Visual Image Processing RAM: Memory Architecture with 2-D Data Location Search and Data Consistency Management for a Multicore Object Recognition Processor

Joo-Young Kim, Student Member, IEEE, Donghyun Kim, Student Member, IEEE, Seungjin Lee, Student Member, IEEE, Kwanho Kim, Student Member, IEEE, and Hoi-Jun Yoo, Fellow, IEEE

Abstract—Visual image processing random access memory (VIP-RAM) is proposed for a real-time multicore object recognition processor. It has two key features for the overall processor: 1) single cycle local maximum location search (LMLS) for fast key-point localization in object recognition, and 2) data consistency management (DCM) for producer-consumer data transactions among the processors. To achieve single cycle LMLS operation for a 3 × 3 window, the VIP-RAM adopts a hierarchical three-bank architecture that finds the maximum of each row in each bank first, then finds the final maximum of the window and its address in the top level. To this end, each memory bank embeds specialized logic blocks, such as three successive data read logic and bitwise competition logic comparator. With the single cycle LMLS operation, the key-point localization task is accelerated by 2.6 × with a 27% reduction of power. For the DCM function, the VIP-RAM includes a valid check unit (VCU) that automatically manages the validity of each 32-bit data. It dynamically updates/checks the validity of the shared data when the producer processor writes the data or the consumer processor reads data. With a customized single-ended memory cell and multibit-line selection logic, the VCU can provide a validity check not only for single data access, but also for multiple data accesses such as burst and LMLS operation. Eliminating data synchronization overhead with the DCM, the VIP-RAM reduces the amount of on-chip data transactions and execution time in producer–consumer data transactions by 22.6% and 15.4%, respectively. The overall object recognition processor that includes eight VIP-RAMs and ten processors is fabricated in 0.18-µm complementary metal–oxide–semiconductor technology with the chip size of 7.7 mm × 5 mm. The VIP-RAM occupies a 1.09 mm × 0.83 mm die area and dissipates 113.2 mW when it performs the LMLS operation in every cycle at 200 MHz frequency and 1.8-V supply.

Index Terms—Data consistency management (DCM), local maximum location search (LMLS), multicore processor, object recognition, visual image processing RAM (VIP-RAM).

I. INTRODUCTION

RECENTLY, visual image recognition applications such as autonomous vehicle control, intelligent mobile robots, and face recognition have been widely studied due to their high market expectations [1]–[7]. For their real-time operation, several application-specific integrated circuit chips and field programmable gate array (FPGA) systems have been implemented [1]–[7]. However, most of these works [1]–[4] required a limited object data set and simple matching, because they only target vehicle vision or face detection. On the other hand, the object recognition processors [6], [7] demand a considerably more sophisticated recognition process, because they are implemented for intelligent robot applications that require general object recognition capability, like human vision. The architectures and processing capability of the previous processors are also determined according to their target applications. The architecture of [1] and [2] employs tens of processing elements, which perform the same operation on multiple pixels to increase data-level parallelism. However, these architectures are not suitable for simultaneous execution of multiple independent tasks, or task-level parallelism, which is essential for a complicated object recognition process. On the contrary, the processor of [3] supports both data-level parallelism and task-level parallelism by employing three very long instruction word processors for media processing. However, its limited number of processors cannot provide sufficient computing power for general object recognition. In the processor proposed in [6], ten independent controlled single instruction multiple data (SIMD) processors and eight visual image processing random access memories (VIP-RAM) are integrated through a network-on-chip (NoC) to provide sufficient computing power for general object recognition. This processor supports not only data-level parallelism but also task-level parallelism by assigning parallel tasks to different SIMD processors. Fig. 1 shows its overall architecture. Fig. 2 shows the overall processing flow of the scale invariant feature transform (SIFT) [8], which is applied to the proposed processor for general artificial vision. It is divided into two stages: 1) key-point localization, and 2) descriptor
Fig. 1. Overall processor architecture.

