Fast Motion and Disparity Estimation With Adaptive Search Range Adjustment in Stereoscopic Video Coding

Zhi-Pin Deng, Yui-Lam Chan, *Member, IEEE*, Ke-Bin Jia, Chang-Hong Fu, and Wan-Chi Siu, *Senior Member, IEEE*

Abstract—Stereoscopic video gives viewers more realistic vision than traditional 2D videos by transmitting two different views simultaneously. It doubles the required bandwidth in comparison with single view videos. Motion and disparity estimation therefore play a key role in reducing the bit rate of stereoscopic videos. However, it brings extremely huge computational complexity to an encoder which obstructs it from practical uses. In the past few years, some fast algorithms were proposed where most of them speed up the coding time based on an accurate estimation in one field (motion/disparity field), and then relieve the computational burden in the other field (disparity/motion field). Nevertheless, the complexities of both motion and disparity estimation cannot be fully reduced. In this paper, an iterative motion and disparity estimation algorithm is proposed. The proposed algorithm can determine motion vectors of the right view and disparity vectors of the current stereo pair in an iterative way. The gain of this iterative search is due to the use of a stereo-motion consistency constraint in which the motion and disparity vectors can be estimated by updating the local optimal vectors iteratively. An adaptive search range adjustment through the confidence measure of the constraint is then designed to further strengthen the reliability of each step for the iterative search. Experimental results show that the speed-up gain of the proposed algorithm is 18.76 ~ 229.19 times compared to the full search algorithm with a negligible quality drop.

Index Terms—Disparity estimation, motion estimation, stereoscopic video coding.

I. INTRODUCTION

STEREOSCOPIC video has gained increasing interest for its wide applications including 3D movie entertainment [1], immersive teleconference [2], medical surgery, etc. To provide viewers with a realistic 3D scene perception, two video sequences (left and right views) are captured by closely located cameras and are then shown to two eyes simultaneously [3], [4]. However, the stereoscopic video leads to a huge amount of data [5]–[7]. Hence, it is of great importance to compress

it efficiently. Many researchers have proposed to employ both temporal (intra-view) and inter-view prediction to improve the coding efficiency of stereoscopic video [8]–[10]. In this approach, the left view is the base view in which pictures are encoded as I-frame followed by P-frames. The right view is the predicted view where frames are estimated by not only the previous frame in the same view but also the frame at the same time instance in the left view. This induces higher computational complexity compared to the case of encoding a single view sequence.

Consequently, a number of fast search algorithms have been proposed, which reduce the computational time by a stereo-motion consistency constraint [11], [12]. This constraint is derived from the relation between motion and disparity vectors in two successive frame pairs as follows:

$$mv_{r,t} + dv_{r,t-1} = dv_{r,t} + mv_{l,t} \quad (1)$$

where $mv_{l,t}$ and $mv_{r,t}$ are the motion vectors of the left and right views at time t , $dv_{r,t-1}$ and $dv_{r,t}$ are the disparity vectors of the right view at $t-1$ and t . Based on the stereo-motion consistency constraint, $mv_{r,t}$ ($dv_{r,t}$) can be computed from $mv_{l,t}$, $dv_{r,t}$ ($mv_{r,t}$) and $dv_{r,t-1}$. Patras *et al.* [13] then suggested a joint algorithm to speed up the motion and disparity estimation, but this algorithm estimates motion and disparity vectors on pixel-by-pixel basis which causes serious accumulation of perturbation errors from the previous vectors and it is also not compatible with the H.264 standard [14]. The use of the stereo-motion consistency constraint was extended to the block-based motion and disparity estimation in MPEG-4 [15]. This algorithm uses the constraint to obtain a candidate disparity vector in the right view at time t given that $mv_{r,t}$ and $dv_{r,t-1}$ are available in the previous step and $mv_{l,t}$ is computed using an independent search algorithm. This candidate disparity vector is then refined to achieve a final $dv_{r,t}$. It can reduce computational complexity with good quality performance. Nevertheless, this algorithm cannot fully utilize the stereo-motion consistency constraint to obtain motion and disparity vectors simultaneously. Ding *et al.* [16] proposed a fast algorithm to reduce the computational complexity of motion estimation by setting the motion vector of the corresponding macroblock (MB) in the left view as the initial motion vector of the current MB in the right view. This algorithm can only speed up the motion estimation while a full search is still needed during the disparity estimation. So it cannot reduce the computational complexity to a sufficient extent. Similarly, a fast motion/disparity search algorithm in [17] was suggested in which a good set of candidate

