A 95mW MPEG2 MP@HL Motion Estimation Processor Core for Portable High Resolution Video Application

Yuichiro Murachi[†], Tetsuro Matsuno[†], Koji Hamano[†], Junichi Miyakoshi[‡],

Masayuki Miyama[†] and Masahiko Yoshimoto[‡]

† Faculty of Engineering, Kanazawa University
2-40-20 Kodatsuno, Kanazawa, Ishikawa, Japan
TEL (FAX) 076-234-4873   E-mail murachi@mics.ee.t.kanazawa-u.ac.jp
‡Faculty of Engineering, Kobe University.
1-1 Rokkodai, Nada-ku, Kobe, Hyogo, Japan

Abstract

This paper describes a 95mW MPEG2 MP@HL motion estimation processor core for portable and high resolution video application like an HD camcorder. It features a novel hierarchical algorithm and a low power ring-connected systolic array architecture. It supports the frame/field and bi-directional prediction with half-pel precision for 1920×1080@30fps resolution video. The search range is ±128×±64. The ME core integrates 2.25M transistors in 3.1mm×3.1mm using 0.18micron technology.
(Keywords: low power, motion estimation, MPEG2, HDTV and IP)

Introduction

The digital video applications are expanding to include high definition TV (HDTV) resolution, and HDTV resolution monitors becoming more widely used in the home. In this trend, portable HDTV systems will continue to gain popularity. An MPEG2 MP@HL encoder requires about 1000GOPS workload for the HDTV resolution video, assuming the conventional sub-sampling method[1] to perform a motion estimation (ME). In this case, the power consumption of the encoder is more than 2400mW. It is prohibitively large for portable products. The motion estimation algorithm occupies more than 90% of the whole workload for the encoding. Therefore it is seen that a highly efficient ME processor core is essential to realize a portable video application.

This paper describes the first sub-100mW motion estimation core IP (MEH) which realizes MPEG2 MP@HL B-picture processing for portable and high resolution video application like an HD camcorder. The IP core has two major features. The first one is a novel hierarchical search algorithm which employs one dimensional diamond search for upper layer image and narrow full search for lower layer image, allowing a remarkable workload reduction to 7% of that in the sub-sampling method assuming +-128 x +-64 maximum search range. The second one is a newly-introduced low power systolic array architecture which realizes both low memory access cycles and low computation cycles in the full search, which results in almost 65% power reduction compared with the conventional systolic array.

Algorithm

A. Hierarchical Search

It is necessary to decrease computation power with no degradation of picture quality for portable and high resolution video application. The gradient descent search (GDS) method[2] can reduce computation power efficiently, but picture quality is sacrificed. In contrast, fine search like the full search (FS) method provides high picture quality with huge computation power. To satisfy the above two requirements at the same time, the hierarchical search method which consists of the coarsely search and the fine search has been utilized.

The proposed hierarchical search method of this work is illustrated in Fig.1. At first, a motion vector is searched coarsely in ±128×±64 of search area in upper layer image, using a one-dimensional diamond search (1D-DS) described later. The vector to be detected there indicates a start position of the succeeding FS in ±8 x ±8 of search area in lower layer image with integer-pel accuracy, where the start position is located at the center of the lower layer. Finally eight half-pel positions surrounding the integer vector obtained by the FS are evaluated. Thus, a computing power can be decreased keeping a wide search range. In addition, the memory size can be reduced because the image of wide search range is decimated by 1/4 sub-sampling technique.
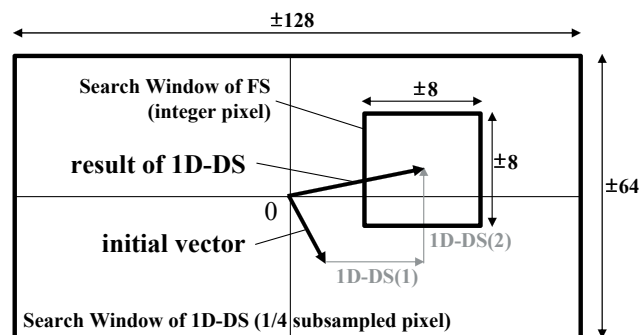

Fig. 1 Proposed Hierarchical Algorithm

B. One-Dimensional Diamond Search

As the first coarse search, we propose the 1D-DS. The 1D-DS shown in Fig.2 is based on a gradient method. At the first step, an initial vector is chosen among the four candidate vectors. At the 2nd step, search direction is decided by calculations of sum of absolute differences (SAD) for the four points surrounding the pixel pointed by the initial vector. This is followed by the 3rd step, in which one dimensional search is executed in the direction toward the point that has the smallest SAD among them. The SADs are computed at several points in the search direction. The point with the minimum distortion in the 1-D search is a temporal solution. Again, new search direction is decided at the temporal solution and new 1-D search begins. The 2nd and 3rd steps are iterated several times until the minimum point is found. Computational complexity of the 1D-DS is quite low because it evaluates only the selected points adaptively.

