

# A 600- $\mu$ W Ultra-Low-Power Associative Processor for Image Pattern Recognition Employing Magnetic Tunnel Junction (MTJ) Based Nonvolatile Memories with Novel Intelligent Power-Gating (IPG) Scheme

Y. Ma<sup>1</sup>, S. Miura<sup>1</sup>, H. Honjo<sup>1</sup>, S. Ikeda<sup>1,4</sup>, T. Hanyu<sup>1,2</sup>, H. Ohno<sup>1,2</sup>, T. Shibata<sup>1</sup> and T. Endoh<sup>1,2,3,4</sup>

<sup>1</sup> Center for Innovative Integrated Electronics Systems, Tohoku University, <sup>2</sup> Research Institute of Electrical Communication, Tohoku University, <sup>3</sup> Graduate School of Engineering, Tohoku University, <sup>4</sup> JST-ACCEL, 468-1 Aramaki Aza Aoba, Aoba-ku, Sendai, 980-0845, Japan, Tel: +81-22-796-3427, Email: tetsuo.endoh@cies.tohoku.ac.jp

## Abstract

An associative processor using magnetic tunnel junction (MTJ) based nonvolatile memories has been proposed and fabricated under 90nm CMOS/70nm perpendicular-MTJ ( $\rho$ -MTJ) hybrid process for achieving the exceptionally low-power performance of image pattern recognition. A 20Kb 4-Transistor 2-MTJ (4T-2MTJ) STT-MRAM was adopted to completely eliminate the standby power. An intelligent power-gating (IPG) scheme specialized for this STT-MRAM is employed to optimize the operation power by only activating currently accessed memory cells. The Operations of prototype chip at 20MHz are demonstrated by measurement. The proposed processor successfully carries out single Bag-of-Feature based texture pattern matching within 6.5 $\mu$ s, and the measured average operation power of entire processor is only 600 $\mu$ W. Comparing to the twin chip designed with 6T-SRAM and the circuit used in recent conventional works, 91.2% and more than 98.4% power reductions are achieved, respectively.

## 1. Introduction

Image pattern recognition plays essential role in various time-critical applications such as automotive vehicle control, human-computer interface, video surveillance, and so forth. For achieving the high-speed performance, a number of dedicated processors have been developed [1,2]. However, in these conventional processors, large-capacity volatile embedded memory is commonly-required to transfer template data from the low-speed stand-alone storage devices for realizing highly concurrent processing. Generally, the higher recognition accuracy is demanded, the larger volume memory is necessary for storing more template data. As a result, the large power consumption remains as the inescapable serious issue, especially for those battery-powered systems such as smartphone and sensor network. The prevalent approach to solve this power issue is power-gating (PG) technique that shuts down idle circuits. However, with the volatile memory, it causes non-negligible extra cost in power and delay for restoring the entire data back to the system after powering on. Recently reported 4T-2MTJ STT-MRAM [3], on the other hand, is regarded as very promising alternative to the volatile memory, which can leverage the PG technique to full advantage with its non-volatility, high access speed and high endurance. In this paper, a nonvolatile associative processor employing IPG scheme based on 4T-2MTJ STT-MRAM is proposed and implemented for the first time as a real-working chip to demonstrate its superiority in ultralow power performance for image pattern recognition.

## 2. Processor Architecture with IPG Scheme

Fig. 1 (a) shows the entire processor architecture, which consists: N template memory units (MUs) coupled with N data-mask/power-gating units (DPUs); K centroid MUs; a word-line (WL) decoder; a multiplexer and a Manhattan Matching Unit (MMU). The proposed processor searches out the candidate most similar to the target pattern from all N templates as the final recognition result utilizing Manhattan distance calculation described in Fig. 1 (b) as similarity evaluation method. Fig. 2 ex-

plains the processor operation with IPG scheme. In the IPG scheme, both the template and centroid stored in MUs are 128-D image feature patterns with 1Kb digital data in each. The N templates are preassigned into K clusters with K cluster IDs (CIDs) stored in corresponding DPUs, and the centroid indicates pre-calculated mean vector of the templates in each cluster. Firstly, the centroid most similar to the target is detected from centroid MUs. Then the final recognition result is found out from the corresponding template MUs with particular CID indicating the detected centroid. Thus, all idle MUs (all template MUs with the other CIDs) can be powered off during the entire operation. It should be noted that N is generally much larger than K in most of the recognition applications. Therefore, operation power can be significantly reduced comparing to the volatile memory based systems that must keep power-on impartibly and constantly. Moreover, fine-grain PG technique is also adopted in the IPG scheme as shown in Fig. 3. Every grain of 8 memory cells in the MU has an independent power line driver (PLD) which is controlled through the clock, the WL and the power enable (PE) signal given from DPU. Thus, only activating the currently accessed memory grain becomes possible. It should be noted that, with this IPG scheme, extending memory capacity (number of templates) used in the proposed processor will not cause increase of power consumption. The detailed circuit configurations of 4T-2MTJ memory cells, PLD, DPU and MMU in the processor are described in Fig. 4.