Manuscript received September 15, 2010; revised September 06, 2011; accepted October 23, 2011. Date of publication December 14, 2011; date of current version February 23, 2012. This work was supported in part by the Center for Signal Processing, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, and a National Natural Science Foundation of China under Grant 30970780.

Z.-P. Deng and K.-B. Jia are with the Department of Electronic Information and Control Engineering, Beijing University of Technology, Beijing 100124, China (e-mail: zhipindeng@gmail.com; kebinj@bjut.edu.cn).

Y.-L. Chan, C.-H. Fu and W.-C. Siu are with the Center for Signal Processing, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong (e-mail: enylchan@polyu.edu.hk; enchfu@hotmail.com; enwcsiu@polyu.edu.hk).

Digital Object Identifier 10.1109/TBC.2011.2174279

Fig. 1. The stereo-motion consistency constraint in block-based motion and disparity estimation.

vectors from the last field (motion/disparity field) is employed for the other field (disparity/motion field). The computational complexity can be saved with a little drop in coding efficiency. However, the complexity of both motion and disparity estimation cannot be removed sufficiently since the candidate vectors for the other field rely on the full-search-based motion/disparity estimation in the last field.

By making use of the stereo-motion consistency constraint with an adaptive search range adjustment, a fast algorithm is proposed in this paper to get motion and disparity vectors simultaneously instead of estimating motion and disparity vectors separately. The candidate disparity vector of the current frame is computed from other available vectors, and it is further used to get a candidate motion vector in the next iteration. This iterative process can obtain significant reduction of complexity. The accuracy of this iterative search strategy highly relies on the search range of each step. To make the proposed algorithm viable and reduce the probability of trapping in the local minimum, a new confidence measure of the stereo-motion consistency constraint is proposed and it becomes a good criterion to determine the search range for each iteration. Experimental results show that our algorithm can remarkably reduce the computational complexity and maintain the coding performance. The remainder of this paper is organized as follows. Section II illustrates the iterative search using the stereo-motion consistency constraint in H.264. The adaptive refinement search range through a confidence measure of the stereo-motion consistency constraint is then presented in Section III. Section IV gives the flowchart of the proposed algorithm. Experimental results are given in Section V. Finally, Section VI concludes this paper.

II. THE ITERATIVE SEARCH USING THE STEREO-MOTION CONSISTENCY CONSTRAINT IN H.264

The use of the stereo-motion consistency constraint has been well studied in the pixel-wise motion and disparity estimation [13]. It can be extended to the block-based H.264 coding standard, as shown in Fig. 1. Let $F_{l,t}$ and $F_{r,t}$ be the left and right pictures at time t , and let $MB_{r,t}$ be the current MB of $F_{r,t}$ to be encoded. In this figure, $F_{l,t}$ and $F_{r,t-1}$ are the inter-view and temporal reference frames of $F_{r,t}$, respectively. In Fig. 1,

Fig. 2. The problem of using the stereo-motion consistency constraint in block-based motion and disparity estimation.

we also assume that $DCMB_{l,t}$ is the best disparity-compensated MB to $MB_{r,t}$, and $MCMB_{l,t-1}$ is the best motion-compensated MB to $DCMB_{l,t}$. Similarly, $MCMB_{r,t-1}$ is the best motion-compensated MB to $MB_{r,t}$, and $DCMB_{l,t-1}$ is the best disparity-compensated MB to $MCMB_{r,t-1}$. When $MCMB_{l,t-1}$ and $DCMB_{l,t-1}$ are coincident, the stereo-motion consistency constraint in the block-based structure can be expressed as:

$$mv_{r,t}(MB_{r,t}) + dv_{r,t-1}(MCMB_{r,t-1}) = dv_{r,t}(MB_{r,t}) + mv_{l,t}(DCMB_{l,t}) \quad (2)$$

