

# A High-Speed and Low-Voltage Associative Co-Processor With Exact Hamming/Manhattan-Distance Estimation Using Word-Parallel and Hierarchical Search Architecture

Yusuke Oike, *Student Member, IEEE*, Makoto Ikeda, *Member, IEEE*, and Kunihiro Asada, *Member, IEEE*

**Abstract**—A high-speed and low-voltage associative co-processor with exact Hamming or Manhattan distance estimation is presented. The word-parallel and hierarchical search architecture is achieved using a logic-in-memory digital implementation. In the bit-serial search architecture, it is important to shorten the search cycle time since the total search time generally increases in proportion to the bit length. The present hierarchical architecture achieves a high-speed operation with a large input number. Furthermore, it provides a result for the data close to the input with a fewer number of clocks. Therefore, it reduces the number of clocks required for nearest-match detection in practical use. The circuit implementation allows unlimited database capacity and achieves a low-voltage operation under 1.0 V for system-on-a-chip applications. The capacity scalability makes it easy to compute a function of Manhattan distance estimation using thermometer encoding. A 64-bit 32-word associative co-processor has been designed using a one-poly-Si five-metal 0.18- $\mu\text{m}$  CMOS process and has been successfully tested. The measurement results show that the operation achieves a speed of 411.5 MHz at a supply voltage of 1.8 V. The worst-case search time is 158.0 ns for a 64-bit 32-word database. In a low-voltage operation, the operation speed achieves 40.0 MHz at a supply voltage of 0.75 V.

**Index Terms**—Associative co-processor, content addressable memory (CAM), Hamming distance, hierarchical search, logic-in-memory architecture, Manhattan distance, word parallel.

## I. INTRODUCTION

SOME applications, such as data compression, pattern recognition, multimedia, and intelligent processing, require considerable memory access and data processing time. Therefore, content addressable memories (CAMs) have been developed to reduce the access and data processing time and to detect completely matched data in a database. In recent years, many advanced applications have required the detection of not only completely matched data but also near/nearest-match data. Conventional associative memories that employ analog circuit techniques have been proposed for quick nearest-match detection [1]–[4]. Generally, their circuit implementations are compact. However, there are difficulties in operating them with faultless precision in a deep-submicron (DSM) process and at a low-voltage supply. Moreover, the feasible database capacity is limited by the analog operation. Therefore, they are not suitable for a system-on-a-chip VLSI in DSM process technologies.

In this paper, we present a high-speed and low-voltage associative co-processor that uses a hierarchical search architecture capable of word-parallel Hamming or Manhattan distance estimation, which has been partially reported in [5]. It has three

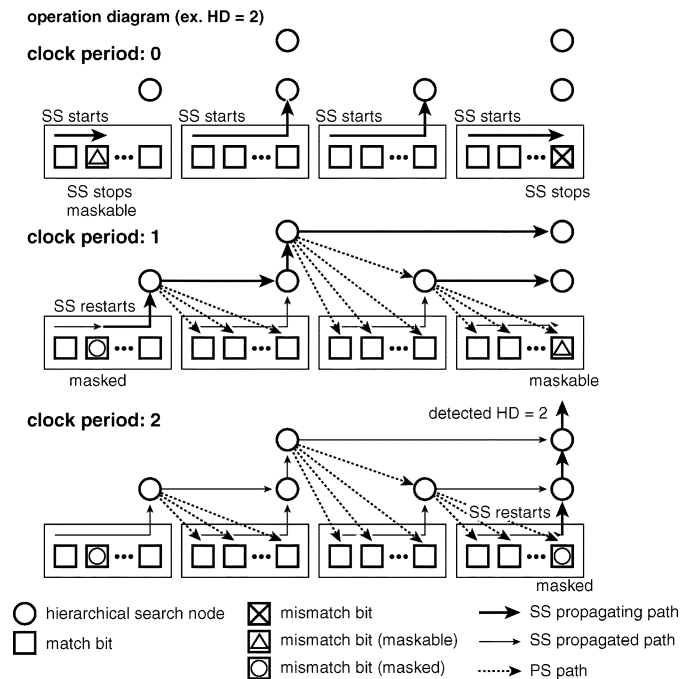


Fig. 1. Hierarchical search structure and operation diagram in the case of  $HD = 2$ .