Fig. 2. Overall processing flow of SIFT. (a) Key-point localization. (b) Descriptor vector generation. For key-point localization, various scale spaces for the input video stream are generated by cascade filtering operations and the key-points are extracted from Difference of Gaussian (DoG) images by a local maximum location search (LMLS) operation with a 3 × 3 search window. The extracted key-points are then converted to descriptor vectors by computing orientation and magnitude histograms over subregions of the surrounding N × N image pixels. From a computational aspect, the Gaussian filtering accounts for over 50% of the overall SIFT processing. To accelerate this, the filtering tasks for different scale spaces are divided into ten SIMD processors and they perform the tasks using SIMD operations. The second most complex operation is the LMLS, which amounts to 27% of the overall processing. The LMLS operation is the process of finding the address of the local maximum value among nine data in a 3 × 3 search window. It is difficult to accelerate this, since it includes nine loads, comparison of nine data, and several conditional branches to decide the location of the maximum data. Fig. 3 shows the microcode of a LMLS based on an advanced RISC machines (ARM) v4 instruction set architecture [19] for further understanding. In this case, the LMLS operation requires 29–53, or 41 cycles on average.

Fig. 4 shows the task mapping of the key-point localization stage in the SIFT with the proposed multicore processor. Since it is a streaming process, there are lots of 1-to-N and M-to-1 producer–consumer data transactions. A processor performs its assigned task using the source data from its former processors (M-to-1) and generates the result data to feed its following processors (1-to-N). In these producer–consumer data transactions, using VIP-RAM as a shared communication buffer can provide a good alternative to remove the redundant data transfers, especially when multiple processors require the same intermediate data. However, the key issue of sharing intermediate data is how to maintain consistency of the data between the producer processor and consumer processors in the time domain. Under the condition that the producer and consumer processors are decided with their communication channel, the consumer processor should receive valid data that are validated from the producer processor.

In this paper, we propose an application-specific memory named VIP-RAM to resolve the aforementioned two problems in a multicore object recognition processor, i.e., complex LMLS operation and the data consistency problem. To accelerate the LMLS operation, the VIP-RAM utilizes a merged memory approach [9], [10], which combines a memory cell array and processing units, over conventional 2-D rank-ordering filters [11]–[14]. While the conventional configuration is composed of specialized logic blocks for sorting networks and a separated memory module, the VIP-RAM merges specialized circuits into a memory cell array to remove data fetch overhead and to perform the role of a data communication channel. For the specialized LMLS operation, the VIP-RAM integrates a 2-D data access unit, an intensive
compare computation unit, and an address generation unit. To address the data consistency problem, the VIP-RAM manages a validity bit array, which stores the dynamic status of every data in the memory. It concurrently updates the validity of the data with a write request and provides the validity of the data with a read request, respectively. For multiple data access cases, such as burst transactions and LMLS operation, a multiple data validity check is also supported with a customized cell array architecture and multibit-line/word-line selection. With the VIP-RAMs data consistency management (DCM) capability, producer-consumer data transactions can be efficiently supported in the proposed multicore processor.

The remainder of this paper is organized as follows. In Section II, the overall architecture of VIP-RAM is introduced. A detailed block design for the LMLS operation and DCM operation is then explained in Section III. The effects of VIP-RAMs two main features are subsequently analyzed in Section IV. Implementation results follow in Section V. Finally, we summarize this paper in Section VI.

