Extended Abstracts of the 2008 International Conference on Solid State Devices and Materials, Tsukuba, 2008, pp. 476-477

P-5-3

A Digital-Pixel-Sensor-Based Global Feature Extraction VLSI for Real-Time Image Recognition

Hongbo Zhu[1] and Tadashi Shibata[2]

[1]Dept. of Electronic Engineering, [2]Dept. of Electrical Engineering and Information Systems, Univ. of Tokyo
7-3-1 Hongo, Bunkyo, Tokyo 113-8656, Japan
Phone: +81-3-5841-6797 Fax: +81-3-5841-6797
E-mail: zhu@if.t.u-tokyo.ac.jp, shibata@ee.t.u-tokyo.ac.jp

1. Introduction

Real-time image recognition is becoming increasingly important in various applications such as automotive car control, video surveillance, robot control, human-computer interfaces and so forth. On one hand, for all these applications, reducing the latency between catching image and final recognition is of prime importance. For time-critical applications such as automotive car control, even a delay of several tens of milliseconds may lead to a serious accident. On the other hand, because the intensity information in an image is massive in quantity, extracting only essential features from the original image usually consumes a lot of time. Dedicated processors that can achieve a large throughput have been developed to resolve this problem. However, as such processors must read image data 1-pixel by 1-pixel from an off chip image sensor, it is difficult to achieve parallel processing. As a result, a large latency occurs, making such approaches not suitable for time-critical applications. Then the purpose of this work is to develop an image feature extraction VLSI compatible with real time image recognition.

In order to break the bottleneck of the low data transfer bandwidth between image sensors and processing circuits, we have integrated parallel feature extraction circuits and image sensors on the same chip. Digital-Pixel-Sensor (DPS) [1, 2] was chosen because of its intrinsic compatibility with digital processing circuits. Block-readout algorithm first introduced in [3] was employed for efficient image data transfer between DPS and processing circuits. As a result, it has become possible to carry out the local feature extraction processing at each pixel in a line-parallel way.

After local feature extraction, the extracted features contain a lot of trivial components and can not be used directly for image recognition. So we have introduced a new global feature extraction algorithm, which retains a particular number of relatively more significant features as essential features for image recognition. As a result, the chip contains three blocks: a 68×68 DPS, 16 groups of parallel local-feature extraction circuits and a global feature extraction circuit. By simulation, it has been shown that the chip is capable of extracting features from a 68×68 image and yielding the results in 0.12 ms when operating at 100MHz, which is more than 700 times faster than software processing running on a 2-GHz general-purpose processor.

2. Feature Extraction Algorithm

Biological research shows that the directional edge information in images is utilized as the most important clue in visual object recognition [4]. Being inspired by this principle, we developed a feature extraction algorithm which is based on four directional edges (horizontal, +45°, vertical, and -45°). The eligibility of this algorithm has been verified by the satisfying performance of the feature vector representation algorithm named Projected Principal-Edge Distribution (PPED [5]) in medical radiograph analysis and handwritten pattern recognition.

The directional edge feature extraction algorithm is performed by three-steps. Fig. 1 illustrates the first two steps for one pixel: firstly, convolutions between the 5×5-pixel region centered by this pixel and four 5×5-pixel filtering kernels (one for each direction) are calculated; then the most significant direction is selected as the direction at the pixel and its convolution value is reserved. Finally, all convolution values are sorted and only a pre-determined number of significant edges in the 68×68 area are retained.

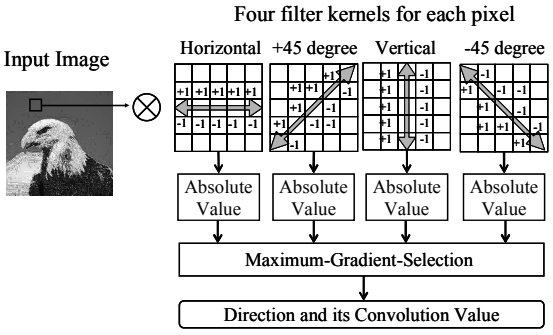

Fig. 1 For each pixel, the direction of edge is calculated by 5×5 spatial filtering. The kernel which gives the maximum value in four directional filters decides the direction of the edge.

3. Hardware Architecture

The VLSI was designed exactly according to the feature extraction algorithm as described in Section 2. Fig. 2(a) shows the overall architecture. It is composed of three main blocks: 68×68 DPS, 16 groups of local feature extraction circuits, and a global feature extraction circuit. Fig. 2(b) shows the circuit in each pixel. The 8-bit memory is controlled by eight separate control signals, which allows us to readout pixel data in a bit serial manner (from LSB to MSB). In addition, all the pixel data in eight consecutive rows can be read out simultaneously, thus making possible the row parallel processing for 5×5 blocks of pixel data. Since the data is processed in a bit-serial way, the inter-

