Visual Image Processing RAM for Fast 2-D Data Location Search

Joo-Young Kim, Donghyun Kim, Seungjin Lee, Kwanho Kim, Seonghyun Jeon, and Hoi-Jun Yoo
School of EECS
Korea Advanced Institute of Science and Technology (KAIST)
Daejeon, Republic of Korea

***Abstract*—Visual Image Processing RAM (VIP-RAM) for fast 2-D data location search is proposed and implemented. It finds the local maximum location of 3x3 size window in single cycle latency using hierarchical 3-bank architecture. Each bank searches intermediate maximum from 3 32-bit data in a row, and top-level logic deduces the final maximum out of 3 32-bit data from 3 banks. Each memory bank includes special logic for 3 consecutive data read and 32-bit 3 input comparator. 8 VIP-RAMs are integrated in multi-core object recognition SoC and fabricated in 0.18μm process. VIP-RAM is measured to operate at 200MHz for 8.2 GOPS peak performance.**

I. INTRODUCTION

Recently, visual image processing applications such as pattern recognition, vehicle automation, robot localization, and face recognition have been studied widely due to the high market demands [1]-[3]. In these applications, high processing performance is generally required to achieve fast object recognition for real-time operation. Many dedicated systems and chips with high computational capability have been implemented [1]-[4]. They actively utilized the data parallelism to enhance the processing performance because their processing stages are sequential and each of them is composed of data-intensive but regular functions.

2-D data location search is one of the most frequent and important operations in visual image processing to find feature points from a number of pixel data [5]. However, this operation consumes a lot of execution cycles in conventional architecture that the processor is separated from memory. For example, to perform the local maximum location search(LMLS) out of 3x3 size window, 9 data and the base address are fetched from memory to the processor. Then, data are compared with the base data one by one, and the initial base data and base address are updated by the larger ones. Its detail algorithm of LMLS can be described in C as shown below.

```
for ( i=1; i<9; i++) {
   if ( Num[0] – Num [i] < 0 ) {
        Num [0] ← Num [i];
        Addr[0] ← Addr [i]; }}
```

Where Num[0] and Addr[0] represent the base data and base address, respectively. Although LMLS algorithm is simple, its number of cycles amounts to 41 in assembly level; 10 cycles to fetch data from memory, 30 cycles to compare and update, and 1 cycle to save the result. To decrease data transfer between processor and memory, most of the previous works have just combined processing units and memory devices without any modification [6]-[8].

In this study, on top of the previous simple integration of memory and processing units, more performance gain can be obtained by the integration of LMLS logic inside of the memory itself. This enables not only to reduce data transfer but also to achieve operational efficiency. For this purpose, we propose a new memory architecture, Visual Image Processing RAM (VIP-RAM), which integrates memory and search logic with small area overhead and fast operation in data location search. In addition, new compare logic, bitwise competition logic (BCL), is proposed to speed up 32-bit 3 input comparison. Its simple architecture gives small area, low-power, and fast operation. With only 22% area overhead, VIP-RAM can complete the local maximum location search operation in a single cycle.

This paper is organized as follows. Chapter II describes a VIP-RAM's architecture. In chapter III, BCL comparator's circuit and operation are explained. Chip implementation and measurement results of VIP-RAM will be shown in chapter IV. Finally, chapter V concludes the paper.

II. VIP-RAM ARCHITECTURE

VIP-RAM has two behavioral modes; normal and local-max modes. In normal mode, VIP-RAM operates as a synchronous dual-port SRAM. It receives two addresses and control signals from two ports, and reads or writes two 32-bit data independently. While in local-max mode, VIP-RAM finds the address of the local maximum out of 3x3 data window when it receives the address of the left and top most data of the window.

