

Low Power Motion Estimation and Motion Compensation Block IPs in MPEG-4 Video Codec Hardware for Portable Applications

Chi-Weon YOON^{†a)} and Hoi-Jun YOO[†], *Nonmembers*

SUMMARY In this paper, two low power hardware structures essential for MPEG-4 video codec are proposed for portable applications. First, an adaptive bit resolution control (ABRC) scheme is proposed for a processing element (PE) in a systolic-array type motion estimator (ME). By appropriately modifying the datapath of PE to exploit the correlations in pixel values, its structure is optimized in terms of both hardware cost and low power consumption. As a result, power is saved up to 29% compared with a conventional PE while the computation accuracy is preserved and the overhead is kept negligible. Second, a low power motion compensation (MC) accelerator is proposed. By embedding DRAM whose structure is optimized for low power consumption, the power consumption for external data I/Os is dramatically reduced. In addition, distributed nine-tiled block mapping (DNTBM) with partial activation scheme in the frame buffer reduces the power for accessing frame buffer up to 31% compared to a conventional 1-bank tiled mapping. With the proposed MC accelerator, MPEG-4 SP@L1 decoding system is fabricated using 0.18 μm embedded memory logic (EML) technology.

key words: *low power, motion estimation, motion compensation, MPEG-4*

1. Introduction

1.1 Motivation

Recently, multimedia-processing capability on the portable terminals such as smart-phones or PDAs (personal digital assistance) is becoming essential with the help of current advanced VLSI technologies and high capacity channel technologies such as IMT-2000. Among various functions, a video processing based on MPEG-4 or H.26x is one of the promising functions that are going to be widely used on the terminals [1]. Clearly, low power consumption is the most important constraint as well as a performance or flexibility when implementing the functions on the portable devices because most of them are battery-driven products. Although a video codec can be implemented in various ways, a hardware implementation is preferred in order to successfully satisfy both the power consumption and the required performance [2]–[7]. It is especially true for the motion estimation (ME) and the motion

compensation (MC) because they require most of the computational power and data I/O operations in video processing [8].

For a low power hardware implementation, various approaches were proposed [2]–[6], [10]–[13]. Since the power consumption of digital circuits is mainly determined by the dynamic power, which is proportional to the switching activities and the capacitive load [9], [14], most of the approaches were focused on reducing them. Previous works [10]–[13] have tried to reduce the switching activities for the low power ME implementation by clock gating [10], by eliminating unnecessary computation using estimation of mathematical inequality [11], or by reducing the required hardware [12], [13]. But they suffer from its large gate count [10], excessive additional hardware amount [11], or image quality degradation with the cost of reduced power consumption [12], [13]. On the other hand, there have been various approaches trying to reduce the power consumption of data I/O operations [5], [6]. Off-chip data I/O operations typically require much more power than that of on-chip operations due to the relatively large capacitive load. By embedding memories (DRAM) into a single chip, the amount of capacitive load can be steeply reduced so that a large amount of power can be saved. But the benefits of embedded DRAM is not fully utilized in their architectures because they just integrated a memory with the logic while maintaining the traditional off-chip memory system architecture and without any considerations of optimization in architectural level for further power reduction.

1.2 Contributions

In this paper, we propose the low power hardware structures for two key blocks (ME and MC) that are essential for the video codec.

First, we propose an adaptive bit resolution control (ABRC) scheme for a processing element (PE) in a systolic array-type motion estimator (ME) [2]. It is a well-known fact that the pixels in a local area of the successive frames are highly correlated. But this nature is not yet successfully utilized for low power consumption in previous works [10]–[13] because it may increase the hardware cost too much or the power reduction

Manuscript received August 26, 2002.

Manuscript revised December 10, 2002.

[†]The authors are with the Korea Advanced Institute of Science and Technology, Seoul, Korea.

a) E-mail: cwyun@eeinfo.kaist.ac.kr

is insignificant if it is not appropriately implemented. However, the optimization in terms of both power and hardware cost was possible by the proposed scheme. By adding simple control circuits, the datapath of PE is modified to adaptively control the hardware amount that is actually required for the computation so that unnecessary transitions in the datapath are reduced. Different from previous works [12], [13], image quality degradation does not occur since the computation accuracy in a PE is preserved. In addition, it can be combined with any kinds of architectural level low power techniques [10], [11].

