We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS BOOK CITATION INDEX INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

2

Low-Power Analog Associative Processors Employing Resonance-Type Current-Voltage Characteristics

Trong Tu Bui[1] and Tadashi Shibata[2]
[1] *The University of Science-HCM City,*
[2]*The University of Tokyo,*
[1]*Vietnam*
[2]*Japan*

1. Introduction

Data-matching function plays an essential role in a number of information processing systems, such as those for voice/ image recognition, codebook-based data compression, image coding, data search applications etc. In order to implement such functions effectively, both proper data representation algorithms and powerful search engines are essential. Concerning the former, robust image representation algorithms such as projected principle edge distribution (PPED) (Shibata et al., 1999; Yagi & Shibata, 2003; Yamasaki & Shibata, 2007) etc. have been developed on the basis of the edge information extracted from original images. Such an algorithm is robust against illumination, rotation, and scale variations, and has been successfully applied to various image recognition problems. Concerning the latter, because search operations are computationally very expensive and time-consuming, it would be better if these operations are carried out by dedicated VLSI associative processors rather than programs running on a general-purpose computer. In this regard, dedicated highly parallel associative processor chips have been developed for the purpose of real-time processing and low-power operation.

It has been demonstrated that associative processors can serve as the basis of humanlike flexible computation, and many examples of flexible pattern perception have been demonstrated that are based on analog and digital technologies as well as mixed signal technologies. Digital approaches are accurate in computation, but often require large chip real estate and often consume large power. Analog implementations are preferred in terms of low-power consumption and high-integration density. In this regard, various distance-calculating circuits, which are used to evaluate the similarity (or dissimilarity) between two vectors, have been proposed. Euclidean distance circuits (Tuttle et al., 1993) utilizing MOSFET square-law cells were employed in an 8-bit parallel analog vector quantization (VQ) chip. Konda et al. (1996) and Cauwenberghs & Pedroni (1997) proposed neuron MOSFET (νMOS)-based and charged-based Manhattan-distance evaluation cells, respectively. A νMOS-based Euclidean distance calculator used in a recognition system for handwritten digits was proposed (Vlassis et al., 2001). Kramer et al. (1997) also proposed an analog Manhattan-distance-based content-addressable memory (CAM) using the analog

Source: Solid State Circuits Technologies, Book edited by: Jacobus W. Swart,
ISBN 978-953-307-045-2, pp. 462, January 2010, INTECH, Croatia, downloaded from SCIYO.COM

non-volatile memory technology. On the other hand, bell-shaped characteristics have been implemented in various analog associative processors (Ogawa & Shibata, 2001; Yamasaki & Shibata, 2003; Hasler et al., 2002; Peng et al., 2005). In such processors, bell-shaped current-voltage (*I-V*) characteristics, or resonance-type *I-V* characteristics, were utilized in building matching cells. This is because such resonance characteristics can represent the correlation between the input data and the template data in the sense that the output current becomes maximum when the input voltage coincides with the peak voltage. The resonance characteristics of single-electron transistors (SETs) were utilized to carry out associative processing for color classification (Saitoh et al., 2004). Since resonance characteristics are the typical nonlinear characteristics often observed in nano devices, such associative processors would be one of the most promising system applications in the coming era of nano devices. Although room-temperature SETs utilizing particular phenomena have been reported (Mastumoto et al., 1996; Uchida et al., 2002; Saitoh et al., 2004), all demonstrations have been reported at the device level or simple circuitry, rather than at realistic system levels. Numerous new developments are now being explored so as to make nano devices applicable to the next-generation integrated circuits. However, because these devices have a higher probability of being defective than conventional CMOS devices, designing reliable digital circuits with such devices is a major challenge. So far, CMOS-based associative processors are still dominant in practical applications. One of the drawbacks in analog implementation, however, is that the matching-cell behavior suffers from the problem of device mismatch. For this reason, architectures that are robust against such problems are desired.

In this chapter, a compact resonance-characteristics matching cell using only NMOS transistors in order to emulate the resonance-type *I-V* characteristics of nano devices and to build a small-area low-power associative processor will be described. In addition, a new calibration scheme (Bui & Shibata, 2008a) that can compensate for matching errors due to device mismatch is presented. System configuration of a single-core architecture and the major circuitries utilized in the prototype chip design as well as measurement results are presented in Section 2. In Section 3, a solution to how the system is hierarchically scaled up to a vast scale integration is presented. For a vast scale integrated system, a large number of template data can be implemented in multiple associative processors, making the recognition system more intelligent. In this regard, a fully-parallel multi-core/ multi-chip scalable architecture of associative processors was developed (Bui & Shibata, 2008b; 2009). Moreover, the problem associated with inter-chip communication delay which is critical in the time-domain WTA operation was resolved by a newly-developed winner-code-decision scheme (Bui & Shibata, 2008b; 2009).

2. Single-core architecture of analog associative processor

2.1 System architecture

Figure 1 shows the block diagram of the single-core associative processor developed in our work (Bui & Shibata, 2008a). It consists of two main parts, the digital memory module and the proposed analog matching-cell module. The memory module employing SRAM is utilized to store template data that represent the past experience or knowledge. The similarity evaluation between the input data and the template data is carried out in parallel by vector-matching circuits in the matching-cell module. All data are represented as 64-dimension PPED vectors compatible with vectors generated from the vector-generation chip

described in the study (Yamasaki & Shibata, 2007). Each vector-matching circuit itself consists of 64 vector-element matching cells (MCs) utilized to evaluate the similarity between vector elements. The matching score between vector elements is given as output current from the matching cell, which has bell-shaped *I-V* characteristics. Consequently, in the conventional manner, the matching scores between the input vector and template vectors are also currents obtained by taking the wired sum of element matching-cell output currents. Current memories are utilized to memorize the peak currents of the bell-shape characteristics and then to generate vector-matching scores by the calibration scheme proposed in Section 2.2.4. Utilizing these vector-matching scores, the winner-take-all (WTA) circuit (Ito et al., 2001) determines the maximum-likelihood template vector and identifies its location, namely, the code of the vector. Serial digital-to-analog converters (SDACs) are used to convert digital values to analog voltages prior to similarity evaluation processing. Once the template data are downloaded from the digital memory module to the matching-cell array via the digital-to-analog converters, the data are temporarily stored in all the matching cells as analog voltages and utilized for a number of parallel pattern matching operations that follow.

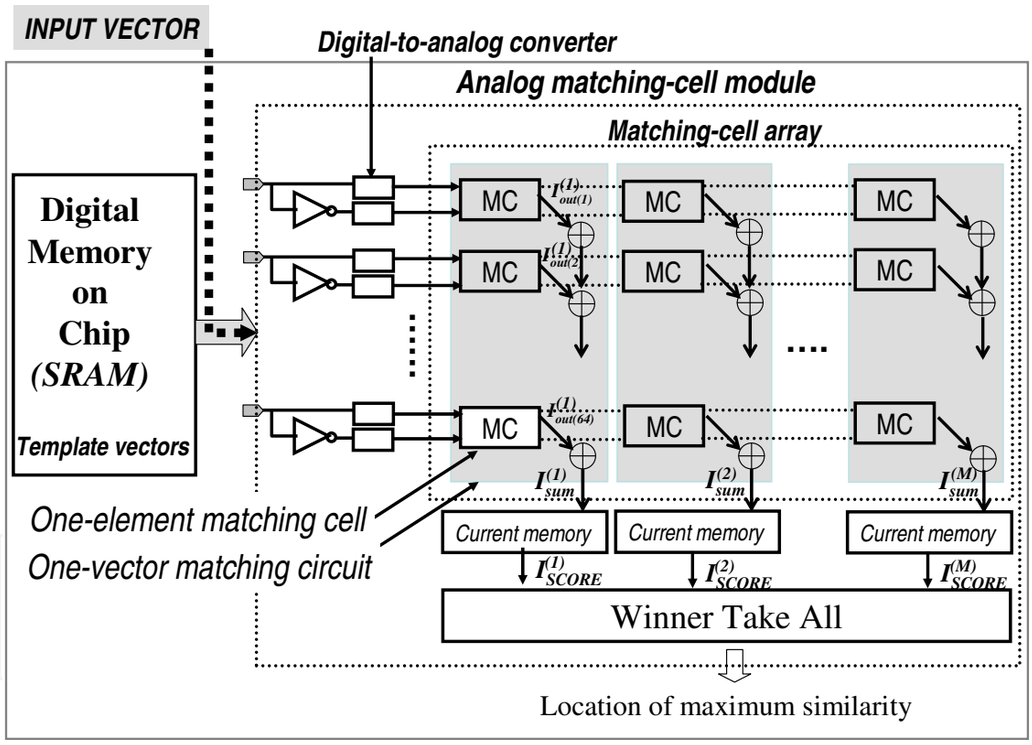


Fig. 1. Block diagram of single-core associative processor employing resonance-type current-voltage characteristics.

In analog associative processor implementations, the storing of analog template data is always a difficult issue. Analog nonvolatile memory technologies (Kramer et al., 1997; Yoon et al., 2000; Yamasaki et al., 2001; Kobayashi et al., 2005) have been developed for such purposes, but they are often very expensive to implement. In the proposed architecture, on the other hand, digital memories such as SRAM, DRAM, and flash can be employed to build

a system that is inexpensive compared with analog nonvolatile memory technologies. By adding an analog matching-cell module to any existing memory system, an associative processor can be easily constructed in the architecture proposed in this work.

2.2 Circuit Implementation

2.2.1 Matching cell

Figure 2 shows the schematic of one element-matching cell, which is used to determine the similarity between each element of the input vector and the corresponding element of the template vector. The cell is composed of only NMOS transistors. This is advantageous in making the cell layout compact because extra areas for N-wells and PMOS transistors are not necessary. In this regard, the present cell is superior to the CMOS cell described in ref. (Yamasaki & Shibata, 2003) as well as the cell described in ref. (Konda et al., 1996).

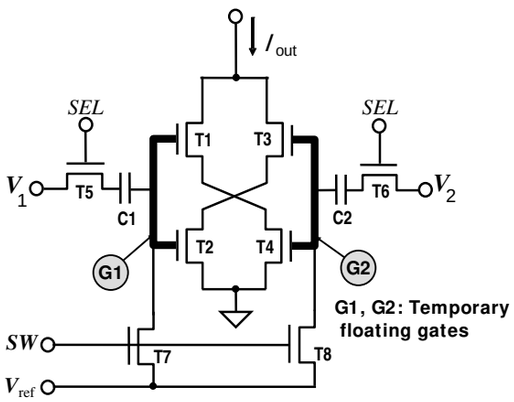


Fig. 2. Schematic of vector-element matching circuit (matching-cell circuit).

