MULTIPLE OUTPUT OF SIMILARITY DATA BY RECALLING TYPE ASSOCIATIVE MEMORY USING NEURON CMOS INVERTER

Toyoki Saho[1], Yujiro Harada[2], Masaaki Fukuhara[1]
and Kuniaki Fujimoto[3]

[1]Graduate School of Information and Telecommunication Engineering
Tokai University
2-3-23, Takanawa, Minato, Tokyo 108-8619, Japan
9bjnm015@mail.u-tokai.ac.jp; fukuhara@tokai.ac.jp

[2]Department of Electrical and Electronic Engineering
National Institute of Technology Kurume College
1-1-1, Komorino, Kurume, Fukuoka 830-8555, Japan
y-harada@kurume-nct.ac.jp

[3]Graduate School of Science and Technology
Tokai University
9-1-1, Toroku, Higashi, Kumamoto 862-8652, Japan
fujimoto@tokai.ac.jp

ABSTRACT. *An associative memory with recalling function using neuron CMOS inverters searches one or multiple reference data which is the most similar to input data from all reference data in the memory, and outputs the nearest reference data itself. The associative memory is using Hamming distance as an index for comparing an input data and multiple reference data. If a plurality of retrieved reference data exists, the memory cannot output these reference data. In this paper, we propose to add a priority encoder to the above associative memory, because the priority encoder reads out the one retrieved reference data in recording position order. The proposed circuit is designed by schematic CAD tool and simulated by HSPICE.*
Keywords: Associative memory, Hamming distance, Neuron CMOS inverter

1. **Introduction.** In recent years, research on big data processing has been actively conducted. The real-time detection process of the stored data that is most similar to the input data is very important. Therefore, an associative memory has attracted attention [1-4]. An associative memory can retrieve stored data that is most similar to input data and perform to not only retrieve an exact matting data but also retrieve multiple similarity data. The associative memory uses distance as an index of similarity, such as Hamming distance and Manhattan distance. A conventional circuit and proposed circuit use Hamming distance. In addition, the conventional circuit and proposed circuit use neuron CMOS inverters that have characteristics similar to human neurons [5,6].

A recalling type associative memory using neuron CMOS inverter was proposed [7] (in this paper, we call conventional circuit). If there is a plurality of reference data, this conventional circuit cannot output expected data. In this paper, by using a priority encoder circuit, the proposed circuit is possible to output expected data. We designed the proposed circuit by schematic CAD tool and simulated by HSPICE.

2. **Configuration of Conventional Circuit.** A configuration of conventional circuit is shown in Figure 1. The conventional circuit detects matching or similarity data by comparing input data N_i ($i = 0, 1, \ldots, 7$) and reference data RF_j ($S_{j,0}, S_{j,1}, \ldots, S_{j,7}$), ($j = 0, 1, \ldots, 31$). The reference data that is the most similar to input data is outputted to O_i ($i = 0, 1, \ldots, 7$). When MA_j which is output voltage of the jth D-FF circuit is set to V_{DD}, SWS_j is the ON state and when MA_j is 0 [V], SWS_j ($j = 0, 1, \ldots, 31$) is the OFF state.
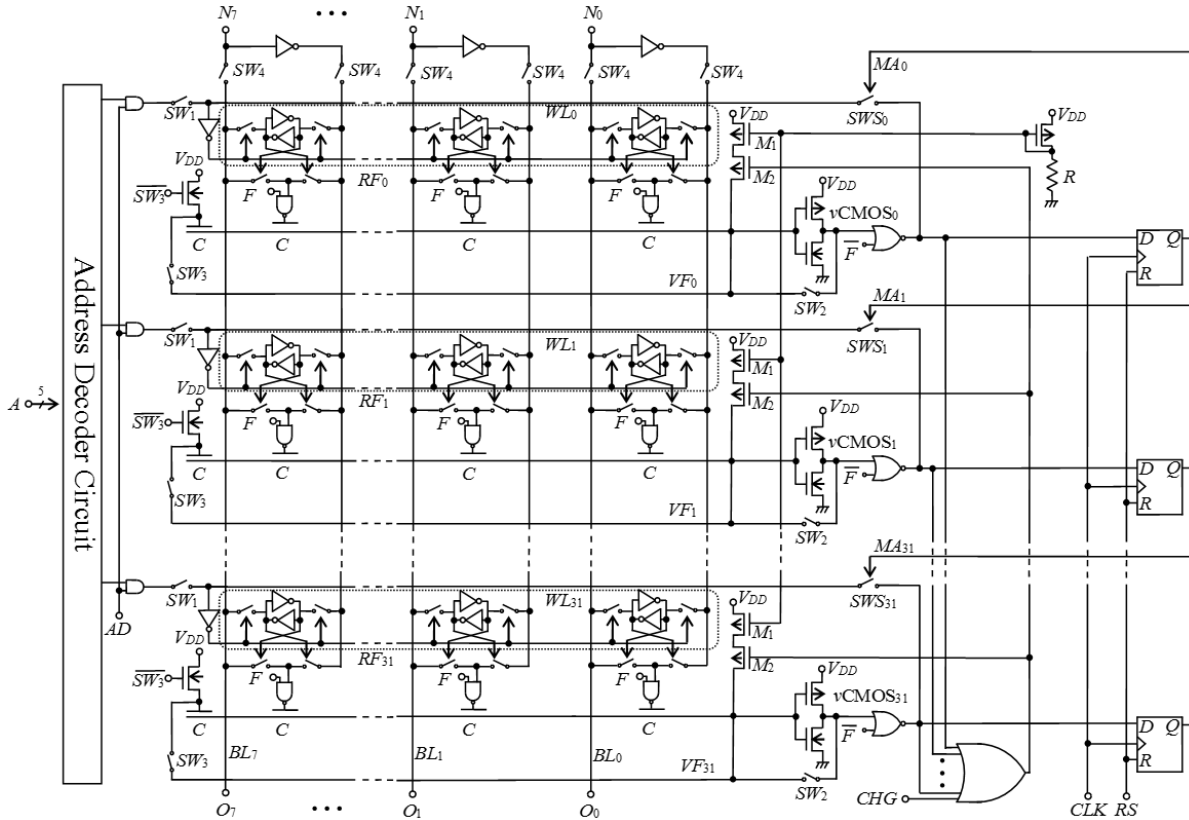