II. VIP-RAM ARCHITECTURE

The VIP-RAM has two behavioral modes, a normal mode and a local-max mode. In the normal mode, the VIP-RAM operates as an ordinary synchronous dual-port static random access memory (SRAM). It has two input address and control signals, and reads or writes two 32-bit data independently. On the other hand, in the local-max mode, the VIP-RAM outputs the address of the local maximum data out of a $3 \times 3$ window while the input address indicates the left and upper most data of the window. In this mode, only one of two inputs is valid. Fig. 5 shows the overall architecture of the VIP-RAM. It consists of two input controllers, three logic-embedded banks, an address generation unit, and a valid check unit. The input controller latches input signals and generates control signals for each bank. Each bank of the VIP-RAM is composed of 128 words, which are organized into 32 rows and four columns. Thus, all three banks can store $12 \times 32$ sized 32-bit pixel data, which amount to 1.5 kB capacity. In each bank, each bit of four columns shares not only write drivers and sense amplifiers.
Fig. 6. Image data mapping between the physical and logical space of VIP-RAM.

for ordinary dual port write and read operation, but also three selection mux pairs and three small-sized sense amplifiers for three successive data read operation. In addition, a compact three-input comparator for 32-bit data is integrated below the memory cell arrays. As a result, each bank of the VIP-RAM is able to read three successive data at once and find the maximum data among them in the same cycle. For the LMLS operation, nine read data out of three banks should form a complete $3 \times 3$ window in image space. For this 2-D data access, the pixel data of the $12 \times 32$ image should be properly mapped into the physical memory of the VIP-RAM, as shown in Fig. 6. Each row of the $12 \times 32$ size image, which occupies eight physical rows in a memory bank, is alternately loaded into each bank of the VIP-RAM. As a result, rows 0, 3, 6, and 9 are stored into bank 0, and rows 1, 4, 7, and 11 are stored into bank 1, and the other rows are stored into bank 2. After that, when the three banks of the VIP-RAM read three successive data from the same internal address, a total of nine data out of three banks form a $3 \times 3$ window in the image space. With the properly mapped pixel data, the LMLS operation is processed hierarchically according to the following three steps. First, three successive data are read simultaneously in each bank and constitute a row of a $3 \times 3$ window. In case the input address indicates the boundary region of the bank column, three successive data are read out of two rows in a memory bank. Second, each bank finds the largest among three data and sends it to the top-level compare logic circuits with its 2-bit address distance. The address distance is the difference value between the address of the maximum data and the input address of the bank. Because the input address indicates the left most data, the possible value of the address distance is 0, 1, or 2. Finally, the top-level address generation unit finds the maximum data of the $3 \times 3$ window among the three bank-level maximum results and deduces its address using the address distance.

For data synchronization in communication among multiple processors, the VIP-RAM includes a valid check unit (VCU) at the top-level, independent of the operation of the three banks. The VIP-RAM initializes the validity array of the VCU to "0" before it is used as the communication channel. Once a VIP-RAM is assigned to the channel, it confirms every write operation of the producer processor by writing "1" to the corresponding bit of the validity array. In the case of reading, the VCU notifies whether the data is valid or not to the consumer processors. If the valid bit is 1, the data is allowed to be read and returned to the processors. On the other hand, in case the valid bit is 0, the data is not allowed to be read and the processors are put on hold until the data becomes valid. This validity management guarantees the timing sequence or consistency, which means the consumer uses the data after the producer updates them. The multibit-line selection logic is added to check the validity of multiple data in the case of a burst or LMLS operation. To respond to these multiple data requests, the validity of every requested data should be checked in a single cycle. In multiple data validity evaluation, the overall validity is determined as invalid if any invalid data exists.

III. Detailed Circuit Design

In this chapter, the detailed circuit design of the VIP-RAM is explained according to its two main functions: the LMLS and DCM. For the LMLS operation, three successive data read logic, a compact three-input comparator, and an address generation unit are described. For the DCM operation, the VCU'S customized cell and multibit-line selection logic are explained.