C. Shared Memory Method

The proposed algorithm also introduces a shared memory method. In this method, the search range is adjusted according to the frame distance. In the case of M=3 and N=15, the frame distance at the P-picture prediction is 3, so the search range has to keep wide. The frame distances at the B-picture are 1 and 2, so the search ranges should be narrow. Thus the two search ranges at the B-picture share a search window memory that also stores image data in the whole search range at the P-picture. Introducing this method reduces memory size and the amount of transferred data for the ME core.
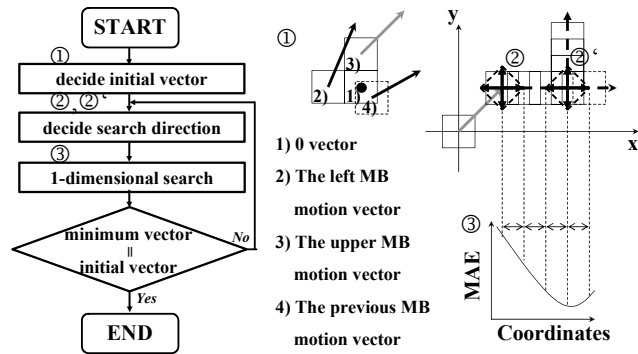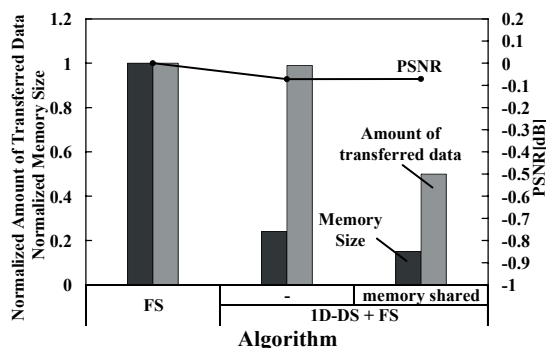

Fig. 2 1D-DS Algorithm

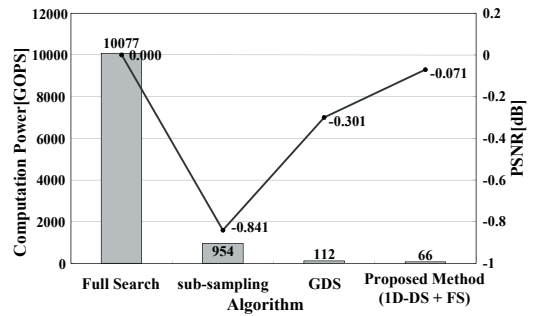Fig. 3 Amount of transferred Data and Memory Size

Fig. 4 Picture Quality and Computational Complexity

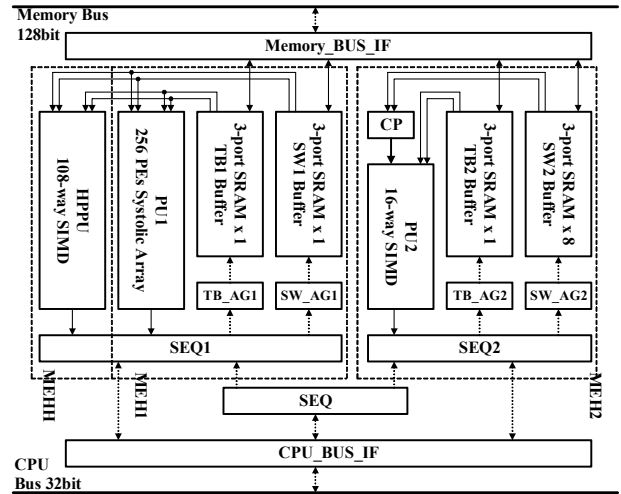Fig. 5 A Block Diagram of MEH

D. Simulation Results

Fig.3 shows the amount of transferred data and memory size of each method. Above two features, the hierarchical search and shared memory method allow not only 50% reduction of the amount of transferred data, but also 85% reduction of memory size. Fig.4 shows comparison of computational complexity and picture quality among several ME algorithms. The computational complexity of the proposed algorithm (1D-DS+FS) is reduced to 7% of the conventional sub-sampling algorithm. It only requires 66GOPS workload. The picture quality of the algorithm exceeds that of the sub-sampling by 0.77dB and the GDS Method by 0.23dB. It is considered that introduction of the FS algorithm for the lower layer image improves the picture quality.