FIGURE 1. Conventional circuit

In the conventional circuit, Table 1 shows a pattern of input data and reference data, Figure 2 shows waveforms of CLK and MA_j, and Figure 3 shows waveforms of BL_i. From Table 1, the retrieved reference data are RF_{11}, RF_{15}, RF_{27}, and RF_{28} that are the most similar to input data. From Figure 2, MA_{11}, MA_{15}, MA_{27}, and MA_{28} are V_{DD} and SWS_{11}, SWS_{15}, SWS_{27}, and SWS_{28} are the ON state, respectively. From Figure 3, reference data RF_{11}, RF_{15}, RF_{27}, and RF_{28} are not outputted. This is because RF_{11}, RF_{15}, RF_{27}, and RF_{28} are shorted and cause malfunction. From Figure 2 and Figure 3, one SWS_j is necessary to ON state.

3. **Configuration and Operation of the Proposed Circuit.**

3.1. **Circuit configuration.** A configuration of the proposed circuit is shown in Figure 4. N_i ($i = 0, 1, \ldots, 7$) is input data, RF_j ($S_{j,0}, S_{j,1}, \ldots, S_{j,7}$), ($j = 0, 1, \ldots, 31$) are reference data, $\nu CMOS_j$ ($j = 0, 1, \ldots, 31$) are neuron CMOS inverters, WL_j ($j = 0, 1, \ldots, 31$) are word line, BL_i ($i = 0, 1, \ldots, 7$) are bit line, $PCLK$ is clock pulse of priority encoder circuit, PMA_j ($j = 0, 1, \ldots, 31$) are voltages that decide the state of SWS_j ($j = 0, 1, \ldots, 31$) to ON or OFF by output of the priority encoder circuit, and O_i ($i = 0, 1, \ldots, 7$) are output voltages that read out the retrieved reference data. C are capacitors of the neuron CMOS inverters, and R is a resistor of the current mirror circuit.