## 3. VLSI Implementation and Fabricated Chip Performance

Fig. 5 (a) shows the waveforms verifying the IPG-based power enable of proposed processor with 3-D user setting templates shown in Fig. 5 (b). The templates of each cluster are shown to be correctly output after the PE signal, where all other MUs are turned "OFF". Fig 6 shows the results of texture recognition using the processor with 16 template textures assigned into 4 clusters as "fur", "water", "stone" and "wood". The most similar texture for the 2 targets of "stone", and "wood" is successfully found. Fig. 7 gives the micrographs of fabricated prototype chip including 12 processor cores with 16 templates and 4 centroids in each. Measurement results demonstrating successful chip operation (VDD=0.9V, @20MHz) before/after power-off period is shown in Fig. 8. The VDD dependence of measured average operation power is analyzed in Fig. 9. The performance comparisons along the proposed processor, the 6T-SRAM-based processor and the other two conventional works [4, 5] are discussed in Table.1 and Fig. 10.

## 4. Conclusions

A 4T-2MTJ STT-MRAM based associative processor employing IPG scheme was proposed and implemented, which achieved much lower power and smaller circuit area than conventional circuits in state-of-the-art researches.

## Acknowledgements

This work is supported by CIES's Industrial Affiliation on STT-MRAM program, R&D Subsidiary Program for promotion of Academic-industry Cooperation of METI, ImPACT of CSTI, ACCEL under JST.

**References** [1] S. Lee, et al., ISSCC, (2010) 332. [2] G. Kim, et al., JSSC, **48**, (2013) 1615. [3] T. Ohsawa, et al., JSSC, **48**, (2013) 1511. [4] J. Y. Kim, et al., JSSC, **45**, (2010) 32. [5] J. Oh, et al., JSSC, **48**, (2013) 33.

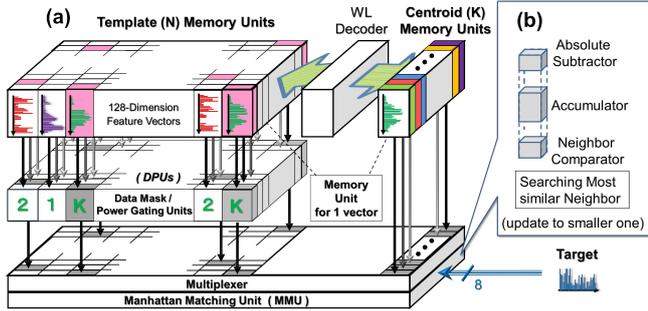


Fig. 1 Entire architecture of the proposed associative processor with IPG scheme.

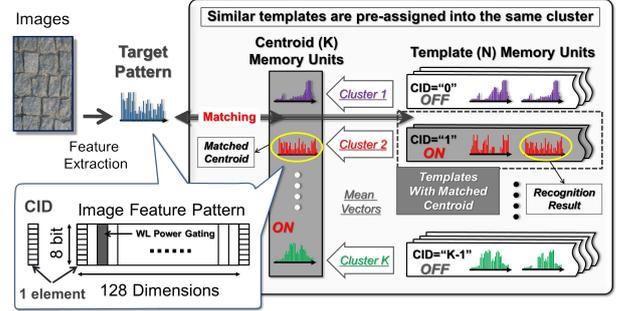


Fig. 2 Basic processor operation with IPG (PG operation among MUs).

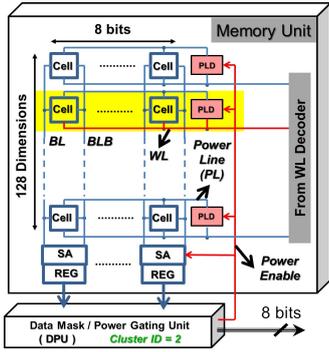


Fig. 3 Fine-grained power gating of IPG Scheme (PG operation inside 1 activated MU) for further power reduction.

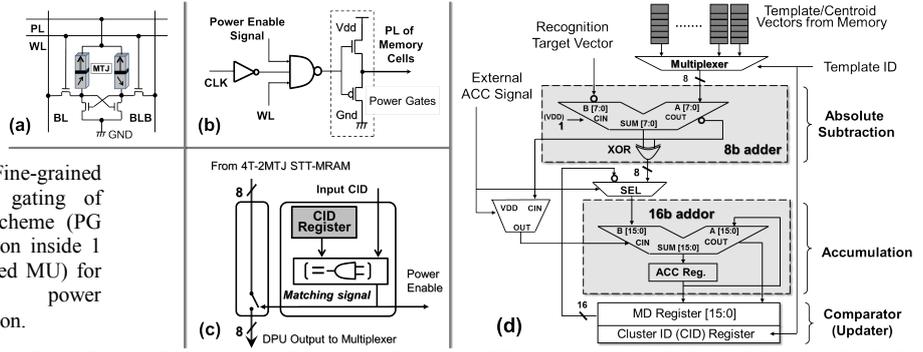


Fig. 4 The detailed circuit configuration of (a) 4T-2MTJ memory cell, (b) PLD, (c) DPU and (d) MMU.