Second, we propose the low power MC accelerator with embedded DRAM frame buffers [2]–[4]. To reduce the power consumption, embedded DRAM structure is adopted, and its architecture is optimized in terms of low power operation based on the frequently used memory access patterns.

1.3 Organization

The organization of this paper is as follows. In Sect. 2, the proposed technique for low power ME and its experimental results are described. The detailed description about low power MC design and its frame buffer optimization is given in Sect. 3, and the conclusions will follow in Sect. 4.

2. Design of Low Power Motion Estimator (ME) for MPEG-4 Encoding

2.1 Adaptive Bit Resolution Control (ABRC) Scheme for Low Power ME

Most of the motion estimation algorithms are developed based on a block matching (BM) algorithm that computes a motion vector on a block-by-block basis, and sum of absolute difference (SAD) is one of the widely used distance criterions for them. SAD is defined as follows:

$$SAD(m, n) = \sum_{i=1}^N \sum_{j=1}^N |x_{i,j} - y_{i+m,j+n}| \quad (1)$$

where N is the block size, and $x_{i,j}$ and $y_{i,j}$ are the pixel values of the current block and the previous block, respectively. In a hardware implementation, a systolic-array type processor based on a processing element (PE) is typically used because SAD operation can be easily parallelized and it is much simpler than other approaches in view of design complexity [10]–[13]. In this structure, basic operation for SAD is mapped on a PE. It mainly consists of three operations; 1) subtraction, 2) absolute operation, 3) accumulation. Basically, these operations are performed in 8-bit resolution, and the hardware with the corresponding bit resolution is required in conventional PE structures.

Table 1 Bit resolution distribution used for ME. (Full search [−16, 15.5])

	0~4 LSEs	5~8 LSEs
Akiyo	53.3	46.7
MissA	<u>67.6</u>	32.4
Carphon	47.6	52.4
Foreman	43.2	56.8
Table	54.9	45.1
Football	<u>27.7</u>	72.3
Stefan	34.0	66.0
Avg	46.9	53.1

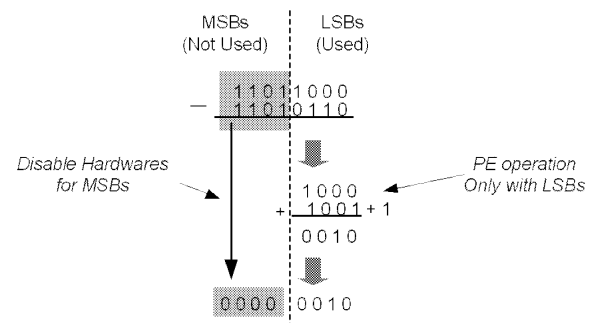


Fig. 1 The concept of adaptive bit resolution control (ABRC) scheme.

By the way, for many operations in a PE, the full bit resolution (8-bit) operation is not actually required. Large correlation in the pixel values at a local area of successive frames means that the values can be very similar so that there are little changes in most-significant bits (MSBs) [8]. The correlation becomes stronger as the scenes become simple or have a constant intensity in a local area (e.g. background). Table 1 shows the number of LSBs that were actually used for ME (Full Search, [−16, 15.5]). As shown in the table, over 50% of the operations can be performed only with 4 LSBs in case of slow scenes (*Miss America*, *Akiyo*) and even for the fast scenes such as *football*, still many operations are performed with lower than 8-bit resolution.

To effectively utilize this, the ABRC scheme adaptively determines the number of LSB bits that are required for the operation according to the operands (Fig. 1). Therefore, only the necessary LSB components in the datapath of PE are enabled, and the redundant transitions in the datapath for MSBs are reduced so that large power saving can be obtained. In addition, the image quality degradation does not occur because the exact computation results are obtained in the PE (Fig. 1).

Figure 2 shows the example implementation of PE with 4-bit granularity control scheme. 4-bit granularity

Table 3 The overall power reduction in PE.