TABLE 1. Pattern of input and reference data of conventional circuit

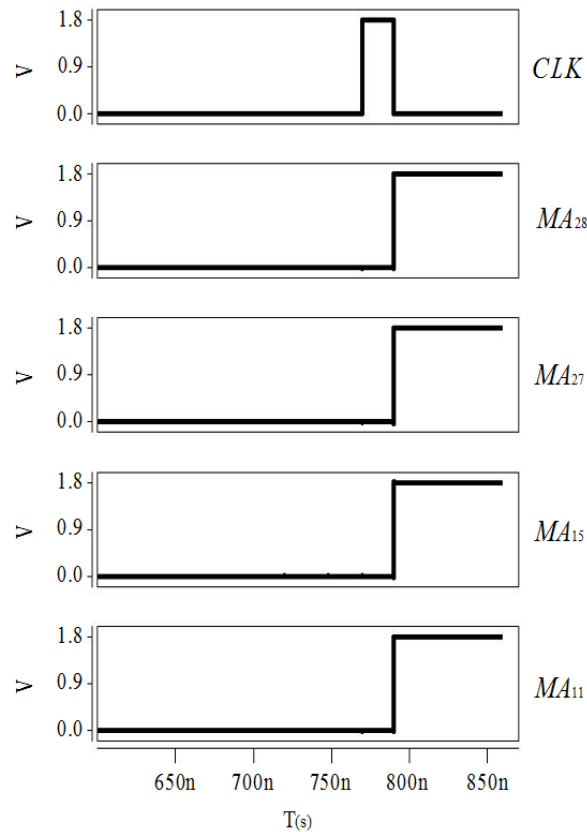| Input data		$N_7:0$	$N_6:0$	$N_5:0$	$N_4:0$	$N_3:1$	$N_2:0$	$N_1:1$	$N_0:0$	Hamming distance
	RF_0	0	0	1	0	0	0	0	0	3
	RF_1	0	0	1	0	0	0	0	1	4
	RF_2	0	0	1	0	0	0	1	0	2
	⋮				⋮					⋮
	RF_{11}	0	0	1	0	1	0	1	0	1
Reference data	⋮				⋮					⋮
	RF_{15}	0	0	0	0	1	1	1	0	1
	⋮				⋮					⋮
	RF_{27}	0	0	1	1	1	1	1	0	1
	RF_{28}	0	0	1	1	1	1	1	1	1
	⋮				⋮					⋮
	RF_{30}	0	0	1	1	1	1	1	0	3
	RF_{31}	0	0	1	1	1	1	1	1	4

FIGURE 2. Waveforms of CLK and MA_j ($j = 11, 15, 27, 28$) in conventional circuit

SW_1, SW_2, SW_3, SW_4, F, H, CLK, AD, and $PCLK$ are control signals. SW_1, SW_4, AD, and CLK are for writing reference data. The write operation to the reference data is performed by an address decoder circuit. The address decoder circuit is controlled by input signal A ($= 5$ bits). SW_2, SW_3, and F are for checking whether input data and reference data match or not. H is for performing a similarity search. $PCLK$ is for performing the priority encoder circuit. Capacitors C in the proposed circuit are designed with the same capacitance.
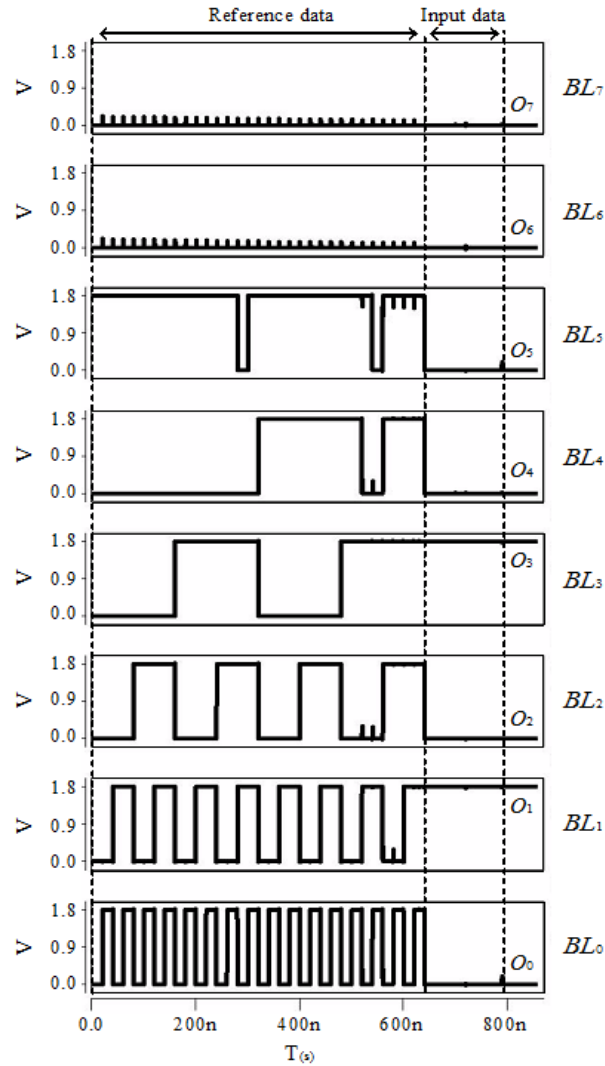
FIGURE 3. Waveforms of BL_i $(i = 0, 1, \ldots, 7)$