A. Circuits for LMLS

In the design of the VIP-RAM, an ordinary 8-transistor dual-port memory cell that contains two independent word-lines and bit-line pairs for simultaneous accesses of two ports is utilized. The highly optimized layout occupies $2.92 \mu m \times 5.00 \mu m$ area in 0.18$\mu m$ complementary metal-oxide-semiconductor (CMOS) technology. For the LMLS operation, each bank of the VIP-RAM can access three
successive row data of a window at once. To this end, each bank contains three successive data read logic in its data read path, as shown in Fig. 7. For the 4 column cell array of the VIP-RAM, four different cases are possible with respect to reading three successive data according to the input address. Among these cases, the cases where the input address indicates the third and fourth columns require data from the corresponding row and the next row at the same time. To read them together, the bank activates both word-lines using two ports of the memory cell. Eight data from the two word-lines are then read through the bit-lines and the correct three data are selected from among them by the three mux pairs. The control signals of the mux pairs are defined as given in the table presented in Fig. 7 according to the four cases. Finally, the obtained three data are amplified by small-sized sense amplifiers.

After each bank reads three successive data, the maximum data among them should be computed for the LMLS operation. For this, each bank of the VIP-RAM merges a 32-bit three-input comparator with a memory cell array and three successive data read logic. The three-input comparator that will be merged in each bank should be compact and fast for easy combination with the memory cell array and further processing outside of the bank, respectively. To satisfy the above conditions, BCL is proposed for long binary number comparison. Different from the conventional comparators that use expanded arithmetic equations for number comparison [15]–[17], the BCL comparator uses the location of the first 1 from the most significant bit (MSB), as shown in the flow chart of Fig. 8(a). First, the two inputs are encoded bit by bit to remove the cases where two bits in the same position are both 1, which lead to logic failure in the BCL. After pre-encoding, only one input has the value 1 in any bit position. After that, a bit comparison is performed from the MSB to the least significant bit (LSB) to detect which input has the first 1 in higher bit, or which input is larger. If no 1 is found...
Fig. 8(b) shows the circuit implementation of the 8-bit BCL comparator. Each bit of the two 8-bit pre-encoded inputs is connected to the A line and B line in parallel with the enabling start signal. The voltages of the A line and B line, initially pre-charged to the VDD, are used as reference values to compare the two input numbers. For the bit comparison process, the enabling START signals are sequentially activated from the MSB to LSB at a constant time interval. Each bit of the pre-encoded signals start is then compared in series by sequentially activated enabling signals. In case the value of the pre-encoded bit is 1, the path from the line to the ground (GND) is connected and the voltage of the line is quickly dropped to zero. Otherwise, the voltage of the line is preserved to the VDD. For the generation of the sequential START signals, a simple inverter chain is used to establish an interval time of about 100 ps, which guarantees the complete transition time of the line voltage. As a result, during the bit comparison process, the line whose voltage is dropped earlier than the other has the first 1 from the MSB in the higher bit. Using this, the BCL comparator can determine the larger input between the two by detecting whose line voltage is dropped earlier. To this end, winner selection logic (WSL) is employed at the end of the lines. The WSL senses the voltage drop of either line and amplifies it to the VDD based on the modified sense amplifier circuits. The N1 and N2 transistors are employed to block further changes of the loser line after the amplification is finished. Over the 8-bit BCL comparator, the 32-bit BCL comparator can be implemented using four parallel 8-bit BCL comparators with 2-stage selection circuits. As a result, the 32-bit 2-input BCL comparator consumes 482 transistors and performs 950 ps of worst case delay, constituting 38% and 16% improvements from the previous comparators [18]. For a 32-bit three-input comparator, three 32-bit 2-input BCL comparators are employed and the worst case delay is about 1.4 ns. This delay is sufficiently small for a 5 ns timing budget of a 200 MHz operating VIP-RAM.