Fig. 1 shows the overall architecture of the proposed VIP-RAM. It has a 1.5K-byte capacity and consists of 3 banks. Each bank is composed of 32 rows and 4 columns,

This research was supported by the MIC (Ministry of Information and Communication) of Korea

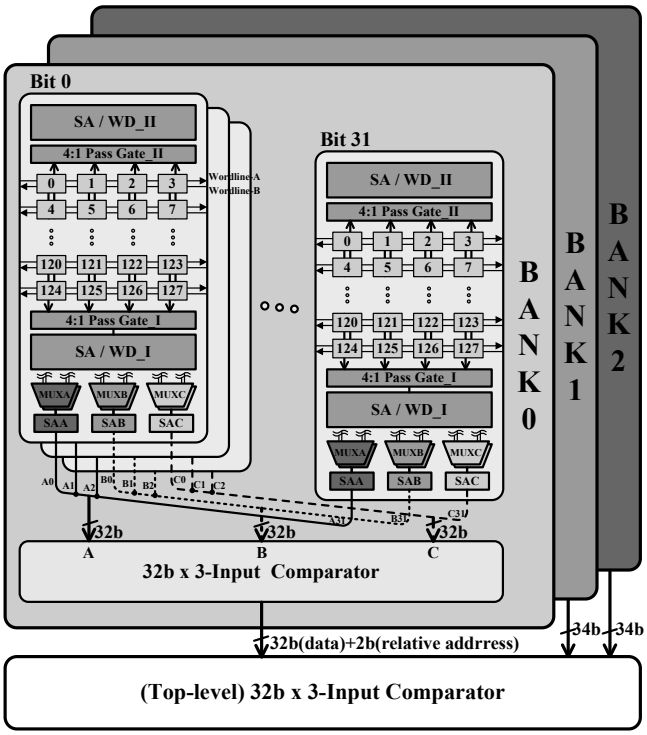

Fig. 1. The Overall Architecture of VIP-RAM

and operates in a word unit. Each bit of 4 columns shares the same memory peripherals such as write driver and sense amplifier, and the logic for LMLS. LMLS logic is composed of muxes and tiny sense amplifiers. And a 3-input comparator for 32-bit number is embedded below the memory arrays.

Before LMLS of a 3x3 window, the pixel data of the image space should have been properly mapped into VIP-RAM. First, the rows of the visual image data are placed into different banks as shown in Fig. 2. Then, the 3 32-bit data from the 3 banks form the 3x3 window. In data mapped VIP-RAM, LMLS operation is processed in 3 steps. First, two successive rows are activated and 3 required data are chosen by muxes. Secondly, the 32-bit 3-input comparators in 3 banks deduce the 3 intermediate maximum values among the

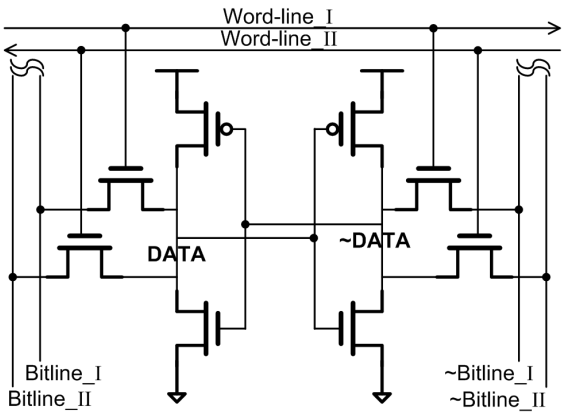

Fig. 3. A Dual-Ported Memory Cell

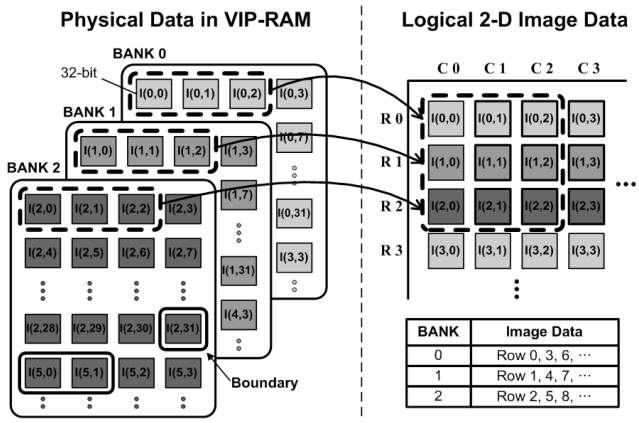

Fig. 2. Image Data Mapping of VIP-RAM

respective three numbers of the respective bank. Finally, the top level 32-bit 3-input comparator finds the final maximum value of the 3x3 window from 3 bank-level intermediate results and outputs the corresponding address.