Note that $dv_{r,t-1}(MCMB_{r,t-1})$ and $mv_{l,t}(DCMB_{l,t})$ have been obtained in the previous step. In a straightforward implementation, $dv_{r,t}(MB_{r,t})$ can firstly be obtained through the full-search algorithm. The relation in (2) can then be used for computing $mv_{r,t}(MB_{r,t})$ from $dv_{r,t}(MB_{r,t})$, $dv_{r,t-1}(MCMB_{r,t-1})$, and $mv_{l,t}(DCMB_{l,t})$. Such a process can only reduce the encoding time for motion estimation, but it does not help disparity estimation. In our work, we aim at calculating the two vectors of $MB_{r,t}$, $mv_{r,t}(MB_{r,t})$ and $dv_{r,t}(MB_{r,t})$, iteratively in order to speed up both motion and disparity estimation. Based on the stereo-motion consistency constraint, $mv_{r,t}(MB_{r,t})$ or $dv_{r,t}(MB_{r,t})$ can be predicted by the other three vectors in two successive frame pairs. However, the coincidence of $MCMB_{l,t-1}$ and $DCMB_{l,t-1}$ in Fig. 1 cannot be simply assumed since $DCMB_{l,t}$ and $MCMB_{r,t-1}$ are not on MB boundary, as depicted in Fig. 2. In other words, $mv_{l,t}(DCMB_{l,t})$ and $dv_{r,t-1}(MCMB_{r,t-1})$ are not available, and hence the stereo-motion consistency constraint does not hold in practice due to the block nature of H.264. Thus (2) is only suitable for generating the base vectors for each iteration, and they have to be refined later. In order to resolve this problem, we propose the following iterative procedure. The i th iteration ($i \geq 1$) is composed of two major steps:

Step 1) In Fig. 3(a), assume that $mv_{r,t}^{i-1}(MB_{r,t})$ has already been obtained in the $(i-1)$ th iteration. Before the 1st iteration, $mv_{r,t}^{i-1}(MB_{r,t})$ is necessary to be initialized and the procedure of initialization will be introduced in Section IV. The base disparity vector ($Bdv_{r,t}^i(MB_{r,t})$) is calculated

Fig. 3. The processes of obtaining (a) the base disparity vector, and (b) the base motion vectors.

from $mv_{r,t}^{i-1}(MB_{r,t})$, $dv_{r,t-1}^{i-1}(MCMB_{r,t-1})$ and $mv_{l,t}^{i-1}(DCMB_{l,t})$ based on the stereo-motion consistency constraint, where the process of how to get $dv_{r,t-1}^{i-1}(MCMB_{r,t-1})$ and $mv_{l,t}^{i-1}(DCMB_{l,t})$ will be elaborated later in this Section. This base disparity vector ($Bdv_{r,t}^i(MB_{r,t})$) is then refined by exhaustively computing all checking points in a new small search window to get the refined $dv_{r,t}^i(MB_{r,t})$.

Step 2) Based on the refined $dv_{r,t}^i(MB_{r,t})$ in step 1, the base motion vector ($Bmv_{r,t}^i(MB_{r,t})$), as depicted in Fig. 3(b), is computed from $dv_{r,t}^i(MB_{r,t})$, $dv_{r,t-1}^{i-1}(MCMB_{r,t-1})$ and $mv_{l,t}^{i-1}(DCMB_{l,t})$ according to the constraint. Refinement search is performed around the base motion vector ($Bmv_{r,t}^i(MB_{r,t})$) with the limited search window only in order to obtain the refined $mv_{r,t}^i(MB_{r,t})$.

In each step of the iteration, one target base vector is generated from (2) using the combination of other three vectors. Therefore, two base vectors are updated during one iteration. These base vectors are used as the search centers, and then the refined vectors around the base vectors are obtained through the refinement with the limited search range. After several iterations, the proposed algorithm can get the optimal $mv_{r,t}(MB_{r,t})$ and $dv_{r,t}(MB_{r,t})$ with low complexity as compared to the independent search.