B. Circuits for DCM

The other functional block of the VIP-RAM is the VCU for DCM in multiprocessor communication. The VCU consists of a validity array, composed of 1-bit validity mapped to each 32-bit data, and multibit-line/word-line selection logic to support the validity check for multiple data access cases such as burst data transmission and the LMLS operation. To minimize the cell area and to facilitate handling multiple validity access cases, the cell of the validity array is customized as shown in Fig. 10. It includes three word-lines, two command lines, and two bit-lines for simultaneous operations of one write and two reads. A single-ended bit-line is exploited to reduce the cell area. The writing operation to the cell is performed separately by two command lines: the write command line and clear command line. When the write command line is activated with the write word-line, the value 1 is memorized to the cell. On the other hand, when the clear command line is activated, the value of the cell is reset to 0 regardless of the word-line activation. The read operation is simply performed by the activations of the two read word-lines. If the stored data is 0, the voltage of the bit-line is lowered to 0 by the pull-down transistor whose gate signal is connected to the complementary value of the cell. In the opposite case, the voltage of the bit-line is preserved to the pre-charged VDD value.

To read multiple validity bits in series, multiple bit-line selection logic is devised. It quickly activates subsequent bit-lines using propagation domino circuits to read out the sequential validity bits in a row, as shown in Fig. 11. Using the start/stop signals and cascaded domino circuits, we can read variable lengths of validity bits in a row by propagating
a bit-line select signal from the bit-line with the start to the bit-line with the stop. As a result, the proposed dynamic propagation circuits generate 16 bit-line select signals within 0.7 ns. The sequential validity bits activated by the bit-line select signals are then read through bit-lines in the same cycle and the final validity output is generated by an AND gate, which aggregates all bit-line signals. If any invalid bit exists in the read validity bits, the overall validity is determined to be invalid. Since the default value of the bit-line is 1, inactivated bit-lines do not affect the AND gate. For the validity check of the LMLS operation, a 2-D data access is required to check whether the nine data in a $3 \times 3$ window are all valid. To this end, word-line activation is additionally necessary for multibit-line selection. For a $3 \times 3$ window data, three word-lines and three bit-lines are activated at the same time. Since the column cells of the validity array share one bit-line and their pull-down transistors are connected in an NOR fashion, the voltage of the bit-line is pulled down if any one of them contains 0. The final validity of the nine data in a $3 \times 3$ window is then generated by the final AND gate, as in the case of the validity check for sequential validity bits.

IV. VIP-RAM Evaluation in Overall System

This chapter evaluates the contributions of the VIP-RAM to the overall system of [6]. First, the advantages of VIP-RAMs LMLS operation are explained with respect to computational aspects, i.e., execution time, power, and energy consumption of object recognition processing. The benefits of the DCM, which contributes to realizing efficient producer–consumer data transactions, are then discussed in terms of the execution time and amount of on-chip traffic.

A. Advantages of LMLS

The merged compare and search circuits of the VIP-RAM replace the 30–50 cycle consuming LMLS operation with a single read operation. To evaluate the advantages of this feature with regard to the overall object recognition processor, the execution time and power consumption are measured from the gate level netlist of the SIMD processor and the transistor level netlist of the VIP-RAM using tools such as Verilog RTL, HSPICE, and Synopsys Power Compiler.

Fig. 12 shows the two processing models for the 3-D LMLS operation of the SIFT for quarter video graphics array (QVGA) (320 × 240) sized DoG images, (a) without VIP-RAM and (b) with VIP-RAM. To determine whether each pixel of the center DoG image is the local maximum of the 3-D cube, three $3 \times 3$ search windows are traversed throughout the top, center, and bottom DoG image and the computed three 2-D local maximum pixels are compared for the final decision. In the case without VIP-RAM, the processor fetches three $3 \times 3$ window data from the top, center, and bottom DoG image to the local memory and computes the maximum data and location among the three $3 \times 3$ window data sets. Since the processor performs the complex LMLS operations, numerous execution cycles and frequent data transactions with the local memory are necessary. With VIP-RAM, on the other hand, the local maximum pixel location of each image can be directly read from the VIP-RAMs, and thus the final result can be easily computed by the processor with three comparisons. The local memory is also rarely used.