3.2. Operation principle.

3.2.1. *Write operation to reference data.* Figure 5 shows the construction of static random-access memory (= SRAM) for storing 1 bit data. WL_j $(j = 0, 1, \ldots, 31)$ are word line, BL_i $(i = 0, 1, \ldots, 7)$ are bit line, $IS_{j,i}$ $(j = 0, 1, \ldots, 31, i = 0, 1, \ldots, 7)$ are an ideal switch, and $S_{j,i}$ $(j = 0, 1, \ldots, 31, i = 0, 1, \ldots, 7)$ are stored data. The ideal switch shown in Figure 1 and Figure 4 is replaced by the real switch (transfer gate) shown in Figure 5. If V_{DD} is applied to real switch, it will be the ON state, and if 0 [V] is applied, it will be the OFF state. In this paper, reference data is composed of 8 SRAM. The operation of the SRAM of Figure 5 will be described. First, when V_{DD} is set to WL_j, one of $IS_{j,i}$ $(i = 0, 1, \ldots, 7)$ becomes the ON state by the WL_j. If BL_i are set to V_{DD}, $S_{j,i}$ is stored 1 $(= V_{DD})$ and if BL_i is set to 0 [V], $S_{j,i}$ is stored 0 $(= 0$ [V]).

The write operation of the reference data will be described. First, SW_1, SW_4, and AD are set to V_{DD}, and CLK is changed from 1 $(= V_{DD})$ to 0 $(= 0$ [V]). From this operation, the output PMA_j of D-FF becomes 0 $(= 0$ [V]). At this time, SWS_j is the OFF state and multiple reference data are enabled possible to write operation. Next, multiple reference data are written, and finally, AD is changed from 1 $(= V_{DD})$ to 0 $(= 0$ [V]).

3.2.2. *Search reference data by Hamming distance.* The Hamming distance is the number of different bits between two bit streams which have the same number of digits. Assuming that D_{Hj} $(j = 0, 1, \ldots, 31)$ is the number of Hamming distances, N_i $(i = 0, 1, \ldots, 7)$ is
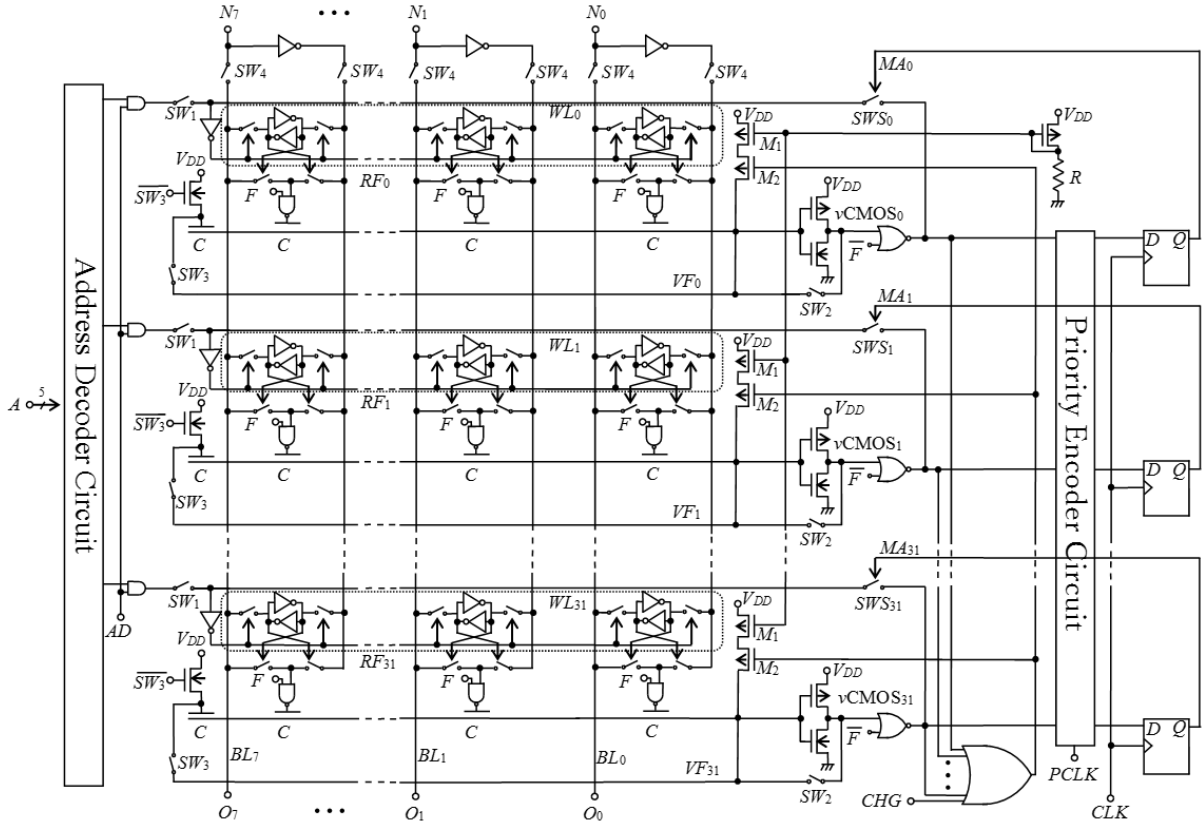
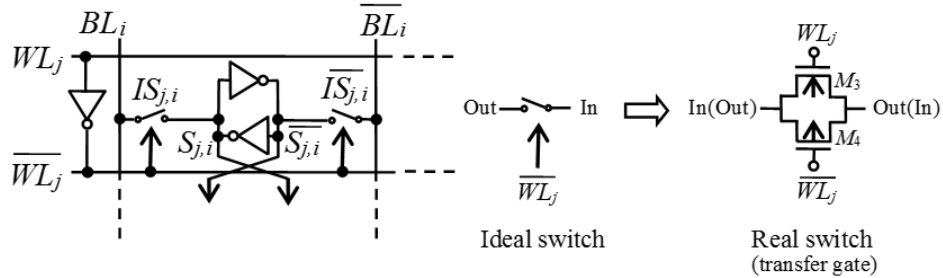FIGURE 4. Proposed circuit with priority encoder circuit