As aforementioned, $dv_{r,t-1}(MCMB_{r,t-1})$ and $mv_{l,t}(DCMB_{l,t})$ in Fig. 2 do not exist. In order to make an approximation of $dv_{r,t-1}(MCMB_{r,t-1})$ and $mv_{l,t}(DCMB_{l,t})$, it is possible to use the disparity and motion vectors of the four overlapping MBs with $MCMB_{r,t-1}$ and $DCMB_{l,t}$ (marked as 1, 2, 3, and 4 in $F_{l,t}$ and $F_{r,t-1}$ of Fig. 3(a) and (b), respectively). Nevertheless, the approximation of $dv_{r,t-1}(MCMB_{r,t-1})$ and $mv_{l,t}(DCMB_{l,t})$ may not exactly be the true vectors of $MCMB_{r,t-1}$ and $DCMB_{l,t}$, and therefore $DCMB_{l,t-1}$ and $MCMB_{l,t-1}$ corresponding to the same content in the 3D space may not locate the same point. Consequently, using the stereo-motion consistency constraint directly may introduce the positional uncertainty. This positional uncertainty would affect the accuracy of the iterative search. To reduce the uncertainty of the iterative search, multiple candidates for $dv_{r,t-1}(MCMB_{r,t-1})$ and $mv_{l,t}(DCMB_{l,t})$ are employed to form the base disparity vector ($Bdv_{r,t}^i(MB_{r,t})$) and the base motion vector ($Bmv_{r,t}^i(MB_{r,t})$), where i is the iteration step ($i \geq 1$). As mentioned in the iteration process, when the motion vector $mv_{r,t}^{i-1}(MB_{r,t})$ in the $(i-1)$ th iteration is fixed, the disparity vectors of the four overlapping MBs with $MCMB_{r,t-1}$ are the possible candidates to compose $Bdv_{r,t}^i(MB_{r,t})$ of the current i th iteration, and they are denoted by $dv_{r,t-1}^{i-1,u}(MCMB_{r,t-1})$, where $1 \leq u \leq 4$, as depicted in Fig. 3(a). Meanwhile, the motion vector of $DCMB_{l,t}$ in the $(i-1)$ th iteration, denoted by $mv_{l,t}^{i-1}(DCMB_{l,t})$, is also achievable. Therefore, according to (2), there would be four possible combinations of $Bdv_{r,t}^{i,u}(MB_{r,t})$, computed by

$$Bdv_{r,t}^{i,u}(MB_{r,t}) = mv_{r,t}^{i-1}(MB_{r,t}) + dv_{r,t-1}^{i-1,u}(MCMB_{r,t-1}) - mv_{l,t}^{i-1}(DCMB_{l,t}) \quad (3)$$

Therefore, the resultant $Bdv_{r,t}^i(MB_{r,t})$ of the i th iteration can then be derived from

$$Bdv_{r,t}^i(MB_{r,t}) = \min_{1 \leq u \leq 4} (RDCost(Bdv_{r,t}^{i,u}(MB_{r,t}))) \quad (4)$$

where $RDCost()$ is the cost function that takes both of the distortion and the bit number consumed for coding MB into account. Note that the corresponding $dv_{r,t-1}^{i-1,u}(MCMB_{r,t-1})$ ($1 \leq u \leq 4$) in (3), which contributes to attain $Bdv_{r,t}^i(MB_{r,t})$ in (4), is updated as $dv_{r,t-1}^{i-1}(MCMB_{r,t-1})$ for the next step to get $Bmv_{r,t}^i(MB_{r,t})$.

