

# 1.15mW Mixed-mode Neuro-Fuzzy Accelerator for Keypoint Localization in Image Processing

Injoon Hong, Jinwook Oh, and Hoi-Jun Yoo

Department of Electrical Engineering  
Korea Advanced Institute of Science and Technology (KAIST)  
Daejeon, Republic of Korea  
E-mail: injoon.hong@kaist.ac.kr

**Abstract**— A mixed-mode neuro-fuzzy accelerator is proposed for keypoint localization of image features of Scale Invariant Feature Transform (SIFT) algorithm. To reduce processing time of keypoint localization with low power consumption, analog Adaptive Neuro-Fuzzy Inference System (ANFIS) and digital controller are implemented together. It is implemented in 0.13 $\mu\text{m}$  CMOS process and achieves 1.15mW power consumption. Compared to the conventional digital standalone system, 0.733mm<sup>2</sup> neuro-fuzzy accelerator achieves 43% processing time reduction and also results in 19.4% time reduction of image feature extraction process.

## I. INTRODUCTION

Recently, the real time object recognition systems have been widely applied in car navigation, surveillance camera, and robot vision system [1]. Most of recognition algorithms used in such applications are based on keypoint matching process which compares keypoints of the input image with keypoints of the database image. To extract keypoints of input images, they usually exploit Scale Invariant Feature Transform (SIFT) [2]. SIFT is composed of three distinctive process to extract image features. First, Difference of Gaussian (DoG) is required to find potential keypoints by center surround filtering. Second, it selects local maximum or minimum points from different DoG image, since such points has high probability to be a keypoint. However, some parts of the extracted keypoint are redundant points such as edge responses or low contrast points. Therefore, the process of eliminating of the redundant points, so called keypoint localization, is performed at last. Conventionally, Hessian function is usually exploited to remove redundant points [2], but it requires derivatives calculation for every candidate points. Using classifying characteristic of localization process, however, neuro-fuzzy system can replace the Hessian function. Its ability of learning makes more optimal system to remove the unnecessary points.

The proposed neuro-fuzzy accelerator is a neuro-fuzzy system which enables high speed inference operation for keypoint localization. It exploits analog implemented ANFIS with digital learning controller so that it can take advantage of

power and area saving characteristic of mixed mode configuration.

This paper is organized as follows. A second chapter explains the role of proposed accelerator in a system level. Through third chapter, inference in ANFIS system and its mixed mode circuits are described. In forth chapter, simulation result of the proposed architecture is reported and conclusion appears in fifth chapter.

## II. SYSTEM ARCHITECTURE

Fig.1 shows the proposed system architecture, consisting of heterogeneous processing elements with neuro-fuzzy accelerator. To take the operation efficiently, a SIMD processing element (SIMD PE) operates on filtering operation which consists of parallel instructions while a MIMD processing element (MIMD PE) performs feature descriptor generation consisting of sequential instructions. In addition to such conventional heterogeneous architecture, neuro-fuzzy accelerator is implemented together for localization.

Without neuro-fuzzy accelerator, SIMD PE not only extracts candidate of keypoints, but also process keypoint

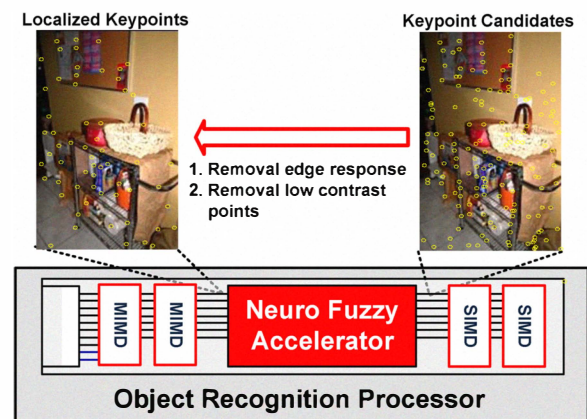


Figure 1. The proposed architecture for heterogeneous processing elements

localization. However, in the proposed architecture, SIMD PE only extracts candidate points, while the neuro-fuzzy accelerator performs the keypoint localization. When the neuro-fuzzy accelerator conducts localization process, SIMD PE can extract another candidate points by task-level pipelining, thus overall throughput is increased. In addition, thanks to learning ability of ANFIS system of the accelerator, it contains the optimal fuzzy rule internally for keypoint localization. Once the optimal rule for localization is settled through pre-learning process, then processing time of localization is significantly reduced compared to localization by Hessian function.

Fig.2 shows performance increases of the proposed architecture. A simulation is conducted on the 66x107 image frame, which contains only total 134 keypoint candidates, and is favorable to be tested repeatedly for performance comparison. As a result, the proposed architecture succeeds to reduce 23% of overall keypoint extraction time, and 42% for localization process.

