A Sub-mW MPEG-4 Motion Estimation Processor Core for Mobile Video Application

Y.Kuroda, J.Miyakoshi, M.Miyama, K.Imamura, H.Hashimoto, and M.Yoshimoto
Faculty of Engineering, Kanazawa University
2-40-20 Kodatsuno, Kanazawa, 920-8667, Japan

Abstract

This paper describes a sub-mW motion estimation processor core for MPEG-4 video encoding. It features a Gradient Descent Search algorithm whose computation power is only 7% of the conventional 1:4-subsampling search, producing higher picture quality. Another feature is an optimized SIMD datapath architecture to decrease a clock frequency and an operating voltage. It has been fabricated with CMOS 5-metal 0.18 um technology. The measured power consumption to process a QCIF 15 fps video is 0.4 mW under 0.85 MHz@1.0 V.

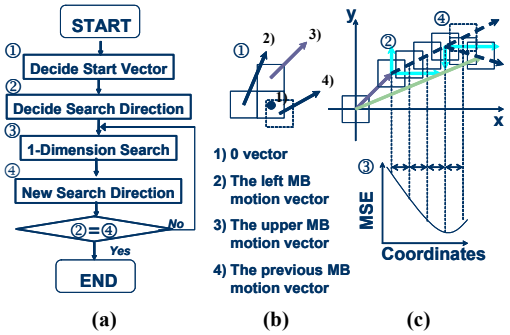
1. INTRODUCTION

A mobile terminal by which people can visually communicate with others continues to gain popularity. To realize an ultra low power and high quality real-time MPEG-4 video codec in the terminal, a highly efficient motion estimation (ME) processor is essential, because the motion estimator with a conventional 1:4-subsampling search using integer-pel and mean absolute difference method (1:4FS) shares more than 70% of the total computational complexity and requires a large amount of computation power, for the MPEG-4 video encoder at QCIF. Power Consumption of a 0.18um motion estimation processor using the conventional method is about 10 mW. This paper describes a sub-mW motion estimation processor core (ME core) for MPEG-4 video encoding.

2. ALGORITHM

2.1. Gradient Descent Search (GDS) Algorithm

The GDS algorithm shown in Fig.1 is a gradient-based method using the steepest decent method. The criterion of a distortional function used in the GDS algorithm is MSE of a Macro Block (MB) indicated by a motion vector. One of the four motion vectors shown in Fig.1 is decided as a start vector. Next a search direction is calculated by differential coefficients of the function at the point indicated by the start vector. The MSEs of motion vectors toward the search direction are calculated step by step in integer pixel width. This search process is called a 1-Dimentional Search (1-DS). The 1-DS is repeated several times to reach minimum distortion. An integer-pel motion vector whose MSE is lowest among search points of the repeated 1-DS is a temporal solution. This is followed by a 3x3 Neighbor Half-pel Search (3x3-NHS), at the position indicated by the temporal solution. Hence, we can reach a final solution with half-pel accuracy. Also, the GDS algorithm includes a hierarchical search to avoid trapping at the local minimum.



(a) (b) (c)

Fig.1. Gradient Descent Search Method

2.2. Optimization for VLSI-implementation

Optimization techniques for VLSI-implementation are described as follows. 1) If the 1-DS search is executed toward arbitrary directions, it requires a quite complex calculation. By limiting the search direction to 8 directions, the calculation becomes more simply, so that the multiplier circuits can be eliminated. 2) Search range should be minimized because Search Window (SW) RAM accounts for a significant amount of power consumption. A simulation result for picture quality indicates that the optimum search range is +-16x16 pixels. Figure 2 shows one of test sequences employed in the simulation. 3) The number of 1-DSs affects on the required computing power. A Simulation result obtained by changing the number indicates that the maximum repeat number of 1-DSs is two.

2.3. Sub Block Search Method

The GDS algorithm was combined with a Sub Block (SB) search method to enhance picture quality. The SB search method is illustrated in Fig.3. First, A MB(16x16 Pixels) indicated by the start vector is divided into 4 SBs(8x8 pixels). Next, the 1-DS for each SB are executed toward a search direction indicated by the differential coefficients. As a result of 1-DS, four SBs as temporal solution indicated by SB vectors are obtained. So four SBs indicated by SB vectors are expanded into MB size in such a way as shown in Fig.3. A final motion vector is decided from 5 vectors obtained by a MB search and four SB searches. The GDS with the SB search decides a motion vector whose MSE is the smallest during the MB search and the SB searches. Therefore, the algorithm always attains higher or equal picture quality comparing with the original algorithm.