Similarly, when the refined disparity vector $dv_{r,t}^i(MB_{r,t})$ is fixed, the motion vectors of the four overlapping MBs with $DCMB_{l,t}$, represented by $mv_{l,t}^{i,v}(DCMB_{l,t})$, where $1 \leq v \leq 4$, are the possible candidates to compose $Bmv_{r,t}^i(MB_{r,t})$, as depicted in Fig. 3(b). The resultant base motion vector $Bmv_{r,t}^i(MB_{r,t})$ of the i th iteration will be

$$Bmv_{r,t}^i(MB_{r,t}) = \min_{1 \leq v \leq 4} (RDCost(Bmv_{r,t}^{i,v}(MB_{r,t}))) \quad (5)$$

where

$$Bmv_{r,t}^{i,v}(MB_{r,t}) = dv_{r,t}^i(MB_{r,t}) + mv_{l,t}^{i,v}(DCMB_{l,t}) - dv_{r,t-1}^{i-1}(MCMB_{r,t-1}) \quad (6)$$

TABLE I
STATISTICAL ANALYSIS OF δ FOR VARIES SEQUENCES UNDER FULL SEARCH

	$\delta = 0$	$\delta < 5$	$\delta \leq 20$
Ballroom	5.39%	30.18%	96.60%
Exit	13.64%	20.43%	91.85%
Vassar	27.15%	62.57%	98.92%
Flamenco2	1.11%	6.87%	90.87%
Race1	1.92%	13.65%	87.38%
Rena	32.26%	74.49%	98.14%

where $dv_{r,t-1}^i(MCMB_{r,t-1})$ is obtained in the last step of the i th iteration.

III. ADAPTIVE SEARCH RANGE THROUGH THE CONFIDENCE MEASURE OF STEREO-MOTION CONSISTENCY CONSTRAINT

In Section II, we have designed a novel search strategy in which both motion and disparity vectors are iteratively updated to minimize the cost function under the stereo-motion consistency constraint. The refined vector obtained in the previous step from motion or disparity estimation can be used to get a new base disparity or motion vector in the current step. As aforementioned, due to the block nature of H.264, $DCMB_{l,t-1}$ and $MCMB_{l,t-1}$ are no longer coincident in $F_{l,t-1}$, as depicted in Fig. 3. This positional uncertainty in $F_{l,t-1}$ leads to a loop difference (δ) attached to the stereo-motion consistency constraint, and (2) can then be re-written as,

$$\delta = \|mv_{r,t}(MB_{r,t}) + dv_{r,t-1}(MCMB_{r,t-1}) - dv_{r,t}(MB_{r,t}) - mv_{l,t}(DCMB_{l,t})\| \quad (7)$$

where $\|v\|$ is the norm of a vector v . Table I shows the statistical analysis of δ for various sequences. δ is obtained via the corresponding four best vectors of the full search algorithm. From Table I, it can be found that the values of δ for most MBs are not always equal to zero. The value of δ surges in some occlusion and ambiguous areas. It implies the use of the stereo-motion consistency constraint directly may introduce the positional uncertainty, which is described by δ in (7). Multiple candidates used in Section II can alleviate this uncertainty in the iterative search. In this section, another technique of dealing with the uncertainty is proposed to assign a confidence measure to the search strategy. For each iteration, this measure is used to adaptively adjust the search range of the refinement process, and can further improve the reliability of the final motion and disparity vectors. It is noted that the refinement process refers to the full search with a limited search range.

In order to have a more accurate vector field for the next vector field, a refinement around each base vector is carried out. The search range of the refinement has a great impact on the performance of the iterative search. It is well known that a large refinement search range can compensate an unreliable base vector and maintain a good RD performance. However, it costs heavy computational load. On the other hand, a small refinement search range only requires less search points resulting in lowering complexity. However, this is easily trapped into a local minimum and leads to incorrect prediction results [20]. This further affects the accuracy of the subsequent iterations for the iterative search.

Fig. 4. Adaptive RSR based on δ .

To avoid being trapped in a local minimum with a reasonable search range, the size of the search range for each MB in each iteration is adaptively controlled according to the reliability of the stereo-motion consistency constraint. In general, a smaller refinement search range is used for the case of the reliable stereo-motion consistency constraint while a larger refinement search range is needed for the case of the unreliable constraint. The loop difference (δ) attached to the stereo-motion consistency constraint can be used as the confidence measure to determine the search range of the refinement. If δ is small, it means that the base vector obtained by the stereo-motion consistency constraint is reliable, and the size of the search range can be reduced to a proper size without affecting the estimation accuracy. On the other hand, if δ is large, it implies that the base vector may be inaccurate, and thus the size of the search range should be larger in order to maximize the possibility for finding the global minimum. This process also benefits the subsequent iterations if more reliable motion and disparity vectors can be obtained in the previous iterations.

