

An Ultra Low Power, Realtime MPEG2 MP@HL Motion Estimation Processor Core with SIMD Datapath Architecture Optimized for Gradient Descent Search Algorithm

M.Miyama*, O.Tooyama*, N.Takamatsu*, T.Kodake*, K.Nakamura*, A.Kato*, J.Miyakoshi*, K.Imamura*, S.Komatsu**, M.Yagi***, M.Morimoto***, K.Taki***, and M.Yoshimoto*

*Faculty of Engineering, Kanazawa University

** VLSI Design and Education Center, The University of Tokyo

***Faculty of Engineering, Kobe University

*2-40-20 Kodatsuno, Kanazawa, 920-8667, Japan

Abstract

This paper describes a motion estimation (ME) processor core for realtime, MP@HL video encoding. It is being fabricated with 0.13um CMOS technology and contains approximately 7 M-transistors on 4.50mm x 3.35mm area. The estimated power consumption is less than 100mW at 81MHz@1.0V. It features a Gradient Descent Search (GDS) algorithm that drastically reduces the required computation power to 7GOPS, an optimized SIMD datapath architecture that decreases the clock frequency and the operating voltage, and a low power 3-port data cache SRAM with a write-disturb-free cell array arrangement. The core can be applicable to a portable HDTV codec system.

1. Introduction

A portable HDTV system such as a MPEG camera becomes more popular if people sends and receives a video mail using the broadband network. To realize a high quality and low power MPEG codec in the system, a highly efficient ME processor is essential, because the conventional ME technique requires more than 90% performance of the codec. Figure1 shows the power consumption trend of a ME processor. The power consumption in the 0.13um processor developed with the conventional technology is more than 1200mW even for 1/4 sub-sampling technique. This is too large for portable products.

Several MPEG codec LSIs have been reported [1-3]. These LSIs perform MP@ML video encoding. These cannot perform ME for HDTV resolution video with a single chip configuration.

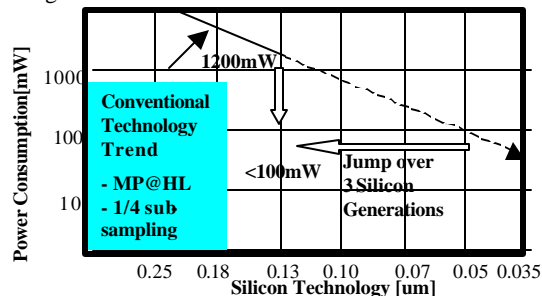


Fig. 1 Power Consumption Trend of ME Processor

Figure2 shows the plot image of the newly developed ME processor. The features are as follows:

- The GDS algorithm [4] is introduced for motion estimation. The GDS algorithm realizes ME for HDTV resolution video only with 7GOPS computation power, though the conventional full search (FS) algorithm requires about 1000GOPS.
- A SIMD datapath architecture optimized for the GDS algorithm is designed. The SIMD datapath contains 32 processing elements (PE). It can calculate mean square error (MSE) in 8 cycles per 1 macro block (MB). Its performance is 10GOPS at 81MHz@1.0V. It operates at low frequency and low voltage, so that its power consumption is quite low.
- The 32Kb 3-port SRAM macro that has a write-disturb-free cell arrangement with a symmetrical memory cell layout is newly introduced. The estimated power consumption of this SRAM macro is 1.32mW at 1.0V and 81MHz.

The above design techniques result in the realization of an ultra low power motion estimator. The characteristics of the motion estimator are as follows:

- It can perform ME for HDTV resolution video in realtime, assuming 1920 x 1080 pixels resolution, 30fps of frame rate, and $\pm 128 \times \pm 64$ of search range.
- The estimated power consumption is less than 100mW at the condition of 1.0V supply voltage and 81MHz clock frequency.

