

One-Pass Computation-Aware Motion Estimation With Adaptive Search Strategy

Ching-Yeh Chen, Yu-Wen Huang, Chia-Lin Lee, and Liang-Gee Chen, *Fellow, IEEE*

Abstract—A computation-aware motion estimation algorithm is proposed in this paper. Its goal is to find the best block-matching results in a computation-limited and computation-variant environment. Our algorithm is characterized by a one-pass flow with adaptive search strategy. In the prior scheme, Tsai *et al.* propose that all macroblocks are processed simultaneously, and more computation is allocated to the macroblock with the largest distortion among the entire frame in a step-by-step fashion. This implies that random access of macroblocks is required, and the related information of neighboring macroblocks cannot be used to be prediction. The random access flow requires a huge memory size for all macroblocks to store the up-to-date minimum distortions, best motion vectors, and searching steps. On the contrary, our one-pass flow processes the macroblocks one by one, which can not only significantly reduce the memory size but also effectively utilize the context information of neighboring macroblocks to achieve faster speed and better quality. Moreover, in order to improve the video quality when the computation resource is still sufficient, the search pattern is allowed to adaptively change from diamond search to three step search, and then to full search. Last but not least, traditional block matching speed-up methods are also combined to provide much better computation-distortion curves.

Index Terms—Adaptive search strategy, block matching, computation-aware, motion estimation, one-pass.

I. INTRODUCTION

MOTION ESTIMATION (ME) is the heart of video encoders to remove temporal redundancy within video sequences. The block-matching algorithm (BMA) is adopted by all existing video-coding standards including the H-series [1]–[3] and the MPEG-series [4]–[6]. Among all BMAs, full-search block-matching algorithm (FSBMA) produces the best quality but demands the most computation. Many fast BMAs, such as three-step search (TSS) [7], one-dimensional full search (1DFS) [8], and diamond search (DS) [9], [10], have been proposed to speed up the FSBMA with acceptable loss of video quality or with sacrifice of simplicity and regularity.

Usually, ME is implemented with a hardware accelerator. The rapid improvements in processors and fast BMAs make the software encoder a feasible solution, too. However, when the encoder has to support a wide range of applications (e.g., QCIF (176×144) and CIF (352×288), 15 frames/s (fps) and

30 fps), traditional BMAs will face two problems. First, a traditional BMA stops only when subsequent search points are all examined, and the searching process of a frame cannot be interrupted when the allowed time interval is passed, so real-time constraints may be violated. Second, once the BMA is finished, it cannot be extended when extra computation is still available, so better video quality cannot be achieved.

Recently, the computation-aware (CA) concept is becoming more and more important. In software implementations, processors may have to support video coding of different frame rates, frame sizes, and search ranges. In hardware implementations, even if the frame rate, frame size, and search range have been clearly determined, the computation resource (e.g., operating frequency) may still be adjusted according to the battery power for portable devices. The goal of CA BMAs is to find the best block matching results in a computation-limited and computation-variant environment.

Tsai *et al.* [11] are pioneers of CA BMAs. They contributed a novel scheme, which allocates more computation to the macroblocks (MBs) with the highest distortion in the entire frame step by step. The main concept is that the larger the initial distortion, the more likely the distortion can be significantly reduced, and thus the more computation should be allocated. It is very simple and effective. Nevertheless, there are three problems in their scheme. First, all the MBs are processed at the same time. Thus, random access of MBs is unavoidable, requiring a huge size of extra memory for all MBs to store the up-to-date minimum distortions, best motion vectors (MVs), and searching steps. Second, the related information of neighboring MBs cannot be available to be prediction, and then the search pattern must be determined in advance. The advantage of MV predictors cannot be applied. For example, the predictive diamond search (PDS) [12] outperforms DS in both speed and quality. Moreover, the advantage of adaptive search strategy cannot be applied, either. For instance, PDS is better in small motion cases, but TSS is better in large motion cases. The third problem is the poor hardware feasibility since it was intended for software. The distortion-sorting operations can be easily implemented as hash tables or lists in software, but they are too expensive in hardware. The random access flow and enormous memory size are also harmful for hardware. Even in the software environment, the random access flow will result in a bottleneck of processing speed.

In this paper, a one-pass CA BMA with adaptive search strategy is presented. The ME procedure is done MB by MB to solve the mentioned problems. The rest of this paper is organized as follows. In Section II, the CA concept is first reviewed and discussed. In Section III, our motion analysis is reported.

Manuscript received October 12, 2004; revised October 18, 2005. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Suh-Yin Lee.

The authors are with DSP/IC Design Lab, Graduate Institute of Electronics Engineering and Department of Electrical Engineering, National Taiwan University, Taipei 106, Taiwan, R.O.C. (e-mail: cychen@video.ee.ntu.edu.tw; yuwen@video.ee.ntu.edu.tw; poohli@video.ee.ntu.edu.tw; lgchen@video.ee.ntu.edu.tw).

Digital Object Identifier 10.1109/TMM.2006.876296

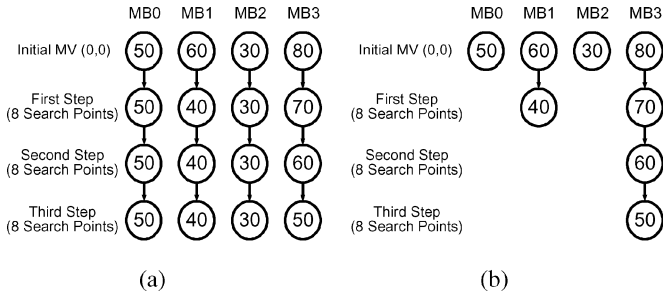


Fig. 1. Examples of the CA concept: (a) original TSS and (b) optimal CA TSS.