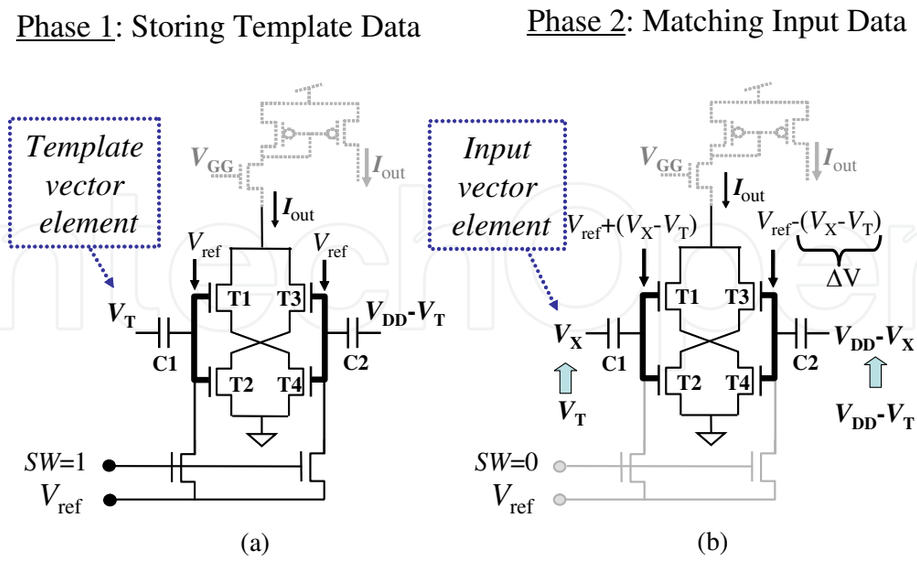


Fig. 3. Operation of matching cell, matching operation, is conducted in two phases. (a) Phase 1, the writing phase; template data are stored in matching cells. (b) Phase 2, the evaluation phase; similarities between template data and input data are evaluated.

Figure 3 illustrates two phases of the operation of the matching cell. In the figure, two NMOS switches (T_5 and T_6 in Fig. 2) connected to input terminals are omitted for simplicity of explanation. In the first phase, as shown in Fig. 3(a), template vector elements are stored temporarily inside matching cells. This phase is also called the writing phase, in which the template element voltage (V_T) and its complement (V_{DD}-V_T) are connected to two input terminals of the matching cell. The floating gates are first connected to the reference voltage, V_{ref}, and then disconnected from that voltage to make them electrically floating. After this phase, template vector elements are memorized as charges on the floating gates inside the corresponding matching cells. Phase 1 is repeated until all the necessary template vectors are downloaded from the memory module. In the second phase (also called the evaluation phase) shown in Fig. 3(b), the input element voltage (V_X) and its complement (V_{DD}-V_X) replace the positions of template elements. As a result, floating gate voltages of $V_{ref} + \Delta V$ and $V_{ref} - \Delta V$ are created. In the figure, ΔV is the difference voltage between the input vector element and the template vector element.

These two voltages create the bell-shaped *I-V* characteristics shown in Fig. 9. Indeed, since the gate voltages of the two serially connected transistors T_1 and T_4 are complementary analog signals, $V_{ref} + \Delta V$ and $V_{ref} - \Delta V$, respectively, they form bell-shaped *I-V* characteristics. Because of the back-gate effect occurring in T_1, these characteristics are slightly asymmetric. Similarly, the T_2-T_3 pair also creates asymmetric characteristics. By cross-coupling four transistors, as shown in Fig. 2, the asymmetry is removed.

The result of the evaluation from each matching cell is given as an output current (I_{out}). A higher current indicates greater similarity. The peak height of the output current I_{out} is also programmable by varying the reference voltage V_{ref} connected to the floating gates. The higher V_{ref} is, the higher the peak current becomes. These characteristics are described clearly in Section 2.3 and Fig. 9. In addition, it should be noted that once all the necessary template data are stored in the matching-cell array, only phase 2 is repeated for each new input vector.

The matching score between the input vector and the template vector is obtained by taking the wired sum of all I_{out}'s from 64 element-matching cells for one vector, as shown in Fig. 1 and eq. (1). In conventional approaches, a higher wired-sum current represents a greater similarity between two vectors.

$$I_{SCORE}^{(k)} = I_{SUM}^{(k)} = \sum_{i=1}^{64} I_{out(i)}^{(k)} \tag{1}$$

2.2.2 Winner-take-all circuitry

The block diagram of the winner-take-all circuit (WTA) is shown in Fig. 4. The matching scores from the vector-matching circuits are first converted to delay times by the current-to-delay-time converter (Yamasaki & Shibata, 2003). This is accomplished by using comparators that compare matching scores and a common ramp voltage signal. The shorter delay time corresponds to the larger matching score. The time-domain WTA circuit (Ito et al., 2001; Yamasaki & Shibata, 2003) utilizes an open-loop OR-tree architecture to sense the first up-setting signal and generates the binary address representing the location of the winner. In this manner, the maximum-likelihood template vector is identified.

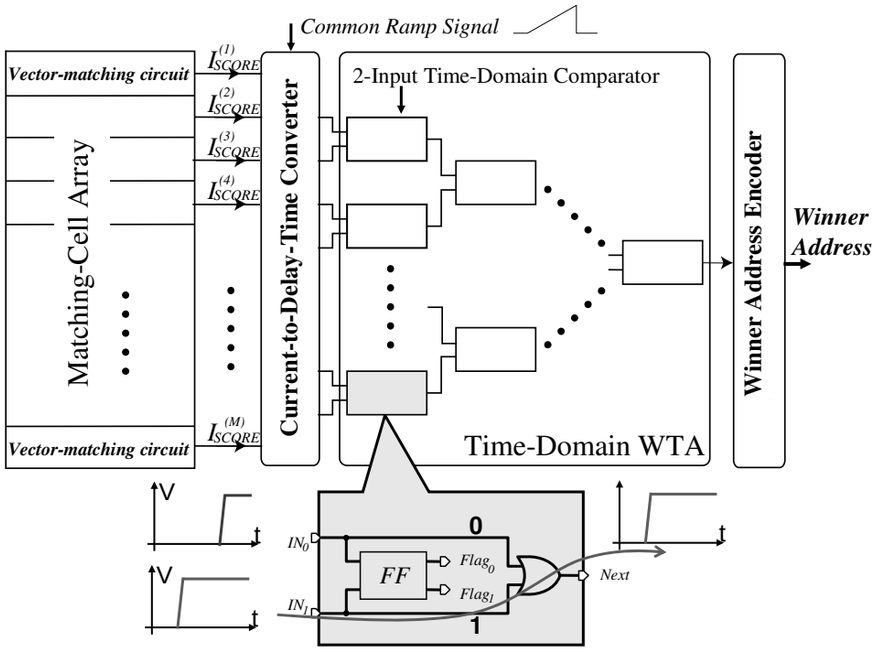


Fig. 4. Block diagram of the time-domain WTA, the flip-flop (FF) compares the timing difference between two input signals and senses the winner. The winner signal is also propagated to the next stage through the OR gate.

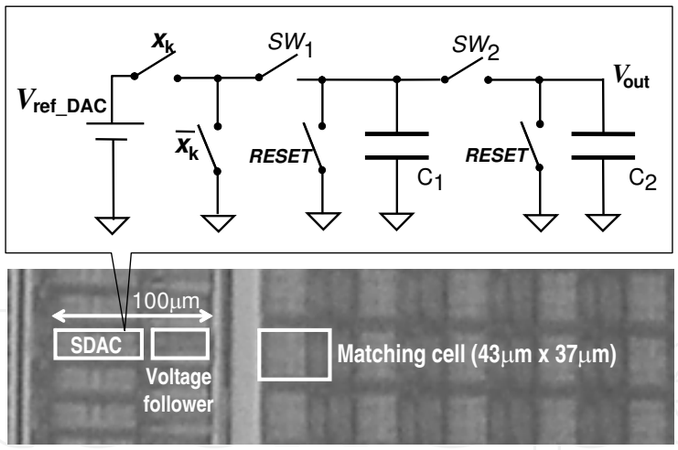


Fig. 5. Simplified schematic of SDAC and its layout area on the chip.

2.2.3 Serial digital-to-analog converter

As shown in Fig. 1, two digital-to-analog converters (DACs) are required for each of the vector elements since each matching cell requires two analog complementary signals; hence, 128 DACs are utilized in the system. Such an on-chip DAC needs to satisfy the requirement of small layout area, low-power dissipation, and small number of interconnects for data input. In this system, a serial digital-to-analog converter (SDAC) is utilized. The simplified schematic of the SDAC is shown in Fig. 5. The key feature of such a SDAC is its simplicity. It requires only two identical capacitors (C_1 and C_2) and a few switches. Basically, the

operation of the SDAC is based on charging and sharing charges between two capacitors. The conversion is done sequentially; one clock cycle is required to convert one bit. Thus, N clock cycles would be required for an N-bit word. The output voltage, V_{out}, is proportional to the serial input data, as illustrated by eq. (2).

$$V_{out} = \frac{1}{2}\left\{\ldots\frac{1}{2}\left[\frac{1}{2}(b_0 V_{ref\_DAC}) + b_1 V_{ref\_DAC}\right] + b_2 V_{ref\_DAC} + \ldots\right\}$$
$$V_{out} = V_{ref\_DAC}\left(\frac{b_0}{2^N} + \frac{b_1}{2^{N-1}} + \ldots + \frac{b_{N-1}}{2}\right) \quad (2)$$

Because of its small size, the SDAC is a much better choice for the proposed architecture. Its layout area compared with the layout area of a matching cell is also shown in Fig. 5.