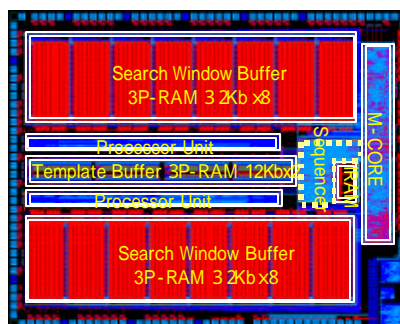


Fig. 2 Plot Image of ME Processor

2. Algorithm

The searching procedure by the GDS algorithm is illustrated in Fig.3(a). The criteria of distortion function used in the GDS algorithm is MSE of a MB indicated by a search vector. The next search starts toward a direction that makes the steepest gradient of the distortion function. The vector whose distortion function is minimum over the search area is a solution of the procedure.

The conventional FS algorithm estimates all vectors in a search range. Then the optimal vector must be found, but the amount of operations is huge. The FS algorithm requires about 1000GOPS to perform ME for HDTV resolution video. So we introduced the new efficient ME algorithm, the GDS algorithm, which requires only 7GOPS as shown in Fig.3(b).

The technical terms are defined here to describe the GDS algorithm. "Template buffer"(TB) is a memory that stores a MB pixel data in a current frame. "Search Window Buffer"(SW) is a memory that stores pixel data in the previous frame. The brightness of the pixel that is located in TB(i,j) is described as TB_{i,j}. The search vector is described as (V_x,V_y). The GDS algorithm is described as follows:

Step.1 Decide start vector

Calculate MSE for the following four vectors. Start searching from the vector that has the smallest MSE among them.

- 1) 0 vector
- 2) The left MB motion vector
- 3) The upper MB motion vector
- 4) The motion vector of the MB that is located in the same position of the previous frame

Here MSE is defined as: $E = \sum_{i,j} (TB_{i,j} - SW_{i+V_x, j+V_y})^2$

Step.2 Decide search direction

Calculate x and y differential coefficients of the distortion function at the point indicated by the start vector.

$$dE/dx = \sum_{i,j} (TB_{i,j} - SW_{i+V_x, j+V_y})(SW_{i+1+V_x, j+V_y} - SW_{i-1+V_x, j+V_y})$$

$$dE/dy = \sum_{i,j} (TB_{i,j} - SW_{i+V_x, j+V_y})(SW_{i+V_x, j+1+V_y} - SW_{i+V_x, j-1+V_y})$$

$$\tan \theta = (dE/dy) / (dE/dx)$$

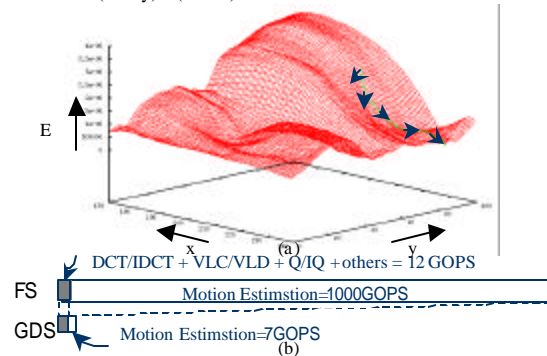


Fig. 3 Gradient Descent Method

Step.3 1-Dimension Search

- Search vectors toward the direction θ with step width Δ .
- Continue to search vectors until MSE increase.
- The vector whose MSE is minimum is a temporary solution.

Step.4 Decide to repeat or not

- Calculate the differential coefficients and new direction θ' at the point obtained in Step.3.
- If θ is not equal to θ' , then go to Step.3, and search in the new direction θ' .
- If θ is equal to θ' , finish the procedure. The latest temporary solution is taken as the final solution..

The simple GDS algorithm as above has a tendency to fall in a local minimum solution. It is effective to add a hierarchical search method in order not to fall in a local minimum, because a smoothing effect in an upper layer removes a noise effect and make it possible to search vectors in a long distance. The hierarchical GDS algorithm is applied for 3 image layers. The layer 1 is the original picture. The layer 2 is obtained by 1/2 sub-sampling of the layer 1, and the layer 3 is obtained by 1/2 sub-sampling of the layer 2. The hierarchical GDS algorithm starts searching in the layer 3. This is followed by searching in the layer 2 and the layer 1.