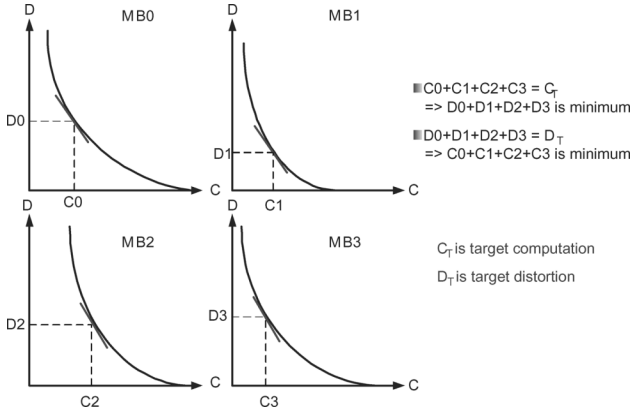


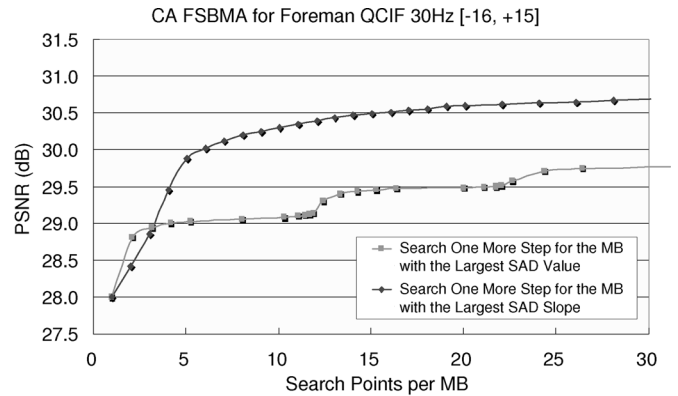
Fig. 2. Computation-distortion optimized truncation by analogy with EBCOT tier two in JPEG 2000.

In Section IV, the proposed algorithm is described according to the analysis. Simulation results are shown in Section V. Finally, Section VI gives a conclusion.

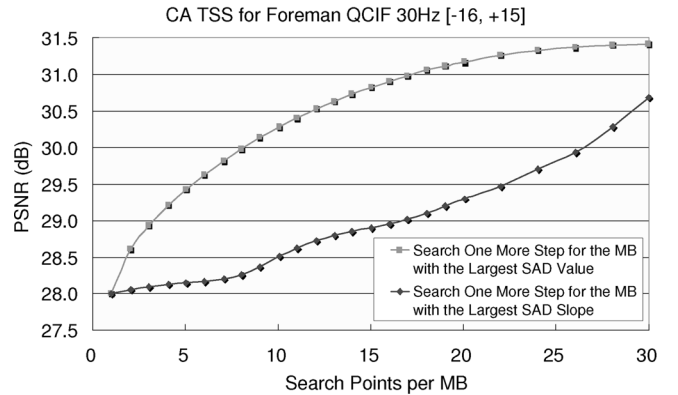
II. CONCEPTS OF COMPUTATION AWARENESS

The CA concept, which was originally proposed in [11], is illustrated in Fig. 1. Assume that there are four MBs in a frame, the number of available search points is 36, and TSS is the search strategy. Fig. 1(a) and (b) shows the original TSS and the optimal CA TSS, respectively. The CA TSS first computes the distortion of the origin ($MV = 0$) for each MB. Afterwards, the MB with the largest distortion is refined by further one step search (eight search points) until the computation resource is exhausted. The distortion of the entire frame for CA TSS is the same as that for the original TSS with much less computation. Different search strategies, such as FSBMA, 1DFS, TSS, and DS have different CA performances. A computation-distortion (C-D) plot can be used to evaluate CA BMAs. On the C-D plot, conventional BMAs are represented as single points, while CA BMAs are expressed as curves, as shown in Fig. 1 of [11].

In fact, the CA scheme considering all MBs of the entire frame with the step-by-step refinement is a little similar to the second tier of embedded block coding with optimized truncation (EBCOT) in JPEG 2000 [13]. Given a search strategy, each MB has an individual C-D curve, as shown in Fig. 2. Assuming the curves are continuous, decreasing, and concave, the optimal decision is obtained when the slopes of tangents at the truncation points for all MBs are the same. Given a target computation, the minimum distortion can be simply found by decreasing the slope until the target computation is reached. On the other hand,



(a)



(b)

Fig. 3. Comparison of two computation allocation methods: (a) CA FSBMA and (b) CA TSS.

given a target distortion, the minimum computation can be allocated in the same way. This implies that spending more computation on the MB with the highest distortion may not always be the best allocation.

Fig. 3(a) shows the comparison between two computation allocation methods applied on CA FSBMA. One is to select the highest distortion, and the other is to select the highest slope. The slope is defined as the amount of reduction in distortion divided by the number of search points in one step. For the initial step, the numerator is changed to the variance of current MB minus the initial distortion. The initial distortion is the sum of absolute differences (SAD) between current MB pixels and candidate MB pixels with $MV = 0$. The variance is the SAD between the average intensity of current MB and each pixel intensity of current MB. It is shown that the method using slopes performs better. However, as shown in Fig. 3(b), when the two methods are applied on CA TSS, the method using slopes becomes worse. This reflects that the assumption of concave is violated under the search pattern of TSS. Therefore, improvement of the search pattern seems to be a more important factor for better C-D performance.

III. MOTION ANALYSIS

In this section, motion analysis is done in four aspects, as described in the following subsections. Four QCIF 30 fps standard video sequences, Foreman, Silent, Stefan, and Weather, will be used in the statistics with search range as $[-16, +15]$.

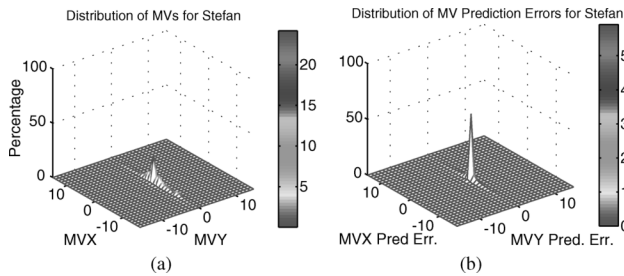


Fig. 4. Statistics of motion for Stefan: (a) MVs and (b) MV prediction errors.

Foreman and Stefan are videos with large motion, while Silent and Weather are videos with small motion.

A. Motion Vector Predictor

MV predictors utilize the spatial correlation of neighboring MBs. Fig. 4(a) and (b) show the distribution of MVs and that of MV prediction errors, respectively. FSBMA and the medium prediction from the left, top, and top right MBs are considered in the statistics. The distribution of MV prediction errors is much more concentrated around the origin than that of MVs, and the peak value at the origin increases from 24% to 59%. Starting from MV predictors makes PDS significantly better than DS in convergence speed and video quality.

Supplementary advantage of MV predictors is to support the rate-distortion optimized mode decision [14], known as Lagrangian method. Not only the distortion but also the MV costs are jointly considered in the mode decision. It is reported that 1-dB PSNR gain can be achieved. However, in our experiments, we only use SAD as the matching criterion for generality because MV costs are dependent on entropy coding and quantization parameters.