### III. NEURO FUZZY ACCELERATOR

#### A. Redundant Keypoint Candidates

Two kinds of redundant keypoint candidates are generated in SIFT algorithm; edge points and low contrast points [2]. For edge responses, the points are placed on the edge of same object so they contain same object information. For low contrast points, they are sensitive to noise but SIFT cannot avoid generating such points.

Conventionally, to resolve the edge responses in SIFT algorithm, Hessian function is usually exploited. For low contrast points, the function value at the extremum is useful to reject such low contrast points. However, such compensation requires many derivatives operation and calculation of determinant of matrix for each candidate.

#### B. ANFIS for keypoint localization

To increase speed of the localization process, the

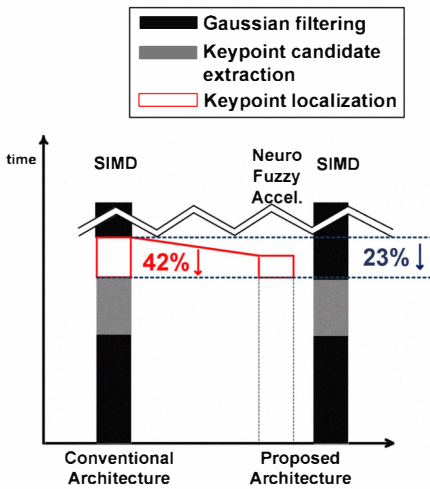


Figure 2. Performance comparison

proposed neuro-fuzzy accelerator uses ANFIS which is one of the neuro-fuzzy classifiers [5][6]. Due to its ability of learning, optimal rule can be settled in the system, so processing time of localization is reduced.

Fig.3 explains Takagi-Sugeno fuzzy model which the proposed ANFIS is based on. It has total 5 layers. The first layer is called fuzzification layer. This is where the crisp inputs are transformed into fuzzy numbers and it contributes membership functions to trace the trait of data sets. The second layer, rule base layer, generate fuzzy rules from the composition of membership functions output. The third layer takes a role to normalize and the forth layer is named as weight multiplication layer for output rule selection. The certain node of this layer computes the normalized firing strengths and linear combination of inputs. The last layer is called defuzzification layer.

Fig.4 shows the analog circuit for the layer 1 and layer 2. For the first layer, making of diverse shape of the membership functions is important. Therefore, Fig 4(a) shows the proposed membership function circuit which has three controllable inputs,  $g_m$ ,  $V_{ref1}$ ,  $V_{ref2}$  to control the slope, high and low boundary of inverted Gaussian functions. These three parameters can be mapped into parameter set  $\{a_i, b_i, c_i\}$  of (1) equation and called premise parameters. The 9 membership functions are implemented to characterize location, scale, and orientation information of each candidate point.

$$\mu_A(x) = \frac{1}{1 + \left| \frac{x-c_i}{a_i} \right|^{2b}} \quad (1)$$

For the fuzzy rule layer, there are several ways to express fuzzy logics in a neuro-fuzzy classifier. The proposed accelerator uses Takagi-Sugeno Model's min-max rule, since it is simple to be implemented in hardware. Therefore, as illustrated in Fig.4(b), the minimum logic is realized for the second layer. The proposed ANFIS uses maximum circuit to make minimum logic for the ease of realization. Equation (2) shows winner-take all circuit operation for minimum operation on fuzzy rule layer.

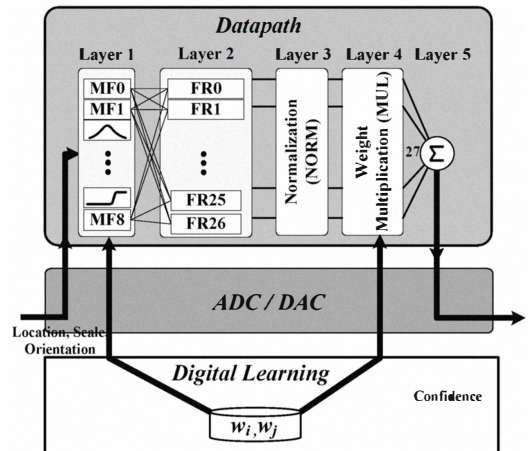


Figure 3. Top-diagram of neuro-fuzzy accelerator for keypoint localization

$$\min(s_1, s_2, \dots, s_M) = \max(s, s_2, \dots, s_M)' \quad (2)$$

$$s' = 1 - s$$

### C. Mixed-mode configuration

To overcome demerits of neural network and fuzzy system in terms of power and area consumption, the proposed neuro-fuzzy accelerator adopts mixed-mode configuration. As illustrated in Fig.3, ANFIS system is implemented with analog circuit for low power and area consumption, while learning controller is digital to take advantage of programmability.