3. Architecture

Figure4 shows the block diagram of the ME processor. The SIMD datapath optimized for the GDS algorithm contains 1 TB, 16 SWs, 32 PEs, and an adder tree. A MemoryBus feeds new image data to TB and SWs. The TB and SWs are image data caches. The TB, SWs and PEs are connected by 2 CrossPaths. The PE executes a calculation for 1 pixel in a cycle. The PEs are followed by an adder tree which completes the calculation. The control part contains a MCore, an instruction RAM (IRAM), a sequencer (SEQ), and address generators (AG). The features of this architecture are as follows.

- 32 PE SIMD datapath
- Concurrent data transfer
- Adaptive control by MCore

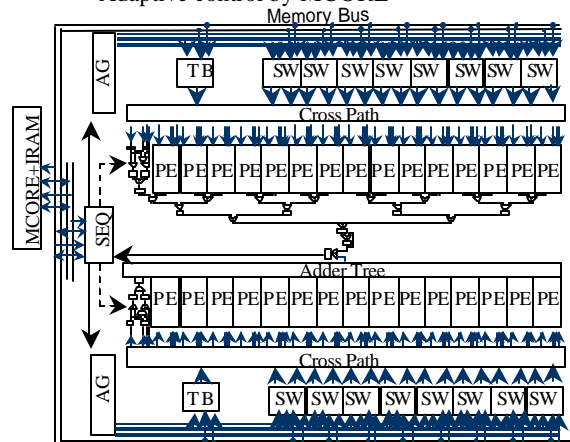


Fig. 4 Block Diagram of ME processor

3.1 32-PE SIMD datapath

The ME chip contains 32 PEs. The PE is optimized for the GDS algorithm. A PE can execute the calculation for MSE and differential coefficients. Figure 5 shows the block diagram of the PE. Figure 5(a) describes the PE configuration for the MSE calculation. The PE receives 1 pixel data from TB and 1 pixel data from SW. Figure 5(b) describes the PE configuration to calculate a differential coefficient in x direction. The PE receives 1 pixel data from TB and 3 pixel data from SWs. The data from S1 terminal is represents a center pixel, and other two data from S0 and S2 are left and right pixels. Figure 5(c) describes the PE configuration for a differential coefficient in y direction. The PE receives 1 pixel data from TB, 2 pixel data from SWs and 1 pixel data from the opposite PE. The data from S3 is a center pixel, and other two data from S1 are upper and lower pixels. Delay buffers are inserted for the delay adjustment for these pixel data. The PEs are followed by adders and accumulators which complete calculation of MSE or differential coefficient for the search vector.

A PE can perform above computation for 1 pixel in a cycle. So 32 PEs can evaluate 1 search vector (or 1 MB) in 8 cycles. A PE can execute 2 subtractions, 1 multiplication and 1 addition simultaneously, so that 32 PEs can execute 128 operations in a cycle. The performance of this SIMD is 10 GOPS at 81 MHz.

3.2 Concurrent data transfer

The TB and SWs receive the next pixel data from the MemoryBus. The TB, SW and PEs are connected by 2 CrossPaths, which sort the sequence of pixel data from SWs corresponding to the sequence of pixel data from TB. The CrossPath is implemented with 1:16 demultiplexers. The TB and SWs feeds pixel data to PEs with the CrossPaths. The MemoryBus and 2 CrossPaths transfer pixel data concurrently. Thus TB and SWs keep supplying pixel data to PEs, so that the pipeline operation of PEs can be kept continuously.