(a) Full search. [-16, 15.5]

Structure	Akiyo	MissA	Car.	Fore.	Table	Foot	Stefan	Avg. (%)	Pwr * Area
<u>4b-4b</u>	13.5	20.8	10.5	8.2	13.8	1.5	3.6	10.3	<u>9601</u>
4b-2b-2b	14.8	21.1	12.8	10.2	15.0	4.4	6.2	12.1	9670
2b-2b-4b	16.4	27.4	12.6	9.1	15.5	0.1	4.0	12.2	10189
<u>2b-2b-2b-2b</u>	19.7	<u>29.5</u>	15.7	13.0	17.7	<u>8.8</u>	8.1	<u>15.6</u>	9827
1b-1b-1b.-1b	2.0	9.4	-2	-4.5	-1	-10.5	-7.0	-1.9	13558

(b) Full search. [-8, 7.5]

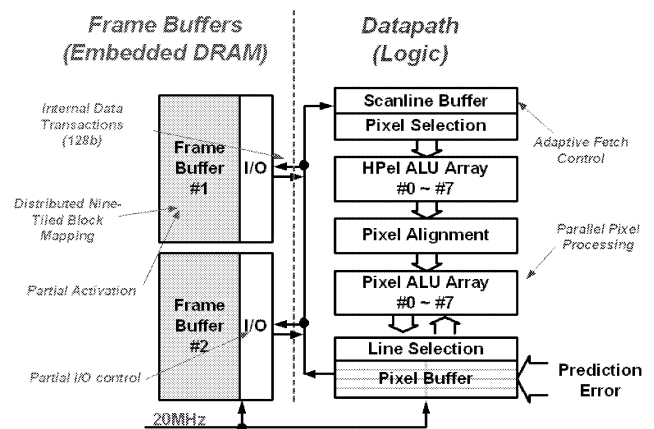
Structure	Akiyo	MissA	Car.	Fore.	Table	Foot	Stefan	Avg. (%)	Pwr * Area
<u>4b-4b</u>	18.6	24.0	16.4	11.6	16.7	3.8	7.0	14.1	<u>9190</u>
4b-2b-2b	20.0	24.0	18.1	16.1	17.5	8.7	9.5	16.3	9207
2b-2b-4b	24.1	31.9	20.6	16.0	19.4	10.5	8.6	18.7	9427
<u>2b-2b-2b-2b</u>	27	<u>34.0</u>	21.3	19.7	21.5	<u>9.1</u>	12.4	<u>20.7</u>	9236
1b-1b-1b.-1b	7.5	12.7	4	0.4	1.4	-8.1	-3.6	2.0	13028

3. Design & Implementation of a Low Power MPEG-4 Video Decoder

3.1 Structure of a Low Power MC Accelerator

The overall structure of the MC accelerator is shown in Fig. 4. It consists of the datapath and two embedded DRAM macros as frame buffers.

The datapath consists of eight processing elements (PEs) which include the half-pel ALUs, the pixel ALUs, the pixel buffers and the shifting logic. Its operation is performed by block (8×8 pixel) granularity, not by macro-block (16×16 pixel) so that it can successfully support advanced block-level prediction features newly provided in MPEG-4 [8]. Since the 128-bit internal bus provides maximum 16 pixels (8 bit/pixel) to the datapath, parallel processing possible in the datapath (8 PEs). Although the length of interconnections in the local datapath area are slightly increased by adopting the parallel architecture, it is possible to lower the operation frequency of the datapath as much as possible (in this design, 20 MHz), and this contributes to reduce the overall power consumption [14]. When adopting wide bus structure, the power consumption of the interconnection may become larger because of its increased overall capacitance; the length may become long according to the physical placement, and the overall coupling capacitance may become larger as the number of interconnections is increased. To prevent this situation, we carefully placed the datapath as close as possible to the frame buffer and widened the wire spaces so that the overall capacitance was maintained to be almost

**Fig. 4** The block diagram of MC accelerator.

the same as that of the narrow bus architecture. In addition, optimization of the bus drivers' size was possible by fully utilizing the relieved speed constraints on the bus so that the power penalty from increased capacitance was successfully compensated.

Finally, aggressive clock gating is adopted in the datapath for the low power consumption.