With the aid of δ , the size of the refinement search range for the current MB, RSR , is adjusted as

$$RSR = \begin{cases} RSR_{MIN} & \delta < T_1 \\ RSR_{MIN} + \frac{\delta - T_1}{T_2 - T_1} \times (RSR_{MAX} - RSR_{MIN}) & T_1 \leq \delta \leq T_2 \\ RSR_{MAX} & \delta > T_2 \end{cases} \quad (8)$$

where RSR_{MIN} and RSR_{MAX} denote the sizes of the minimum and maximum search range, respectively, in the proposed algorithm. From (8), two confidence thresholds, T_1 and T_2 ($T_1 < T_2$), are given to discriminate RSR for the three regions of the search range, as depicted in Fig. 4. If δ is larger than T_2 , it is highly probable that the stereo-motion consistency constraint is no longer workable for the current MB, then RSR is set to RSR_{MAX} . Since the constraint does not work well in this case, RSR_{MAX} is equal to the search range of the full search algorithm. When $T_1 \leq \delta \leq T_2$, RSR varies in direct proportion to δ . If δ is smaller than T_1 , the constraint should be reliable and it is confident to use a very small RSR , i.e. RSR_{MIN} . In order to fix the value of RSR_{MIN} , the vector difference between the final optimal vector and the initial base vector of the proposed algorithm is shown in Fig. 5, where the X-axis and Y-axis represent the horizontal and vertical values of the vector difference, respectively, and Z-axis denotes the number of MBs. It can be found that more than 90% of MBs with vector difference are less than 2. It infers that performing a

Fig. 5. The histogram of vector difference between the final and initial base vectors of “Ballroom” at QP37.

Fig. 6. The flowchart of the proposed algorithm.

2×2 refinement around the initial base vector is good enough to ensure the prediction accuracy for these MBs. Therefore, in the proposed iterative search algorithm, RSR_{MIN} is set to 2.

IV. THE FLOWCHART OF THE PROPOSED ALGORITHM

The flowchart of the proposed fast algorithm is shown in Fig. 6, where i denotes the iteration step. It is mainly divided into four steps including an initialization, a refinement search range adjustment, an iterative search, and a stopping process. Let us provide a summary of the proposed algorithm in the following:

1. Initialization:

- (i) Set $i = 0$ and $\delta = 0$.
- (ii) Choose an initial base motion vector ($Bmv_{r,t}^0(MB_{r,t})$) and an initial base disparity

Fig. 7. Vector sets for selecting initial base motion and disparity vectors.

vector ($Bdv_{r,t}^0(MB_{r,t})$) from the following vector sets:

$$Bmv_{r,t}^0(MB_{r,t}) : \{mv_{l,t}, mv_{med}, mv_a, mv_b, mv_c, \vec{0}\}$$

,and

$$Bdv_{r,t}^0(MB_{r,t}) : \{dv_{r,t-1}, dv_{med}, dv_a, dv_b, dv_c, \vec{0}\}$$

where mv_a/dv_a , mv_b/dv_b and mv_c/dv_c are motion/disparity vectors from the neighboring left, top, and top-right blocks of the current block. mv_{med} and dv_{med} are the medians of the motion and disparity vectors, respectively, marked as mv_a/dv_a , mv_b/dv_b and mv_c/dv_c in Fig. 7. Note that either motion or disparity vector is available [18], [19] for the spatially neighboring blocks. Thus, if block b is predicted from the inter-view reference frame, mv_b is not available. In this case, (0, 0) is used to replace mv_b for deriving mv_{med} . For each initial base vector, the one that has the minimum value of $RDCost$ is selected.