Fig. 13 presents performance comparisons between the two models. The execution time, power dissipation, and energy consumption are measured when the mentioned 3-D LMLS tasks are applied for a QVGA sized image. With VIP-RAM, the execution speed is improved by 2.6× times due to its single cycle LMLS operation. Furthermore, the power consumption
is reduced by 27.6%. Since the VIP-RAMs perform the LMLS operations on behalf of the processors, the power to perform the LMLS operations can be saved. In addition, the LMLS operation also reduces the data transaction power of the local memory, because it eliminates data transactions between the memory and processor. It only dissipates 1 mW to store temporal variables. Instead of these power savings, the VIP-RAM dissipates an additional 3 mW to perform a single cycle of LMLS. However, the VIP-RAM contributes to reducing the overall power consumption due to the large power savings from the processors. With respect to energy consumption, the VIP-RAM achieves a 4.62 times improvement to process an image frame.

The negative aspect of the VIP-RAM is area overhead. Since each bank of the VIP-RAM includes additional logic blocks such as successive data read logic and BCL comparators, it requires additional bank area under the memory cell array. In this VIP-RAM design, these logic blocks amount to 22% of the area of the bank. Since the required area for the embedded logic blocks is constant, their area will become smaller as the size of VIP-RAM increases.

B. Advantages of DCM

To evaluate the advantages of the DCM, the key-point localization stage of the SIFT is mapped on the proposed multicore processor, as shown in Fig. 4. In this case, a massive amount of producer–consumer data transactions occur between the processors of the adjacent stages, and frequent data synchronizations are required among them. Without VIP-RAMs DCM capability, data synchronization between the producer and consumer processor should be performed using additional synchronization variables, as shown in Fig. 14(a). After a producer processor writes a certain amount of data in a shared memory, it also updates synchronization variables to inform the consumer processors about the amount of available data. Accordingly, the consumer processors check the synchronization variables before reading the data to access whether they are ready or not. In this producer–consumer data synchronization, the drawbacks are additional data transactions and execution time to update/check the synchronization variables. On the other hand, with VIP-RAMs DCM, these data synchronizations are automatically performed by the VIP-RAM, as shown in Fig. 14(b). When a producer processor writes data, the VIP-RAM automatically updates its corresponding validity bits in the VCU. When the consumer processors read data, the validity bits of the requested data are also automatically read from the VCU to the consumer processors. Since the VIP-RAM dynamically updates/checks the validity of the requested data in the same cycle, the data transactions and execution time for data synchronization can be eliminated. The advantages of the DCM are shown in Fig. 15. It compares the on-chip data transactions and their execution time for the task of Fig. 4. The vertical axis of the graph represents the amount of on-chip data transactions or
Fig. 17. Comparisons with other works. (a) overall recognition processor. (b) 2-D ordering filter.

the execution time without the DCM, which are normalized to those obtained with the DCM. The horizontal axis depicts the data synchronization unit, which represents the size of the synchronization variables updated or read by the processors. For example, a 128-byte data synchronization unit means a producer processor updates the synchronization variable every read of 128-byte data and the consumer processor checks the synchronization variable every read of 128-byte data. As shown in Fig. 15, for small data synchronization units less than 128 bytes, the overhead of frequent accesses to the synchronization variables degrades the performance and the large data synchronization unit increases unnecessary polling to wait for valid data from the producer processor. In the case of this SIFT task, the 128-byte data synchronization unit appears to be optimal. However, none of these cases outperforms the computation based on the DCM. The DCM reduces the amount of on-chip data transactions and execution time by more than 22.6% and 15.4%, respectively, compared to the best results for a synchronization scenario without the DCM.