As the frame buffers, two embedded DRAM macros are integrated with the logic. Each macro can store all of the color information for one video object plane (VOP). Since a MPEG-4 simple profile (SP) is expected to be widely used in most portable video systems, only 2 frames (one for a current frame, and the other for a previous VOP) are stored. The structure of the frame buffer is depicted in Fig. 5 [2]–[4]. It has $512 \text{ bit} \times 128 \text{ row} \times 9 \text{ bank}$, and 128 bit shared I/O. The cell core in each bank is divided into 4 segments. Since

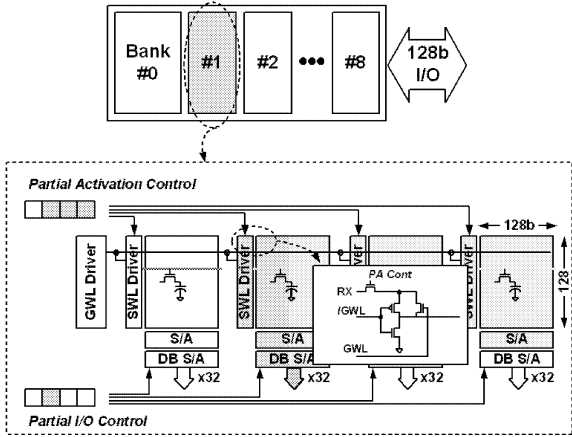


Fig. 5 Structure of the frame buffer.

the segments are driven by the dedicated sub-wordline drivers, the cell core can be partially activated according to the partial activation control (PAC) signals. In addition, partial I/O control scheme is adopted so that data I/Os for each segment can be controlled independently.

The proposed frame buffer structure is tightly coupled with the characteristics of the memory access patterns frequently used in the motion compensation for the low power consumption. Since the video scenes typically used in portable applications have relatively small motion vector displacements, most of their motion vectors are confined in 8×8 integer pixel boundary as shown in Fig. 6(a). Therefore, there exists large spatial locality among the required blocks for the successive frame buffer accesses (Fig. 6(b)).

To effectively utilize these characteristics for the low power consumption, we propose a distributed nine-tiled block mapping (DNTBM). This mapping maximizes the reusability of the previously used block data so that the power consumption for accessing the frame buffer is minimized. Figure 7 explains the mapping. A block in a frame is mapped onto a row whose size is tuned to accommodate a color block ($8 \text{ bit/pixel} \times 64 \text{ pixels/block} = 512 \text{ bit/row}$). It is a natural choice because the data processing is mainly performed by block or macroblock granularity. For a macroblock reconstruction, the required data is distributed in a maximum of nine adjacent blocks. Since a frame consists of 594 blocks (22×18 for luminance blocks, $11 \times 9 \times 2$ for the chrominance blocks) in case of 4:2:0 quarter-common interface format (QCIF), each bank have to store more than 66 blocks ($594/9$). Therefore, the number of rows is set to 128, which is the minimum number among the “power of two” values that is greater than 66. If these blocks are all located in one bank, there must be nine row changes to access all of them. Since all the adjacent blocks are mapped to be located in different banks by the DNTBM, the required blocks can always be controlled independently, and the

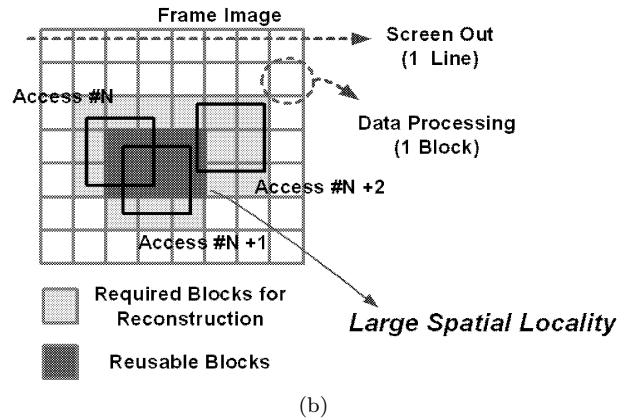
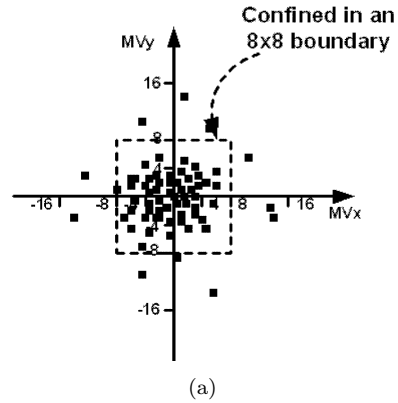


Fig. 6 (a) Distribution of motion vectors. (b) Large spatial locality.