2.2.4 Calibration circuitry

Process variations influence device parameters, and hence matching-circuit behaviors. The matching result, therefore, may lead to errors. The new calibration scheme shown in Figs. 6 and 7 has been developed to mitigate the errors caused by device mismatch. According to the International Technology Roadmap for Semiconductors (ITRS-2008), transistor

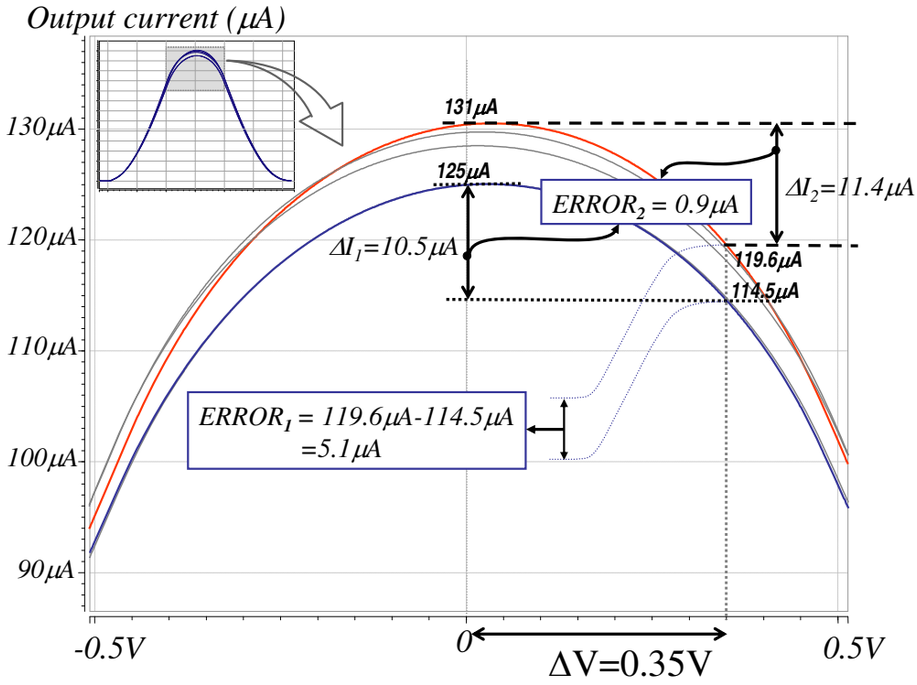


Fig. 6. Two distance-evaluating methods. Curves were generated by a 5-interation post-layout Monte Carlo simulation of a matching cell having random changes of 10% in transistors' length and width. The simulation was carried out at V_{DD} = 3.3 V and V_{ref} =1.65 V. Highest and lowest current curves were focused on. For the same distance between the input vector element and the template vector element, ΔV = 0.35 V, for example, the conventional distance-evaluating method and the proposed method are demonstrated.

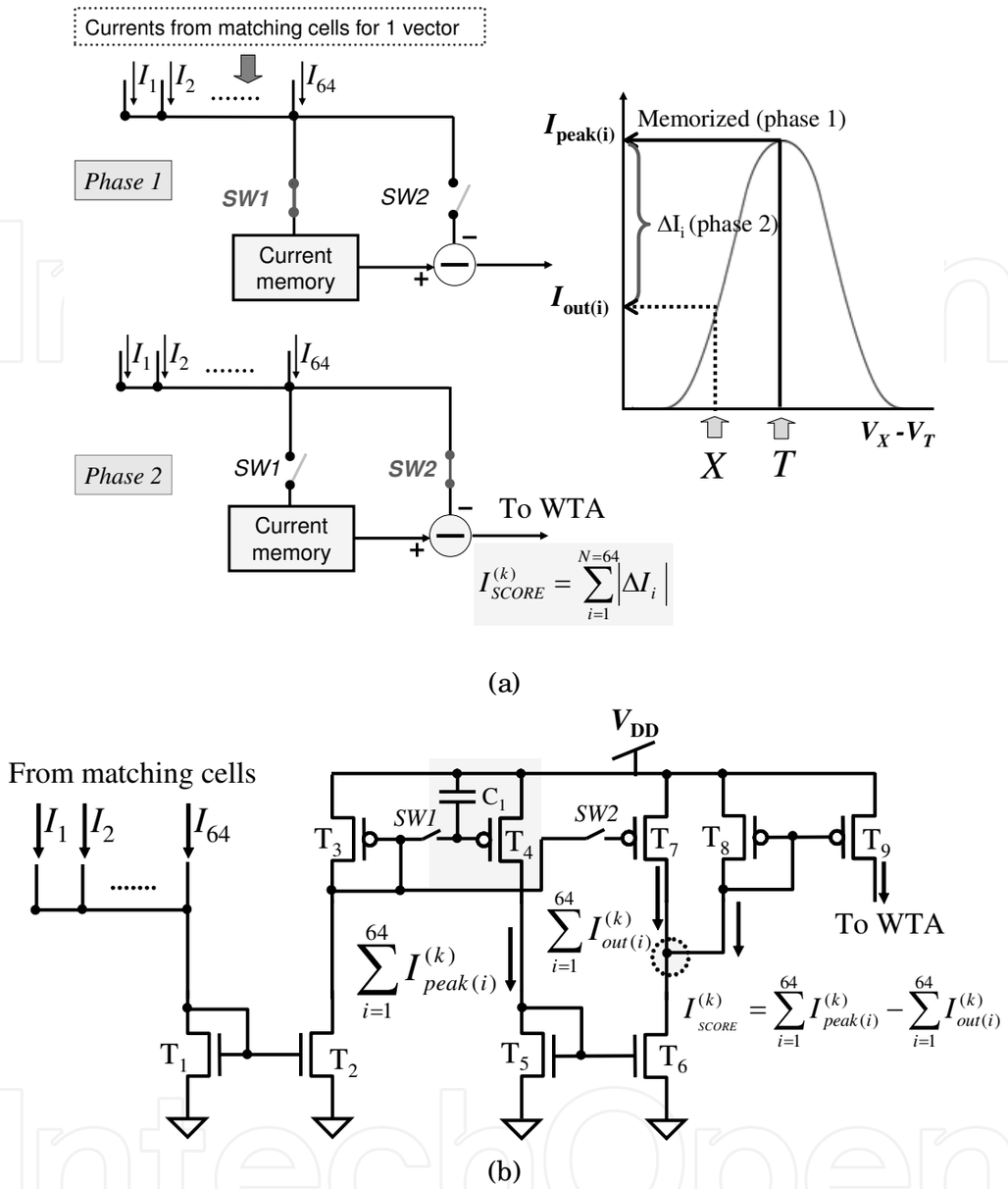


Fig. 7. Calibration scheme. (a) Calibration scheme operation. In phase 1, all peak output currents are memorized in current memories. In phase 2, the similarities between the input vector and the template vectors are evaluated. Only one current memory is required for one vector-matching circuit. (b) Circuit diagram of the current memory and subtractor.

dimensions may vary above 10%. The small figure at the top left of Fig. 6 illustrates matching-cell characteristics where the widths and the lengths of NMOS transistors of the matching cell vary randomly up to 10% as a result of process variations. These characteristics were obtained by a post-layout extracted circuit Monte Carlo simulation, and we focus on the highest and the lowest current curves. For the same distance between the input vector element and the template vector element, ΔV = 0.35 V, for example, two distance-evaluating methods are shown in the remaining part of Fig. 6, which is an enlarged

image of the small rectangle at the top left. In the proposed method, the similarity is determined by the difference between the peak current and the output current at the moment of data matching. In the previous conventional approaches (Delbruck, 1991; Hasler et al., 2002; Yamasaki & Shibata, 2003; Ogawa & Shibata, 2001; Peng et al., 2005), the output current itself was utilized as the matching result. $ERROR_2$ (0.9 μA) and $ERROR_1$ (5.1 μA) in the figure refer to errors caused by the former method and the latter one, respectively. It is clearly shown that the proposed differential current method offers a better result. In order to implement this method, peak currents are stored in current memories in phase 1 (the writing phase), namely, at the time of template data download to matching cells. In phase 2 (the evaluation phase), differences between currents are obtained. Only phase 2 is repeated for each new input vector. This scheme is shown in Fig. 7(a), and the circuit diagram of the current memory and subtractor is presented in Fig. 7(b). The matching scores between input vector and template vectors are calculated by eq. (3).

$$I_{\mathrm{SCORE}}^{(k)} = \sum_{i=1}^{64} \left| I_{\mathrm{peak}(i)}^{(k)} - I_{\mathrm{out}(i)}^{(k)} \right| = \sum_{i=1}^{64} I_{\mathrm{peak}(i)}^{(k)} - \sum_{i=1}^{64} I_{\mathrm{out}(i)}^{(k)} \tag{3}$$

According to this scheme, the greater similarity corresponds to the lower current rather than the higher one in the previous approaches.

2.3 Experimental results

2.3.1 Chip fabrication

The proof-of-concept chip was designed and fabricated using 0.35-μm 2P3M CMOS technology. The proposed matching-cell module includes 32 template vectors for the purpose of demonstration. The mechanism is preserved even in the case of a larger number of template vectors. The chip micrograph is shown in Fig. 8. The chip size is 4.9×4.9 mm², and the features of the chip are summarized in Table 1.

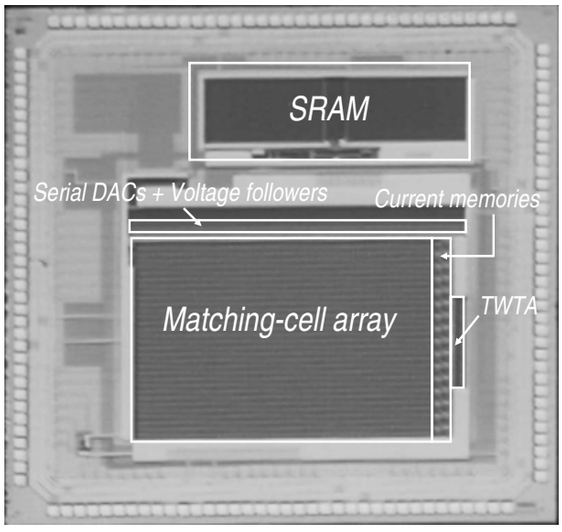


Fig. 8. Micrograph of the proof-of-concept chip fabricated using 0.35-μm CMOS process.

2.3.2 Measurement results and discussion

The measured characteristics of the vector element matching cell with various values of the reference voltage are illustrated in Fig. 9. Since the NMOS threshold voltage is around 0.6 V in the 0.35-μm CMOS technology in which the test chip was fabricated, it is shown that by varying V_{ref} from high to low values, the operation of the matching cell is altered from the above-threshold regime to the subthreshold regime, respectively. When operating in the subthreshold regime, the peak output current becomes as low as 80 nA at V_{ref} of 0.4 V. The results suggest an opportunity for building very low-power information processing systems.

Technology	2P3M 0.35-μm CMOS Process
Power supply (V)	3.3 (maximum)
Die size (mm²)	4.9 × 4.9
Number of vectors	32 vectors, 64 dimensions
Frequency (MHz)	33.3
Power consumption (mW)	21 at V_{ref} = 0.55 V, V_{DD} = 3 V, Clk = 33.3 MHz
Matching time (μs)	2.2 at 33.3 MHz

Table 1. Specifications of the proof-of-concept single-core chip.