V. CHIP IMPLEMENTATION

The proposed multicore object recognition processor is fabricated in a TSMC 0.18µm standard CMOS technology. The chip micrographs, summary, and power breakdown are shown in Fig. 16. The chip has 7.7 mm × 5 mm die area while the VIP-RAM accounts for 1092 m× 825 m area. The operating frequency is 400 MHz for interconnection NoC and 200 MHz for the other parts. The overall chip integrates a total of 30 KB on-chip memories including the 12 KB VIP-RAM, and 838,000 gates for logic. Its peak performance amounts to 81.6 giga operation per second (GOPS), which consists of 16 GOPS from the 10 SIMD processors and 65.6 GOPS from eight VIP-RAMs. Each VIP-RAM accounts for 8.2 GOPS performance, that is, when the 41-cycle equivalent LMLS operation is performed every cycle. Fig. 16(c) presents the power breakdown of the overall processor under a scenario where all components are fully loaded for peak performance; all eight VIP-RAMs are operated in the LMLS mode and all ten processors and NoC are fully loaded. Under these conditions, the power consumption by the eight VIP-RAMs amounts to more than half of the total consumption, because the LMLS operation, which activates all three banks, reads 3 × 3 window data, and compares nine data in the same cycle, consumes more than three times higher power dissipation than that of a normal dual-port memory operation.

To evaluate the proposed processor relative to existing approaches, we perform two comparisons (Fig. 17). First, we compare the proposed recognition processor with previous recognition processors [1]–[5]. While the processors of [1]–[3] and the proposed processor select a very large scale integration solution with large programmability, the processors of [4], [5] are implemented using FPGA with dedicated logic blocks. For power efficiency, which is measured by dividing the peak performance (GOPS) by the peak power (W), the proposed processor achieves more than 3.2× improvement compared to that of previous chips. For applications, the processor of [5] and the proposed processor aim at general object recognition, specifically SIFT, while the other processors are targeting for vehicle or face recognition. The proposed processor achieves 16 frame per second (fps) application performance for QVGA...
The VIP-RAM achieves single cycle max/min search operation. It eliminates data fetch overhead and provides a communication buffer function simultaneously. Therefore, in VIP-RAM area, 1.5 KB memory cells that can be used for shared data are also included. With the help of the merged architecture and specially designed compare circuits, the VIP-RAM achieves single cycle max/min search operation for a 3 × 3 32-bit data window at a 200 MHz operating frequency, which amounts to 200 M max/min searches per second. While the previous works also performed single or few cycle search operations, they only processed 8-bit or 11-bit data, not 32-bit data as in the case of VIP-RAM.

VI. Conclusion

The design and implementation of the VIP-RAM were described for real-time object recognition processing and data synchronization in multiprocessor communication. To achieve a single cycle LMLS operation for a 3 × 3 sized local window, the VIP-RAM exploits a hierarchical 3 bank architecture that finds the maximum of each row in each bank and the final maximum data of the window in the top-level logic. To this end, each bank includes three successive data read logic in its data read path and 32-bit three-input comparator at the bottom of the memory array. To implement a compact and fast 32-bit three-input comparator, a new compare logic named BCL is proposed. As the data communication channel, the VIP-RAM includes a VCU that manages the 1-bit validity of each 32-bit data to provide DCM for producer-consumer data transactions. To support a multiple data validity check in the case of burst and LMLS operations, a customized single-ended memory cell and multibit selection logic were employed. As a result, eight VIP-RAM cores are integrated into the object recognition system-on-chip with ten processors within a 7.7 mm × 5.5 mm area in 0.18 μm technology. The VIP-RAM occupies 1092 m × 825 m and its LMLS and DCM operations were successfully evaluated in the overall recognition system.