III. BITWISE COMPETITION LOGIC (BCL) COMPARATOR

The SRAM of the proposed VIP RAM is composed of dual-ported storage cell which has 8 transistors, as shown in Fig. 3. A word-line and a pair of bit-lines are added to the conventional 6-transistor cell. Pull-down NMOS transistors are large compared to other minimum-sized transistors for the purpose of the data stability. A cell layout occupies 2.92 μm x 5.00μm in 0.18 μm process.

Bitwise Competition Logic (BCL) is devised to implement a fast, low-power, and area efficient 32-bit 3-input comparator. It just locates the first 1 from the MSB to the LSB of two 32-bit numbers and decides the lager number between the two 32-bit numbers without complex logics. The BCL enables 32-bit 3-input comparator to be compactly embedded to the memory bank. Fig. 4 describes its circuit diagram and operation. Before input to BCL comparator, each bit of two numbers are pre-encoded from A[i] and B[i] into (A[i] · ~B[i]) and (~A[i] · B[i]), respectively. Pre-encoding prevents the logic failures of BCL occurred when both inputs have 1 at the same bit position.

In the BCL, A line and B line are pre-charged to VDD initially. Then, START signals are activated to trigger each bit of pre-encoded signals sequentially from the MSB to LSB. If any triggered signal is 1, the path from the corresponding line to GND is opened and its voltage goes down immediately. Then, decision logic, at the right end of the lines, detects the line that firstly goes down, and keeps the result until the bit comparisons end. As shown in Fig. 5, the circuit of decision logic is the same as the sense amplifier except transistor N1 and N2. N1 and N2 receive the feedback signals and disable input of the small number to preserve only the first decision or the large number. For example, the timing diagram of Fig. 4 illustrates its operation in case that two pre-encoded input A* and B* are 00000010 and 00000001. The gray boxes represent the transitions of a

few important events in BCL operation. At box (1), all lines are precharged when the START signal is low. Triggering starts but both lines stay in VDD by the 7th start signal. At the 7th triggering of box (2), A line is dropped to GND because 7th bit of A is 1. The drop of A line forces the decision logic to turn down Bwin signal of box (3). Finally, Awin and Bwin, which represent the comparison results, are kept until the end of the cycle irrespective of B line voltage as shown in box (4). 32-bit data comparator is composed of the 4 parallel 8-bit BCLs. The 32-bit comparison results can be obtained from 4 results of the 4 parallel BCLs by setting higher priority to the result of the MSB part BCL. As a result, 32-bit 2-input comparator with BCL uses only 482 transistors, which are 38% reduced from the lowest transistor count comparator reported [9]. The 32-bit 3-input comparator is designed using 3 of 32-bit 2-input BCL comparators. Its worst case delay is 1.4 ns, which is sufficiently small considering the 5-ns timing budget of VIP-RAM when operating at 200MHz.