VLSI Core Architecture

A. Multi Processor Configuration

Fig.5 illustrates a block diagram of the MEH. The MEH includes three types of ME processors (MEH2, MEH1 and MEHH). It performs the 1D-DS on the upper layer image with the MEH2, the FS on the lower layer image with the MEH1 and the half-pel search with the MEHH. Fig.6 shows timing chart of MEH. It indicates the MEH2, the MEH1 and the MEHH can operate concurrently. The MEH2 is composed of a 16-way SIMD data path and 3-port (2

read-ports 1 write-port) SRAMs as the search window buffer (SW2) and the template buffer (TB2). The MEH1 introduces a newly developed systolic array architecture[3] constructed by 256 ring-connected processor elements (PEs) and shift resistors (SRs) to perform the FS algorithm. The MEH1 shares the search window buffer (SW1) and the template buffer (TB1) with the half-pel processing unit (HPPU) which is composed of 108-way SIMD data path.

A memory bandwidth of the MEH is 13 Gbps assuming that the memory bus operates at 108 MHz with 128 bit width. The total amount of transmission of MEH1 is 3.5 Gbps; that of MEH2 is 1.5 Gbps. In addition, the total amount of transmission of other modules (DCT, IDCT, etc) is 3.3 Gbps. Therefore, the bandwidth of memory I/F of MEH is fully secured.

B. Systolic Array Architecture

Even with narrow range (±8 x ±8), the full search in HDTV resolution requires a significant larger workload so that it has been very hard to realize low power VLSI circuit. So the important break through of this paper is a novel architecture to implement the FS algorithm for the lower layer with enough low power consumption. The systolic array is illustrated in Fig.7. It consists of the N-by-N PEs (PE-array), the N-by-N SRs (SR-array) and an adder tree.
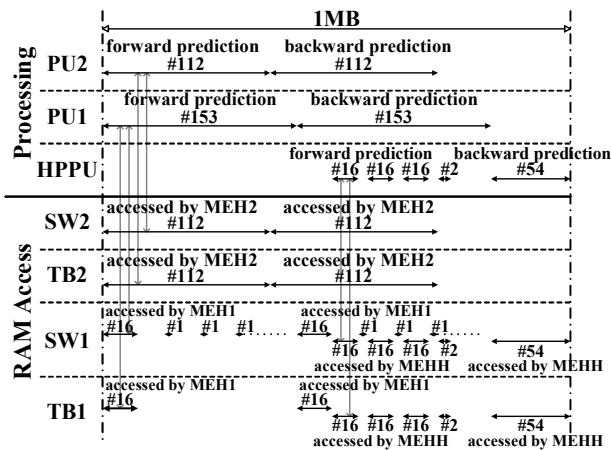
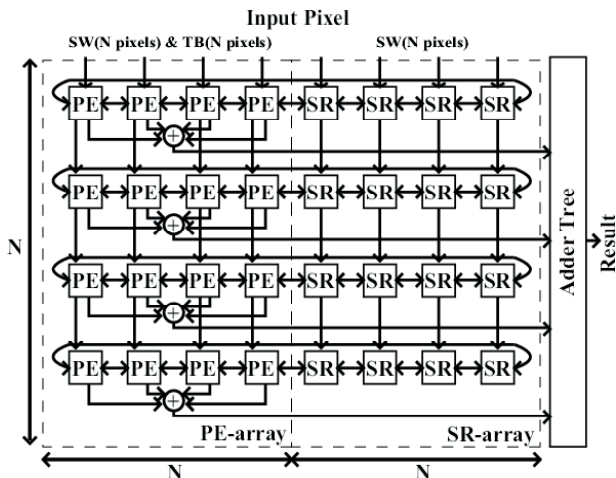The N is equal to 16 in this case; the number of PEs and SRs is 256 each. Sixteen pixels (128bits) in the search window and 16 pixels (128bits) in the template block are loaded into the PE-array from the SW1 and TB1, respectively. The SR-array receives 16 pixels from the SW1. The key feature is that the PEs and SRs in one row are connected with a ring-buffer scheme so that the PE-SR-array can shift pixels toward right direction or left direction.

