfigured from each other. As the PE array is smaller than the pixel array in size, dynamically-mapping relationships with different pixel sampling intervals can be established between the two arrays. The PE and RP array processors perform low- and mid-level processing, such as image filtering, image segmentation, mathematical morphology and feature extraction. The SOM network recognizes the image features with a speedup of 16×16, avoiding the serious bottleneck in the high-level recognition processing. The dual-core MPU performs other simpler non-recognition high-level processing tasks, as well as the overall chip management.

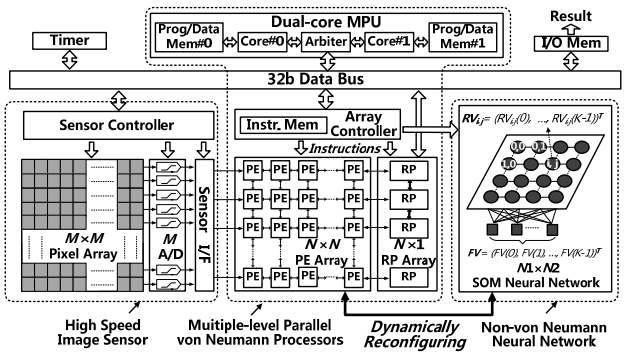

Figure 1 Vision chip architecture.

3. Circuit design and algorithms

Fig. 2 shows the reconfigurable PE circuit. The PE circuit contains: a 1-bit ALU, two multiplexers op1_Mux and op2_Mux for ALU operands selection, a carry register C, a 1-bit temporary register T, a 1-bit wide PE Memory and eight reconfiguration multiplexers (shaded in Fig. 2). The ALU can perform the operations of full adder, inverter, AND gate and OR gate. The reconfiguration multiplexers can switch the topological connections between neighboring PEs for different reconfiguration modes. In the PE array processor mode, the signal op1_sel selects the first ALU operand from its own PE Memory, or from the PE Memorys in its east, south, west and north neighboring PEs. The register C stores the ALU carry-out in current cycle and acts as the ALU carry-in for next cycle. Although only a simple 1-bit ALU is used, the PE can still realize multiple-bit operations by bit-serial processing.

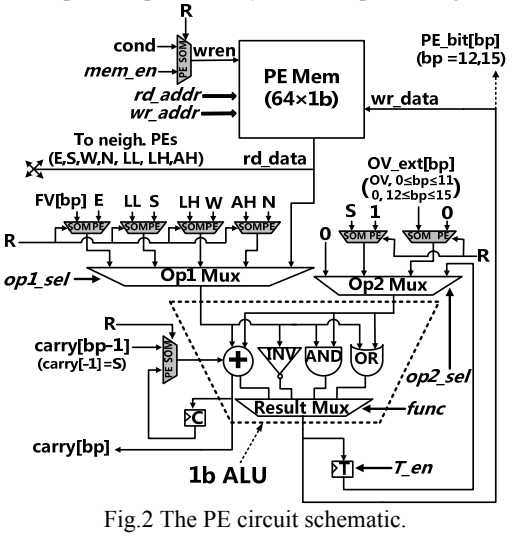

Fig.2 The PE circuit schematic.

The RP circuit consists of an 8-bit ALU, an 8-bit buffer, an 8-bit wide RP memory constructed by a bank of registers, three multiplexers and some condition flag registers, as shown in Fig. 3. Each RP can access its nearest upper and lower RPs through the first ALU operand selection. Moreover, every 4th RP in the RP array has extra access to its 4th neighboring upper and lower RPs. Thus an RP skip chain is formed for global operation acceleration. The buffer is used as the interface between the RP and PEs on the same row, as well as the interface between the RP array processor and the MPU. The flag P enables the PE-condition operation by establishing a more efficient data path between the PE and RP to accelerate feature extraction. The ALU can perform addition, subtraction, as well as max/min extraction for non-linear operations like medium filtering, grayscale morphology and SOM winner neuron determination.

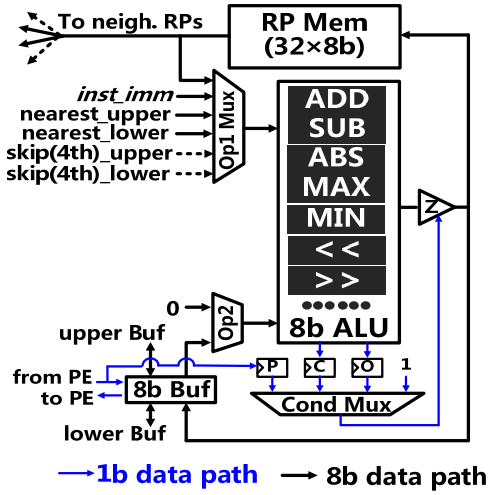

Fig. 3 The schematic of the row processor.

The chip can realize high speed hand gesture recognition by the following algorithm, as shown in Fig. 4. The images captured by the image sensor are sent to the PE and RP arrays to extract a PPED feature vector, which is finally recognized by the SOM network. The algorithm used 16 dimensional feature vectors extracted on a segmented binary hand image.