Fig. 6 shows simulated waveforms of VIP-RAM when it searches the local maximum address of the 3x3 window data where 9 window data are 0x02, 0x01, 0x00, and 0x00, 0x00, 0x00 and 0x00, 0x00, 0x00. The MSB bits of test data are set to 0xFFFFFF. The waveforms include BCL comparator signals, bank max address signal, and final max address signal. Bank max address signal represents the relative address of the bank-level maximum data, and the final max address represents the final top-level result. Their transition sequence verifies that LMLS operation is correctly

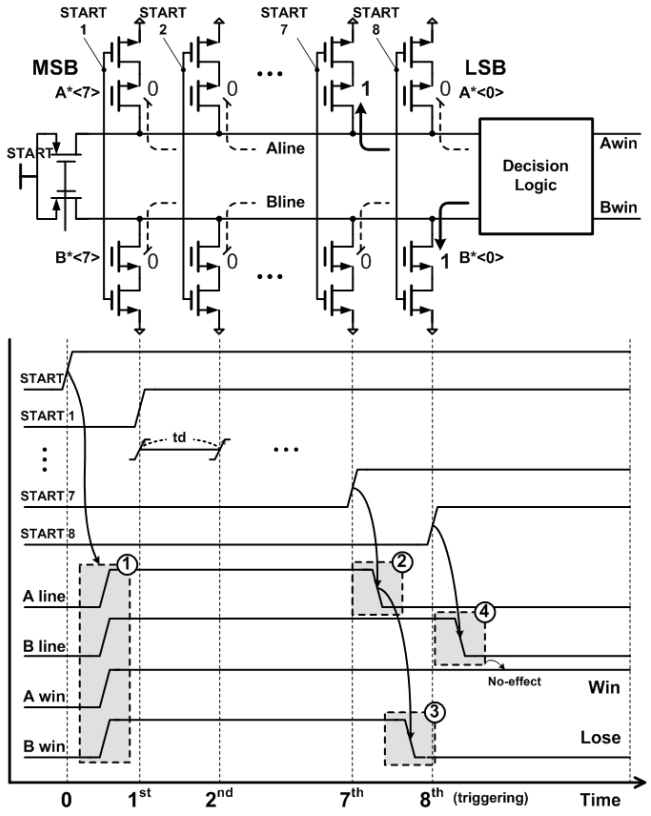

Fig. 4. BCL Comparator

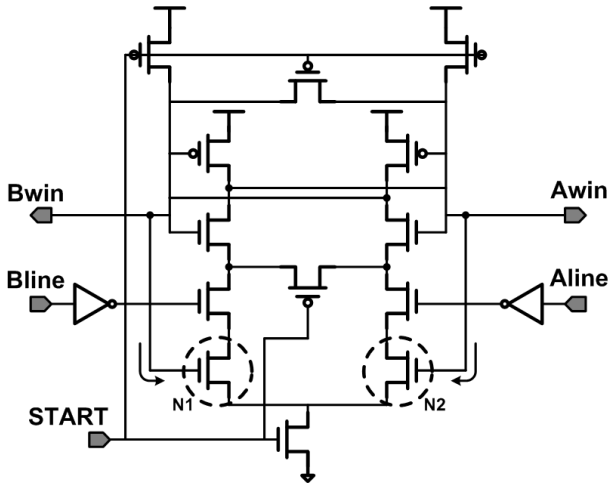

Fig. 5. Decision Logic

performed in VIP-RAM.

IV. Chip Implementation And Results

Fig. 7 shows the fabricated chip photograph and core layout of VIP-RAM in 0.18μm process. For implementation of VIP-RAM, all peripheral logic and 3-input comparator are layouted in a memory bank. As a result, all required logic blocks for LMLS operation was embedded adding only 75.4μm height to the bank and this accounts 22% of the bank height. The portion of additional 75.4μm height will be negligible as the memory size increases. In overall, a VIP-RAM occupies 1092μm x 825μm area including input buffers, control and decoding logic, and interconnection. 8 VIP-RAMs are integrated into the multi-core chip for object recognition processing [10].