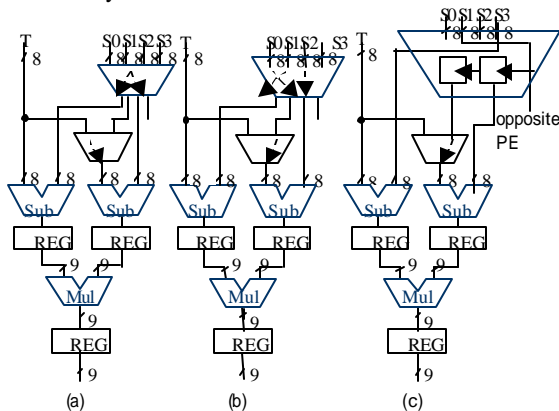


Fig. 5 PE Block Diagram

3.3 Adaptive control by MCore

The control part of the ME processor consists of a MCore, an instruction RAM (IRAM), a sequencer (SEQ), and address generators (AG). The MCore is a RISC processor developed by Motorola for embedded systems. It executes instructions in IRAM. It sets commands and parameters in SEQ registers and can control SIMD datapath indirectly. The command set includes initial vector evaluation, search vector evaluation, differential coefficients calculation, and 1-dimension search. The parameter set includes the number of cycles, the image layer number, the size of TB, the size of SW, initial vectors, search vector, search direction and search width. The SEQ supplies control signals to PEs according to the commands and parameters. It supplies a search vector and start signal to AG. The AG translates a search vector from SEQ to memory addresses and supplies them to TB and SWs.

Adaptive and efficient control of ME processing is realized by a combination of MCore and SEQ. For example, an appropriate search width can be calculated by MCore using MSE and differential coefficients at the start point of 1-dimension search. The MCore makes the control system adaptive. On the other hand, in a 1-dimension search, if MCore accesses SEQ registers every time the evaluation completes, overhead by MCore is not negligible. The SEQ has been designed so as to execute a series of evaluations and stop searching when it finds a temporary solution without control by MCore. The SEQ makes the control system efficient.

4. Circuit Design

The block diagram of 3-port SRAM macro which is utilized for SWs is illustrated in Fig. 6. The macro has concurrent 3-port access capability (2R1W) and a 4Kword by 8bit configuration. The ME integrates 16 pieces of the macro for about 500Kbit storage as SWs. So a low power design for the 3-port SRAM macro is essential for realization of sub-100mW ME LSI.

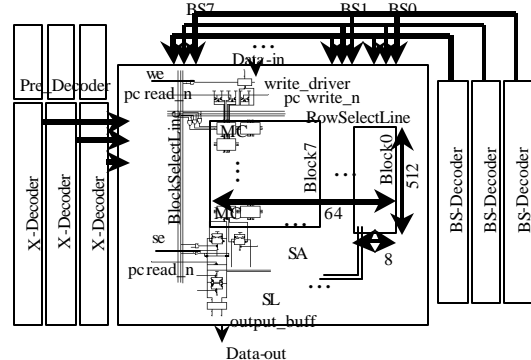


Fig. 6 Block Diagram of 3-port SRAM Macro

The 3-port SRAM macro has three major features to reduce the power dissipation. A symmetric 3-port memory cell layout has been introduced to avoid influence to the cell ratio by misalignment and processing issues. This enhances the cell stability particularly under low voltage condition less than 1V. Also the write-disturb problem which is frequently appeared in the operation of the conventional multi-port RAM is completely eliminated by a newly developed cell-array arrangement. This is realized by a combination of full divided wordline structure for all of 3-port circuit and a wordline scheme which is connected to only 1 row of 8 memory cell (1 pixel). The above two features enables 1V operation allowing low power characteristics. Moreover, the divided wordline structure drastically reduces the bitline current, which occupies a significant amount of total power consumption of the macro, to 1/8. Thus a macro of 4Kx8bit consumes only 1.32mW at 1V and 0.36mW at 0.7V under 81MHz operation condition. Hence, the power dissipation of total search window buffer is suppressed to 25mW under 1V operation, which is about one-third of that using conventional design technique.

5. Estimation

5.1 Performance Estimation

The computation power required by the FS algorithm for HDTV resolution video is shown as follows.

$((1920 \times 1080) / (16 \times 16)) \times (16 \times 16) \times 2 \times (128 \times 64) \times 30 = 1080 \text{ GOPS}$
 The 1920 x 1080 pixels resolution, $\pm 128 \times \pm 64$ of the search range, 2 operations to calculate mean absolute error (MAE) are assumed in the equation.