principal advantages. The first advantage is that the hierarchical search architecture enables a high-speed search in a large database. The search cycle time is limited by  $O(\sqrt{N})$  or  $O(\log M)$  at an  $N$ -bit  $M$ -word data capacity. Although the total search time increases in proportion to the bit length, it reduces the number of clocks for nearest-match detection in practical use since it provides a result for the data close to the input with a fewer number of clocks. In addition, theoretically there are no limitations on the data patterns  $M$ , the bit length  $N$ , and the data distance  $D$ . The second advantage is a low-voltage operation in a DSM process. The circuit implementation has a tolerance for device fluctuation and allows a low-voltage operation of less than 1.0 V, which is difficult to attain using the conventional analog approaches. The third advantage is that it provides additional functions for associative processing. The present architecture provides data addresses with the exact Hamming or Manhattan distance sorted in order of the distance. Therefore, it enables high-speed data sorting in addition to nearest-match detection for conventional use. We have designed a 64-bit 32-word associative co-processor using a one-poly-Si five-metal (1P5M) 0.18- $\mu\text{m}$  CMOS process and have successfully demonstrated the high-speed distance estimation and low-voltage operation with faultless precision.

Manuscript received January 8, 2004; revised April 27, 2004.

The authors are with the University of Tokyo, Tokyo 113-8656, Japan (e-mail: y-oike@silicon.u-tokyo.ac.jp).

Digital Object Identifier 10.1109/JSSC.2004.831805

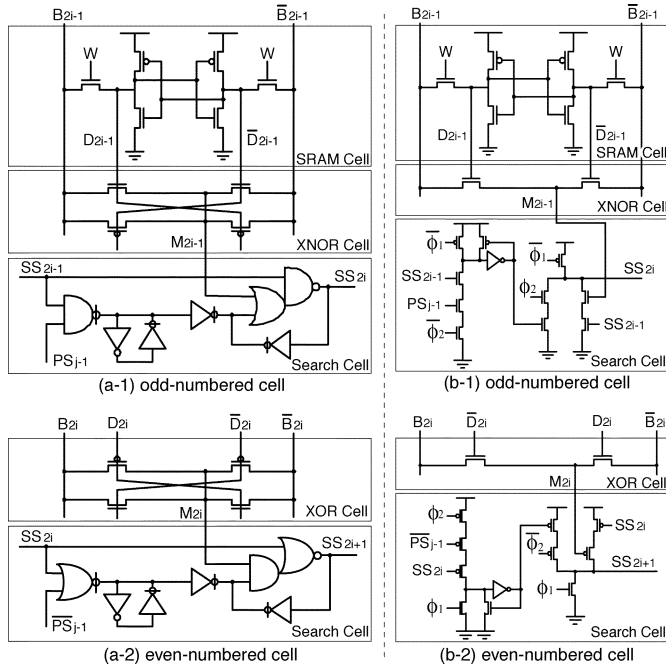


Fig. 2. Circuit configuration of the associative memory cell: (a-1) odd-numbered cell, (a-2) even-numbered cell of static circuit implementation; (b-1) odd-numbered cell, (b-2) even-numbered cell of dynamic circuit implementation (SRAM part of even-numbered cell is omitted).

## II. WORD-PARALLEL AND HIERARCHICAL SEARCH ARCHITECTURE

We propose a logic-in-memory architecture using search signal propagation via chained search circuits in word parallel. The Hamming distance (HD) search operation includes data comparison, search signal propagation, and mismatch masking. First, the input is compared with all the template data using an XOR gate in bit parallel. Then, the mismatch bits are counted in word parallel by the chained search circuits as shown in Fig. 1. The template data are divided into blocks and connected by hierarchical nodes since the search cycle time is limited by the search signal propagation via chained search circuits. The hierarchical node provides permission signals to the next block and the next hierarchical node. The permission signal makes a mismatch bit maskable.