FIGURE 5. SRAM in the proposed circuit

input data, and $S_{j,i}$ are stored data, the Hamming distance equation is expressed as follows:

$$D_{Hj} = \sum_{i=0}^{7}(N_i \oplus S_{j,i}), \quad (j = 0, 1, \ldots, 31). \tag{1}$$

When input data and reference data are equal, V_{DD} is applied to the capacitor C through the NAND circuit. Otherwise, 0 [V] is applied. If all input gates of νCMOS$_j$ are applied to V_{DD}, exact matching search is performed, and the search operation is finished. Otherwise, a similarity search is performed and a constant current flows into the floating gate.

3.2.3. *Operation of priority encoder circuit.* Table 2 is a truth table of a priority encoder circuit and Figure 6 shows a configuration of the circuit [8]. In this paper, we use 32-to-32 bit priority encoder circuit. IN_j ($j = 0, 1, \ldots, 31$) are input data, O_j ($j = 0, 1, \ldots, 31$) are output data, $PCLK$ is clock pulse, L-PE is 4-bit the lookahead priority encoder circuit, and D-PE is 8-bit data priority encoder circuit. The priority encoder circuit is controlled by clock pulse $PCLK$.

TABLE 2. Truth table of priority encoder circuit

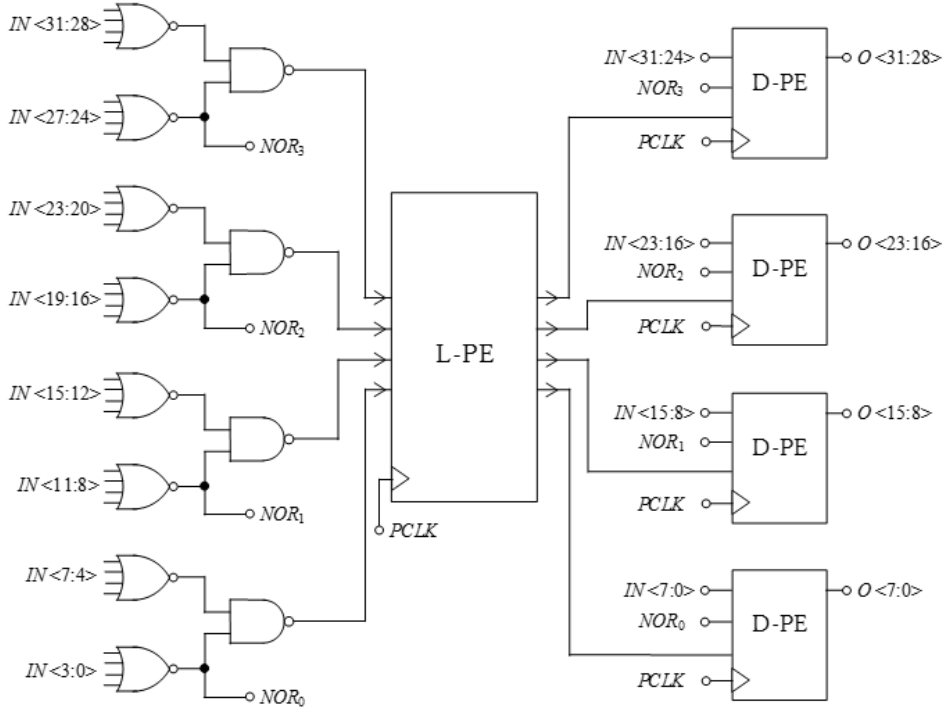| Input data					Output data				
IN_{31}	...	IN_2	IN_1	IN_0	O_{31}	...	O_2	O_1	O_0
1		1	1	1	0		0	0	1
1	...	1	1	0	0	...	0	1	0
1		1	0	0	0		1	0	0
1		0	0	0	1		0	0	0

FIGURE 6. Circuit configuration of priority encoder circuit

The operation of the priority encoder circuit will be described. If multiple input data are given at the same time, the input data having the highest priority will be to take precedence. The highest priority is given to input data that have the smallest number of j. Likewise, output data having the smallest number of j is given to the highest priority.