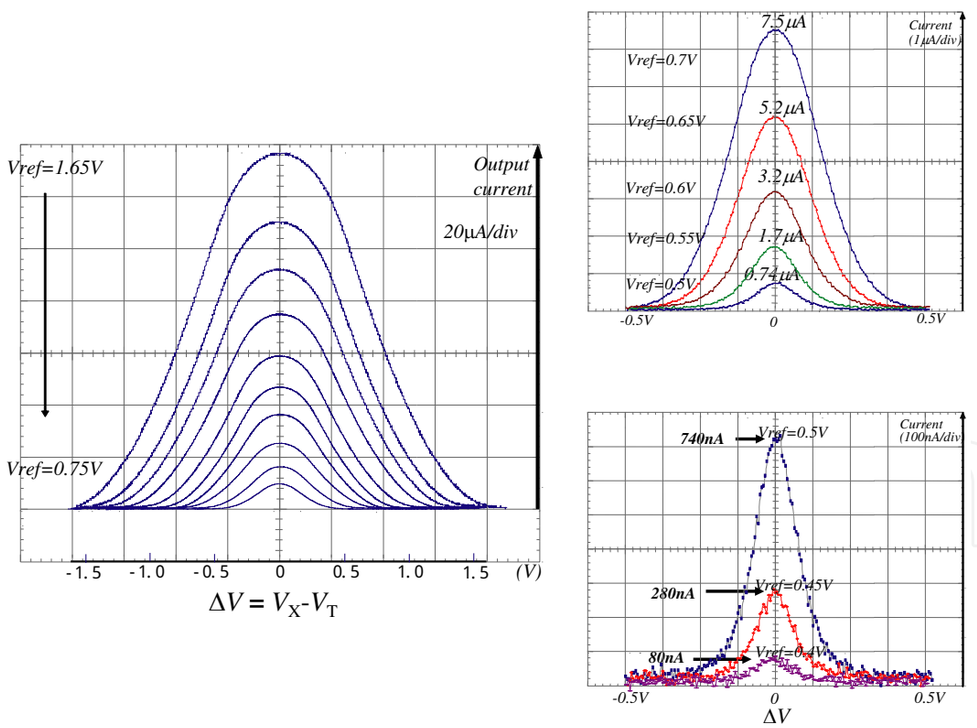
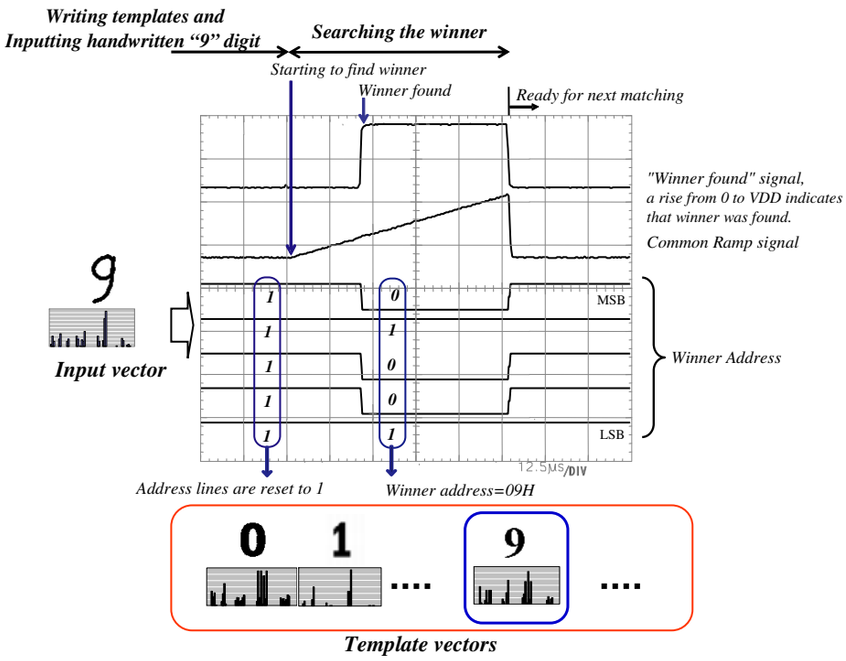


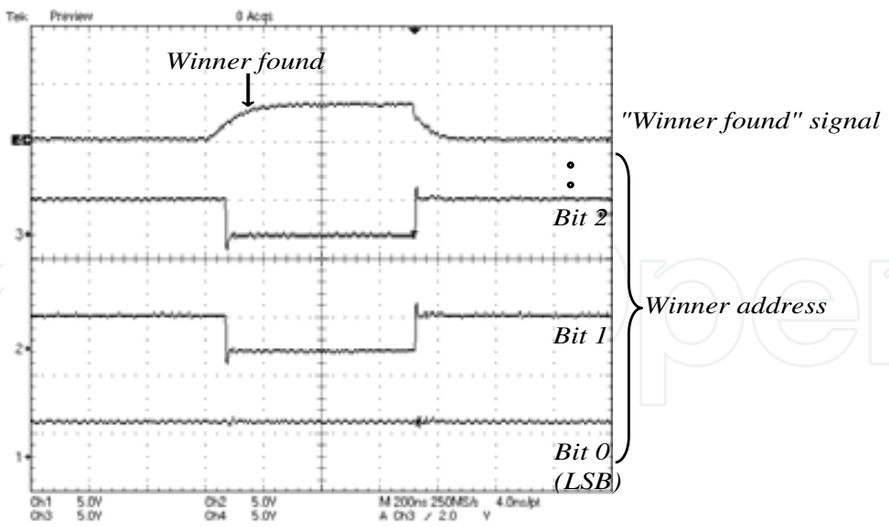
Fig. 9. Measured characteristics of the matching cell with various values of the reference voltage.

Figure 10 illustrates the experimental results for handwritten digit recognition utilizing the proposed architecture, as a simple demonstration. The digits "0"-"9" were converted to

PPED vectors so as to play the role of template vectors. The twenty-two other template vectors were dummy vectors. Then, the PPED vector of the handwritten digit "9" was employed as the input vector. The winner address shown in Fig. 10(a) corresponds to the location of the digit "9". This result verifies correct chip operation. Figures 10 (a) and 10 (b)



(a)



(b)

Fig. 10. Demonstration of the whole system operation. (a) Waveforms obtained with a logic scope describe the chip operation at 1 MHz for the purpose of illustration. The operating frequency is low because of the resolution limitation of the logic scope. (b) Waveforms obtained using an oscilloscope verify the chip operation at the frequency of 33.3 MHz

show the waveforms captured from a logic scope and an oscilloscope, respectively. Since 72 clock cycles, comprising 8 cycles for SDAC and 64 cycles for an off-chip digital- to-analog converter utilized as the ramp-signal generator for the WTA circuit, are required to finish a template-matching cycle, the search time in this experiment is 2.2 μs at the frequency of 33.3 MHz and depends strongly on the speed of the ramp-up voltage signal employed in the current-to-delay-time converter. The system was set up to operate at the supply voltage of 3.0 V and the reference voltage of 0.55 V. As a result, the average power dissipation of the whole chip was about 21 mW.

Moreover, in Fig. 11, the average supply current of the whole chip, including the matching-cell array, the SRAM module, SDACs, voltage buffers, current memories, the WTA circuit, and I/ O pads, measured with various V_{ref}'s is reported. The curves inherit the NMOS *I-V* characteristics owing to the NMOS-based matching-cell architecture. It can be observed that low supply currents are obtained with values of V_{ref} below the threshold voltage. These low reference voltages enable matching cells to operate in the subthreshold regime, in which the matching cell output currents drop exponentially with decreasing V_{ref}. As a result, the matching-cell array consumes very low power. Since the measured currents are for both the matching-cell array and the other parts, the supply currents in the subthreshold region remain at certain values rather than very low ones. These currents are mainly for the other parts whose power dissipations are reduced when lowering the supply voltage, and are independent of V_{ref}. Consequently, the supply currents are approximately constant values in the subthreshold region, as shown in Fig. 11.

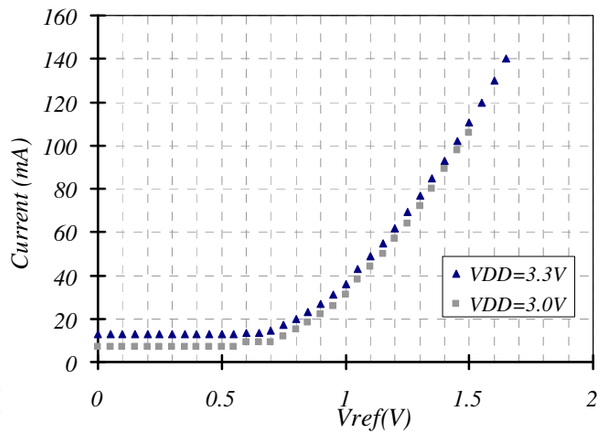


Fig. 11. Relationship between V_{ref} and supply current.

The performance of the associative processor is summarized with some others from the literature in Table 2. Because the time-domain WTA is utilized in this work because of its simple architecture, the search time is quite long compared with those of digital implementation (Nakata et al., 1999) and mixed signal implementation (Abedin et al., 2007). In addition to the matching-cell array, the WTA plays an important role in the power-saving scheme because the power consumption of the WTA increases significantly upon increasing the number of template vectors. In the present chip, the optimization of the speed and power dissipation of the WTA has not been considered. In order to make the proposed architecture practical and much better than digital approaches, a low power WTA would be considered in future studies. Furthermore, although analog flash implementation (Kramer et

al., 1997) offers very low power consumption, such an implementation requires particular mechanisms in the template-writing phase, making the flash implementation difficult to control and hence, flexible programmability difficult to realize.

	Technol.	Power consumption (mW)	Matching time (μs)	Estimated power/ MC (mW)
This work	Analog	21 (32 vectors, 64 elements)	2.2	0.01
Tuttle, *et al.* 1993	Analog	50*) (256 vectors, 16 elements)	2	0.012
Kramer, *et al.* 1997	Analog flash	195 (4K vectors, 64 elements)	4.6	0.00074
Oike, *et al.* 2004a	Digital	320.7 at V_{DD}=1.8V 15.1 at V_{DD}=0.9V (64 vectors, 32 elements)	2 ~8.12	0.157 0.0074
Nakada, *et al.* 1999	Digital	290 (256 vectors, 16 elements)	1.1	0.071
Abedin, *et al.* 2007	Mixed signal	195 (64 vectors, 16 elements)	0.16	0.19

Table 2. Performance comparison.
*)Not including power for memory and D/ A converters.

3. Extension to a multi-core/multi-chip architecture of associative processors

3.1 Multi-core/Multi-chip configuration

In this session, a solution to how the system is hierarchically scaled up to a vast scale integration is presented. For a vast scale integrated system, a large number of template data can be implemented in multiple associative processors, thus making the recognition system more intelligent. In this regard, a multi-core/ multi-chip architecture of associative processors has been developed (Bui & Shibata, 2008b; 2009).