Search signals (SS in the figure) are simultaneously injected to all blocks. The search signal passes through match bits via the search circuits. Some propagations are interrupted at the first-encountered mismatch bit. The others pass to the hierarchical nodes and update the permission signals for the next block and hierarchical node as shown by the clock period 0 in Fig. 1. In this period, the data with HD = 0 are detected since the search signal is provided from the last hierarchical node without any interruption. Only one mismatch bit, which interrupts the search signal propagation and receives a permission signal from the previous hierarchical node, becomes maskable in each word. During the next clock period, the search signal restarts from the masked bit and updates the permission signals again. It must be noted that the consumed clock cycles represent the Hamming distance of the detected data. For example, the data with HD = 2 are detected in the clock period 2 as shown in Fig. 1. In this manner, the data with HD =  $n$  are detected in

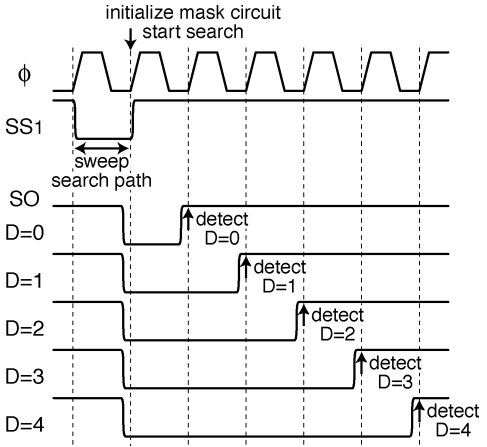


Fig. 3. Timing diagram of search circuit.

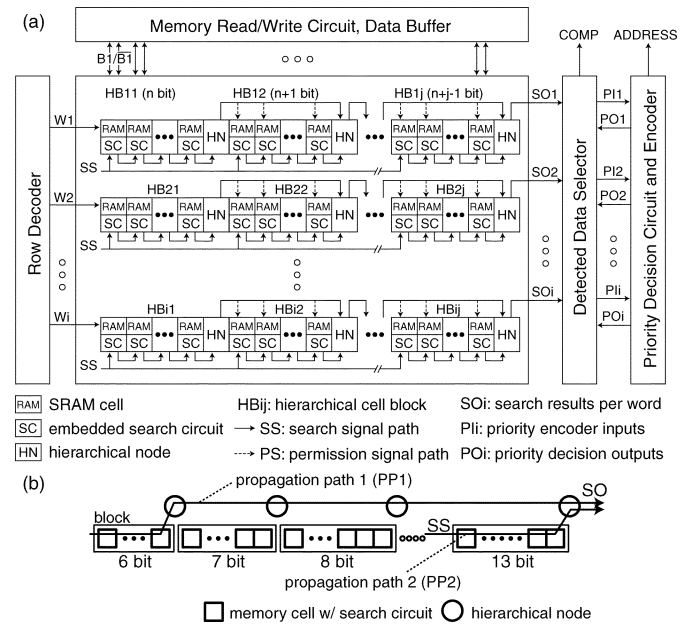


Fig. 4. Block diagram. (a) Associative co-processor. (b) Word structure.

the  $n$ -th clock period. Thus, the search operation can detect not only the nearest-match data but also all data sorted in order of Hamming distance in synchronization with the clock cycle.

All associative memories with Hamming distance estimation can deal with Manhattan distance estimation using thermometer encoding as reported in [4]. In general,  $k$ -bit binary data are translated to  $2^k - 1$  bit data using thermometer encoding. Hardware reusability for a wide variety of applications is important as an associative co-processor. The present architecture also has the capability of Manhattan distance estimation using thermometer encoding in the same operation as Hamming distance estimation.

## III. CIRCUIT CONFIGURATION

Fig. 2 shows a schematic of the associative memory cell implemented by static circuit implementation and dynamic circuit implementation. It is composed of an SRAM cell, an XOR/XNOR circuit for comparison with the input data, and a chained search circuit. The even-numbered and odd-numbered search circuits