B. Different Search Patterns

Different search patterns have different merits and thus should be combined into one CA BMA. Fig. 5 compares FSBMA, TSS, and PDS. For all frames, FSBMA gives the best quality (motion compensated PSNR). On average, PDS is better than TSS. However, when the camera pans very fast, TSS is better than PDS. The results are quite reasonable. When the motion field is small and regular, MV predictor works well, and the diamond pattern can quickly find a good match. As for TSS, the first step search points are dispersed, making final results tend to be trapped in local minima. On the contrary, when the motion field is large and complex, MV predictors do not work well, and the diamond pattern moves slowly toward the best MVs with a high probability of being trapped in local minima. In this case, TSS first glances the entire search area and has better chances to focus on the vicinity of global minimum.

C. PDS Versus FSBMA

When the allocated computation for an MB has not been used up, a CA BMA will continue. However, if the global minimum distortion has been reached, searching more candidates is a waste. Therefore, there should be some detection to check if the optimal MV is reached for early termination of an MB. Thus,

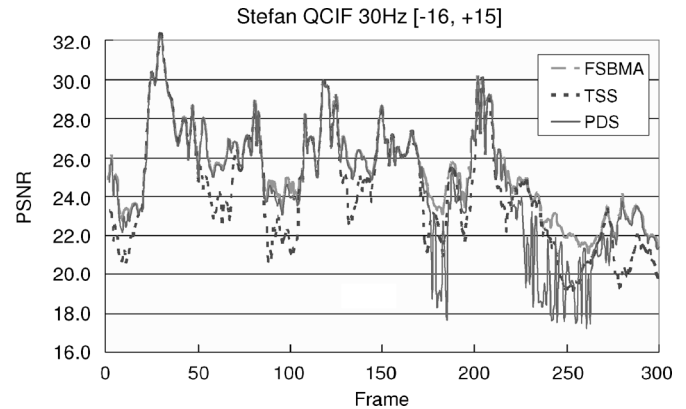


Fig. 5. Comparison of different search patterns.

TABLE I
PERCENTAGES OF IDENTICAL MVs BETWEEN PDS AND FSBMA

Sequence	MVD ≤ 0	MVD ≤ 1	MVD ≤ 2	MVD ≤ 3
Foreman	97.41	94.17	79.84	80.27
Silent	99.51	97.66	92.24	91.30
Stefan	97.11	91.49	80.55	79.87
Weather	99.96	99.31	90.44	96.56

$$\text{MV difference} = \text{MVD} = |\text{MV}_x - \text{MVP}_x| + |\text{MV}_y - \text{MVP}_y|$$

$$\text{MV} = (\text{MV}_x, \text{MV}_y), \text{MV predictor} = \text{MVP} = (\text{MVP}_x, \text{MVP}_y)$$

TABLE II
PERCENTAGES OF IDENTICAL MVs BETWEEN TSS AND FSBMA

Sequence	MV1st=0	MV1st!=0
Foreman	92.76	7.24
Silent	99.17	0.83
Stefan	93.97	6.03
Weather	99.72	0.28

MV1st: best MV after 1st step search

the saved computation can be utilized for later MBs. Table I lists the conditional probabilities of identical MVs between PDS and FSBMA. The smaller the distance from the MV predictor to the final MV, the more likely the global distortion minimum is reached. Therefore, the MV differences (MVDs) defined in Table I can be used to skip BMA operations after PDS.

D. TSS Versus FSBMA

Table II lists the conditional probabilities of identical MVs between TSS and FSBMA. After the first step search, if the best MV is the origin, it is very possible that the optimal MV will be found. Hence, the best MV right after the first step search can be used to stop the BMA operations after TSS.

E. Summary

The motion analysis is summarized as follows.

- MV predictors can be used to achieve faster speed and better quality.
- PDS is suitable for small and regular motion fields.
- TSS is suitable for large and complex motion fields.
- PDS tends to reach the global minimum distortion when the MV predictor is close to the final MV.
- TSS tends to reach the global minimum distortion when the best MV of the first step is the origin.

- Frame layer computation allocation
- Initialize $AccMinSAD$, $DoneMB$, $LeftMB$, and $Left_FM_Tar_SPts_EL$
- Loop MBs
 - Initial block matching (MV=MV Predictor) and find $InitSAD$
 - MB layer computation allocation
 - Block matching motion estimation
 - Adaptive search strategy (PDS, TSS, FSBMA)
 - Terminate when $MB_Actual_SPts \geq MB_Alloc_SPts$
 - Terminate when quasi-optimal MV is found
 - Update $AccMinSAD$, $DoneMB$, $LeftMB$, and $Left_FM_Tar_SPts_EL$

Fig. 6. Macroblock procedure.

- User definition
 - MB_Tar_SPts
 - $MB_Tar_SPts_BL$
- Frame level computation allocation
 - $FM_Tar_SPts = MB_Tar_SPts * TotalMB$
 - $FM_Tar_SPts_BL = MB_Tar_SPts_BL * TotalMB$
 - $FM_Tar_SPts_EL = FM_Tar_SPts - FM_Tar_SPts_BL$
- Macroblock level computation allocation
 - $AvgMinSAD = AccMinSAD / DoneMB$
 - $MB_Alloc_SPts = MB_Tar_SPts_BL + (Left_FM_Tar_SPts_EL / LeftMB) * (InitSAD / AvgMinSAD)$

Fig. 7. Proposed computation allocation.

IV. PROPOSED ALGORITHM

In this section, our one-pass CA BMA will be introduced from top to bottom viewpoints as the following subsections.

A. Macroblock Procedure

Fig. 6 shows the macroblock procedure of our proposed one-pass CA BMA. The one-pass flow denotes that BMA is processed one MB by one MB in the raster scan. Before entering the loop of MBs, frame-level computation allocation and the initialization of variables are required. Inside the loop, the first step is to compute the SAD at the MV predictor, which is the medium of the motion vectors of neighboring MBs and become available because of one-pass scheme with the raster scan, to find $InitSAD$ for MB-layer computation allocation. Then, the proposed adaptive search strategy determines the next search points. As long as the number of actual searched points reaches MB_Alloc_SPts , or the quasi-optimal MV is found, which is judged by the detection of the global minimum distortion, the CA BMA is terminated, and some variables are updated for the next MB.

B. Computation Allocation

For real-time bidirectional communication applications in which low latency is required, ME must be finished in time for every frame, and the frame computation pool must not exceed the reciprocal of frame rate (e.g. 1/15 s for 15 fps videos). Therefore, we focus on the MB-level computation allocation. The frame computation pool is taken as a given parameter.