In the literature, several multi-chip architectures based on all-digital technology have also been introduced (Nakata et al., 1999; Oike et al., 2004b). Although these systems offer accuracy, they occupy large chip real estate and usually have complicated structures. On the contrary, analog-technology-based system employing time-domain winner-take-all (WTA) is introduced in this study. The multi-core/ multi-chip architecture inherits the architecture developed for the fully parallel single-core associative processor described in the previsous session. The problem associated with inter-chip communication delay which is critical in the time-domain WTA operation has been resolved by a newly-developed winner-code-decision scheme. In addition, switched-current technology has been utilized so as to further reduce the power consumption.

The block diagram of a multi-core/ multi-chip associative system is shown Fig. 12. In general, the system includes many chips, and each chip itself has many cores. For the purpose of demonstration, the poof-of-concept system in this study is composed of four associative chips, namely, one master chip and three slave chips. Each chip consists of four

32-vector cores. (Each vector has 64 elements of 8-bit numbers.) As a result, a 512-vector associative system is constructed as a demonstration. The master chip and the slave chips are designed in the same configuration. They play master/ slave roles when they are combined to form the whole system and operate in parallel. The master chip is distinguished from other slave chips by activating an additional majority-code-decision circuit described in the following section. Employing many cores on a single chip reduces the time required for downloading the information of template vectors stored in SRAMs to analog matching-cell arrays. In addition, four cores are activated separately, thus they can do matching operations independently or as a whole large system.

The 32-vector single-core architecture was already described in the previous section. In each core, template vectors are stored in on-chip digital memory, namely SRAM in the design. Employing digital memories is an inexpensive solution instead of using high-cost analog nonvolatile memory technologies. And compact serial digital-to-analog converters (SDACs) are used to convert digital values to analog voltages prior to similarity evaluation processing. The similarity evaluation between the input vector and template vectors is carried out in parallel by vector-matching circuits, each of which consists of 64 bell-shaped vector-element matching cells (MCs), a current memory, and a current subtractor as shown in Fig. 13. Signals *WR* and *RD* in Fig. 13 correspond to *WRITE* control signal and *READ* control signal, respectively. These signals permit to store matching results represented by currents into the current memories and to read out the matching scores from the subtractors. As mentioned in Section 2, current memory plays an important role in the device-mismatch calibration scheme in which the similarity is determined by the difference between the peak current and the output current at the moment of similarity evaluation. In the study, switched-current technology is employed to control *RD* and *WR* signals in order to cut-off currents flowing in the vector-matching circuits as well as the current memories except moments of downloading template vectors to the matching-cell arrays and evaluating similarities. As a result, the power dissipation is reduced further as compared with the design in Section 2.

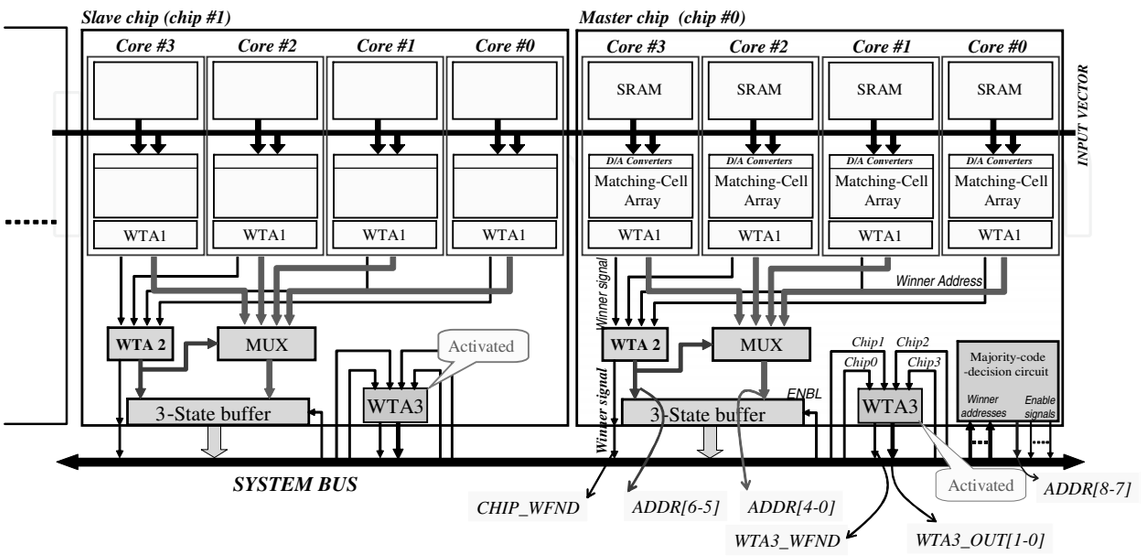


Fig. 12. Block diagram of the multi-core/ multi-chip architecture.

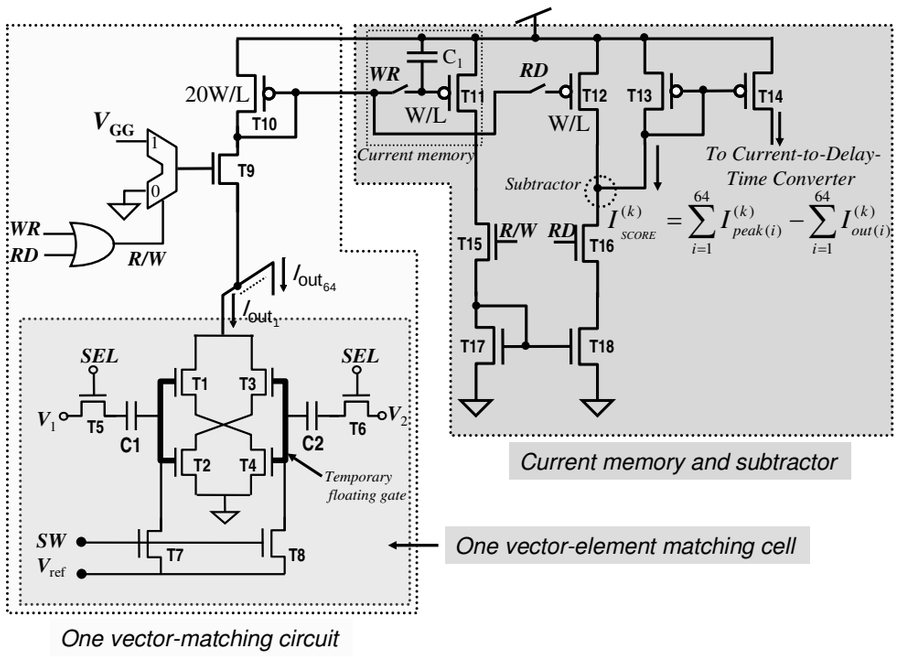


Fig. 13. Schematic of a vector-matching circuit.

Multi-core /Multi-chip configuration

The global winner, namely the template vector having the minimum distance to the input vector is searched for through a three-stage WTA circuit. Each WTA stage employing a time-domain WTA (Ito et al., 2001) senses the first up-setting signal among inputs and generates the binary address representing the location of the winner. The winner signal is also passed to the next WTA stage. In this manner, WTA1 searches for the local winner inside the 32-vector matching-cell array, WTA2 searches for the winner of one chip, and WTA3 searches for the global winner which is the winner when combining various chips together. All three WTA stages and the majority-code-decision circuit described below are layouted on each chip. The configuration is illustrated in Fig. 12.

However, when integrating several chips to form a larger system, signal propagation delays occurring in long inter-chip interconnects may lead to errors in time-domain signals. This will result in the decision error of the final WTA's (WTA3's). In order to deal with this problem, a balanced architecture should be satisfied to equalize delay times of inter-connection signals. However, even though with the balanced architecture, different propagation delays may still occur. Because of this problem, a redundant circuit following the final stage WTA, called the majority-code-decision circuit, has been developed. This circuit is only activated on the master chip. The circuit makes the decision based on the winner address codes generated by all WTA3's. The block diagram of the circuit is shown in Fig. 14. Basically, it consists of a binary counter, binary comparators, and a majority voting circuit (MVC). In the proof-of-concept chip, they are a 2-bit counter, 2-bit comparators, and a three-of-four MVC, respectively. As a result, the global result becomes more reliable than the architecture without a majority-code-decision circuit. In the case of a 2-bit 4-input majority-code-decision circuit like that in this study, the circuit can be constructed by combining two three-of-four MVCs whose outputs form the 2-bit majority winner code; but

it is not the general case. It means that such architecture is not correct for other cases whose winner codes are larger than two bits. On the contrary, the method developed in this study is general and suitable for any case. The counter counts up from zero when it is activated; the winner-indicating-signal (*ADDR_FND*) indicates whether the majority winner code is found. This signal goes high when output of the counter coincides with the majority winner code.

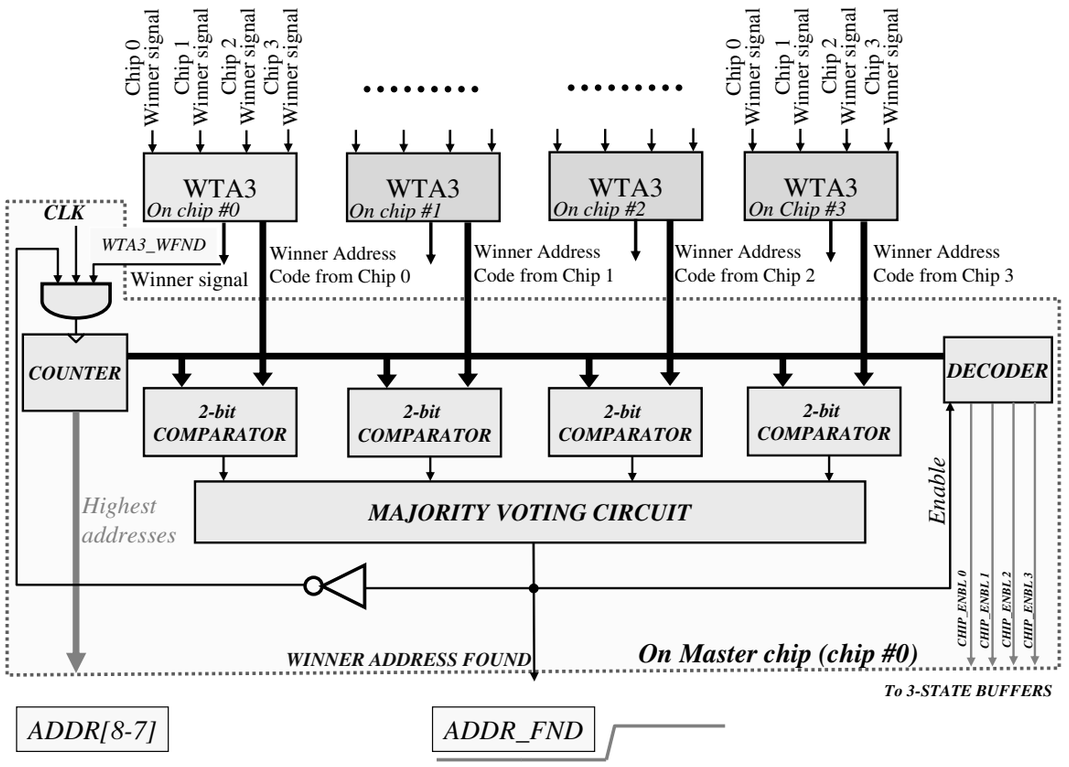


Fig. 14. WTA3 and the majority-code-decision circuitry.