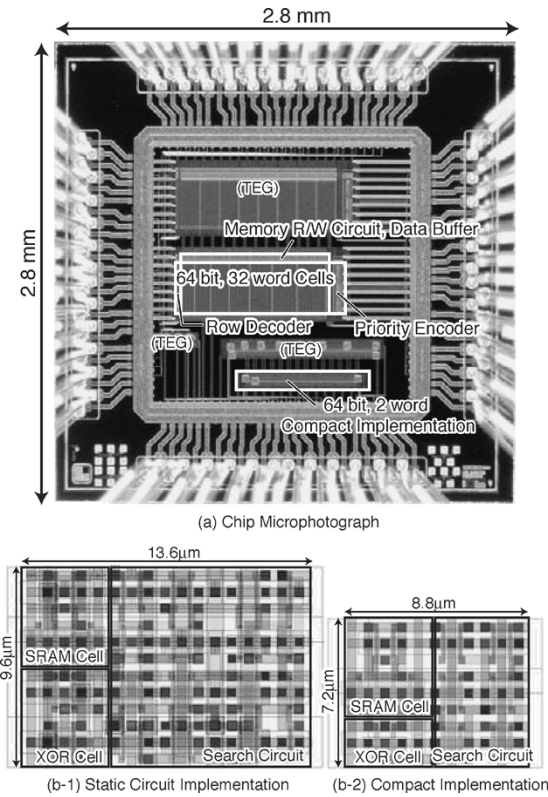


Fig. 5. Chip implementation: (a) chip microphotograph; (b-1) cell layout of static circuit implementation, (b-2) cell layout of dynamic circuit implementation.

are complementary in order to reduce the critical path and circuit area. The static circuit implementation enables a low-voltage operation and a high tolerance for device fluctuation though it occupies a large circuit area. On the other hand, the dynamic circuit implementation enables a small cell area and a large data capacity, however, it has lesser tolerance for power supply noise, cross-talk noise, and leakage current especially in a low-voltage operation.

Fig. 3 shows a timing diagram of the search circuit. All search paths are swept by setting the search signal to 0. Then, all mask registers are initialized before the search operation starts. For a match bit, the search signal passes to the next bit since the comparison result ( $M$ ) is true. For a mismatch bit, the search signal stops and waits for the next clock ( $\phi$ ). A false result of  $M$  results in masking by the next clock, and the search signal restarts from the masked cell where both the search signal and the permission signal ( $PS$ ) are true. Therefore, all data are detected in order of Hamming or Manhattan distance ( $D$ ) in word parallel as shown in Fig. 3. In dynamic circuit implementation, all search circuits are charged by  $\phi_1$  prior to the search operation. Then, a mismatch bit is masked by  $\phi_2$  in a similar manner as in the static circuit implementation.

#### IV. CHIP IMPLEMENTATION

We have designed and fabricated a 64-bit 32-word associative co-processor with the static circuit implementation using a 1P5M 0.18- $\mu\text{m}$  CMOS process. Fig. 4 illustrates a block diagram of the associative co-processor. Fig. 5 shows the chip

(a) Hamming distance estimation

**Input data (64 bit)**  
1011110011010110000010110001111000111010111101001010100100011101

search clock period	detected data address	Hamming distance	
0	none	-	
1	none	-	
2	none	-	
...	...	...	...
23	20	HD=23	Nearest matched data
23 (2)	none	-	
24	15	HD=24	2nd matched data
24 (2)	23	HD=24	2nd matched data
24 (3)	none	-	
25	14	HD=25	4th matched data
25 (2)	none	-	

**Detected data address**

20	0101110101010010100000110100001000011000111010001001110101111111
15	101111101011010010011110011110011100001111101111110000110001100
23	010111101111000110101011101000100011101010111001101100100100100
14	1110110111001100011010110000101000011001011100101100111101110000

(b) Manhattan distance estimation

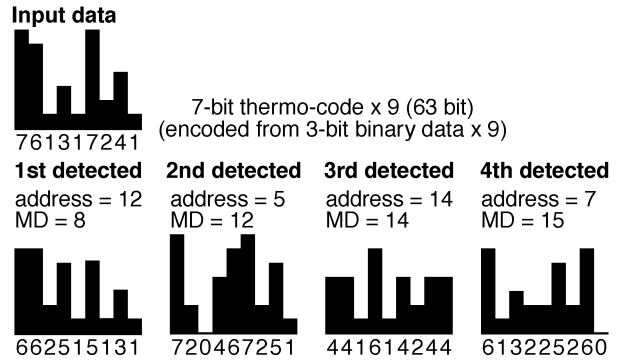


Fig. 6. Functional test results of (a) Hamming distance estimation and (b) Manhattan distance estimation.

microphotograph and the cell layouts. The associative co-processor is composed of a 64-bit 32-word associative memory array, a memory read/write circuit with data buffers, a word address decoder, and a 32-input priority encoder with detected data selectors. A two-stage hierarchical structure is implemented as shown in Fig. 4(b). A hierarchical node is achieved by a two-input AND gate. In the two-stage hierarchical structure, the number of hierarchical nodes on each propagation path is different. Therefore, the number of blocks and each bit length need to be optimized for the minimum critical path. A priority encoder employs a binary tree structure as reported in [5]. We have also designed a 64-bit 2-word associative memory using dynamic circuit implementation for feasibility and performance evaluation.