Fig. 7 is the pseudo code of our computation allocation program. The new concept is to divide the computation resource into a base layer and an enhancement layer. The base layer guarantees the least computation for each MB. The enhancement layer allows each MB to receive additional computation according to the MB-level adjustment and early stop criteria. As shown in Fig. 7, the target search points per MB (MB_Tar_SPts) and that in the base layer ($MB_Tar_SPts_BL$) are user-defined. Afterwards, the frame target search points (FM_Tar_SPts)

and that in the base layer ($FM_Tar_SPts_BL$) can be obtained from multiplying MB_Tar_SPts and $MB_Tar_SPts_BL$, respectively, with the total number of MBs in one frame ($TotalMB$). The frame target search points in the enhancement layer ($FM_Tar_SPts_EL$) is the result of subtracting $FM_Tar_SPts_BL$ from FM_Tar_SPts .

At the MB-level in Fig. 7, the concept of allocating more resources to MBs with larger distortions is still adopted. The average minimum SAD of previous MBs ($AvgMinSAD$) in the current frame is obtained as the accumulated minimum SAD ($AccMinSAD$) divided by the number of processed MBs ($DoneMB$). The allocated search points for an MB (MB_Alloc_SPts) is the base-layer part ($MB_Tar_SPts_BL$) plus the enhancement-layer part which is a product of two items. The first item denotes the future average search points per MB in the enhancement layer, and is the left available computation pool of the enhancement layer ($Left_FM_Tar_SPts_EL$) divided by the number of MBs that have not been processed ($LeftMB$). The second item denotes the ratio of initial distortion of current MB ($InitSAD$) to $AvgMinSAD$.

In short, the base-layer computation is user defined to guarantee the least computation for each MB, and the enhancement-layer computation is in proportional to the ratio of initial SAD to the average SAD of previous MBs to dynamically allocate the computation resources. The SAD slope cannot be applied in one-pass CA BMA since the computation resources of an MB must be allocated before block matching. However, other computation allocation methods still can be tried.

C. Adaptive Search Strategy

Fig. 8(a) illustrates one of our adaptive search strategies. First, PDS is selected as the initial search pattern for an MB. Second, when the PDS ends with available computation left for current MB, the search pattern is switched to TSS. Finally, FSBMA will be adopted if TSS is finished with extra computation resource left. In general, PDS is better than TSS in speed and quality, except for scenes with large and complex motion. In addition, CA DS and CA TSS performs better than CA FSBMA in the C-D plots, as stated in [11]. When the system is relatively abundant in computation resource, FSBMA still can improve the results. Based on the above reasons, we combine the three search strategies in this way.

Fig. 8(b) is the other search strategy, which is modified from the previous one. For large motion sequences, the ‘‘arm’’ of the diamond pattern is not long enough to quickly move toward the global minimum distortion. Hence the initial search pattern may be changed from PDS to TSS, and the PDS is skipped. We use variance of neighboring MVs as the criterion of selecting initial search strategy. The variance of neighboring MVs is defined as the sum of MV distances between each neighboring MV and the medium MV predictor. The neighboring MVs are from left, top, and top right MBs. The computation of medium MV predictor is not an overhead because the initial search point of our algorithm is the MV predictor regardless of the initial search strategy.

Note that because the numbers of search candidates in one step of PDS and TSS are too large for the computation allocation of an MB, we define the processing order of the candidates

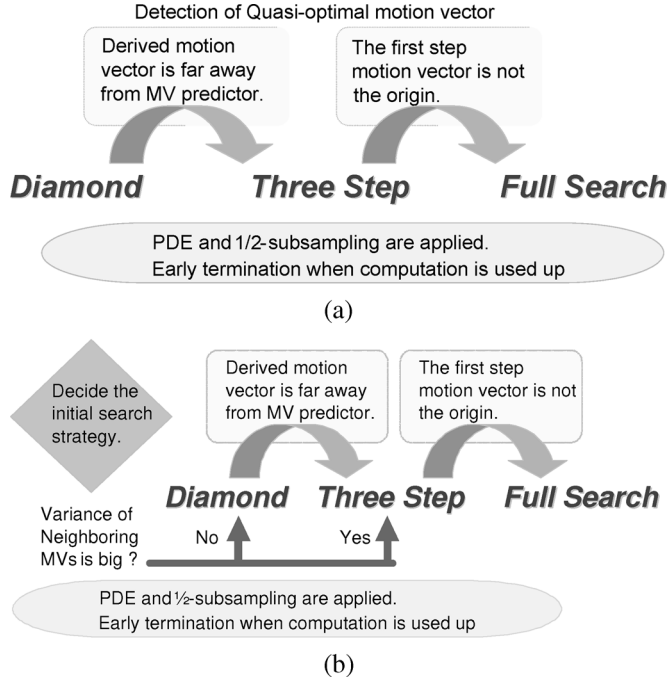


Fig. 8. Proposed adaptive search strategies: (a) strategy 1 and (b) strategy 2.

in one step to provide a fine-grain computation allocation. The left part of Fig. 9(a) shows the flow of PDS which moves the large diamond until the center position of the large diamond has the smallest distortion among nine candidates and further uses the small diamond to refine the result. Therefore, we defined the processing orders in the large and small diamonds of PDS, as shown in the right part of Fig. 9(a). Similarly, we also introduce the procedure and the processing order in each step of TSS in Fig. 9(b) where the left part is an example of TSS and the right part is the corresponding processing order in each step of TSS. The number of steps (NoS) in TSS with searching range, $[-P, P]$, is $\log_2(P + 1)$, the pixel interval of the n -th step is $2^{(NoS-n)}$, and in each step, TSS calculates the distortions of nine candidates and moves the center to the position with the smallest distortion for the next step. As for the FSBMA, we also define a processing order which is like the spiral scan, as shown in Fig. 9(c). Therefore, even if the allocated computation resource is not enough for one step of PDS or TSS or the whole FSBMA, we still can process the candidates based on these defined processing orders until the given computation resource for an MB is consumed.