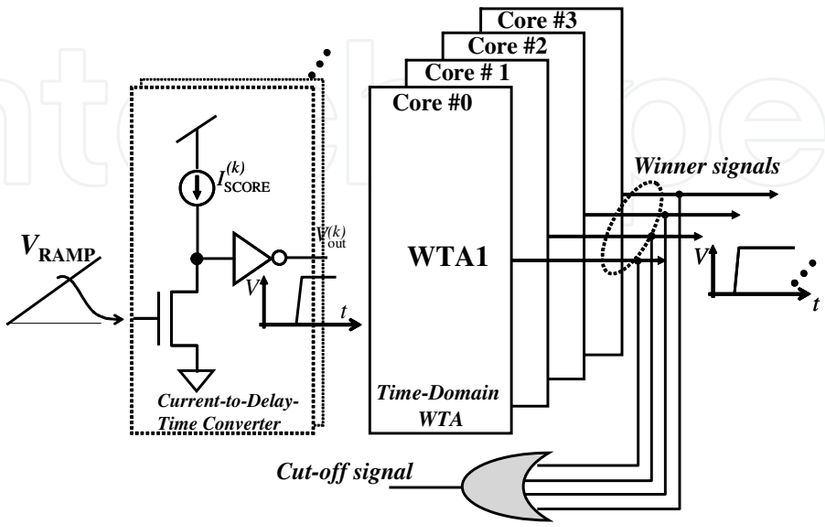


Fig. 15. Current-to-delay-time converter.

In addition, in order to further reduce the power dissipation, switched-current technology is also utilized in the current-to-delay-time converters by the method illustrated in Fig. 15. Winner signals obtained by WTA1's are combined by an OR-gate; the output signal is employed as a cut-off signal disconnecting both the common ramp voltage signal and score currents from current-to-delay-time converters. In this manner, once the winner signal is found by one of WTA1's, all current-to-delay-time converters are deactivated, thus further reducing the power consumption. This method can be applied to any large matching-cell array by dividing the array into several smaller blocks.

3.2 Experimental results

3.2.1 Chip fabrication

Measurement results obtained from the previous single-core chip fabricated in a 0.35-μm double-poly triple-metal CMOS technology have been discussed in Section 2. As an extended research, a proof-of-concept chip consisting of four cores was designed and fabricated in a 0.18-μm 5-metal CMOS technology. Figure 16 shows a micrograph of the test chip, and layout of a matching cell is shown in Fig. 17. Each core including a memory module and a matching-cell module occupies an area of 1760 μm × 570 μm. The size of matching cell is 19.7 μm × 7 μm. It should be noted again that the CMOS inverter-based matching cell presented in (Yamasaki & Shibata, 2003) is larger than the present cell due to the N-well region required for implementing PMOS transistors. This is an advantage of pure NMOS configuration. However, the present matching cell size is still large due to the large area required for capacitor layout. The specifications of the proof-of-concept chip are summarized in Table 3.

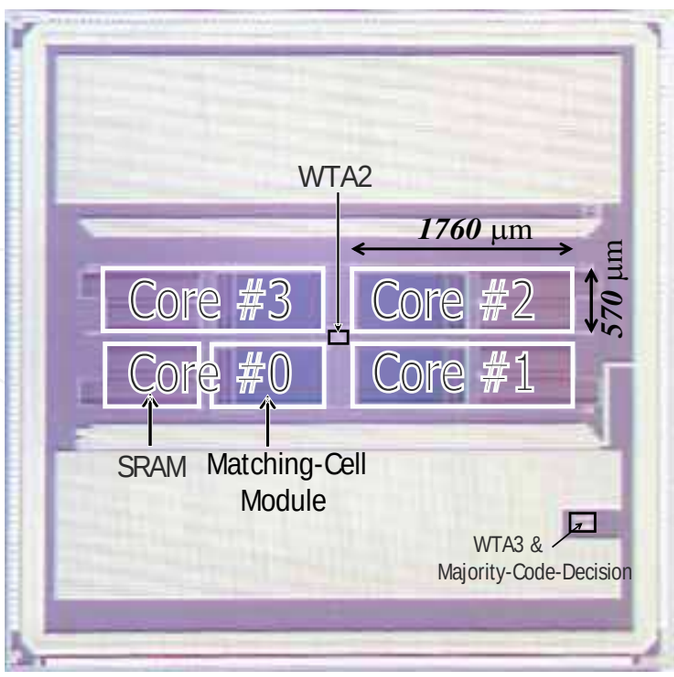


Fig. 16. Micrograph of the proof-of-concept chip fabricated using 0.18-μm CMOS process.

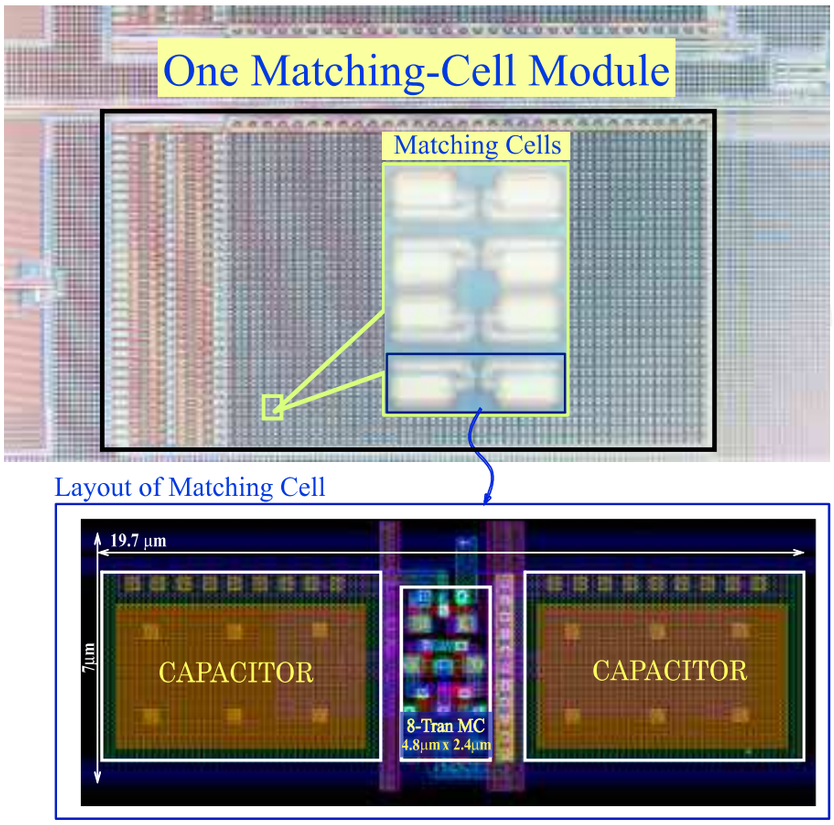


Fig. 17. Micrograph of a matching-cell module and layout of a matching cell (MC).

Technology	1P5M 0.18-μm CMOS
Power supply (V)	1.8
Core size (mm^2)	1.76×0.57
Matching cell size (μm^2)	19.7×7
Search time (μs)	8.16 at clock frequency = 16.7MHz (Incl. 8 clocks for SDAC and 128 clocks for the ramp voltage)
Power consumption (mW)	1.17 mW/ 32-vector matching-cell module; 6.48 mW/ chip when operating in the subthreshold region with V_{DD}=1.8 V.
Function	128 vectors/ chip, 512 vectors/ 4-chip system. Nearest match identification.

Table 3. Specifications of the proof-of-concept chip.

3.2.2 Measurement results

Figure 18(a) shows the characteristics of matching-cell measured with some small reference voltages. For the 0.18-μm CMOS technology in which the prototype chip has been fabricated, the threshold voltage of NMOS is around 0.45 V. As can be seen in the figure, in the subthreshold regime, the peak current of the matching cell characteristics is reduced to only several tens of nA. This is an important issue in power-saving schemes. The entire

curve of peak output current with respect to the reference voltage shown in Fig. 18(b) has the shape of NMOS transistor characteristics.

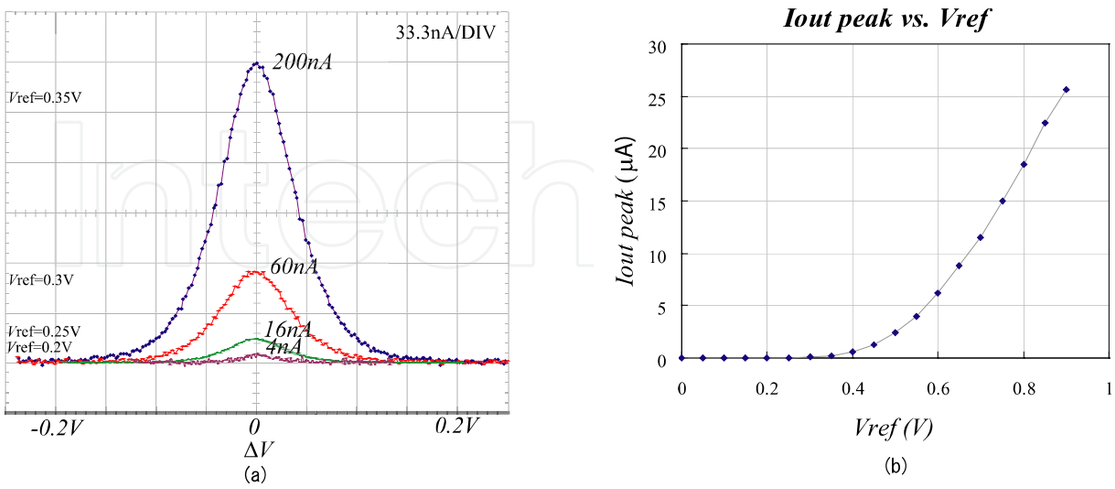


Fig. 18. Measured matching cell characteristics.