#### V. MEASUREMENT RESULTS AND DISCUSSIONS

##### A. Functionality

Fig. 6(a) and (b) shows the functional test results of Hamming and Manhattan distance estimation, respectively. In the Hamming distance estimation, 32-word temporary data are generated randomly and stored in the memories. At first, the co-processor provides the detected address in the clock period 23. That is, the detected data has a 23-bit Hamming distance

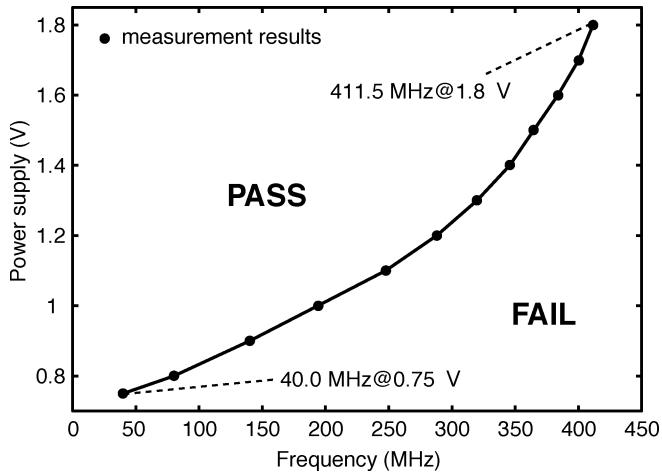


Fig. 7. Operation frequency versus power supply voltage.

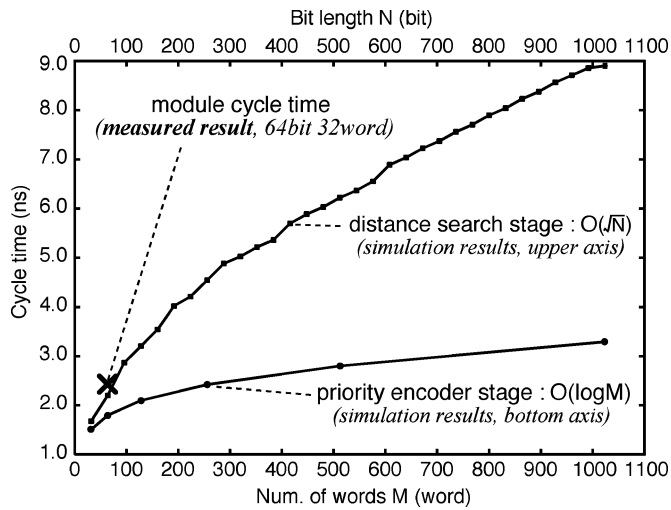


Fig. 8. Cycle time and data capacity.

and this signifies the nearest-match data. The search operation is suspended to acquire other data having the same distance. The search operation starts again in case of no remaining data of the same distance. For example, two data with  $HD = 24$  are detected as shown by the clock period 24 in Fig. 6(a). The associative co-processor has the capability of Manhattan distance estimation using thermometer encoding in a manner similar to the Hamming distance estimation. The 3-bit binary code is encoded to a 7-bit thermometer code. Each word has 7-bit times 9-element thermometer codes (i.e. 63-bit data) as shown in Fig. 6(b). In the functional test, as shown in Fig. 6(b), the 12th word with 8-bit Manhattan distance is detected as the nearest match in the clock period 8. Further, the second and third match data are also detected in order. The present associative co-processor provides the detected data addresses with the strictly exact Hamming or Manhattan distance regardless of the bit length, the number of words, and the data distance. This feature is important to ensure high capacity scalability and high reliability in distance estimation, which has not been achieved by the conventional fully parallel architectures based on analog techniques [1]–[4].