In this paper, the priority encoder circuit is used to select one of SWS_j. After the similarity search, V_{DD} is applied to the multiple output through the NOR circuit of the neuron CMOS inverter in Figure 4. The multiple output will be input signal of the priority encoder. The priority encoder circuit preferentially outputs one data having the smallest jth.

3.2.4. *Operation of floating gate voltage.* The waveform of the floating gate voltage is shown in Figure 7. In the proposed circuit, the maximum Hamming distance is 8; however, the floating gate voltage goes up to 4. This is because the threshold voltage V_{TH} is designed to half of the power supply voltage V_{DD}. Figure 7 shows the floating gate voltage when performing the similarity search. The Phase 1 is at the time of the operation described in 3.2.1. Assuming that the floating gate voltage is V_{Fj}, V_{Fj} is expressed as follows:

$$V_{Fj} = V_{TH}, \quad (j = 0, 1, \ldots, 31). \tag{2}$$

In Phase 2, by controlling SW_2 and SW_3, V_{Fj} increases a little. Assuming that the sum of capacitors in one-word reference data is C_T, V_{Fj} at this time is expressed as follows:

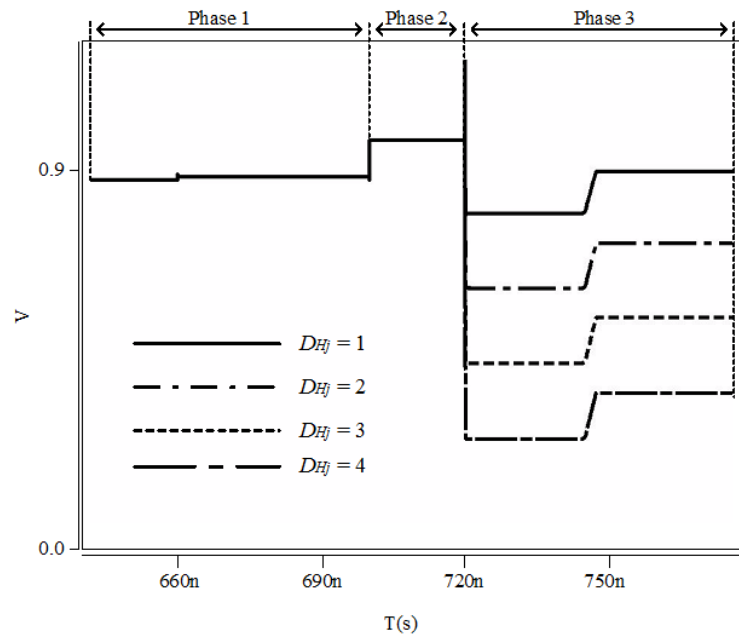$$V_{Fj} = V_{TH} + \frac{C}{C_T}(V_{DD} - V_{TH}), \quad (j = 0, 1, \ldots, 31). \tag{3}$$

FIGURE 7. Waveform of the floating gate voltage

The Phase 3 is at the time of the operation described in 3.2.3. By rising the control voltage F, input gates of vCMOS_j are changed corresponding to D_{Hj} through the NAND circuit. If the input data and the reference data are matched, V_{DD} is applied, and if the input data and the reference data are unmatched, 0 [V] is applied. The floating gate voltage drops by the number of mismatches between the input data and the reference data. V_{Fj} at this time is expressed as follows:

$$V_{Fj} = V_{TH} + \frac{C}{C_T}(V_{DD} - V_{TH}) - D_{Hj}\frac{C}{C_T}V_{DD}, \quad (j = 0, 1, \ldots, 31). \tag{4}$$

4. **Simulation Result.** The proposed circuit was designed by schematic CAD tool and simulated by HSPICE. In the proposed circuit, Table 3 shows parameters of input data and reference data, Table 4 shows device parameters, Figure 8 shows waveform of CLK and