Fig.2. Mobile & Calendar

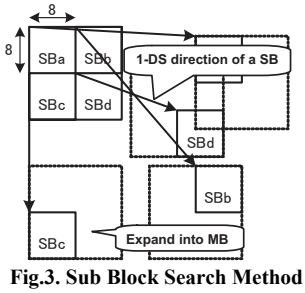


Fig.3. Sub Block Search Method

2.4. Simulation results

The PSNR between the predicted picture and the original picture is measured through simulations. Figure 4 shows a comparison of average PSNR and computing power among various ME algorithm. It indicates that the GDS without the SB search reduces the computing power to 7% and produces higher picture quality comparing with the 1:4-subsampling search. The quality of the algorithm without the SB search is higher than that of the cote method [1], although the computing power is almost equal to that of the cote method. In addition, the GDS with the SB search method attains superior picture quality to the GDS without the SB search.

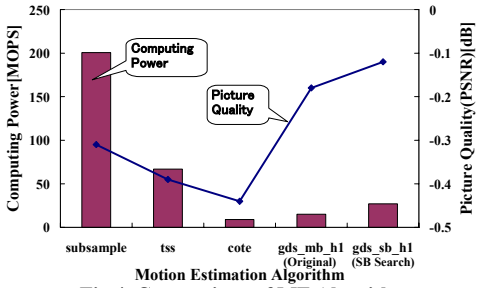


Fig.4. Comparison of ME Algorithms

3. ARCHITECTURE

3.1. SIMD Datapath Architecture

Figure 5 shows the block diagram of the ME core. The ME core is connected to a 32 bit CPU Bus(CPU_BUS) and a 32 bit Memory Bus(MEM_BUS). The SIMD datapath architecture contains 2 Template Buffers(TB), 8 SW Buffers and a Processing Unit(PU). The PU contains 16 Processing Elements(PE), an Adder Tree(AT), and an Accumulator (ACC). The SW Buffer and Template Buffer have 3-port access capability(2read/1write). Memory data mapping is optimized for the SB search. The TBs, SWs and PEs are connected by a Cross Path to sort pixel data. The PE executes a calculation for 1 pixel in one cycle. The PEs are followed by an AT which completes the summation. The control part consists of a sequencer(SEQ), an address generator(AG) and a vector generator(VG). The VG is a circuit to decide a search direction and a motion vector.

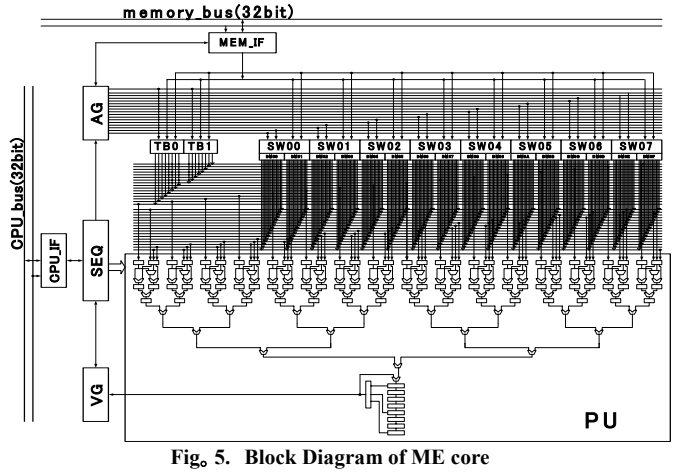


Fig. 5. Block Diagram of ME core

3.2. Processing Element(PE)

A PE was newly developed to perform efficiently the GDS algorithm optimized for the SB search. The PE can calculate MSE and differential coefficients, and it can execute the 3x3-NHS by using Half Pixel Blender(HPB) for both a MB and a SB. The HPB generates half-pel data by filtering operation among integer-pel data. Fig.6 shows the block diagram of the PE.