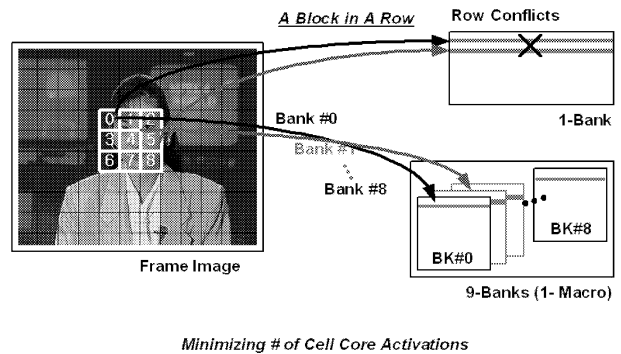


Fig. 7 Distributed nine tiled block mapping (DNTBM).

previously activated rows do not need to be activated. As a result, the number of cell core activations in the frame buffer can be minimized.

Although the tiled mapping method is better for data processing, it is inappropriate for some cases in view of power consumption. For normal pixel reconstruction, only a part of a whole block is used for the processing while the rest are discarded, as shown in Fig. 8(a). The situation worsens when frame buffer to serial access memory (SAM) transfer operation for screen display is performed (Fig. 8(b)). In this case, it is clear that the activation of a whole wordline will

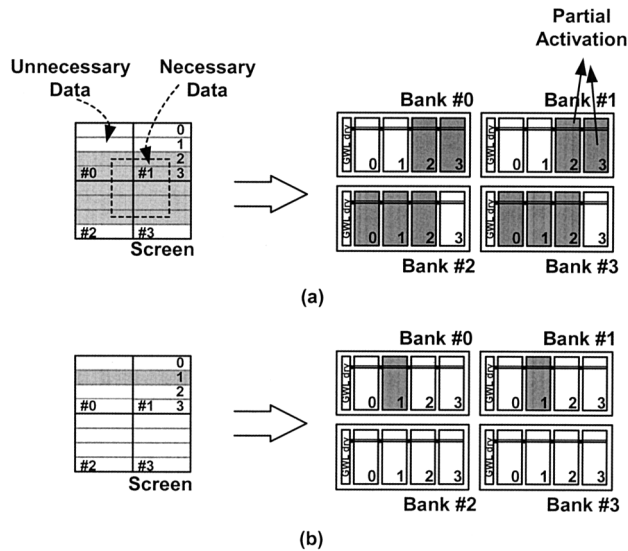


Fig. 8 Partial activation of cell core; (a) Data processing, (b) SAM transfer.

Table 4 Hardware cost.

	1 Seg.	2 Seg.	4 Seg.	8 Seg.
Akiyo	1	0.73	<u>0.69</u>	0.92
MissA	1	0.79	<u>0.75</u>	0.94
Fore	1	0.85	<u>0.78</u>	0.94
Car	1	0.82	<u>0.77</u>	1.03
Table	1	<u>0.81</u>	0.82	1.12

be a waste of power. A sub-wordline scheme with partial activation control (PAC) enables the cell core to be activated by segments. Therefore, it is possible to activate and use only the necessary parts selectively. When dividing the cell core, the trade-off in determining the number of cell cores is required. For example, if we divide the cell core into 8 segments, power can be maximally saved in case of SAM transfer because only the necessary pixels can be selectively activated. But the overall area is increased due to the additional sub-wordline drivers, and this scheme may consume more power than that of one-segment structure due to the additional circuits. The optimization results based on a hardware cost is summarized Table 4. A hardware cost is defined as the product of the power reduction ratio and the area overhead normalized to that of the 1-segment structure. As shown in the table, 4-segment structure is optimal when both the power reduction and the area overhead are considered together.