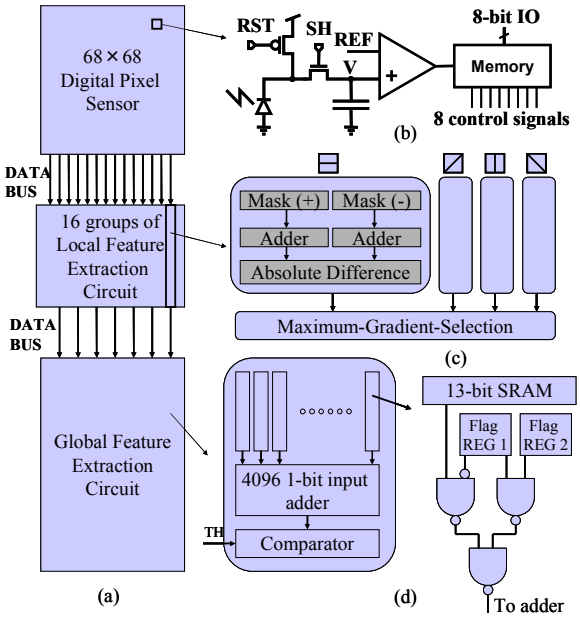

Fig. 2 Architecture of global feature extraction VLSI.

connection is greatly simplified. There are 16 groups of processing elements included in this system for parallel local feature extraction. As shown in Fig. 2(c), each element contains eight masks, two masks for each kernel: one for the plus component; the other for the minus component. After the summation of the data passed through the mask, four absolute difference calculation circuits are employed for these four pairs of a plus component and a minus component. The largest result remains with its direction determined by the maximum-gradient-selection circuit. Fig. 2(d) shows the global feature extraction circuit. Because the size of each filter kernel is 5×5, a 68×68 original image becomes a 64×64 directional edge map with the convolution value for each pixel. Rank-order-filter algorithm has been employed so that the global thresholding process of all 4096 11-bit data can be achieved in only 11 steps of comparisons, in contrast to those complex sort algorithms used in software.

4. Design and Simulation Results

A VLSI chip was designed in a 0.18-μm 5-metal technology. Fig. 3 shows the entire layout. The specification of this chip is summarized in Table I. This design is ready for fabrication and we hope it is finished at the time of the conference.

By simulation using NANOSIM, the function of the system was verified. As shown in Fig. 4, we converted an image into analog signals and assigned these signals into corresponding nodes in the photodiode array. The global threshold was set to be 2000 which means it should retain 2000 edges out of 4096 pixels. The simulation results give 2005 edges as shown in Fig. 4(b). The extra 5 edges are due to the pixels that have the same convolution value. The remaining edges for each direction are shown in Fig. 4(c-f). If distribution histograms are produced from these edge maps, they play essential roles in intelligent image recognition applications as shown in [5].

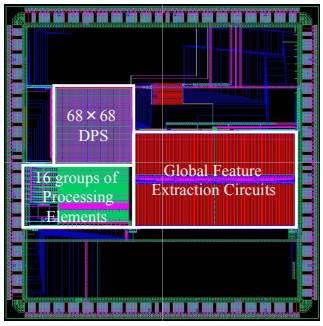

Fig. 3 Layout of global feature extraction VLSI, without the top metal, which is used for light shielding.

Table I Chip Specification

Process Technology	0.18um 5-Metal technology
Core Size	4.5mm×2.5mm
No. of Pixels	68×68
Pixel Pitch	18μm×18μm
Fill Factor	7.4%
Transistors Count	1.6 million

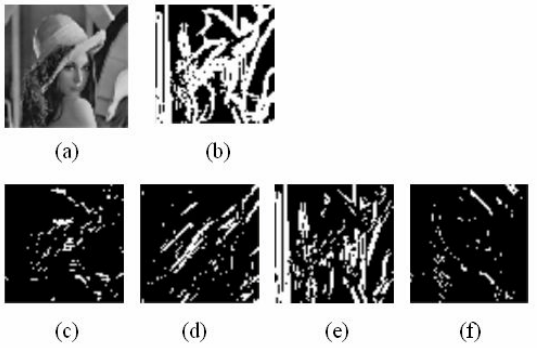

Fig. 4 Simulation result of global feature extraction VLSI. Global Threshold: 2000; (a) original image; (b) all edge map; (c) edge map in horizontal; (d) edge map in +45°; (e) edge map in vertical; (f) edge map in -45°.

5. Conclusions

A digital-pixel-sensor-based global feature extraction VLSI for real-time image recognition has been proposed. The simulation results show that this VLSI can extract features more than 700 times faster than software processing running on a 2-GHz general-purpose processor.

Acknowledgements

The VLSI chip in this study will be fabricated in the chip fabrication program of VLSI Design and Education Center (VDEC), the University of Tokyo in collaboration with Rohm Corporation, Toppan Printing Corporation, Synopsys, Inc., Cadence Design Systems, Inc. and Mentor Graphics, Inc..

References

[1] S. Kleinfelder, S. Lim, X. Liu, and A. E. Gamal, *IEEE J. Solid-State Circuits*, vol. 36, pp. 2049–2059, Dec. 2001.

[2] A. E. Gamal, *SSDM* pp. 236-237 Sep. 2002

[3] B. Tongprasit, K. Ito, and T. Shibata, *ISCAS*, pp. 2389-2392, May 2005.

[4] G. J. Hubel and T. N. Wiesel, *J. Physiol*., vol. 148, pp. 574-591, Oct. 1959.

[5] T. Shibata, M. Yagi, and M. Adachi, *Proc. Int. Conf. Information Fusion*, vol. 1, pp. 648–656, 1999.