TABLE I  
SPECIFICATIONS OF THE FABRICATED ASSOCIATIVE CO-PROCESSOR

Process	1P5M 0.18 $\mu\text{m}$ CMOS process
Power Voltage Supply	0.7 V – 1.8 V
Organization	64 bit $\times$ 32 word memory cells 32-input priority encoder
Functions	Nearest-match detection Distance ordering
Module Size	475 $\mu\text{m}$ $\times$ 1160 $\mu\text{m}$ (0.55 $\text{mm}^2$ )
Num. of Transistors	88.5k transistors
Memory Cell Size	9.6 $\mu\text{m}$ $\times$ 13.6 $\mu\text{m}$ (130.56 $\mu\text{m}^2$ ) $\dagger$ 7.2 $\mu\text{m}$ $\times$ 8.8 $\mu\text{m}$ (63.36 $\mu\text{m}^2$ )
Search Time Order	$O(\sqrt{N})$ (@ $N$ -bit capacity)
Encoding Time Order	$O(\log M)$ (@ $M$ -word capacity)
Operation Speed	411.5 MHz (@ 1.8V, measured) 454.5 MHz (@ 1.8V, simulated) 40.0 MHz (@ 0.75V, measured) 41.4 MHz (@ 0.75V, simulated)
Worst-Case Search Time	158.0 ns (0-bit to 64-bit distance)
Power Dissipation	51.3 mW (@ 1.8V, 400MHz) 1.18 mW (@ 0.75V, 40MHz)

$\dagger$ designed using dynamic circuits.

### B. Area and Capacity

The designed 64-bit 32-word associative co-processor occupies an area of 475  $\mu\text{m}$   $\times$  1160  $\mu\text{m}$  (0.55  $\text{mm}^2$ ). The area of a memory macro cell with a static search circuit is 9.6  $\mu\text{m}$   $\times$  13.6  $\mu\text{m}$  (130.56  $\mu\text{m}^2$ ) as shown in Fig. 5(b-1). In the static circuit implementation using the 0.18- $\mu\text{m}$  process, the cell area is six and three times as large as a 6T SRAM cell and a complete-match CAM cell, respectively. Fig. 5(b-2) presents a layout of the dynamic circuit implementation. It occupies an area of 7.2  $\mu\text{m}$   $\times$  8.8  $\mu\text{m}$  (63.36  $\mu\text{m}^2$ ). In this case, the cell area is three and two times as large as a 6T SRAM cell and a complete-match CAM cell, respectively. The number of transistors in the present memory cell is larger than that possible by applying the conventional analog approaches [1]–[4]. All the analog approaches make device scaling difficult to achieve while maintaining the performance and marginal capacity. The present approach can achieve device scaling and operate at a low supply voltage because of the synchronous digital search logics embedded in the memories. Also, it has no capacity or search distance limitations. Therefore, in comparison with the conventional designs, the associative co-processor has greater potential for practical use and a larger capacity.

### C. Operation Speed

The measurement results show that the operation speed is 411.5 and 40.0 MHz at a supply voltage of 1.8 and 0.75 V, respectively. Fig. 7 shows the operation speed as a function of the supply voltage from 0.75 V to 1.8 V. The total search time increases in proportion to the distance of detected data in cases in which a target application requires the nearest-match data. For example, the nearest-match detection is completed in 17 clock periods (i.e., 41.3 ns) when the nearest-match data has a 16-bit distance from the input. The worst-case operation requires 65 clock periods in cases in which the nearest-match data has a maximum distance of 64 bits. Therefore, it takes 158.0 ns in the worst case. Fig. 8 shows the relation between the search cycle time and data capacity. The search cycle time is limited by the search signal propagation or priority encoding.

TABLE II  
PERFORMANCE COMPARISON

	Nagata et al.[2]	Ikeda et al.[3]	Mattausch et al.[4]	Proposed
Architecture	Pulse width (PWM) (analog/digital)	Delay time (analog/digital)	Current summation (analog/digital)	Digital bit-serial (fully digital)
Functions	Nearest match	Nearest match	Nearest match	Exact Nearest/Ordering
Reported Speed	160 ns <sup>(1)</sup> @ 5V	61 ns @ 5V	70 ns @ 3.3V	4.9 ns–158.0 ns <sup>(2)</sup> @ 1.8V
Capacity	256 bits/3 words	96 bits/128 words	768 bits/32 words	64 bits/32 words
Power	N/A	N/A	43 mW	51.3 mW
Low Power Mode	N/A	N/A	N/A	1.18 mW @ 0.75V <sup>(3)</sup>
Estimated Speed	~320 ns <sup>(1)</sup> @ 5V	~488 ns @ 5V	70 ns @ 3.3V	15.2 ns–5844.4 ns <sup>(2)</sup> @ 1.8V
Capacity	768 bits/32 words	768 bits/32 words	768 bits/32 words	768 bits/32 words
Power	N/A	N/A	43 mW	205.2 mW
Low Power Mode	N/A	N/A	N/A	4.72 mW @ 0.75V <sup>(3)</sup>
Error Possibility	Possible	Possible	Possible (< 1.13%)	None
Reported Process	0.8 $\mu$ m process	0.6 $\mu$ m process	0.6 $\mu$ m process	0.18 $\mu$ m process

<sup>(1)</sup> Assuming that it is applied to Hamming distance estimation.