In Fig. 19, the average current of the whole chip including four cores and chip-I/ O buffers and the current in a single 32-vector matching-cell module measured with various V_{ref}'s are reported. As can be seen from the figure, the curves have the shape of the NMOS *I-V* characteristics owing to the NMOS-based matching-cell architecture. In the subthreshold region, the current of the entire chip and that of one matching-cell module are 3.6 mA and 0.65 mA, respectively. As a result, the power consumption per matching cell is reduced to as small as 0.79 μW.

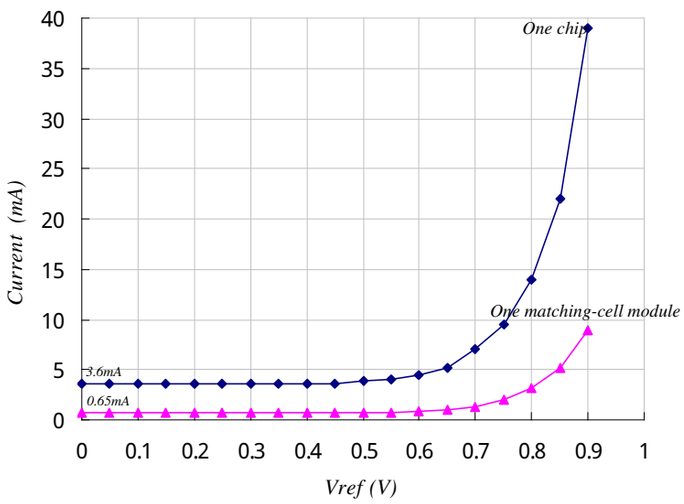


Fig. 19. Measured current as a function of the reference voltage V_{ref}.

Figure 20 (a) shows measured signals *CHIP_WFND* and *WTA3_WFND* generated by the WTA2 and WTA3 on the master chip, respectively. Waveforms at the output of the majority-code-decision circuit measured by an oscilloscope are shown in Fig. 20(b). The signal *WTA3_WFND* generated by WTA3 is employed as the control signal enabling the operation

of the counter in Fig. 14. When the winner is found by the WTA3 on the master chip, the counter is activated, and begins to count up. When the counter output, *ADDR[8-7]*, coincides with the majority winner code, *ADDR_FND* signal goes high, indicating that the majority code was found and available on address lines *ADDR[8-7]*. This signal also stops the counter counting. In the demonstration, the majority winner code is 10_2 corresponding to chip #2. Majority-making-decision principle plays an important role not only in this design of a multi-chip architecture but also in miniscule-device-based designs where the device parameter variability is an important issue.

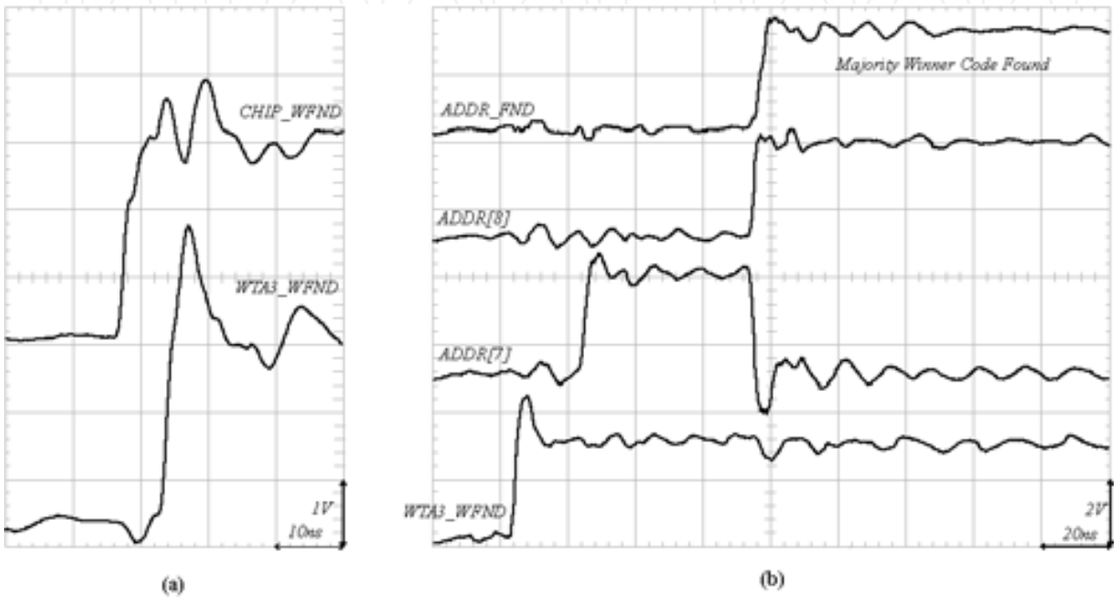


Fig. 20. Measured waveforms of the majority-code-decision circuit operating at a clock frequency of 20 MHz.

Demonstration of the whole system operation is illustrated in Fig. 21. All vectors of the test chip were assigned with given data. Required signals were connected to illustrate a system consisting four chips. After all template vectors were temporarily memorized inside matching-cell arrays, two input vectors were applied to the system input successively for matching. In Fig. 21, which is the measurement result captured from a logic scope, address lines *ADDR[4-0]*, *ADDR[6-5]*, and *ADDR[8-7]* represent winner address codes generated by WTA1, WTA2, and the majority-code-decision circuit, respectively. Namely, they are the winner template vector inside the winner core, the winner core inside the winner chip, and the winner chip of the multi-chip configuration, respectively. As a result, the global winner address is the combination of these three address codes. In this demonstration, the global winner addresses captured on the system bus are "100000101_2" representing the global winner is vector #5 (00101_2) of core #0 (00_2) in chip #2 (10_2) and "101010111_2" representing the global winner is vector #23 (10111_2) of core #2 (10_2) in chip #2 (10_2), respectively. *WTA_EVAL* signal enables the operation of the three-stage WTA circuitry. When this signal goes high, it also enables an off-chip ADC to generate the common ramp voltage used in current-to-delay-time converters. *GLOBAL_WFND* signal indicates that the winner template vector has been found and its address is available on the system bus. This signal also latches the global winner addresses on the system bus. The experimental results verify the correct operation of the system.

A searching cycle finishes in 136 clock cycles including eight clocks for on-chip D/ A conversion of an input vector and 128 clocks for off-chip ramp voltage generation. In addition, employing many cores on a single chip reduces the time required for downloading the information of template vectors to analog matching-cell arrays.

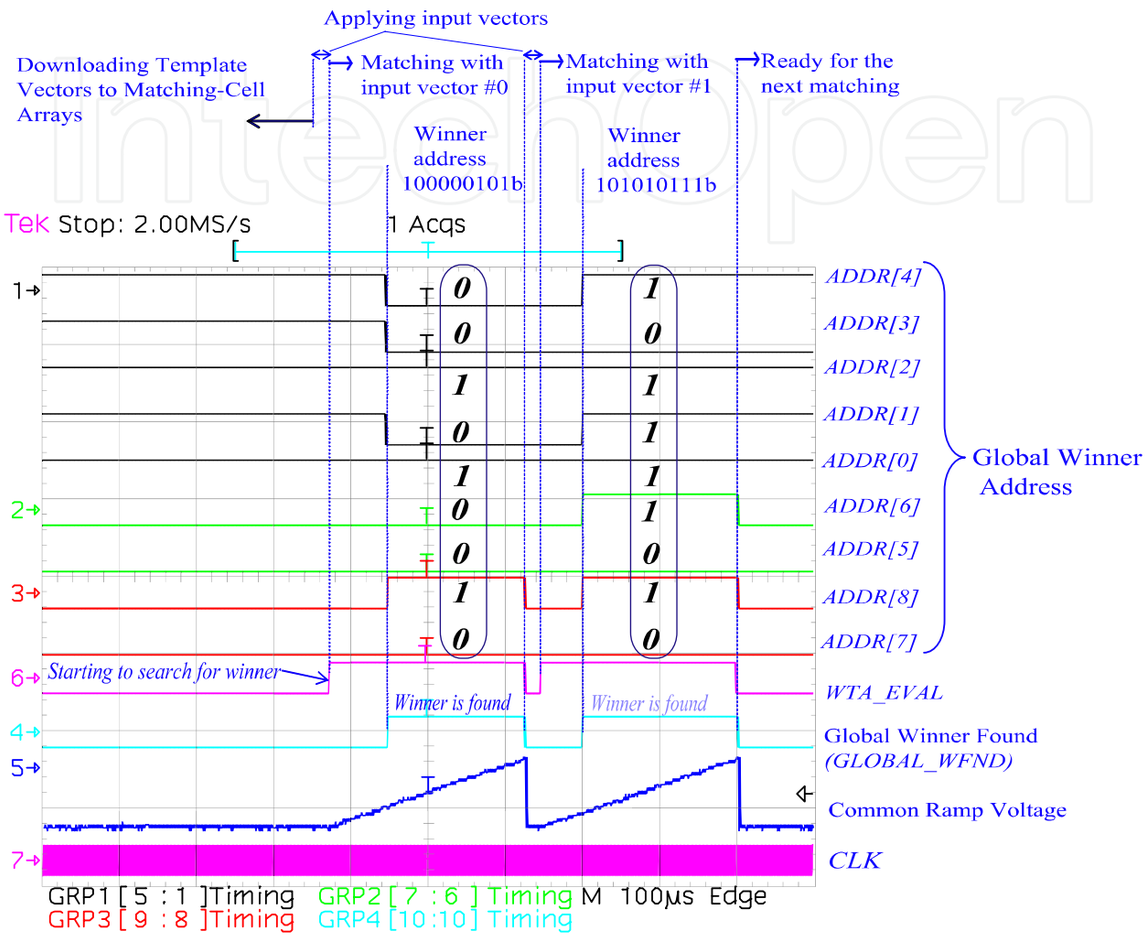


Fig. 21. Demonstration of the whole system operation by waveforms captured by a logic scope.

