

AN AREA/TIME-EFFICIENT MOTION ESTIMATION MICRO CORE*

Sau-Gee Chen
 Department of Electronics Engineering
 National Chiao Tung University
 Hsinchu, Taiwan, ROC

Abstract- *An efficient micro architecture for motion estimation is proposed. It achieves better time and area performance than the existing structures. Through pipelining and effective manipulation of 2's complement arithmetic, the adder complexity is kept to its lowest, while speed for a combined subtraction, absolute and accumulation operations is made as fast as a carry-save addition (CSA). Together with a new DCT algorithm, the micro structure is further expanded and tailored to facilitate efficient execution of other video operations like DCT and filtering operations.*

I. INTRODUCTION

Video encoding demands huge amount of real-time computation, especially for the motion estimation (ME) operation with full search block matching, and the multiplication-intensive discrete cosine transform (DCT) operation. Other multiply-and-accumulate type of operations such as filtering, and matrix operations are also common in general signal processings. For real-time video signal processing such as HDTV encoding/decoding to become a reality, design of efficient architecture for those operations is indispensable.

Among those operations, motion estimation is the most time-consuming one. For instance, assumed that the minimum absolute error (MAE) cost function is used, and a 16x16 block with a block matching of search range from +M to -M in both spatial directions, then there will be total $512 \times (2M + 1)^2$ addition operations for every

motion vector generated. On the other hand, total numbers of multiplication and addition required for a 16x16 block are about 4096, based on direct row-column matrix multiplication. From the given figures we can show that time required for ME rapidly exceeds that of DCT for even moderate M. Note that this comparison does not consider the existence of more efficient DCT algorithms and high memory bandwidth required for ME operations. For more efficient encoding, M has to be as large as possible. Consequently, the ultimate bottleneck for real-time video encoding lies in time-consuming ME operation.

The known architectures for ME in the literature can be divided into two classes. The first class is the ASIC array processors that utilizes the regular and repetitive properties of ME operations. The second class is designed by optimizing data paths between different function blocks of a general-purpose video signal processors (VSP). Here in this paper, we adopt the general-purpose approach for the consideration of programmability and practical implementation.

An ideal ME core is required to efficiently perform various functions, including addition, subtraction, absolute and accumulation operations. Figure 1 depicts a conventional ME core. The structure provides parallel adders for fast absolute, but is area consuming. It requires two selectors, two adders, and two 1's complement circuits. Goto et al. [1] eliminates one of the selectors at the expense of one more adder. Kikuchi et al. [2] provided a relatively efficient hardware for absolute operation only, but still requires two adders and one selector. Moreover, all the mentioned designs do not include the accumulator for total sum of these absolute differences. They did not take advantage of the correlation between the adder for absolute and its accumulator.

* This work is supported in part by National Science Council under Contract NSC82-0404-E009-154, and Computer and Communication Laboratory of ITRI.

Unlike the mentioned general-purpose approach that optimized data paths between dif-

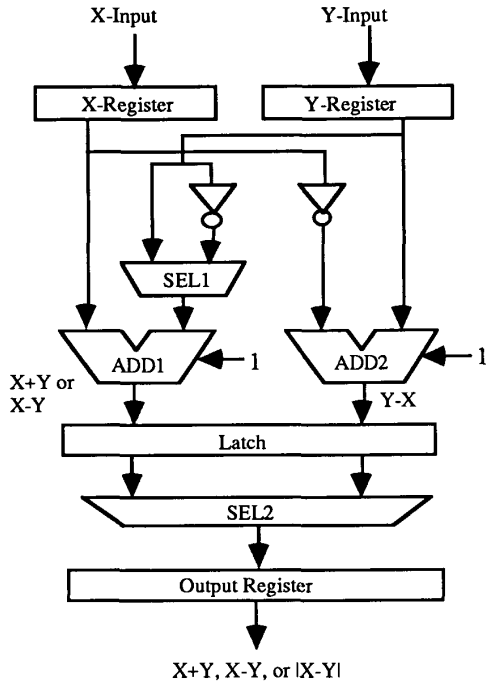


Figure 1. Conventional micro structure for motion estimation.

ferent function blocks, we optimize the ME function block from within the bottom circuit level. Then the optimized design is extended to the higher block-level optimization, by taking into account of other operations like DCT and filtering.