As the analysis of Section III summarizes, the detection of global minimum is employed. If the final MV of PDS is close to the MV predictor, the final MV of PDS is taken as the quasi-optimal MV, and TSS will not be continued. Similarly, if the best MV of the first step in TSS is the origin, the MV is taken as the quasi-optimal MV, and FSBMA will not be processed. To sum up, PDS, TSS, and FSBMA are selected as search patterns in our CA BMA. For stationary videos, the switching order is PDS followed by TSS and then by FSBMA. For high motion videos, PDS may be skipped. As a matter of fact, other search strategies still can be used to replace the PDS and TSS with corresponding motion analysis to achieve better C-D performances. The main

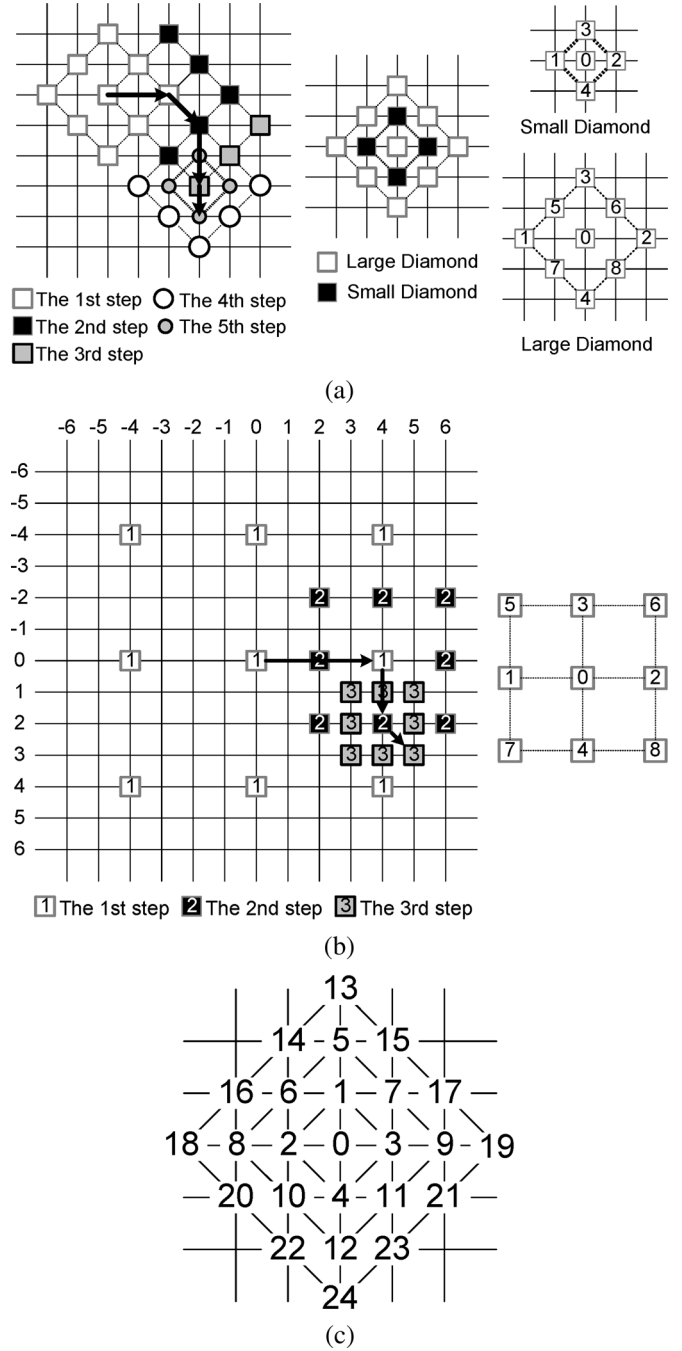


Fig. 9. The search pattern and processing order of three stages. (a) Predictive diamond search. (b) Three-step search. (c) Full search.

purpose of this subsection is to address the advantage of adaptive search strategy, which will be clarified in the Section V.

D. Combination With Traditional Speed-Up Methods

Simplification of matching criterion is often used to speed up the BMA. According to our experiences, 1/2-, 1/4-, and 1/8-sub-sampling cause unnoticed (<0.05 dB), slight (<0.2 dB), and unacceptable (>0.8 dB) video quality losses, respectively. In our experiments, 1/2-sub-sampling is adopted, and the SAD computation only considers half of the MB pixels. Furthermore, partial distortion elimination (PDE) [15], [16] is applied to eliminate redundant SAD calculations. As long as the partial SAD