To learn the ANFIS system precisely, the digital learning block controls analog implemented ANFIS system. In particular, as shown in Fig.3, perturbation learning scheme is exploited. The implemented perturbation scheme is based on not only the negative gradient update but also measurement of output error gradient for digital controller so that make it easy to implement neural network learning hardware. This algorithm is well suited to circuit implementation for two reasons. First, error gradient measurement rather than error calculation is easily implemented in circuits and complex error trace and calculation algorithm are unnecessary. The second reason of perturbation-based learning algorithm for the proposed system is its rule-base of fuzzy structure so as to exploit the known fuzzy rules of the data.

Therefore, it is easy to learn with high speed and accuracy. Since it reduces the number of weights to be adjusted during learning sequences; an accurate fine perturbation process can be achieved in order to obtain more refined fuzzy rule

representation for improved accuracy with high speed.

As illustrated in Fig.5, digital learning block consists of a shift and exponential multiplier, a sign conversion unit and an adder for updating of parameters. In addition, bit selection controller or parameter memory is also implemented for other digital processes of learning controller.

With high programmability of digital learning block, most of the neuro-fuzzy layers of ANFIS system are implemented with analog circuit. When reading a stored value in digital memory or controlling analog circuits by digital controller, digital to analog conversion is needed. When analog ANFIS circuit updates new parameters to digital controller, analog to digital conversion should be used. Thus, additional conversion time and area or power consumption of conversion block are overhead from mixed mode configuration. Therefore, resolution optimization of conversion is essential without performance degradation like classification accuracy [9] and control prediction accuracy [10].

Two conversion paths of mixed mode design are output conversion at output of weight multiplication layer into digital value and parameter conversion from digital value into analog at first layer of ANFIS system. Illustrated in Fig.6, the optimization simulation is carried out for the both cases. As long as the hardware cost of conversion units exponentially varies along with its resolution, each bit width can be optimized for classification accuracy with respect to the normalized power consumptions. For analog output bit optimization, the normalized power is rapidly decreased as the bit width decreases. However, accuracy of the

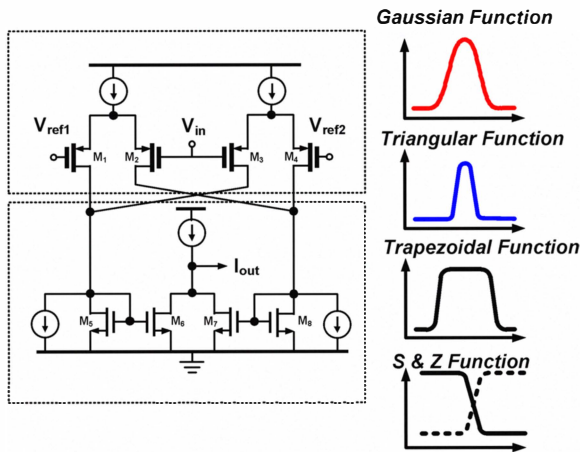


Figure 4(a). Membership functions

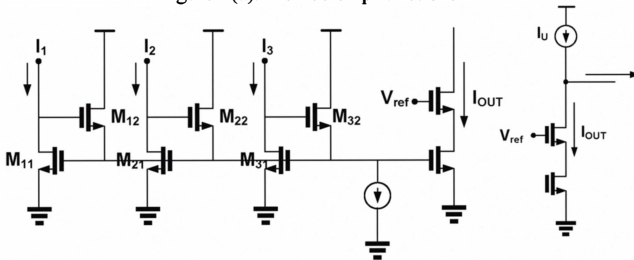


Figure 4(b). Winner-Take-All circuit for Fuzzy-rule layer

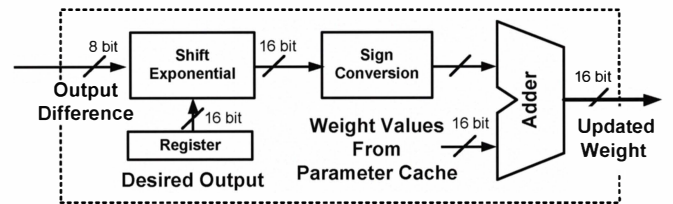
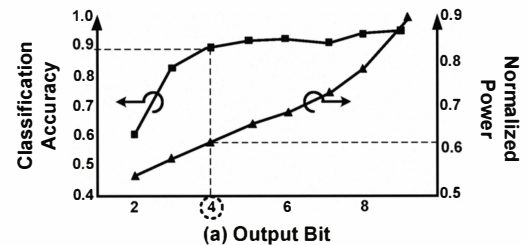
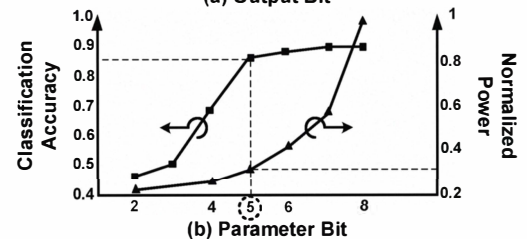