The data flow is shown in Fig.8. Four phases execute the full search. The first is an initialize phase (Init-phase). Pixels of the template block and the search window are loaded into all PEs and SRs in this phase. The second one is a calculation phase (Calc-phase). The SADs are calculated by the PE array in this phase. The pixels that were loaded into PEs and SRs in the previous phase are shifted to the right side or left side. The third one is an input phase (Input-phase). In this phase, the subsequent pixels are transferred from SW1 to a PE array and SR array. The full-search can be executed by iterating the Calc-phase and the Input-phase in turn after the Init-phase. The Calc-phase and the Input-phase are iterated until reaching a boundary of the search area. The subsequent pixels are loaded by vertical shift operations at that time (Idle-phase). The systolic array receives 16 pixels from SW1 every Input-phase. It is necessary to swaps more significant 128 bits and less significant 128 bits. By using the SW1 constructed by 3-port SRAMs with 2 read-ports, the cross path is eliminated. The pixels can be swapped by controlling the SRAM addresses.

Fig. 6 Timing Chart of MEH
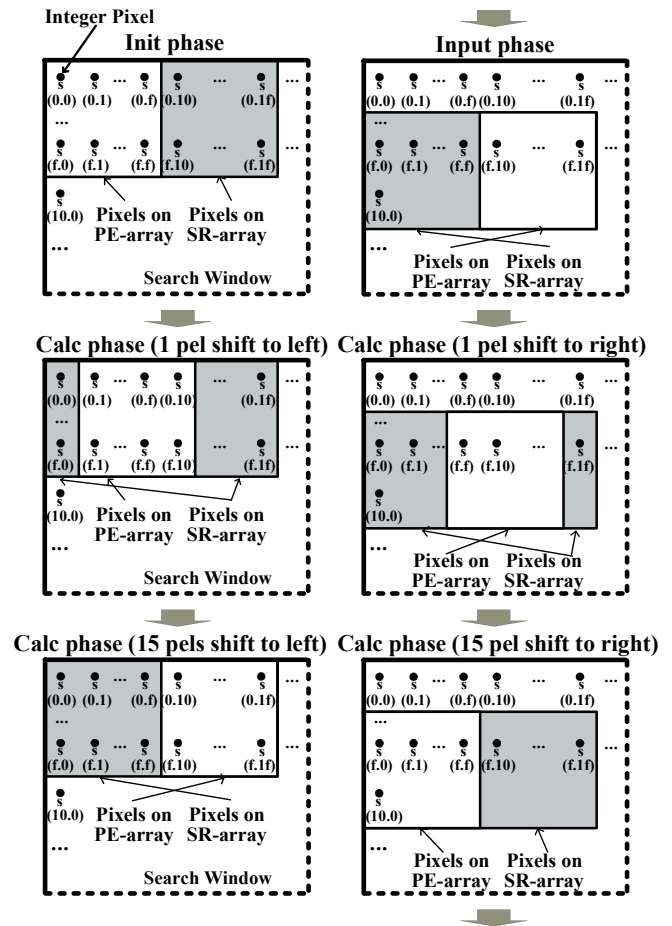

Fig. 7 Block Diagram of Systolic Array (N=4)

Fig. 8 Data flow of Systolic Array

The advantages of the systolic array architecture are as follows. The first one is a reduction of computing cycles. The second one is a reduction of memory access cycles. The third one is adaptability for field/frame mode. This is because the architecture allows various summation schemes in adder tree for absolute difference values. The last one is to enable a concurrent operation of HPPU with no extra cache. HPPU can access the SW1 and the TB1 during the Calc-phase, in which the SW1 and the TB1 are not accessed by the systolic array.

Implementation

Table I shows MEH specification. The MEH has been designed and fabricated with 0.18μm CMOS technology. The MEH integrates 192 signal I/Os and 2.25 million transistors in 3.1mm×3.1mm. It operates at 108 MHz@1V supply voltage. Fig.9 shows a chip photomicrograph of the MEH.