Through pipelining and effective manipulation of 2's complement arithmetic, complexity of the micro architecture is kept to its lowest, while time spent for a combined subtraction, absolute and accumulation operations is made as fast as a CSA (carry-save addition).

The efficient micro architecture for ME (to be detailed later) is small in area and fast in speed, while it is flexible enough to execute various different operations for video signal processings. In particular, we will also present a new matrix-based DCT algorithm, that can be efficiently computed by this hardware. The new DCT algorithm costs only 288 multiplications for an 8x8 subblock.

II. THE ME CORE

Figure 2 shows the data paths of the ME core. The ME core basically consists of two simple adders. The first adder is the ADD module that executes $x - y$ operation for difference of frame displacement (DFD), and $x + y$ operation for motion compensation. Operation $|x - y|$ for MAE is not explicitly computed, but are carried out altogether with the operation

$$\sum_{n=1}^{256} |x_n - y_n|$$

in the second adder module CSACC.

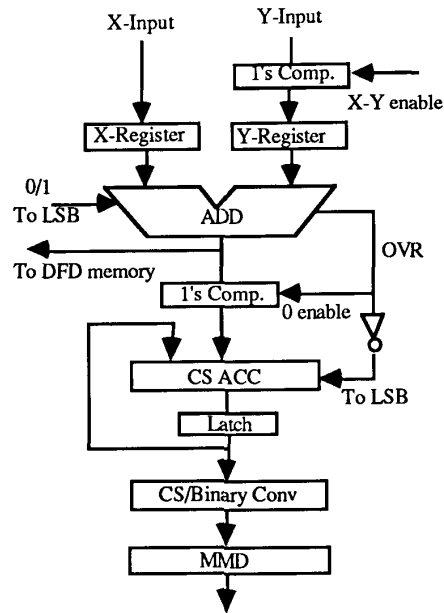


Figure 2. Proposed micro core for motion estimation.

This configuration replaces the slow 2's complement operation for absolute value of $x - y$ with the simple 1's complement operation and addition of 1 to LSB in the CSACC module. Signal OVR is the output carry generated from ADD module, indicating a negative result of $x - y$ if $OVR = 0$, positive otherwise.

Due to repetitive nature of block matching operations, a highly pipelined adder is desirable. Figure 3 shows a fully pipelined parallel adder realization for ADD module. The adder's cycle time can be made as short as a CSA time, or as large as a carry-ripple addition time without re-

sorting to CLA or any other fast addition schemes that are area consuming.

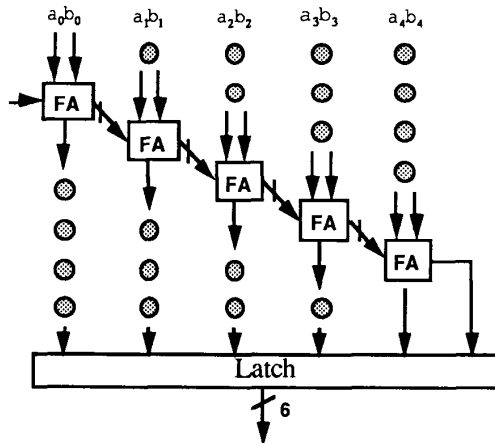


Figure 3. Fully pipelined bit parallel adder with 6-bit wordlength.

Since CSACC module is based on CSA, it is free of carry-propagation delay. As such, time spent for ME is greatly reduced. The accumulator outputs an absolute error every 256 CSA cycles for a 16×16 block matching. Thus, conversion of output from CSA to binary is accomplished in the converter, that is only a simple bit-serial full adder. The conversion is finished in N out of 256 CSA addition cycles for an N -bit binary output. The MMD module updates the minimum absolute error by comparing the currently minimum error with the error of the newly matched block, bit serially starting from LSB. Similarly, MMD module can be realized by a bit-serial full adder.

Assume that one CSA time consumes at most several nano seconds, say 4ns, then a 250MHz clock rate will be required to exploit the speed advantage. Modest process and circuit design technology may not achieve smooth operation under such speed. Alternatively, we can construct the pipelined adder based on a simple m -bit carry-ripple adder (CPA), instead of the 1-bit full adder. The length of CPA is the one that the result CPA time best matches the shortest cycle time the design environment allows and guarantees smooth operation. The benefits of this approach are the reduction of pipeline latches and elimination of a complicated adder.