By adopting these techniques, up to 45% of the required rows were re-used and overall power is reduced up to 31% in accessing frame buffers compared with the conventional 1-bank tiled mapping frame buffer structure (Fig. 9). Most of the power reduction was through

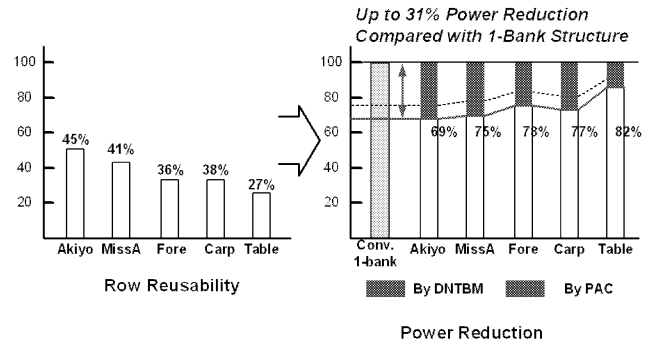


Fig. 9 Power reduction from the proposed techniques.

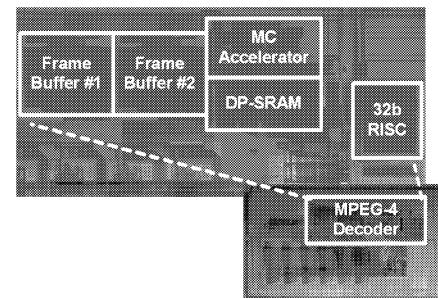


Fig. 10 Die photograph.

Table 5 Implementation results. (Summary)

Technology	0.18um EML technology with 3-poly, 6-metal	
	Logic	DRAM
Area (mm ²)	2.3 mm ²	5.25 mm ²
Power Supply	1.5 V	2.5 V
Clock Frequency	20 MHz	20 MHz
Power	4.6 mW	11.7 mW
DRAM Capacity	1.125 Mbit 2 x (512bit x 128row x 9bank)	
Target Functionality	MPEG-4 SP@L1, H.263, H.263+	

distributed nine tiled block mapping, and about 5 to 10% additional power reduction were achieved from the partial activation control scheme.

3.2 Implementation Results

Using the proposed MC accelerator, an example MPEG-4 video decoding system is implemented [2]–[4]. It is a HW/SW mixed solution that is tuned for MPEG-4 SP@L1 video decoding. 32 bit, 80 MHz RISC processor with 70 MIPs executes all operations except motion compensation, and its results are transferred to internal dual-port SRAM buffer. MC with embedded frame buffer receives the data via 512bit wide bus. Since operations in RISC and MC can be parallel, pipeline architecture with macro-block granularity is possible.

A chip was implemented using $0.18\ \mu\text{m}$ CMOS EML process with 3-poly, 6 metal layers. Figure 10 shows its photograph and Table 5 summarizes its implementation results. Overall power consumption for MC accelerator is 16.2 mW including the power consumption for DRAM frame buffers, which is very low compared with other designs [5]–[7]. The MPEG-4 video decoder is implemented as a part of the low power multimedia processor for portable applications [4].

4. Conclusions

We proposed the low power structure of two hardware blocks (ME and MC with embedded DRAM frame buffer) for the implementation of 2D video processing on portable devices.

First, we proposed the adaptive bit-resolution control (ABRC) scheme for a low power systolic-array type ME. The bit resolution in PE operation is adaptively adjusted according to the operands so that unnecessary transitions in the datapath are minimized. As a result, up to 29.5% of power is reduced compared with the conventional PE designs with negligible area overhead and without any sacrifice of the encoding quality.

In addition, we implemented a low power motion compensation accelerator for portable applications. For low power consumption, two DRAM frame buffers are integrated with the datapath. The architecture of the frame buffer is optimized in terms of low power consumption. In addition, various low power techniques such as distributed nine-tiled block mapping (DNTBM), partial activations control scheme (PAC) are adopted. Using the proposed MC accelerator, MPEG-4 SP@L1 video decoder is fabricated using $0.18\ \mu\text{m}$ EML technology.