REFERENCES


Joo-Young Kim (S’05) received the B.S. and M.S. degrees in electrical engineering and computer science from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2005 and 2007, respectively. He is currently working toward the Ph.D. degree in electrical engineering and computer science at the Department of Electrical Engineering and Computer Science, KAIST. Since 2006, he has been involved with the development of parallel processors for computer vision. Currently, his research interests include parallel architecture, sub-systems, and very large scale integration implementation for bio-inspired vision processors.
Donghyun Kim (S’03) received the B.S. degree in electrical engineering and computer science from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2003. He is currently working toward the Ph.D. degree in electrical engineering and computer science at the Department of Electrical Engineering and Computer Science, KAIST. His research interests include network-on-chip design, multiprocessor design, and parallel image processing. He is currently working on an analysis of on-chip data transactions and task mappings of applications with streamed data flow into multiprocessor system-on-chip architectures.

Seungjin Lee (S’06) received the B.S. and M.S. degrees in electrical engineering and computer science from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2006 and 2008, respectively. He is currently working toward the Ph.D. degree in electrical engineering and computer science at the Department of Electrical Engineering and Computer Science, KAIST. His previous research interests included low-power digital signal processors for digital hearing aids and body area communication. Currently, he is investigating parallel architectures for computer vision processing.

Kwanho Kim (S’04) received the B.S. and M.S. degrees in electrical engineering and computer science from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2004 and 2006, respectively. He is currently working toward the Ph.D. degree in electrical engineering and computer science at the Department of Electrical Engineering and Computer Science, KAIST. In 2004, he was with the Semiconductor System Laboratory, KAIST, as a Research Assistant. His research interests include very large scale integration design for object recognition, architecture, and implementation of network-on-chip-based system-on-chip.

Hoi-Jun Yoo (M’95–SM’04–F’08) received the B.S. degree from the Department of Electronics, Seoul National University, Seoul, Korea, in 1983, and received the M.S. and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 1985 and 1988, respectively. His Ph.D. work concerned the fabrication process for gallium arsenide vertical optoelectronic integrated circuits. From 1988 to 1990, he was with Bell Communications Research, Red Bank, NJ, where he invented the 2-D phase-locked vertical cavity surface-emitting laser array, the front-surface-emitting laser, and the high-speed lateral heterojunction bipolar transistor. In 1991, he was the Manager of the Dynamic Random Access Memory (DRAM) Design Group, Hyundai Electronics, Ichonu, Kyungsung-do, Korea, and designed a family of 4-m, 16-m, 64-m, and 256-m synchronous DRAMs. In 1998, he joined the Faculty of the Department of Electrical Engineering, KAIST, where he is now a Full Professor. From 2001 to 2005, he was the Director of the System Integration and Internet Protocol (IP) Authoring Research Center, funded by the Korean Government to promote worldwide IP authoring and its system-on-chip (SoC) application. From 2003 to 2005, he was the Full-Time Advisor to the Minister of Korea, Ministry of Information and Communication, and the National Project Manager for SoC and Computers. In 2007, he founded the System Design Innovation and Application Research Center, KAIST, to research and develop SoCs for intelligent robots, wearable computers, and bio systems. He is the author of the books DRAM Design (Seoul, Korea: Hongleung, 1996 in Korean), High Performance DRAM (Seoul, Korea: Sigma, 1999 in Korean), and chapters of Networks on Chips (New York: Morgan Kaufmann, 2006). His current research interests include high-speed and low-power network-on-chips, 3-D graphics, body area networks, biomedical devices and circuits, and memory circuits and systems.

Dr. Yoo received the Electronic Industrial Association of Korea Award for his contribution to DRAM technology in 1994, the Hynix Development Award in 1995, the Korea Semiconductor Industry Association Award in 2002, the Best Research of KAIST Award in 2007, the 2001 Asia and South Pacific Design Automation Conference Design Award, and the Outstanding Design Awards in the 2005, 2006, and 2007 Asian Solid-State Circuits Conferences (A-SSCC). He is a Member of the Executive Committee of the International Solid-State Circuits Conference, the Symposium on Very Large Scale Integration, and A-SSCC. He was the Technical Program Committee Chair of the A-SSCC in 2008.