<sup>(2)</sup> The search time depends on the distance between the input and the detected data.

<sup>(3)</sup> The search time becomes slower in a low-voltage operation as shown in Fig.7.

The search signal propagation takes  $O(\sqrt{N})$  at  $N$ -bit length due to the two-stage hierarchical structure. On the other hand, the priority encoding takes  $O(\log M)$  at  $M$ -word length due to a binary tree structure. Therefore, the present architecture allows for a high-speed operation in a large database as shown in Fig. 8. Finally, the total search time  $T_s$  is given by

$$T_s \propto O(\sqrt{N}) \times O(D) + O(\log M) \quad (1)$$

where  $N$  and  $M$  are the bit length and the number of words of the database, respectively, and  $D$  is the distance between the input and the detected data. The distance estimation has no data capacity limitation, as mentioned above.

#### D. Power Dissipation

The power dissipation of the associative co-processor is <51.3 mW at 1.8 V power supply and 400 MHz operation. In a low-voltage operation, it is 1.18 mW at 0.75-V power supply and 40-MHz operation. The search accuracy of the conventional analog approach is unstable in a low-voltage operation and can sometimes be ineffective. The present search results are strictly exact regardless of the power supply voltage. This feature contributes to not only a low-power operation but also the suitability to a system-on-a-chip application. The specifications of the fabricated co-processor are summarized in Table I. The feature and performance comparisons are summarized in Table II.

## VI. CONCLUSION

We have proposed a new concept and circuit implementation for a high-speed and low-voltage associative co-processor with exact Hamming or Manhattan distance estimation. It suffers no data capacity limitation and maintains a high-speed operation

in a large database due to the hierarchical search architecture and a synchronous search logic embedded in a memory cell. The circuit implementation enables a high tolerance for device fluctuation in a DSM process and a low-voltage operation under 1.0 V. The associative co-processor provides the exact distance of the detected data, hence it is capable of data sorting in the order of Hamming or Manhattan distance as well as the traditional nearest-match detection. A 64-bit 32-word associative co-processor has been designed using 1P5M 0.18- $\mu$ m CMOS process and successfully tested. The operation speed attains a speed of 411.5 and 40.0 MHz at a supply voltage of 1.8 and 0.75 V, respectively.

## ACKNOWLEDGMENT

The VLSI chip in this study has been fabricated through VLSI Design and Education Center (VDEC), University of Tokyo, in collaboration with Hitachi Ltd. and Dai Nippon Printing Company.

## REFERENCES

- [1] T. Yamashita, T. Shibata, and T. Ohmi, "Neuron MOS winner-take-all circuit and its application to associative memory," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 1993, pp. 236–237.
- [2] M. Nagata, T. Yoneda, D. Nomasaki, M. Sato, and A. Iwata, "A minimum-distance search circuit using dual-line PWM signal processing and charge-packet counting techniques," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 1997, pp. 42–43.
- [3] M. Ikeda and K. Asada, "Time-domain minimum-distance detector and its application to low-power coding scheme on chip-interface," in *Proc. Eur. Solid-State Circuits Conf. (ESSCIRC)*, 1998, pp. 464–467.
- [4] H. J. Mattausch, N. Omori, S. Fukae, T. Koide, and T. Gyohten, "Fully-parallel pattern-matching engine with dynamic adaptability to Hamming or Manhattan distance," in *Symp. VLSI Circuits Dig. Tech. Papers*, 2002, pp. 252–255.
- [5] Y. Oike, M. Ikeda, and K. Asada, "A high-speed and low-voltage associative co-processor with hamming distance ordering using word-parallel and hierarchical search architecture," in *Proc. IEEE Custom Integrated Circuits Conf. (CICC)*, 2003, pp. 643–646.