TABLE 3. Pattern of input and reference data of proposed circuit

Input data		$N_7:0$	$N_6:0$	$N_5:0$	$N_4:0$	$N_3:1$	$N_2:0$	$N_1:1$	$N_0:0$	Hamming distance
Reference data	RF_0	0	0	1	0	0	0	0	0	3
	RF_1	0	0	1	0	0	0	0	1	4
	RF_2	0	0	1	0	0	0	1	0	2
	$\vdots$				$\vdots$					$\vdots$
	RF_{11}	0	0	1	0	1	0	1	0	1
	$\vdots$				$\vdots$					$\vdots$
	RF_{15}	0	0	0	0	1	1	1	0	1
	$\vdots$				$\vdots$					$\vdots$
	RF_{27}	0	0	1	1	1	1	1	0	1
	RF_{28}	0	0	1	1	1	1	1	1	1
	$\vdots$				$\vdots$					$\vdots$
	RF_{30}	0	0	1	1	1	1	1	0	3
	RF_{31}	0	0	1	1	1	1	1	1	4

SWS_j $(j = 11, 15, 27, 28)$, and Figure 9 shows BL_i waveform. From Table 3, the reference data RF_{11}, RF_{15}, RF_{27}, and RF_{28} are most similar to the input data N_i $(i = 0, 1, \ldots, 7)$. From Figure 8, one SWS_{11} is V_{DD}. From Figure 9, the reference data RF_{11} in Table 3 is outputted to O_i. From this result, it can be confirmed that one SWS_{11} can be ON state by the priority encoder circuit, and the correct output is obtained.

TABLE 4. Device parameters

Device parameter	Gate length	Gate width
pMOS		3 μm
nMOS		1 μm
M_1 (Figure 4)	180 nm	
M_2 (Figure 4)		1 μm
M_3 (Figure 5)		
M_4 (Figure 5)		
C	16 fF	
R	40 kΩ	
V_{DD}	1.8 V	
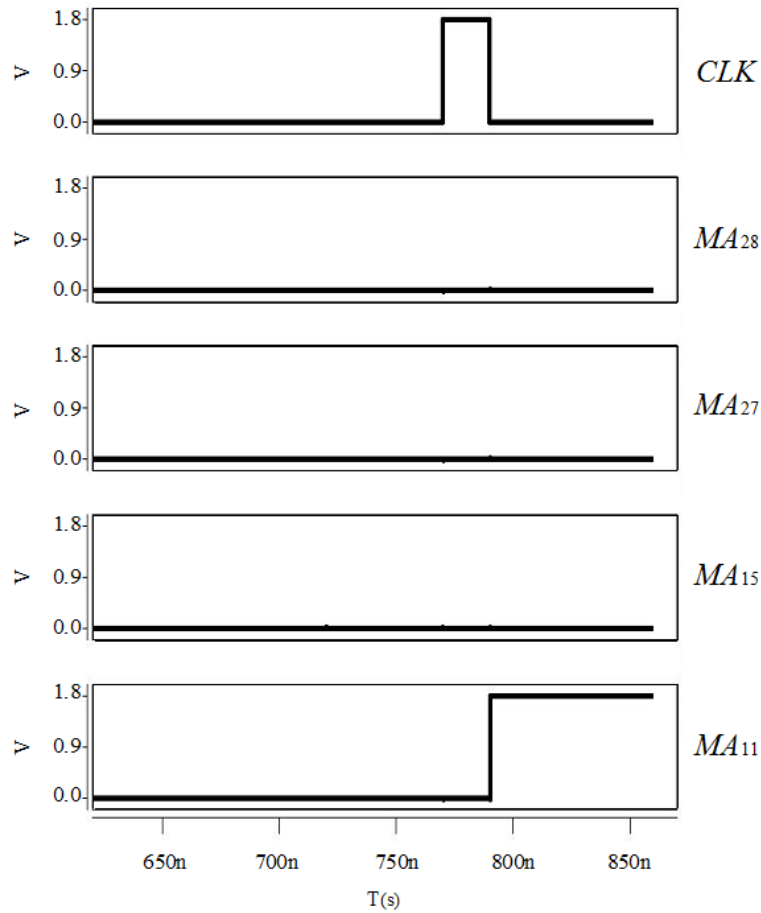
FIGURE 8. Waveform of CLK and MA_j $(j = 11, 15, 27, 28)$ in proposed circuit
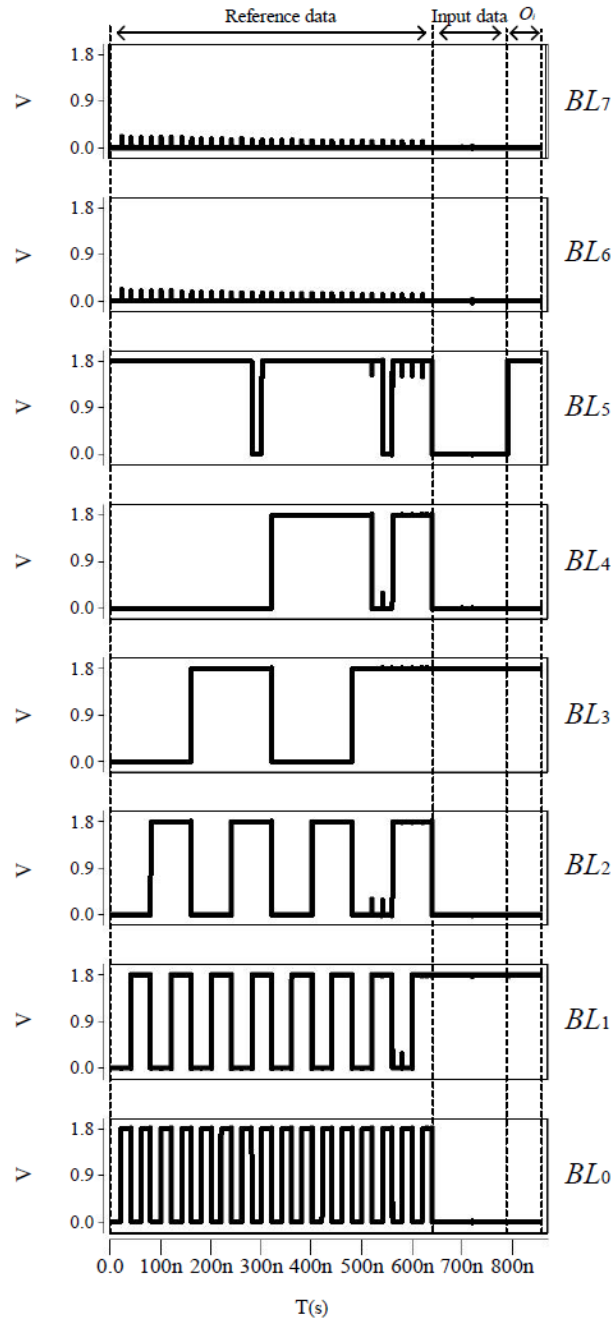
FIGURE 9. Waveforms of BL_i $(i = 0, 1, \ldots, 7)$ of the proposed circuit