Figure 5. Digital Learning Block



(a) Output Bit



(b) Parameter Bit

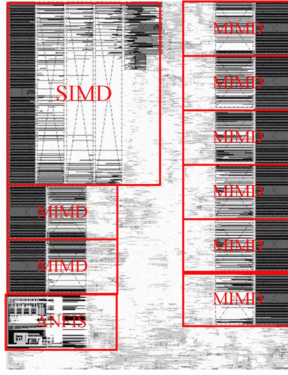
Figure 6. Optimization of output/parameter bit width

## V. CONCLUSION

For keypoints localization in SIFT algorithm, mixed-mode neuro-fuzzy accelerator is proposed. Using analog ANFIS system, high speed and low power consumption are achieved. It is implemented in 0.13 $\mu$ m CMOS process and consumes 1.15mW power consumption with 0.733mm<sup>2</sup>. Also, neuro-fuzzy accelerators achieves 43% processing time reduction and 19.4% reduction time of image feature extraction in SIFT algorithm.

## REFERENCE

- [1] Y. Hirano et al., "Industry and Object Recognition: Applications, Applied Research and Challenges," in LNCS, Springer, 2006
- [2] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol.60, no.20, pp. 91-110, 2004.
- [3] S. Lee, et al., "A 345mW Heterogeneous Many-Core Processor with an Intelligent Inference Engine for Robust Object Recognition," ISSCC Dig. Tech.Papers, pp. 332-333, Feb. 2010.
- [4] J.-Y. Kim, et al., "A 201.4GOPS 496mW Real-Time Multi-Object Recognition Processor with Bio-Inspired Neural Perception Engine," ISSCC Dig. Tech.Papers, pp. 150-151, Feb. 2009.
- [5] J. Oh, et al. "A 1.2mW On-Line Learning Mixed Mode Intelligent Inference Engine for Robust Object Recognition", SOVC, 2010
- [6] J.S.R Jang, ANFIS: adaptive-network-based fuzzy inference system, IEEE Trans. System, Man, Cybernetics 23 (1993) 665-685.
- [7] James C. Spall, An overview of the simultaneous perturbation method for efficient optimization, Johns Hopkins APL Technical Digest 19 (1998) 482-492.
- [8] A. R-Vazquez, F. V-Verdu, A modular programmable CMOS analog fuzzy controller chip, IEEE Trans. Circuits and Systems 46 (1999) 251-265.
- [9] Minsu Kim, Joo-Young Kim, Seungjin Lee, Jinwook Oh, Hoi-Jun Yoo, A 22.8GOPS 2.83mW neruo-fuzzy object detection engine for fast multi-object recognition, IEEE Symp. VLSI Circuits (2009) 260-261.
- [10] Renee St. Amant, Daneal A. Jimenez, Doug Burger, Mixed-signal approximate computation: A neural predictor case study, IEEE Micro Jan./Feb. (2009) 104-115.



Item	Specification	
Process Technology	0.13 $\mu$ m 1P8M Logic CMOS	
Area	900 $\mu$ m $\times$ 811 $\mu$ m	
Power Supply	1.2V	
Operating Frequency	200MHz	
Average Power Consumption	NF accelerator	1.15mW
	SIMD	7mW
	MIMD x8	28mW
	Total	36.15mW

Figure 7. Layout photograph

classification has been sustained until 4 bit by the error tolerance characteristic of ANFIS. For parameter bit optimization, classification accuracy became saturated at 5 bit for the same reason. Therefore, D/A conversion block has 4 bit-width, while A/D conversion unit has 5 bit-width as optimal bit-width.

## VI. SIMULATION RESULTS

Fig.7 shows the layout photograph of the proposed architecture. All the power consumption reported in Fig.7 are based on the simulation result. It is tested on a 0.13 $\mu$ m CMOS process. The performance is tested on a 60s 30fps video sequence of a scene containing large edge responses and low contrast points. The proposed neuro-fuzzy accelerator achieved 1.15mW power at 1.2V and 0.733mm<sup>2</sup> area consumption. Compared to typical heterogeneous architecture of processing elements with 1 SIMD PE and 8 MIMD PEs [3] the proposed architecture achieved 43% reduction of localization processing time, and 19% reduction for image feature extraction time at the expense of additional 1.15mW additional power consumption as shown in Fig. 8.

	Without NF accelerator	With NF accelerator	Comparison
Average Keypoint localization time	52 $\mu$ s	29 $\mu$ s	43% ↓
Average power consumption	35mW	36.15mW	1.03% ↑
Average Keypoint generation time	210 $\mu$ s	170 $\mu$ s	19% ↓

Figure 8. Performance Evaluation