Table 1 summarizes the occupied area for the proposed and known architectures. As shown, the proposed architecture has the lowest hardware complexity. Moreover, adders of the whole ME core are simple CSA, CPA adders and bit-serial adders.

	Proposed	Conventional	Goto's	Kikuchi's
adders	1	2	3	2
1's compl. cks	2	2	2	2
selectors	0	2	1	1
Functions	$ x-y $ $x \pm y$	$ x-y $ $x \pm y$	$ x-y $ $x \pm y$	$ x-y $

Table 1. Comparison of area complexities of the new structure and existing structures.

III. EXTENDED MICRO CORE FOR VIDEO ENCODING

By taking advantages of the micro architecture, we can duplicate as many modules as possible, subject to design constraints, to facilitate efficient executions of other encoding operations. Figure 4 depicts such a general-purpose VSP structure based on four ME cores. The architecture is tailored to both ME and multiply-and-accumulate (MAC) type of operations such as DCT/IDCT, quantization, and filtering.

Supporting circuitry such as multiplexors are added to program the data paths for different operations. Multiplication operations are carried out in the unit based on the shift-and-add operations with the supports of shifters and Booth's recoding. The architecture executes one MAC operation at a time. Module CSACC2 is for summation of products produced by CSACC1.

Figure 5 shows the data paths of the function unit when it is programmed to ME function, as is highlighted by bold lines. The ME function is to find the minimum of $\sum |X-Y|$. Four ME modules operate independently and simultaneously. In order to execute $|X-Y|$, the control signals S1 and S4 of the first 1'S COMP module are set to "0". Meanwhile, signals S2, S3, S5, and S6 are set to "1" and selected by multiplexers M1, M12, and M13 respectively. On the other hand,

M2, M3, M14, M7, M8, M18, M19 and M15 all select the upper inputs. The carries out from ADD modules are used to control the second 1's complementers. The carry and the sum vectors of CSACC1 are pass through modules SHIFTER and CSACC2 unchanged. Whereas M10, M22, and M23 select the lower input data "0".

When doing DCT, quantization, or filtering operations, all four ME cores are connected together as shown in Figure 6. This configuration supports 16-bit by 16-bit and 24-bit by 24-bit MAC operations. Precisions of results can be as high as 64 bits. The left CSACC1 is cascaded to the right CSACC1 as a 64-bit accumulator. The carry and the sum vectors from CSACC1 are shifted right by 2 bits then accumulated in every cycle. CSACC2s are used to accumulate the results of the multiplication results from CSACC1.

IV. AN EFFICIENT DCT ALGORITHM FOR THE MICRO ARCHITECTURE

An efficient DCT algorithm is proposed here, which can be effectively executed by the micro architecture. Given $N \times N$ 2-D video signal X and its 2-D DCT in matrix form

$$Z = C^T X C$$

where elements of C are defined as

$$c_{k,l} = \sqrt{\frac{2}{N}} \cos\left[\frac{(2k-1)(l-1)\pi}{2N}\right]$$

for $k = 1, 2, \dots, N, l = 2, 3, \dots, N,$

and $c_{k,1} = N^{-1/2}$ for $l = 1$

It can be computed by the well-known row-column approach. It is also widely known that the matrix multiplication $Y = X C$ can be simplified as

$$y_{k,l} = \sum_{m=1}^{N/2} u_{k,m} c_{m,l}$$

for $l = 1, 3, \dots, N-1$, with $u_{k,m} = x_{k,m} + x_{k,N-m+1}$,
and

$$y_{k,l} = \sum_{m=1}^{N/2} v_{k,m} c_{m,l}$$

for $l = 2, 4, \dots, N$, with $v_{k,m} = x_{k,m} - x_{k,N-m+1}$

where $y_{k,l}$ is the element of Y .