5. **Conclusions.** In this paper, we proposed RAM type associative memory using neuron CMOS inverter that can detect multiple similar stored data. We also simulated by HSPICE and confirmed that the expected stored data was obtained. In this paper, by using priority encoder circuit, the expected stored data was obtained. However, it did not select stored data when there was a plurality of retrieve reference data. This is because choice of stored data is determined by the design of the priority encoder circuit. In the future, we would like to solve this problem and try to further improve the circuit performance.

REFERENCES

[1] F. K. Gurkaynak, Y. Leblebici and D. Mlynek, A compact high-speed hamming distance comparator for pattern matching applications, *Proc. of 1998 European Signal Processing Conference*, 1998.

[2] H. J. Mattausch, T. Gyohten, T. Soda and T. Koide, Compact associative-memory architecture with fully parallel search capability for the minimum Hamming distance, *IEEE Journal of Solid-State Circuits*, vol.37, no.2, pp.218-227, 2002.

[3] Y. Oike, M. Ikeda and K. Asada, High-speed content-addressable memory using synchronous Hamming distance search circuits, *IEICE*, vol.102, pp.19-24, 2002 (in Japanese).

[4] T. Koide, T. Gyoten, Y. Soda and H. J. Mattausch, Architecture for compact and fast associative-memories with all-parallel nearest-match Hamming-distance search, *IECIC*, vol.101, pp.27-34, 2001 (in Japanese).

[5] T. Shibata and T. Ohmi, A functional MOS transistor featuring gate-level weighted sum and threshold operations, *IEEE Trans. Electron Devices*, vol.39, no.6, pp.1444-1455, 1992.

[6] T. Shibata and T. Ohmi, New Architecture Logic Integrated Circuits Using Neuron MOS Transistors, *IEICE*, ICD93-6, pp.39-46, 1993 (in Japanese).

[7] Y. Harada, K. Kuniaki, M. Fukuhara and M. Yoshida, A minimum Hamming distance search associative memory using neuron CMOS inverters, *The Institute of Electrical Engineers of Japan C*, vol.136, no.1, pp.36-42, 2016.

[8] D. Balobas and N. Konofaos, Low-power, high-performance 64-bit CMOS priority encoder using static-dynamic parallel architecture, *International Conference on Modern Circuits and Systems Technologies*, pp.1-4, 2016.