References

- [1] G. Weinberger, "The new millennium: Wireless technologies for a truly mobile society," *ISSCC Dig. Tech. Papers*, pp.20–24, Feb. 2000.
- [2] C.-W. Yoon and H.-J. Yoo, "A low power MPEG-4 video codec hardware for portable applications," *Proc. COOLCHIPS V*, pp.77–89, April 2002.
- [3] C.-W. Yoon, J. Kook, R. Woo, S.-J. Lee, K. Lee and H.-J. Yoo, "Low power motion compensation block IP with embedded DRAM macro for portable multimedia applications," *Symp. VLSI Circuits Dig. Tech. Papers*, pp.99–102, June 2001.
- [4] C.-W. Yoon, R. Woo, J. Kook, S.-J. Lee, K. Lee, Y.-D. Bae, I.-C. Park, and H.-J. Yoo, "A 80 MHz/20 MHz multimedia processor integrated with embedded DRAM, MPEG-4 accelerator and 3D rendering engine for mobile applications," *ISSCC Dig. Tech. Papers*, pp.142–143, Feb. 2001.
- [5] T. Hashimoto, et al., "A 90 mW MPEG4 video codec LSI with the capability of core profile," *ISSCC Dig. Tech. Papers*, pp.140–141, Feb. 2001.
- [6] T. Nishikawa, et al., "A 60 MHz 240 mW MPEG-4 videophone LSI with 16 Mb embedded DRAM," *ISSCC Dig. Tech. Papers*, pp.230–231, Feb. 2000.
- [7] S. Kurohmaru, et al., "A MPEG-4 programmable code DSP with an embedded pre/post-processing engine," *CICC Proceedings*, pp.69–72, 1999.
- [8] P. Kuhn, *Algorithms, Complexity Analysis and VLSI Architectures for MPEG-4 Motion Estimation*, pp.8–10, Kluwer Academic, 1999.
- [9] T. Enomoto, "Low power design technology for digital LSIs," *IEICE Trans. Electron.*, vol.E79-C, no.12, pp.1639–1649, Dec. 1996.
- [10] J.-F. Shen, T.-C. Wang, and L.-G. Chen, "A novel low-power full-search block-matching motion-estimation design for H.263+," *IEEE Trans. Circuits Syst. Video Technol.*, vol.11, no.7, pp.890–897, July 2001.
- [11] V.L. Do and K.Y. Yun, "A low-power VLSI architecture for full-search block-matching motion estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol.8, no.4, pp.393–398, Aug. 1998.
- [12] B. Natarajan, V. Bhaskaran, and K. Konstantinides, "Low complexity block-based motion estimation via one-bit transformation," *IEEE Trans. Circuits Syst. Video Technol.*, vol.7, no.4, pp.702–706, Aug. 1997.
- [13] Z.-L. He, C.-Y. Tsui, K.-K. Chan, and M.L. Liou, "Low-power VLSI design for motion estimation using adaptive pixel truncation," *IEEE Trans. Circuits Syst. Video Technol.*, vol.10, no.5, pp.669–678, Aug. 2000.
- [14] A.P. Chandrakasan, S. Sheng, and R.W. Broderson, "Low-power CMOS digital design," *IEEE J. Solid State Circuits*, vol.27, no.4, pp.473–484, April 1992.



Chi-Weon Yoon was born in Pusan, Korea, in 1976. He received the B.S. degree and M.S. degree in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Korea, in 1997 and 1999, respectively. He is currently working toward Ph.D. degree in the same department. His research interests include application specific embedded memory logic design, VLSI architecture for video processing, and mixed mode

VLSI design.



Hoi-Jun Yoo graduated from the Electronic Department of Seoul National University in 1983 and received the M.S. and Ph.D. degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Seoul, in 1985 and 1988, respectively. His Ph.D. work concerned the fabrication process for GaAs vertical optoelectronic integrated circuits. From 1988 to 1990, he was a Visiting Researcher at Bell Communica-

tions Research, Red Bank, NJ, and invented the two-dimensional phase-locked VCSEL array, the front-surface-emitting laser, and the high-speed lateral HBT. In 1991, he became Manager of a DRAM design group at Hyundai Electronics and designed a family of fast-1M DRAM's and synchronous DRAM's including 256M SDRAM. From 1995 to 1997, he was a faculty member of Kangwon National University. In 1998, he joined the faculty of the Department of Electrical Engineering at KAIST and currently leads a project team on RAMP (RAM Processor). In 2001, he founds a national research center, SIPAC (System Integration and IP Authoring Research Center), funded by Korean government to promote word-wide IP authoring and its SOC application. His current interests are SOC design, IP authoring, high-speed and low-power memory circuits and architectures, design of embedded memory logic, optoelectronic integrated circuits, and novel devices and circuits. He is the author of the books DRAM Design (in Korean, 1996) and High Performance DRAM (in Korean, 1999). Dr. Yoo received the 1994 Electronic Industrial Association of Korea Award for his contribution to DRAM technology.