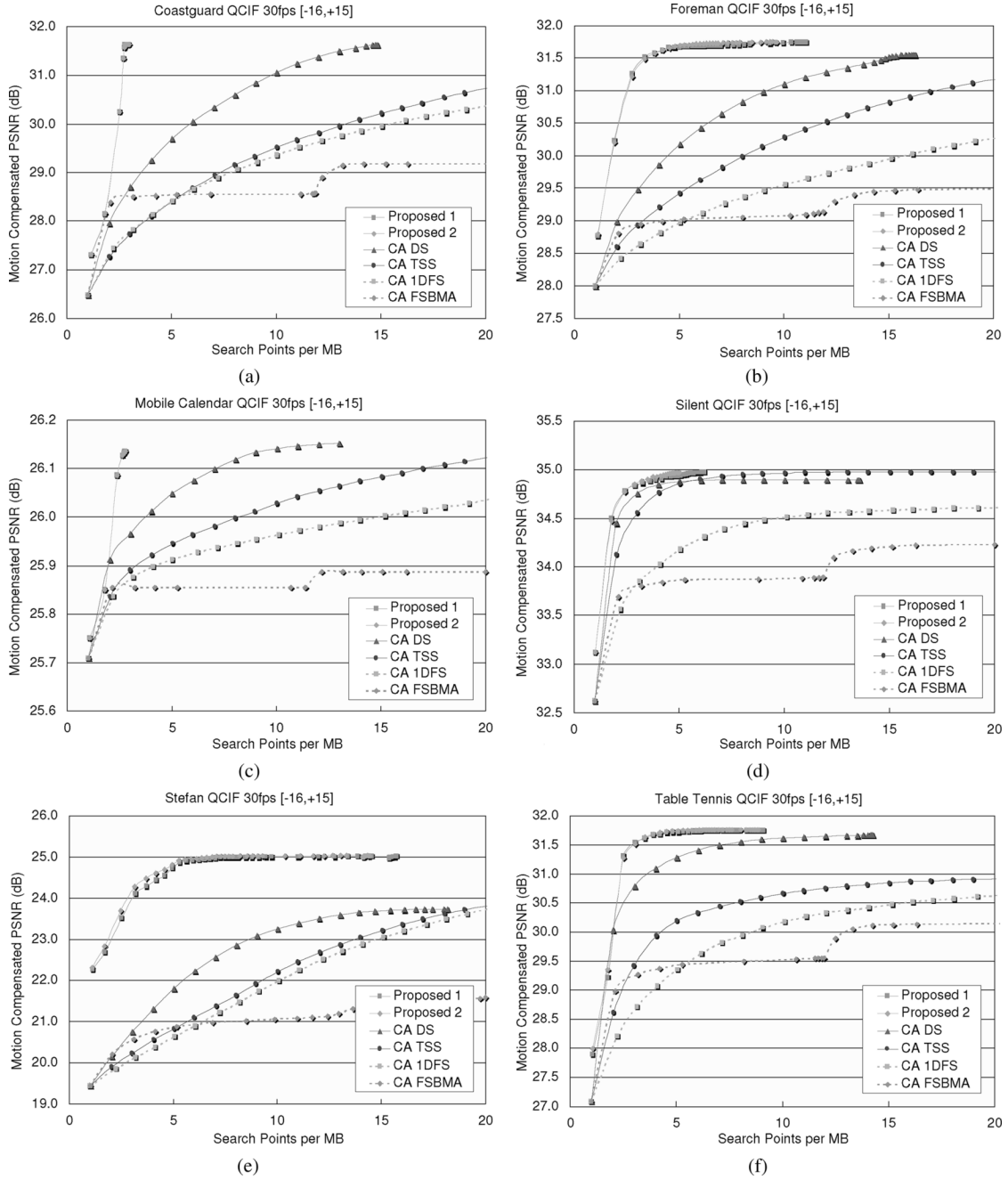


Fig. 10. Comparison of computation-distortion curves between the proposed CA BMAs and the prior CA BMAs: (a) coastguard; (b) Foreman; (c) mobile calendar; (d) silent; (e) Stefan; and (f) table tennis.

of a candidate MB is larger than the up-to-date minimum SAD, the remaining accumulation of pixel differences can be skipped. For the sake of simplicity, we compare the partial SAD with the minimum SAD after every row of pixel differences is generated. Therefore, one search point in the computation resource of our proposed one-pass CA BMA means that 16×2 rows of pixel differences with 1/2-subsampling can be computed.

V. SIMULATION RESULTS

Fig. 10 shows the C-D curves of the proposed algorithms and the CA DS, CA TSS, CA 1DFS, and CA FSBMA stated

in [11]. The “proposed 1” and “proposed 2” denote the search strategies shown in Fig. 8(a) and (b), respectively. Many sequences were tested with the same settings of variables, including the allocated search points for an MB in the base-layer part ($MB_{Tar_SPts_BL}$) and the criteria of quasi-optimal MV, but only Coastguard, Foreman, Mobile Calendar, Silent, Stefan, and Table Tennis are shown due to the limited space and similar trends of C-D curves. The C-D performances of the proposed algorithms are significantly better than those of others. Roughly speaking, the ranking from the best to the worst is “proposed 2”, “proposed 1”, CA DS, CA TSS, CA 1DFS, and CA FSBMA. Most of the time, the two proposed algorithms are

TABLE III
THE ACHIEVED QUALITIES AND REQUIRED COMPUTATIONS OF
CA BMAS AT CONVERGENCE FOR FOREMAN

Algorithm	Achieved Quality (PSNR in dB)	Required Computation (Search Points per MB)
CA FSBMA	31.911	1024
CA 1DFS	31.170	92
CA TSS	31.431	33
CA DS	31.555	16
Proposed 1	31.742	12
Proposed 2	31.748	10

QCIF, 30fps, [-16,+15]

TABLE IV
THE ACHIEVED QUALITIES AND REQUIRED COMPUTATIONS OF
CA BMAS AT CONVERGENCE FOR STEFAN

Algorithm	Achieved Quality (PSNR in dB)	Required Computation (Search Points per MB)
CA FSBMA	25.191	1024
CA 1DFS	25.057	92
CA TSS	24.099	33
CA DS	23.736	18
Proposed 1	25.015	16
Proposed 2	25.036	15

QCIF, 30fps, [-16,+15]

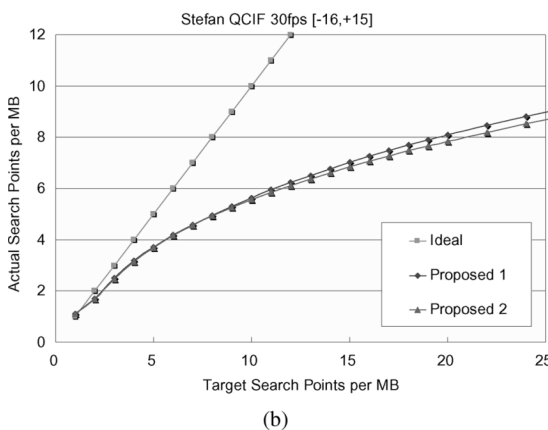
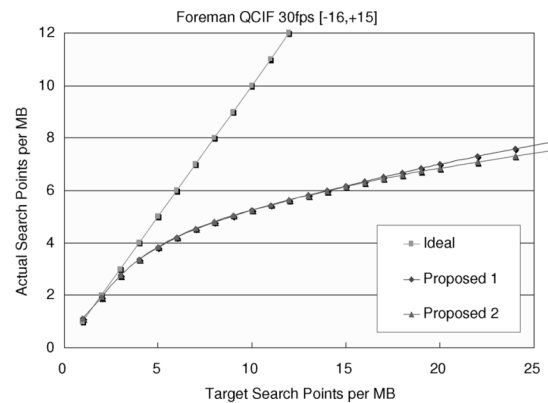


Fig. 11. Capability of the proposed computation control. (a) Foreman. (b) Stefan.

competitively the same in C-D performance, but for high motion video sequences, such as Stefan, “proposed 2” is better due to the proper bypass of the PDS.

The average actually used computation of our algorithms cannot exceed a certain value for each sequence because our

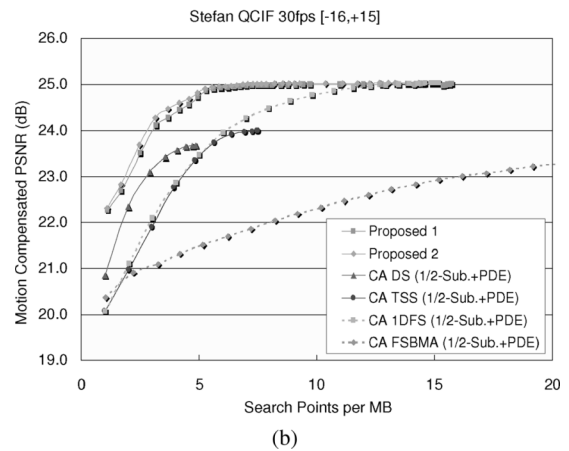
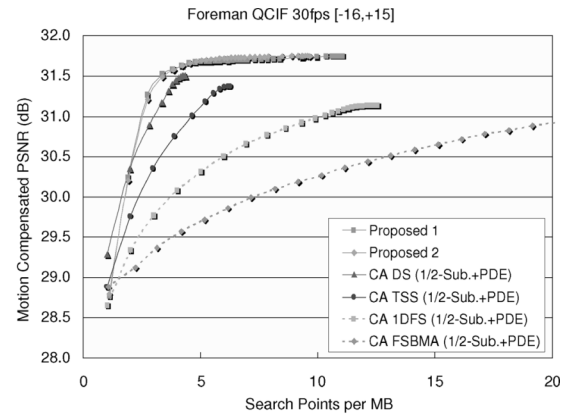


Fig. 12. Comparison of computation-distortion curves between the proposed CA BMAs and the prior CA BMAs with 1/2-sub-sampling and PDE. (a) Foreman. (b) Stefan.

CA BMAs early terminate the operations when detecting that all MBs have reached the optimal MVs. Therefore, further increasing MB_{Tar_Spts} will not increase the actual search points. Furthermore, the best video qualities of our CA BMAs are only 0.1-0.2dB lower than that of CA FSBMA, and is much better than those of remaining CA BMAs. However, this cannot be represented by Fig. 10 because CA FSBMA reaches the best quality with many more search points. Tables III and IV show the best video qualities for Foreman and Stefan, respectively, when the CA BMAs achieve convergence ($MB_{Tar_Spts} = 1024$). The advantage of adaptive search strategy, which can further improve video quality when the computation resource is very abundant, is thus clarified.

Fig. 11 shows the capability of the proposed computation control. The number of actual search points is never larger than that of target search points, which meets the real-time constraints. When the computation resource is little, the available computation will be exhausted. When the computation resource is rich, the resource may not run out due to the detection of global minimum distortion.

In fact, if PDE and 1/2-sub-sampling are applied to [11], our algorithm cannot win so much, and even a small part of the CA DS C-D curve may move to the upper left side of the proposed curves, as shown in Fig. 12. The information of the entire frame is indeed good for computation allocation. However,

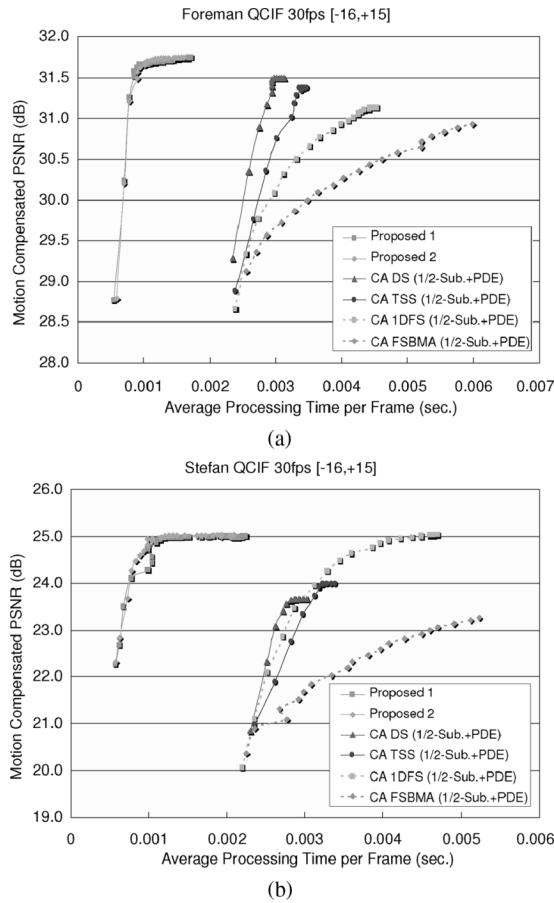


Fig. 13. Use processing time as the unit of computation to compare the computation-distortion curves between the proposed CA BMAs and the prior CA BMAs with 1/2-subsampling and PDE on a PC platform with a 2.5 GHz CPU. (a) Foreman. (b) Stefan.

only our one-pass method can be benefited from Lagrangian mode decision, which enhances a lot of quality. Our strength also includes high hardware feasibility and much less memory requirement.

In the above of this paper, we use “search points per MB” as the unit of computation to make the results independent of different platforms. Now we change the unit of computation to “processing time” in order to be more practical. The simulation platform is a PC with Intel Pentium IV 2.5-GHz CPU and 333-MHz 1-GB DDR DRAM running Microsoft Windows 2000. The program is written in C language. The more realistic C-D curves are drawn in Fig. 13, which uses “processing time” as the horizontal axis. Please note that the prior CA BMAs are all improved with 1/2-subsampling and PDE for fairness. It is shown that the C-D performance of our one-pass flow becomes significantly better than those of prior CA BMAs, in contrast to the C-D curves shown in Fig. 12 where “search points per MB” is the unit of the horizontal axis. The main reason is that prior CA BMAs requires a huge size of memory in proportional to the frame size while our CA BMAs need much less memory to store information of one MB. The cache miss rate of the prior random access flow is very high, degrading the system performance considerably.

VI. CONCLUSION

In this paper, we presented a computation-aware motion estimation which has four features. The first one, the main idea, is to convert the processing flow from random access to one-pass for hardware feasibility. The second feature is to divide the computation allocation into two part, the base layer and the enhancement layer. The former is to guarantee the least computation for each MB, and the latter is to dynamically allocate the computation resources for some MBs with the larger distortions. Thirdly, because of our one-pass scheme, adaptive search strategy and motion vector predictors can be utilized for the faster speed and better quality. Finally, the detection of global minimum distortion is proposed, and traditional speed-up methods are also applied to early stop the unnecessary computation. Simulation results show that the provided computation-distortion performance is relatively better.

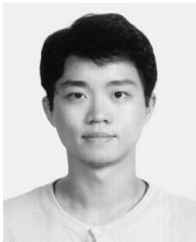
REFERENCES

- [1] *Video Codec for Audiovisual Services at $p \times 64$ Kbit/s*, ITU-T Rec. H.261, Mar. 1993.
- [2] *Video Coding for Low Bit Rate Communication*, ITU-T Rec. H.263, Feb. 1998.
- [3] *Draft ITU-T Rec. and Final Draft International Standard of Joint Video Specification*, ITU-T Rec. H.264 and ISO/IEC 14496-10 AVC, Joint Video Team, May 2003.
- [4] *Information Technology—Coding of Moving Pictures and Associated Audio for Digital Storage Media at Up to About 1.5 Mbit/s—Part 2: Video*, ISO/IEC 11172-2, 1993.
- [5] *Information Technology—Generic Coding of Moving Pictures and Associated Audio Information: Video*, ISO/IEC 13818-2 and ITU-T Rec. H.262, 1996.
- [6] *Information Technology—Coding of Audio-Visual Objects—Part 2: Visual*, ISO/IEC 14496-2, 1999.
- [7] T. Koga, K. Iinuma, A. Hirano, Y. Iijima, and T. Ishiguro, “Motion compensated interframe coding for video conferencing,” in *Proc. Nat. Telecommunications Conf.*, 1981, pp. C9.6.1–C9.6.5.
- [8] M. J. Chen, L. G. Chen, and T. D. Chiueh, “One-dimensional full search motion estimation algorithm for video coding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, no. 5, pp. 504–509, Jun, 1994.
- [9] S. Zhu and K. K. Ma, “A new diamond search algorithm for fast block matching motion estimation,” in *Proc. IEEE Int. Conf. Image Processing (ICIP’97)*, 1997, pp. 292–296.
- [10] J. Y. Tham, S. Ranganath, M. Ranganath, and A. A. Kassim, “A novel unrestricted center-biased diamond search algorithm for block motion estimation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 4, pp. 369–377, Aug. 1998.
- [11] P. L. Tsai, S. Y. Huang, C. T. Liu, and J. S. Wang, “Computation-aware scheme for software-based block motion estimation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 9, pp. 901–913, Sep. 2003.
- [12] A. M. Tourapis, O. C. Au, and M. L. Liu, “Highly efficient predictive zonal algorithms for fast block-matching motion estimation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 10, pp. 934–947, Oct. 2002.
- [13] *Information Technology—JPEG 2000 Image Coding System—Part 1: Core Coding System*, ISO/IEC JTC1/SC29/WG1, 2000.
- [14] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, “Rate-constrained coder control and comparison of video coding standards,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 688–703, Jul. 2003.
- [15] , ITU-T Rec. H.263 software implementation, Digital Video Coding Group, 1995, Telenor R&D.
- [16] C. K. Cheung and L. M. Po, “Normalized partial distortion search algorithm for block motion estimation,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 3, pp. 417–422, Apr. 2000.



Ching-Yeh Chen was born in Taipei, Taiwan, R.O.C., in 1980. He received the B.S. degree from the Department of Electrical Engineering, National Taiwan University (NTU), Taipei, Taiwan, R.O.C., in 2002. He is currently pursuing the Ph.D. degree at the Graduate Institute of Electronics Engineering, NTU.

His research interests include intelligent video signal processing, motion estimation, scalable video coding, and associated VLSI architectures.



Yu-Wen Huang was born in Kaohsiung, Taiwan, R.O.C., in 1978. He received the B.S. degree in electrical engineering and the Ph.D. degree from the Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan, in 2000 and 2004, respectively.

He joined MediaTek, Inc., Hsinchu, Taiwan, in 2004, where he develops integrated circuits related to video coding systems. His research interests include video segmentation, moving-object detection and tracking, intelligent video coding technology,

motion estimation, face detection and recognition, H.264/AVC video coding, and associated VLSI architectures.



Chia-Lin Lee was born in Taipei, Taiwan, R.O.C., in 1980. She received the B.S. degree from the Department of Electrical Engineering, National Taiwan University (NTU), Taipei, Taiwan, R.O.C., in 2005. She currently is pursuing the Master's degree at the Graduate Institute of Electronics Engineering, NTU.

Her research interests include computation-aware motion estimation and associated VLSI architectures.



Liang-Gee Chen (S'84–M'86–SM'94–F'01) was born in Yun-Lin, Taiwan, R.O.C., in 1956. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from National Cheng Kung University, Tainan, Taiwan, in 1979, 1981, and 1986, respectively.

He was an Instructor (1981–1986), and an Associate Professor (1986–1988) in the Department of Electrical Engineering, National Cheng Kung University. While in the military service during 1987–1988, he was an Associate Professor in the

Institute of Resource Management, Defense Management College. In 1988, he joined the Department of Electrical Engineering, National Taiwan University (NTU), Taipei, Taiwan. During 1993–1994, he was a Visiting Consultant with the Digital Signal Processing (DSP) Research Department, AT&T Bell Labs, Murray Hill, NJ. In 1997, he was a Visiting Scholar with the Department of Electrical Engineering, University of Washington, Seattle. During 2001 to 2004, he was the first Director of the Graduate Institute of Electronics Engineering (GIEE), NTU. Currently, he is a Professor with the Department of Electrical Engineering and GIEE at NTU. He is also the Director of the Electronics Research and Service Organization, Industrial Technology Research Institute, Hsinchu, Taiwan. His current research interests are DSP architecture design, video processor design, and video coding systems.

Dr. Chen has served as an Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS for Video Technology since 1996, as Associate Editor of IEEE TRANSACTIONS ON VLSI SYSTEMS since 1999, and as Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—II since 2000. He has been the Associate Editor of the *Journal of Circuits, Systems, and Signal Processing* since 1999, and a Guest Editor for the *Journal of Video Signal Processing Systems*. He is also an Associate Editor of the PROCEEDINGS OF THE IEEE. He was the General Chairman of the 7th VLSI Design/CAD Symposium in 1995 and of the 1999 IEEE Workshop on Signal Processing Systems: Design and Implementation. He is the Past-Chair of Taipei Chapter of IEEE Circuits and Systems (CAS) Society, and is a member of the IEEE CAS Technical Committee of VLSI Systems and Applications, the Technical Committee of Visual Signal Processing and Communications, and the IEEE Signal Processing Technical Committee of Design and Implementation of Signal Processing Systems. He is the Chair-Elect of the IEEE CAS Technical Committee on Multimedia Systems and Applications. During 2001–2002, he served as a Distinguished Lecturer of the IEEE CAS Society. He received the Best Paper Award from the R.O.C. Computer Society in 1990 and 1994. Annually from 1991 to 1999, he received Long-Term (Acer) Paper Awards. In 1992, he received the Best Paper Award of the 1992 Asia-Pacific Conference on circuits and systems in the VLSI design track. In 1993, he received the Annual Paper Award of the Chinese Engineer Society. In 1996 and 2000, he received the Outstanding Research Award from the National Science Council, and in 2000, the Dragon Excellence Award from Acer. He is a member of Phi Tan Phi.