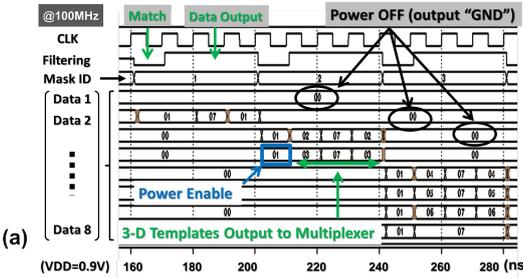


Fig. 5 SPICE waveforms of IPG-based power control with eight 3-D template data.

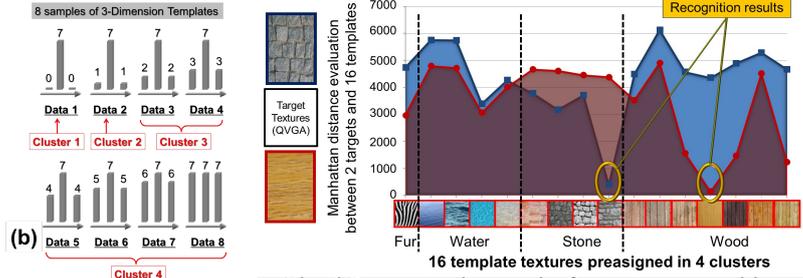


Fig. 6 Demonstration results for texture recognition. (Performing 1 BoF texture pattern matching in 6.5  $\mu$ s)

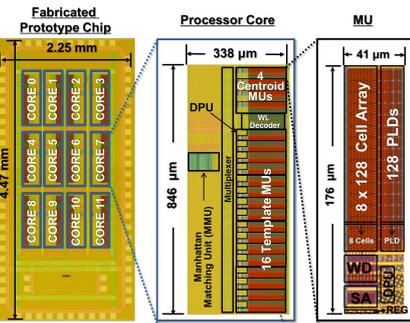


Fig. 7 Photograph of the fabricated prototype chip.

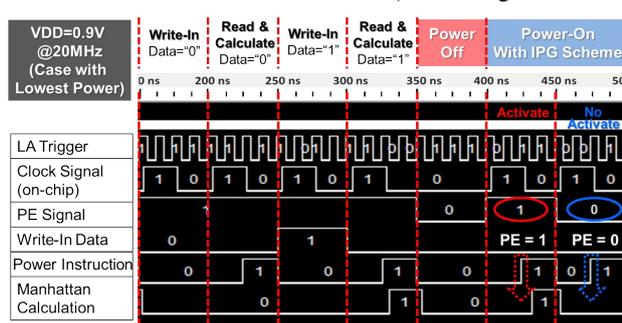


Fig. 8 Measured waveforms of associative operation with IPG

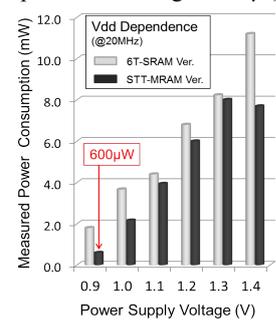


Fig. 9 VDD dependence of power.

TABLE I Comparison of Specification and Performance

	JSSC 2010 [7]	JSSC 2013 [8]	6T-SRAM-Based Chip	This Work
Process	130 nm	130 nm	90 nm	90 nm
Core Size	1mm x 2.5mm	3.2mm x 0.8mm	0.3mm x 0.9mm	0.3mm x 0.9mm
Supply Voltage	1.2 V	1.2 V	1.2 V	0.9 V
Data Format (Data Size)	16D [SIFT] (32-Byte)	128D [SIFT] (128-Byte)	128D [BoF] (128-Byte)	128D [BoF] (128-Byte)
Average Power of Recognition Core	37 mW	86 mW	6.8 mW	600 $\mu$ W
Throughput	4 cycles/vector *	8 cycles/vector	128 cycles/vector	128 cycles/vector
Frequency	200 MHz	200 MHz	20 MHz	20 MHz

\* Estimated matching throughput with uniform processing data format (128D 128-Byte pattern).

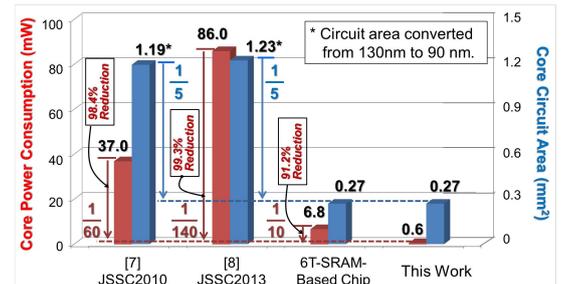


Fig. 10 Comparisons of power and circuit area of processor core.