4. Conclusion

In this chapter, a methodology for building a low-power high-capacity associative system has been presented. Device mismatch problems as well as decision errors associated with inter-chip communication delays have been resolved by introducing the calibration scheme and the majority-code-decision circuit. Because of employing bell-shaped matching cell as similarity/ dissimilarity-evaluation element, this study, therefore, provides an intermediary stage connecting CMOS designs and the coming era of nano devices. This is because such resonance-type current-voltage characteristics are typical characteristics often observed in nano-scale devices. The system also has the possibility of a large database capacity by employing the multi-core/ multi-chip architecture. In principle, search time is independent of the number of cores as well as the number of chips. The operation of the systems as well

as the concept of design have been verified by measurement results of the proof-of-concept chips designed in 0.35-μm and 0.18-μm CMOS processes.
The number of cores per chip as well as the number of chips in the system can be enlarged to a larger capacity without changing the methodology, making the system more intelligent by increasing the number of template vectors, i.e., giving a larger amount of knowledge in making decisions. Moreover, another method to enlarge the system is to employ WTA3's as extended WTA stages in a tree-like architecture. In principle, the system can be enlarged to any extent, making such an approach promising in giga-scale integration of nano devices.
In addition, as for associative processors developed in this chapter, template data are temporarily stored as voltages on capacitors. Due to design rules provided from the foundry, most of the matching-cell area is occupied by capacitors. If we assume a high-k MIM capacitance technology, such as the technology used in DRAM, to be available in the technologies used in this chapter's designs, the matching-cell array can be constructed in a very small area. It can be concluded that the combination of the high-k MIM capacitance technology with the architectures developed in this study would be a promising technology for analog implementations of associative processors in future.

5. Acknowledgements

The VLSI chips in this paper were fabricated as part of the chip fabrication program of the VLSI Design and Education Center (VDEC), the University of Tokyo, in collaboration with Rohm Corporation and Toppan Printing Corporation.

6. References

Abedin, Md. A.; Tanaka, Y.; Ahmadi, A.; Koide, T. & Mattausch, H. J. (2007). "Mixed Digital-Analog Associative Memory Enabling Fully Parallel Nearest Euclidean Distance Search", *Japanese Journal of Applied Physics (JJAP),* Vol. 46, 2007, pp. 2231-2237.

Bui, T.T. & Shibata, T. (2008a). "Compact Bell-Shaped Analog Matching-Cell Module for Digital-Memory-Based Associative Processors," *Japanese Journal of Applied Physics (JJAP)*, vol. 47, No. 4, April 2008, pp. 2788–2796.

Bui, T.T. & Shibata, T. (2008b). "A Multi-core/ Multi-chip Scalable Architecture of Associative Processors Employing Bell-Shaped Analog Matching Cells," *Proceedings of the 2008 9th International Conference on Solid-State and Integrated-Circuit Technology (ICSICT)*, pp. 1819-1822, Oct 20-23, 2008, Beijing.

Bui, T.T. & Shibata, T. (2009) "A Scalable Architecture of Associative Processors Employing Nano Functional Devices", *Proceedings of the 10th International Conference on Ultimate Integration of Silicon (ULIS)*, pp. 213-216, Mar.18-20, Aachen, Germany.

Cauwenberghs, G. & Pedroni, V. (1997). "A Low-Power CMOS Analog Vector Quantizer", *IEEE Journal of Solid-State Circuits,* Vol. 32, August 1997, pp. 1278-1283.

Delbruck, T. (1991). "Bump Circuits for Computing Similarity and Dissimilarity of Analog Voltages", *Proceedings of International Joint Conference on Neural Networks (IJCNN-91),* pp. 475-479, June 1991, Seattle.

Hasler, P.; Smith, P.; Duffy, C.; Gordon, C.; Dugger, J. & Anderson, D. (2002). "A Floating-Gate Vector-Quantizer", *Proceedings of the 45th Midwest Symposium on Circuits and Systems (MWSCAS-2002)*, pp. 196-199, August 2002, Oklahoma.

Ito, K.; Ogawa, M.; & Shibata, T. (2001). "A High-Performance Ramp-Voltage-Scan Winner-Take-All Circuit in an Open Loop Architecture", *Japanese Journal Applied Physics (JJAP),* Vol. 41, 2001, pp. 2301-2305.

Kobayashi, D.; Shibata, T.; Fujimori, Y.; Nakamura, T. & Takasu, H. (2005). "A Ferroelectric Associative Memory Technology Employing Heterogate FGMOS Structure", *IEEE Transactions on Electron Devices,* Vol. 52, Oct. 2005, pp. 2188-2197.

Konda, M.; Shibata, T. & Ohmi, T. (1996). "Neuron-MOS Correlator Based on Manhattan Distance Computation for Event Recognition Hardware", *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS 1996),* pp. 217-220, May 1996, Atlanta, Georgia.

Kramer, A.; Canegallo, R.; Chinosi, M.; Doise, D.; Gozzini, G.; Navoni, L.; Rolandi, P. L. & Sabatini, M. (1997). "55GCPS CAM Using 5b Analog Flash", in *IEEE International Solid-State Circuits Conference(ISSCC) Digest of Technical Papers*, pp. 44-45, Feb. 1997, San Francisco.

Matsumoto, K.; Ishii, M.; Segawa, K. & Oka, Y. (1996). "Room Temperature Operation of a Single Electron Transistor Made by the Scanning Tunneling Microscope Nanooxidation Process for the TiOx/ Ti System", *Applied Physics Letter,* Vol. 68, Jan. 1996, pp. 34-36.

Nakada, A.; Shibata, T.; Konda, M.; Morimoto, T. & and Ohmi, T. (1999). "A Fully Parallel Vector-Quantization Processor for Real-Time Motion-Picture Compression", *IEEE Journal of Solid-State Circuits,* Vol. 34, June 1999, pp. 822-830.

Ogawa, M. & Shibata, T. (2001). "NMOS-based Gaussian-Element-Matching Analog Associative Memory", *Proceedings of the 27th European Solid-State Circuits Conference ESSCIRC 2001*, pp. 257-260, Sept. 2001, Villach, Austria.

Oike, Y.; Ikeda, M. & Asada, K. (2004a). "A Word-Parallel Digital Associative Engine with Wide Search Range Based on Manhattan Distance", *Proceedings IEEE Custom Integrated Circuits Conference (CICC),* pp. 295-298, Oct. 2004, Orlando.

Oike, Y.; Ikeda, M.; & Asada, K. (2004b). "Hierarchical Multi-Chip Architecture for High Capacity Scalability of Fully Parallel Hamming-Distance Associative Memories", *IEICE Trans. Electron*, E87-C, Nov. 2004, pp. 1847-1855.

Peng, S.-Y.; Minch, B. A.; & Hasler, P. (2005). "A Programmable Floating-Gate Bump Circuit with Variable Width", *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS 2005)*, pp. 4341-4344, May 2005, Kobe.

Saitoh, M.; Harata, H. & Hiramoto, T. (2004). "Room-Temperature Demonstration of Integrated Silicon Single-Electron Transistor Circuits for Current Switching and Analog Pattern Matching", *Technical Digest IEEE International Electron Devices Meeting (IEDM),* pp. 187-190, Dec. 2004, San Francisco.

Shibata, T.; Yagi, M. & Adachi, M. (1999). "Soft-Computing Integrated Circuits for Intelligent Information Processing", *Proceedings of the Second International Conference on Information Fusion*, pp. 648-656, July 1999, Sunnyvale, California.

Tuttle, G. T.; Fallahi, S. & Adibi, A. A. (1993). "An 8b CMOS vector A/ D converter", in *IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers*, pp. 38-39, Feb. 1993, San Francisco.

Uchida, K.; Koga, J; Ohba, R. & Toriumi, A. (2002). " Programmable Single-Electron Transistor Logic for Low-Power Intelligent Si LSI", *IEEE International Solid-State*

Circuits Conference (ISSCC) Digest of Technical Papers, pp. 206-207, Feb. 2002, San Francisco.

Vlassis, S.; Fikos, G. & Siskos, S. (2001). "A Floating Gate CMOS Euclidean Distance Calculator and Its Application to Hand-Written Digit Recognition", *Proceedings of International Conference on Image Processing,* pp. 350-353, Oct. 2001, Thessaloniki, Greece.

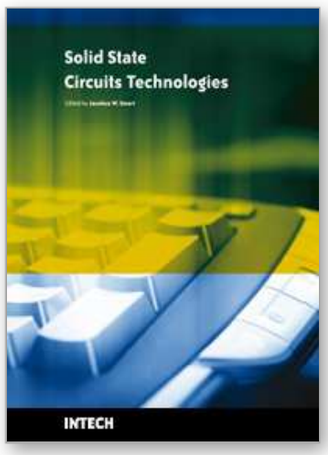
Yagi, M. & Shibata, T. (2003). "An Image Representation Algorithm Compatible with Neural-Associative-Processor-Based Hardware Recognition Systems", *IEEE Transactions on Neural Networks,* Vol. 14, Sept. 2003, pp. 1144-1161.

Yamasaki, H. & Shibata, T. (2007). "A Real-Time Image-Feature-Extraction and Vector-Generation VLSI Employing Arrayed-Shift-Register Architecture", *IEEE Journal of Solid-State Circuits,* Vol. 42, Sept. 2007, pp. 2046-2053.

Yamasaki, T. & Shibata, T. (2003). "Analog Soft-Pattern-Matching Classifier Using Floating-Gate MOS Technology", *IEEE Transactions on Neural Networks,* Vol. 14, Sept. 2003, pp. 1257-1265.

Yamasaki, T.; Suzuki, A.; Kobayashi, D.; & Shibata, T. (2001). "A Fast Self-Convergent Flash-Memory Programming Scheme for MV and Analog Data Storage", *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS 2001)*, pp. 930-933, May 2001, Sydney.

Yoon, S. M.; Tokumitsu, E.; & Ishiwara, H. (2000). "Ferroelectric Neuron Integrated Circuits Using $SrBi_2Ta_2O_9$-Gate FET's and CMOS Schmitt-Trigger Oscillators", *IEEE Transactions on Electron Devices,* Vol. 47, Aug. 2000, pp. 1630-1635.



Solid State Circuits Technologies

Edited by Jacobus W. Swart

ISBN 978-953-307-045-2
Hard cover, 462 pages
Publisher InTech
Published online 01, January, 2010
Published in print edition January, 2010

The evolution of solid-state circuit technology has a long history within a relatively short period of time. This technology has lead to the modern information society that connects us and tools, a large market, and many types of products and applications. The solid-state circuit technology continuously evolves via breakthroughs and improvements every year. This book is devoted to review and present novel approaches for some of the main issues involved in this exciting and vigorous technology. The book is composed of 22 chapters, written by authors coming from 30 different institutions located in 12 different countries throughout the Americas, Asia and Europe. Thus, reflecting the wide international contribution to the book. The broad range of subjects presented in the book offers a general overview of the main issues in modern solid-state circuit technology. Furthermore, the book offers an in depth analysis on specific subjects for specialists. We believe the book is of great scientific and educational value for many readers. I am profoundly indebted to the support provided by all of those involved in the work. First and foremost I would like to acknowledge and thank the authors who worked hard and generously agreed to share their results and knowledge. Second I would like to express my gratitude to the Intech team that invited me to edit the book and give me their full support and a fruitful experience while working together to combine this book.

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen