# A VLSI Processor with a Configurable Processing Element Array for Balanced Feature Extraction in High Resolution Images

Hongbo Zhu<sup>1</sup> and Tadashi Shibata<sup>2</sup>

 <sup>1</sup> VLSI Design & Education Center (VDEC), Univ. of Tokyo, Tokyo, Japan
<sup>2</sup> Center for Innovative Integrated Electronic Systems, Tohoku University, Sendai, Japan E-mail: zhu@vdec.u-tokyo.ac.jp, shibata@jsap.or.jp

## Abstract

A balanced feature extraction VLSI processor employing a configurable processing element array is developed and fabricated. Operating at 45 MHz, the chip can extract balanced features from 1920×1080 images at a speed of 55-fps.

# 1. Introduction

Feature extraction is a basic process in many intelligent image processing tasks such as object recognition, object tracking and action recognition. Most feature extraction algorithms [1] have the problem of complex computation which limits their operating speed. It is also difficult to develop highly efficient VLSI processors for acceleration due to the complexity. As a result, even with specific processors [2], the real-time performance still remains as a problem. In particular, the problem is more serious for high resolution images.

Hardware friendly feature extraction algorithms based on binary salient edge maps are proposed and their performance is demonstrated [3]. By developing highly efficient VLSI processors [4], object recognition system with a processing speed of 906 us per frame is achieved [5]. However, there are two problems in the feature extraction processor in [4]. Firstly, the resolution of the image is limited by the number of processing elements (PEs), which is only  $64 \times 64$ . Secondly, extracting features only based on the absolute edge gradient in each pixel site usually results in unbalanced features. For example, as shown in Fig. 1(a) and (b), pixel sites having relatively high edge gradients (see the region in the solid ellipse) in a local region of the image are lost, while pixel sites having high gradients (see the region in the dotted circle) are blurred due to that too much pixel sites are retained. The problem of unbalanced features becomes more serious in high resolution images because there are usually heterogeneous local regions in the image that need be processed separately.

In this paper, a VLSI processor for extracting balanced features in high resolution images is proposed. The core of the processor is a 2D reconfigurable processing element array (PEA). By hierarchical grouping the PEs, this processor can extraction features in different local regions of the images. In addition, the resolution of the image is not limited by the number of PEs. A test-of-concept chip has been fabricated in a 0.18um CMOS technology with a  $32\times32$  PEA. By measurement results from a fabricated chip, a speed of 7.5-Kfps has been achieved when it processes a

 $128 \times 128$  image. The clock frequency is 45 MHz and the power consumption is 108 mW. A speed of 55-fps can be acquired for feature extraction in  $1920 \times 1080$  images using this chip.

## 2. Balanced Feature Extraction

Fig. 1(c) shows the basic flow of the balanced feature extraction algorithm which contains two steps: directional filtering and salient feature selection. In the first step, the input image is full scanned by four directional filtering kernels and in each pixel site there are four convolution values. The maximum value is recorded as the edge gradient of this pixel site and its direction is preserved. These edge gradients form a gradient image used for salient feature extraction. In the second step, the gradient image is divided into several regions for salient feature extraction. Fig. 1(d) shows the features when the gradient image is divided into  $4 \times 3$  square local regions. Because the number of pixel sites in each region (local threshold) can be specifically tuned, these features are well balanced.

#### 3. VLSI Architecture

# Overall Architecture

Fig. 2 shows the overall architecture of the processor. The core unit is the PEA which contains  $32 \times 32$  PEs. In order to achieve a flexible architecture for images with different resolutions, all PEs are symmetrically grouped hierarchically into three levels. Four PE blocks in a lower level compose a PE block in the higher level. The system controller configures and controls the PEA for performing feature extraction.

#### Elementary PE Block

Fig. 3(a) illustrates the architecture of the elementary PE block using a binary sorting algorithm [4]. Each column SRAM stores the gradients from multiple pixels so that high resolution images can be processed. Column SRAMs and PEs (Fig. 3(b)) are full connected. A mask is added for region selection. When extracting features for the 8×8 region, the sum of the adder is compared with the local threshold and the comparison result (FB\_in) is selected as the feedback (FB) to all these 8×8 PEs. When extracting features for larger regions, the sum is transferred to the upper PE block, from which the feedback signal (FB\_out) is used as FB.

### High Level PE Blocks

Fig. 4(a) shows the  $16 \times 16$  PE block which contains four  $8 \times 8$  PE blocks and a feedback generator (FBG). The FBG adds four sums from  $8 \times 8$  PE blocks. The summation results can be used directly for feature extraction in the  $16 \times 16$  region or be transferred to the  $32 \times 32$  PE block. Fig. 4(b) shows the detail of FBG, which generates the feedback signal (FB) to be broadcasted to PEs in four 8×8 PE blocks. The  $32 \times 32$  PE block is composed by four  $16 \times 16$  PE blocks just like the way that the  $16 \times 16$  PE block is composed by the  $8 \times 8$  PE blocks. The only difference is that the adder in the FBG of the 32×32 PE block can hold the summation result for adding the data from multiple 32×32 image regions so that images with a resolution that much larger than  $32 \times 32$  can be processed.

## 4. Experimental Results

Fig. 5 shows the photomicrograph and specification of the fabricated chip. The chip was designed in a 0.18 um CMOS technology. With the 32×32 PEA, this chip can extract features for 128×128 images at the speed of 7.5K-fps and for 1920×1080 images at the speed of 55-fps. Fig. 6 shows the results when the features of a 128×128 image were extracted from the gradient image. The power consumption is 108 mW at the frequency of 45 MHz.

#### Acknowledgements

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

### References

- [1] D. G. Lowe, Int. J. Comput. Vis. 60 (2004) 91.
- [2] S. Lee et al., ISSCC Dig. Tech. Papers, (2010) 332.
- [3] M. Yagi and T. Shibata, IEEE Trans. Neural Netw. 14 (2003) 1144.
- [4] H. Zhu and T. Shibata, Jpn. J. Appl. Phys. (2009) 04C080-1.
- [5] H. Zhu and T. Shibata, Proc. 35<sup>th</sup> ESSCIRC, (2009) 248.







Fig. 2 Architecture of feature extraction processor.



Fig. 3 (a) Elementary PE block contains 8×8 PEs, (b) circuits inside PE.



Fig. 4 (a) 16×16 PE block is composed of four 8×8 PE blocks, (b) feedback generator (FBG).

	Specification of test chip	
	Technology Die Size (mm²) Core Size (mm²) Number of PEs Image Resolution (Max) Supply Voltage (V) Clock Frequency (MHz) Speed at Max Resolution (fps)	0.18 um CMOS, 1P5M 5 × 5 4.5 × 4.5 32 × 32 128 × 128 1.8 45 7.5K
← 5 mm →	Power Consumption (mW)	108
(a)	(b)	

Fig. 5 (a) Chip photomicrograph and (b) specification.



Original Image

Features

Fig. 6 Measurement results from a fabricated chip.