As a further step, the two equations can be further decomposed as

$$y_{k,l} = \sum_{m=1}^{N/4} (u_{k,2m-1} + c_{2m,l})(u_{k,2m} + c_{2m-1,l}) - \sum_{m=1}^{N/4} u_{k,2m} u_{k,2m-1} - \sum_{m=1}^{N/4} c_{2m-1,l} c_{2m,l}$$

and

$$y_{k,l} = \sum_{m=1}^{N/4} (v_{k,2m-1} + c_{2m,l})(v_{k,2m} + c_{2m-1,l}) - \sum_{m=1}^{N/4} v_{k,2m} v_{k,2m-1} - \sum_{m=1}^{N/4} c_{2m-1,l} c_{2m,l}$$

for odd and even l respectively. The total number of multiplications required is $O(N^3/4)$. For typical 8×8 block, only 288 multiplications are required. This reduces the time for DCT considerably. Although this algorithm consumes more operations than the existing optimal algorithms, it is free of complicated data routing for butterfly operations like the latter algorithms.

Since DCT is performed either on input video data or on DFD which are integer values, the addition of signal data to DCT constants which are purely fractional, may greatly increase the dynamic range required by hardware when implementing this DCT algorithm. The problem can be easily solved by scaling the input data to be purely fractional. Final results are then scaled up to its correct value.

V. CONCLUSION

In this work, we proposed an micro structure for video encoding, and an efficient regular DCT algorithm for the architecture. Although efficient in its function alone, its overall performance is subject to good designs of the supporting memory systems. This is a crucial issue that is currently under simulation and investigation

REFERENCES

- [1] Goto, et al., "250 MHz BiCMOS Super-High-Speed Video Signal Processor," IEEE JSSC, Dec. 1991, pp. 1876-1884.
- [2] Kikuchi, et al., "A Single-Chip 16-bit 25-ns Real Time Video/Image Signal Processor," IEEE JSSC, Dec. 1989, pp. 1662-1667.

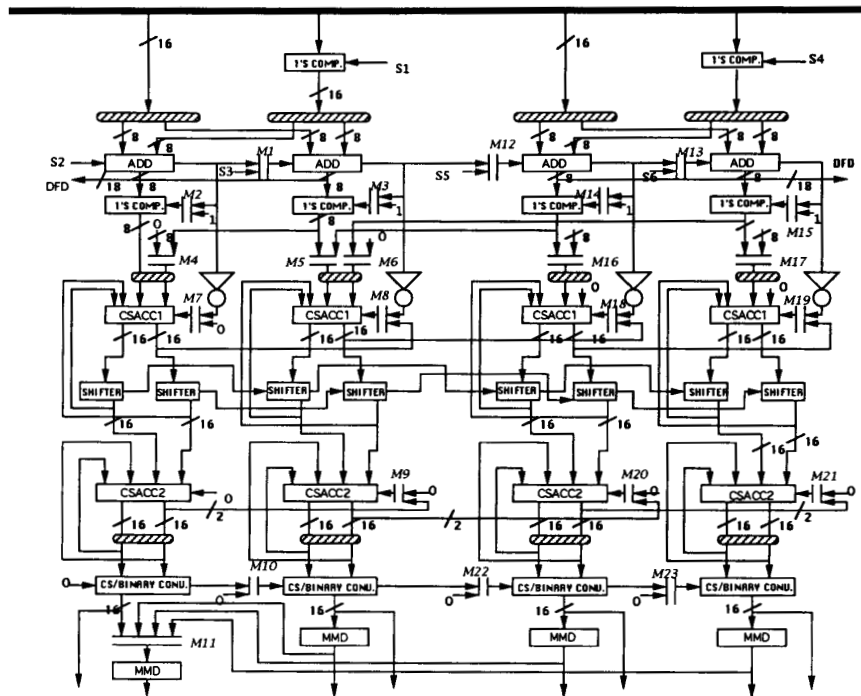


Figure 4. Extended micro architecture for video encoding.