The power consumption is 95mW as shown in Fig.10, which is 3.8% of VLSI1 with the sub-sampling method and 41% of VLSI2 with the GDS technique. The figure also indicates that introducing the proposed systolic array architecture for the FS in the lower layer reduces power consumption of the MEH to 30% as compared with the most basic one-dimensional systolic array[4] (1D-SA0) and 60% as compared with the one-dimensional systolic array using a template-pixel mapping method with broadcasting of search window data[5] (1D-SA1).

TABLE I
MEH Specification

Tech.	0.18μm 6-layer metal
Core Size	3.1mm×3.1mm
Supply	1.0V
Freq.	108MHz
Transistor#	2.25M
Power Consumption	95mW

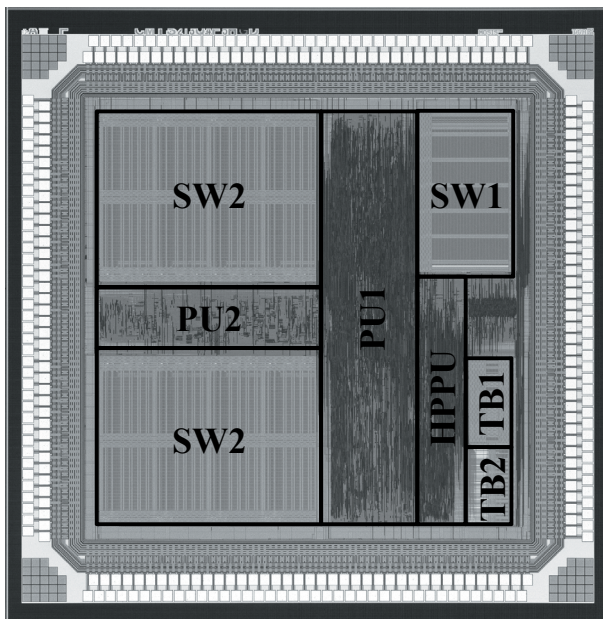

Fig. 9 Chip Photomicrograph
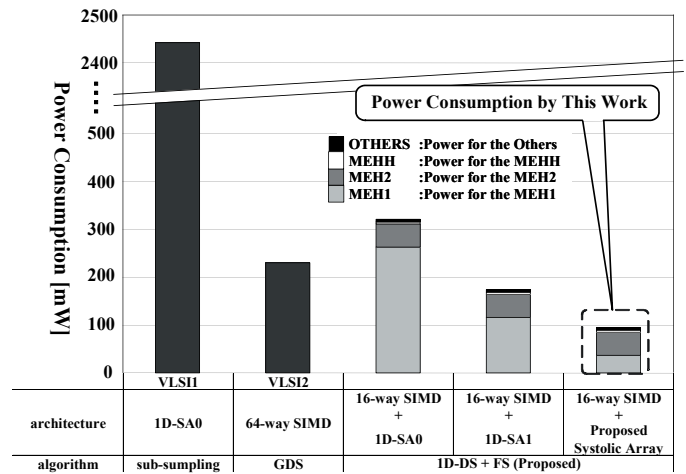

Fig. 10 Power Consumption

Conclusion

This paper presents a 95mW MPEG2 MP@HL motion estimator, which can execute frame/field and bi-directional prediction with half-pel precision using 1920×1080@30fps resolution video. The search range is up to ±128×±64. The MEH is expected to be applicable to SOCs with MPEG codec function for energy aware potable video applications.

Acknowledgments

References

[1] T. Matsumura, et. al., "A single-chip MPEG2 422@ML video, audio, and system encoder with a 162MHz media-processor and dual motion estimation cores," IEICE Trans. Electron., vol.E84-C, pp.202-211, Jan. 2001.

[2] M. Miyama, et. al., "An ultra low power motion estimation processor for MPEG2 HDTV resolution video", IEICE Trans. Electron., pp.561-569, Vol. E86 C, NO. 4 APR. 2003.

[3] Junichi Miyakoshi, et. al., "A low power systolic array architecture for block-matching motion estimation", IEICE Trans. Electron., APR. 2005. (in press)

[4] S. Uramoto, A. Takabatake, M. Suzuki, H. Sakurai, and M. Yoshimoto, "A Half-Pel Precision Motion Estimation Processor for NTSC-Resolution Video", IEICE Trans. Electron., vol.E77 C, no.12 Dec. 1994.

[5] T. Minami, T. Kondo, K. Suguri, and R. Kasai, "A proposal of a one-dimensional array architecture for the full-search block matching algorithm", IEICE Trans., pp.913-925, Vol. J78-D-I, NO. 12 DEC. 1995.