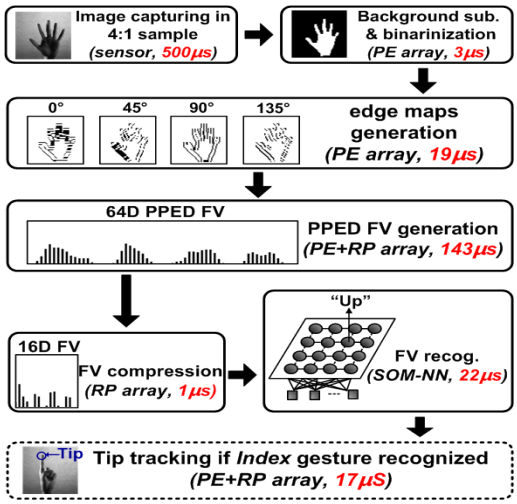

Fig. 4. : Algorithm of hand gesture recognition

3. Implementation and measurement results

A prototype vision chip was implemented in a 1P5M 0.18μm CMOS image sensor technology. The hand gesture recognition algorithm was implemented to demonstrate the real-time recognition capability of the SOM neural network. Seven types of hand gestures are used, defined as Grasp, Yeah, Up, Fist, Palm, Index and Down, respectively, as shown in Fig. 5.

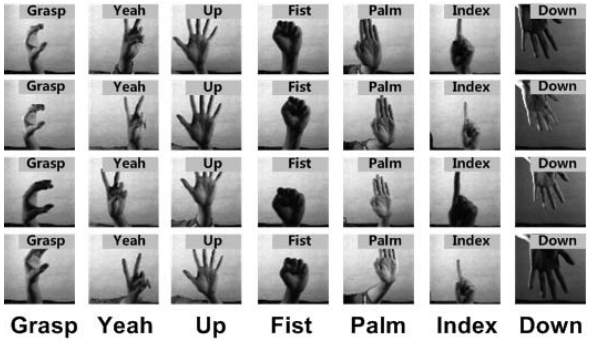

Fig. 5. 1000fps hand gesture recognition

The vision chip can also realize face recognition among a small group of persons based on similar algorithm. The face recognition can be used in identification systems, which would require the human faces appear in a dedicated location such as the central 128 × 128 pixel region. So we can use 64 × 64 images captured by the image sensor with 2:1 sample manner on the central 128 × 128 region. We used 100 face images for each person to train the SOM neural network online. These face images were captured under different head rotations, different distances and different expressions. Fig. 6 shows some face samples from the target person group for training and some recognition results.

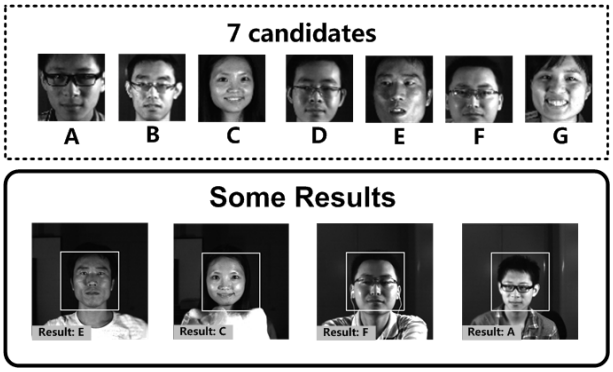

Fig. 6. Some results of human face recognition.

4. Conclusion

This paper implemented the high speed vision chips based on multiple levels of parallel processors. It integrated an image sensor, three parallel processors and a SOM neural network. The SOM network can be reconfigured from the PE array processor. The chip achieved 1000fps system-level performance from image acquisition to high-level feature recognition processing.

References

[1] W. Miao, et al., "A Programmable SIMD Vision Chip for Real-Time Vision Applications", IEEE J. Solid-State Circuits, vol. 43, no. 6, pp. 1470-1479, Jun. 2008.

[2] W. Jendernalik, et al., "An Analog Sub-Miliwatt CMOS Image Sensor with Pixel-Level Convolution Processing", IEEE Trans. Circuits and Systems-I: Regular Papers, vol. 60, no. 2, Feb. 2013.

[3] C. Cheng, et al., "iVisual: An Intelligent Visual Sensor SoC with 2790fps CMOS Image Sensor and 205GOPS/W Vision Processor", ISSCC Dig. Tech. Papers, pp. 306-308, Feb. 2008.

[4] W. Zhang, et al., "A Programmable Vision Chip Based on Multiple Levels of Parallel Processors", IEEE J. Solid-State Circuits, vol. 46, no. 9, pp. 2132-2147, Sept. 2011.

[5] C. Shi, J. Yang, Y. Han, et al., "A 1000 fps Vision Chip Based on a Dynamically Reconfigurable Hybrid Architecture Comprising a PE Array processor and Self-Organizing Map Neural Network", IEEE Int Solid State Circuits Conf (ISSCC), San Francisco, CA, 2014. 128–129