The fabricated VIP-RAM is measured to verify a single cycle LMLS operation. Fig. 8 shows the measured waveforms of VIP-RAM when the 9 test input data are the same as those of the simulation of Fig.6. The waveforms include clock signal, Bwin signal, bank max address signal, and final max address signal. The measured results are well-matched with the simulation waveforms. It takes 2.53ns,

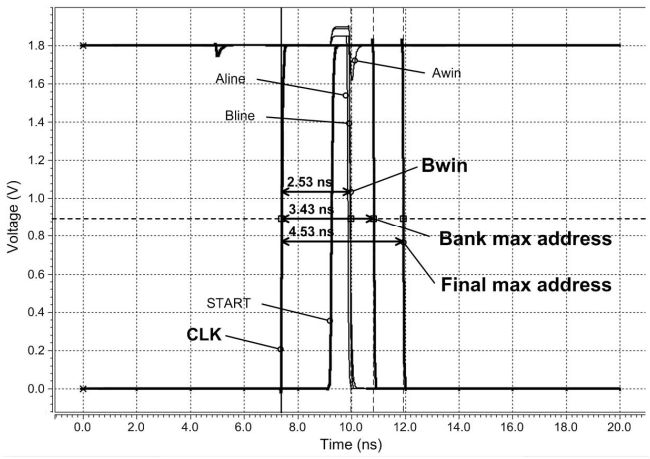

Fig. 6. Simulation Waveforms of VIP-RAM

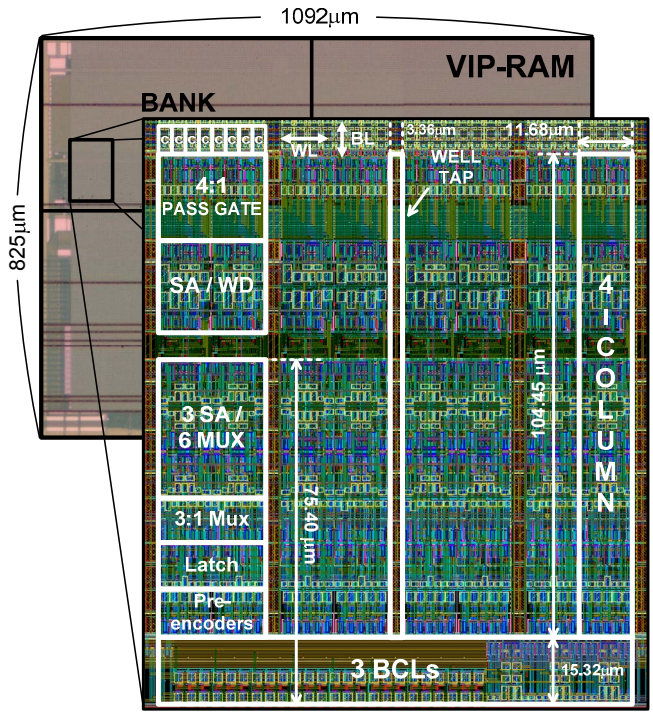

Fig. 7. Photograph of the Fabricated Chip and Core Layout

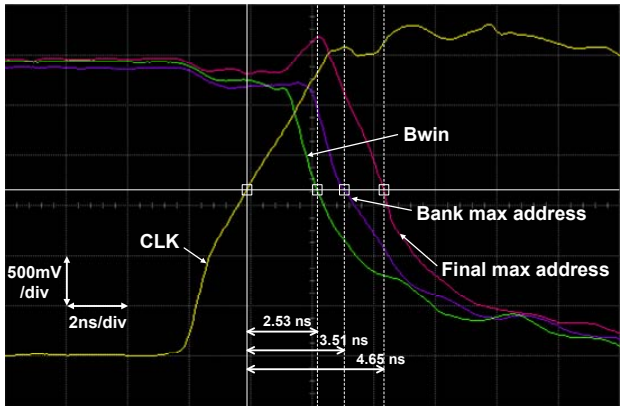

Fig. 8. Measured Waveforms of VIP-RAM

3.51ns and 4.65ns to finish two-number comparison, bank-level operation and the final operation, respectively. These are almost the same values as the simulated delay times with only the negligible error due to devoice parameter variation. And the slow slope of each waveform is caused by loading of the measurement probe and parasitic ingredients of the chip board. The measured data verifies 200MHz operation of LMLS. And 8.2GOPS peak performance is achieved in VIP-RAM when it carries out LMLS operation in every cycle. In the implemented object recognition chip, VIP-RAM is utilized in the feature point extraction stage of the well-known object recognition algorithm SIFT [5], for a 320x240 size image. VIP-RAM takes 0.137 M cycle counts to complete the task, while a general CPU running an optimized algorithm written in C takes 7.185 M cycle counts to complete the same task. Table 1 summarizes the VIP-RAM's features.