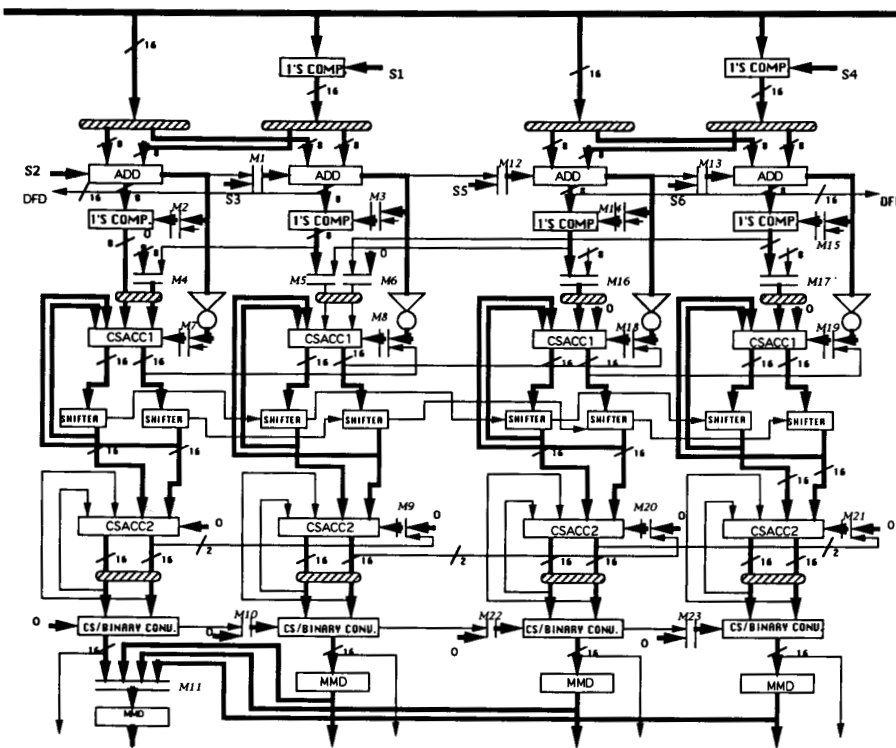


Figure 5. Data path for motion estimation.

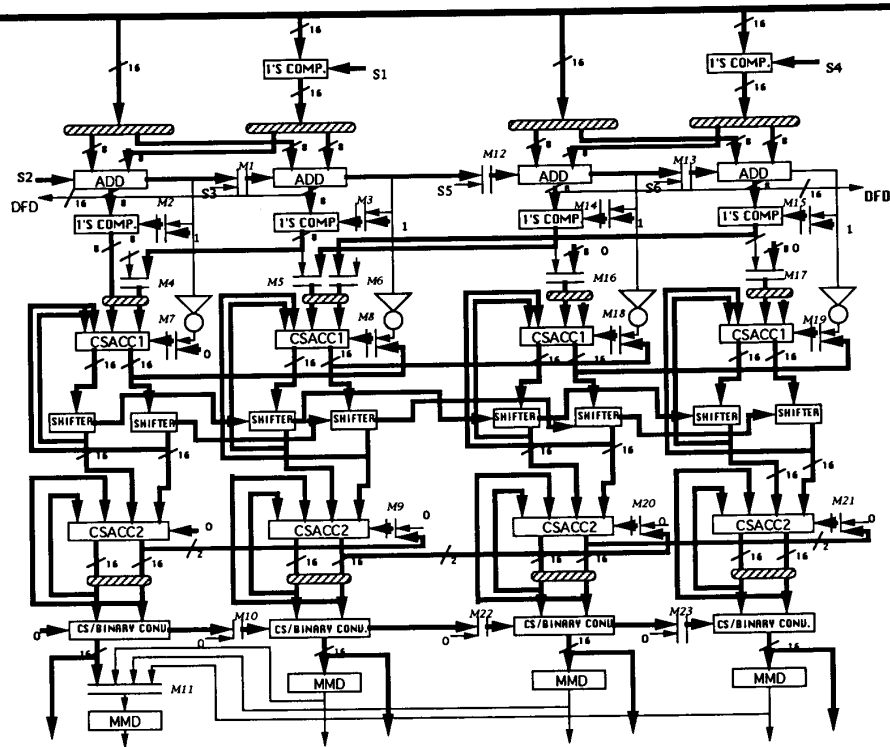


Figure 6. Data path for DCT operation



Sau-Gee Chen received the B.S. degree in Nuclear Engineering from the National Tsing Hua University, Taiwan, in 1978, the M.S. degree and Ph.D. degree in electrical engineering, from the State University of New York at Buffalo, NY, in 1984 and 1988 respectively.

He currently holds a joint appointment as an Associate Professor of the Microelectronics and Information Science Research Center, and the Department of Electronics Engineering, National Chiao Tung University, Taiwan. His research interests include digital arithmetic, digital signal processing, and VLSI system design.