The computation power required by the GDS algorithm is investigated with the simulation. "Football" is chosen as a sample because it has a large motion displacement and requires high performance to process it. The sample resolution is 720 x 480 pixels, and the frame number is from 1 to 150. The GDS algorithm required only 1 GOPS to process this sample. Then the computation power required by the GDS algorithm for HDTV video is estimated at 6~7 GOPS. As described in section 3.1, the SIMD performance is 10 GOPS at 81 MHz. Then the ME processor can perform ME for HDTV resolution video.

Figure 7 demonstrates the performance of the SIMD datapath compared to the conventional RISC processor. The SIMD datapath reduces the number of cycles from 28806 cycles to 333 cycles per one MB.

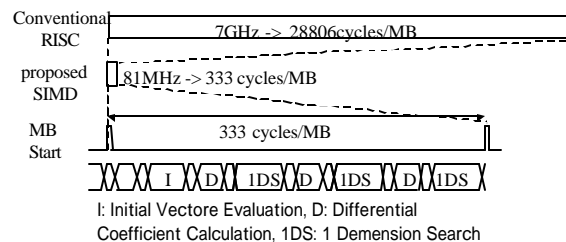


Fig. 7 Performance of SIMD Datapath

Using circuit simulation and static timing analysis, it is verified that the ME processor operates at 81 MHz@0.7V. Then it can operate at 81MHz@1.0V certainly.

5.2 Picture Quality Estimation

The SNR between the original picture and the predicted picture obtained by the GDS algorithm is measured using the simulation. "Football" is applied as a sample again. Figure 8 shows the simulation result. The SNR obtained by the GDS algorithm lowered by 0.3~0.4dB compared to the FS algorithm with the search range ± 32 . The ME processor can execute ME processing with little degradation of video quality.

5.3 Power Consumption Estimation

The power consumption is estimated by circuit simulation. The estimated power consumption of RAM part, standard cell part, and interconnection part is 25mW, 20mW and 20mW respectively under 81MHz@1.0V condition. So less than 100mW of power consumption for the ME core is attained.

6. Conclusion

A motion estimation processor for MP@HL video encoding is newly designed. The estimated power consumption is less than 100mW at 81MHz@1.0V, which is equal to less than 10% of the power dissipation realized by 1/4 sub-sampling technique. This low power characteristic is obtained through the development of the GDS algorithm whose required computation power is about 7% of the FS algorithm and the VLSI architecture at the expense of little degradation of video quality. Consequently it can be applicable to portable HDTV systems.

Acknowledgement

The authors would like to thank STARC (Semiconductor Technology Academic Research Center) for giving us the opportunity to develop the 0.13um LSI. This study has supported by VLSI Design and Education Center(VDEC), the University of Tokyo with the collaboration by Nippon Motorola LTD. The VLSI chip in this study is designed with Avant! CAD tools.

References

- [1]S. Kumaki, et al, "A Single-Chip MPEG2 [422@ML](#) Video, Audio, and System Encoder with a 162MHz MediaProcessor and Dual Motion Estimation Cores," Proc. CICC99, pp.7.2.1-7.2.4, May 1999.
- [2]A. Harada et al., "An Architectural Study of an MPEG-2 [422P@HL](#) Encoder Chip Set" IEICE Trans., vol. E83-A, no.8, pp.1614-1623, Aug. 2000.
- [3]S. Kumaki, et al, "A 99-mm², 0.7-W, Single-chip MPEG-2 [422P@ML](#) Video, Audio, and System Encoder with a 64-Mbit Embedded DRAM for Portable [422P@HL](#) Encoder System," Proc. CICC2001, May 2001.
- [4]M. Takabayashi, et al, "A Fast Motion Vector Detection based on Gradient Method," Technical Report of IEICE, IE2001-74, Sep. 2001.

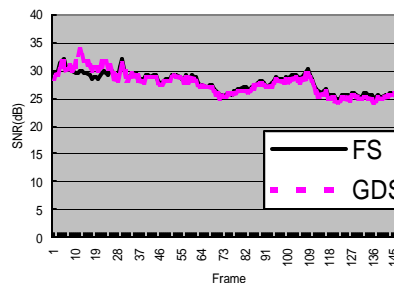


Fig. 8 Picture Quality