- (iii) Perform refinement on both $Bmv_{r,t}^0(MB_{r,t})$ and $Bdv_{r,t}^0(MB_{r,t})$ with RSR equal to RSR_{MIN} to get $mv_{r,t}^0(MB_{r,t})$ and $dv_{r,t}^0(MB_{r,t})$. Save $RDCost(mv_{r,t}^0(MB_{r,t}))$ and $RDCost(dv_{r,t}^0(MB_{r,t}))$, and set $i = i + 1$.
2. Refinement search range adjustment: Based on δ computed in the previous step (either in step 1 or step 4), set the refinement search range RSR via (8).
3. Iterative search:
 - (i) Fix $mv_{r,t}^{i-1}(MB_{r,t})$ and compute the base disparity vector $Bdv_{r,t}^i(MB_{r,t})$ through (3) and (4). Perform disparity vector refinement with RSR on $Bdv_{r,t}^i(MB_{r,t})$ to obtain $dv_{r,t}^i(MB_{r,t})$. Save $RDCost(dv_{r,t}^i(MB_{r,t}))$.
 - (ii) Fix $dv_{r,t}^i(MB_{r,t})$ and calculate the base motion vector $Bmv_{r,t}^i(MB_{r,t})$ using (5) and (6). Determine $mv_{r,t}^i(MB_{r,t})$ by carrying out motion vector refinement with RSR on $Bmv_{r,t}^i(MB_{r,t})$. Save $RDCost(mv_{r,t}^i(MB_{r,t}))$.
4. Stopping process: if $RDCost(mv_{r,t}^i(MB_{r,t})) \geq RDCost(mv_{r,t}^{i-1}(MB_{r,t}))$ and