V. Conclusion & Further Remark

VIP-RAM for fast 2-D data location search is proposed and implemented. It finds the local maximum location of 3x3 size window in single cycle latency using hierarchical 3-bank architecture. Each memory bank includes special logic for 3 consecutive data read and 32-bit 3-input comparator. For the implementation of area efficient and fast 32-bit 3-input comparator, a new compare logic, BCL is devised. As a result, fabricated VIP-RAM is measured to operate at 200MHz for 8.2 GOPS peak performance. VIP-RAM architecture can be widely applied to other 2-D window operations for visual image processing with modifications to the internal embedded logic.

Table 1. VIP-RAM summary

Ports	Dual port
Special feature	2-D Data Location Search
Architecture	3 banks (logic embedded) Hierarchical structure
Physical size / Process	1092 μm x 825 μm / 0.18 μm
Access time (normal)	2.2ns
Access time (local-max)	4.65ns
Operating Frequency	200 MHz
Peak Performance	8.2 GOPS

References

[1] Wolfgang Raab, Nico Bruels, Ulrich Hachmann, Jens Harnisch, Ulrich Ramacher, Christian Sauer and Axel Techmer, "A 100-GOPS programmable processor for vehicle vision systems," IEEE Design & Test of Computers, Vol. 20, Issue 1, pp. 8-15, Jan.-Feb. 2003

[2] Juan Nieto, Jose Guivant, Eduardo Nebot, "Real time data association for FastSLAM," Proceedings of IEEE International Conference on Robotics and Automation, Vol. 1, pp. 412 – 418, Sept. 2003

[3] Yuichi Hori, Kenji Shimizu, Yutaka Nakamura and Tadahiro Kuroda, "A Real-Time Multi Face Detection Technique Using Positive-Negative Lines-of-Face Template," Proceedings of International Conference on Pattern Recognition, Vol. 1, pp. 765 – 768, Aug. 2004

[4] Hideyuki Noda, et el., "The Design and Implementation of the Massively Parallel Processor Based on the Matrix Architecture," IEEE Journal of Solid-State Circuits, Vol. 42, Issue 1, pp. 183 – 192, Jan. 2007

[5] David G. Lowe, "Object Recognition from Local Scale-Invariant Features," IEEE International Confernce on Computer Vision, Sept. 1999, pp. 1150-1157

[6] D. G. Elliott, et al., "Computational RAM: a memory-SIMD hybrid and its application to DSP," IEEE Custom Integrated Circuits Conference, May 1992, pp. 30.6.1-30.6.4.

[7] D. Patterson, et al., "Intelligent RAM (IRAM): chips that remember and compute," IEEE International Solid-State Circuits Conference, Feb. 1997, pp. 224-225

[8] N. Yamashita, et al., "A 3.84 GIPS integrated memory array processor with 64-processing elements and 2Mb SRAM," IEEE Journal of Solid-State Circuits, Vol. 29, Issue 11, Nov. 1994, pp.1336-1343.

[9] Shun-Wen Cheng, "A High-Speed Magnitude Comparator With Small Transistor Count," IEEE Proceedings of International Conference on Electronics, Circuits and Systems, Vol. 3, Dec. 2003, pp. 1168 – 1171

[10] Donghyun Kim, Kwanho Kim, Joo-Young Kim, Seung-Jin Lee, and Hoi-Jun Yoo, "Solutions for Real Chip Implementation Issues of NoC and Their Application to Memory-Centric NoC," unpublished.