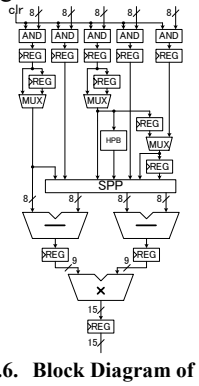


Fig.6. Block Diagram of PE

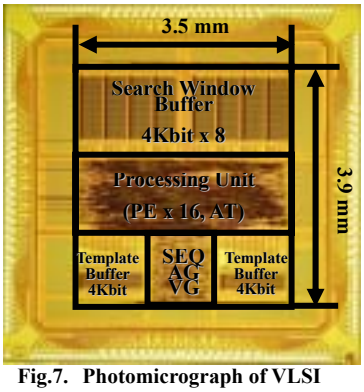


Fig.7. Photomicrograph of VLSI

4. VLSI FABRICATION & EVALUATION

The ME core characteristics are summarized in Table 1. The LSI was fabricated with 5-metal 0.18 um CMOS technology. A photomicrograph of ME core is shown in Fig.7. The ME core incorporates about 1M transistors. Ten pieces of 4Kbit 3-port SRAMs are integrated in the core as Search Window Buffers(SW) and Template Buffers(TB). Table 2 shows the power consumption measured using picture data. The power consumption in the GDS without SB search is 0.4mW under 0.85MHz@1.0V. In the case of the GDS combined with SB search method, the power dissipation is 0.8 mW under 1.70MHz@1.0V. The Shmoo plot obtained by VLSI tester is shown in Fig.8.

Table 1 The ME Core Characteristics

Chip size	**5.9mm x 5.9mm**
Core size	**3.9mm x 3.5mm**
Process technology	**CMOS 0.18um**
Operating voltage	**1.0V**
Number of Transistors	**1M**

Table 2 Power Consumption

With the SB search		
QCIF 15fps	**1.70 MHz**	**0.8mW(@1.0V)**
CIF 30fps	**13.5 MHz**	**7.2mW(@1.2V)**
Without the SB search		
QCIF 15fps	**0.85 MHz**	**0.4mW(@1.0V)**
CIF 30fps	**6.75 MHz**	**2.5mW(@1.0V)**

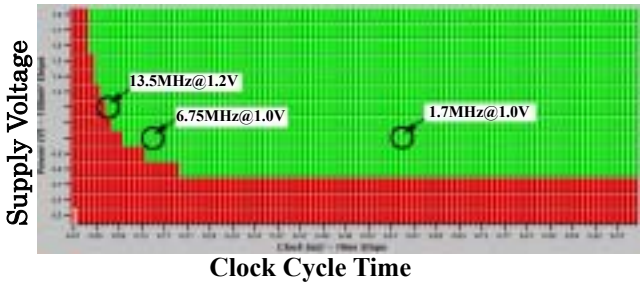


Fig.8. Shmoo Plot of ME core

5. CONCLUSION

A motion estimation processor for MPEG-4 video encoding was developed. The GDS algorithm reduces the computing power approximately to 7%, and produces higher picture quality comparing with the conventional 1:4-subsampling method. The ME core contains 16-way SIMD datapath architecture and 3-port SRAM for highly parallel operation. A clock frequency and an operating voltage were reduced by above techniques.

Hence, the ME core attains ultra low power dissipation less than 1mW at a QCIF 15 fps with high picture quality. Also, the ME core supports wide range of resolution from QCIF 15 fps to CIF 30 fps.

ACKNOWLEDGEMENT

This study was supported by STARC (Semiconductor Technology Academic Research Center) and VDEC (VLSI Design and Education Center. The VLSI chip in this study is designed with Cadence and Synopsys CAD tools.

REFERENCES

[1] P.Kuhn,"Algorithms, Complexity Analysis and VLSI Architectures for MPEG-4 Motion Estimation." Kluwer Academic Publishers, 1999

[2] M.Miyama, J.Miyakoshi, and M.Yoshimoto, et al.,"An Ultra Low Power Motion Estimation Processor for MPEG2 HDTV Resolution Video,"IEICE Trans. Electron., Vol.E86-C, No.4 April 2003

[3]J.Miyakoshi,Y.kuroda,M.Miyama,K.Imamura,H.Hashimoto,and M.Yoshimoto,"A Sub-mW MPEG-4 Motion Estimation Processor Core for Mobile Video Application." in CICC2003 9-3, 2003