TABLE II
COMPARISON OF COMPUTATIONAL COMPLEXITY AND THE SPEED-UP RATIO OF THE TESTED ALGORITHMS FOR DIFFERENT SEQUENCES

Sequence	Algorithm	Average number of search points per MB				Total motion and disparity estimation time in seconds			
		QP22	QP27	QP32	QP37	QP22	QP27	QP32	QP37
Ballroom	Full search	65616	65692	65707	65745	47188	46695	46454	46300
	Algorithm in [17]	32906 (1.99)	32925 (2.00)	32934 (2.00)	32935 (2.00)	23815 (1.98)	23541 (1.98)	23401 (1.99)	23303 (1.99)
	TZ search	3503 (18.73)	3471 (18.92)	3356 (19.58)	3201 (20.54)	2680 (17.61)	2604 (17.93)	2518 (18.45)	2347 (19.73)
	Proposed	3513 (18.68)	3138 (20.93)	2619 (25.09)	2297 (28.62)	2515 (18.76)	2222 (21.01)	1849 (25.12)	1620 (28.58)
	Full search	65343	65532	65588	65621	45548	45035	44855	44768
Exit	Algorithm in [17]	32858 (1.99)	32884 (1.99)	32883 (1.99)	32883 (2.00)	23000 (1.98)	22686 (1.99)	22571 (1.99)	22507 (1.99)
	TZ search	3663 (17.84)	3524 (18.60)	3329 (19.70)	3112 (21.09)	2665 (17.09)	2517 (17.89)	2369 (18.93)	2198 (20.37)
	Proposed	3376 (19.35)	1436 (45.63)	1150 (57.04)	1051 (62.46)	2427 (18.77)	1058 (42.57)	853 (52.58)	780 (57.39)
	Full search	65738	65751	65752	65742	50077	48729	47889	47442
	Algorithm in [17]	32924 (2.00)	32921 (2.00)	32919 (2.00)	32912 (2.00)	25168 (1.99)	24503 (1.99)	24095 (1.99)	23878 (1.99)
Vassar	TZ search	3242 (20.28)	3210 (20.48)	2884 (22.80)	2424 (27.12)	2601 (19.25)	2498 (19.51)	2199 (21.78)	1835 (25.85)
	Proposed	296 (222.17)	263 (249.59)	207 (318.26)	183 (359.48)	315 (158.97)	277 (175.92)	229 (209.12)	207 (229.19)
	Full search	65699	65703	65675	65660	52295	52201	52129	52042
	Algorithm in [17]	39555 (1.99)	39563 (1.99)	39570 (1.99)	39574 (1.99)	26375 (1.98)	26335 (1.98)	26310 (1.98)	26294 (1.98)
	TZ search	4745 (16.63)	4684 (16.84)	4580 (17.22)	4486 (17.58)	3242 (16.13)	3219 (16.22)	3127 (16.67)	3073 (16.94)
Flamenco2	Proposed	1321 (59.71)	1348 (58.52)	1270 (62.08)	1178 (66.94)	965 (54.19)	977 (53.43)	923 (56.48)	861 (60.44)
	Full search	65585	65606	65604	65619	53761	53496	53251	53018
	Algorithm in [17]	39503 (1.99)	39508 (1.99)	39505 (1.99)	39495 (2.00)	27268 (1.97)	27134 (1.97)	26997 (1.97)	26055 (2.03)
	TZ search	4186 (18.81)	4183 (18.83)	4135 (19.05)	3980 (19.80)	2935 (18.32)	2915 (18.35)	2868 (18.57)	2757 (19.23)
	Proposed	2441 (32.26)	2340 (33.66)	2207 (35.69)	1955 (40.31)	1703 (31.57)	1630 (32.82)	1537 (34.65)	1368 (38.76)
Race1	Full search	65748	65758	65757	65771	55094	54779	54520	54182
	Algorithm in [17]	39531 (2.00)	39531 (2.00)	39529 (2.00)	39531 (2.00)	27765 (1.98)	27619 (1.98)	27461 (1.99)	26611 (2.04)
	TZ search	2412 (32.74)	2419 (32.64)	2345 (33.67)	2073 (38.09)	1799 (30.62)	1790 (30.60)	1729 (31.53)	1527 (35.48)
	Proposed	1688 (46.77)	1594 (49.54)	1377 (57.36)	865 (91.31)	1249 (44.11)	1175 (46.62)	1020 (53.45)	671 (80.75)
	Full search	65748	65758	65757	65771	55094	54779	54520	54182
Rena	Algorithm in [17]	39531 (2.00)	39531 (2.00)	39529 (2.00)	39531 (2.00)	27765 (1.98)	27619 (1.98)	27461 (1.99)	26611 (2.04)
	TZ search	2412 (32.74)	2419 (32.64)	2345 (33.67)	2073 (38.09)	1799 (30.62)	1790 (30.60)	1729 (31.53)	1527 (35.48)
	Proposed	1688 (46.77)	1594 (49.54)	1377 (57.36)	865 (91.31)	1249 (44.11)	1175 (46.62)	1020 (53.45)	671 (80.75)
	Full search	65748	65758	65757	65771	55094	54779	54520	54182
	Algorithm in [17]	39531 (2.00)	39531 (2.00)	39529 (2.00)	39531 (2.00)	27765 (1.98)	27619 (1.98)	27461 (1.99)	26611 (2.04)
	TZ search	2412 (32.74)	2419 (32.64)	2345 (33.67)	2073 (38.09)	1799 (30.62)	1790 (30.60)	1729 (31.53)	1527 (35.48)
	Proposed	1688 (46.77)	1594 (49.54)	1377 (57.36)	865 (91.31)	1249 (44.11)	1175 (46.62)	1020 (53.45)	671 (80.75)

$RDCost(dv_{r,t}^i(MB_{r,t})) \geq RDCost(dv_{r,t}^{i-1}(MB_{r,t}))$, set $mv_{r,t}^{i-1}(MB_{r,t})$ and $dv_{r,t}^{i-1}(MB_{r,t})$ as the final motion and disparity vectors, stop the iteration. Otherwise, set $i = i + 1$, calculate and update δ from (7), go to step 2.

Once the final motion and disparity vectors are calculated, the selection between motion and disparity vectors of $MB_{r,t}$ is based on $RDCost(mv_{r,t}^{i-1}(MB_{r,t}))$ and $RDCost(dv_{r,t}^{i-1}(MB_{r,t}))$. If $RDCost(mv_{r,t}^{i-1}(MB_{r,t}))$ is smaller, $mv_{r,t}^{i-1}(MB_{r,t})$ is used to compute the prediction error of $MB_{r,t}$. Otherwise, $dv_{r,t}^{i-1}(MB_{r,t})$ is selected.

V. EXPERIMENTAL RESULTS

We evaluated the performance of the proposed algorithm using a large variety of multiview sequences. Six public multiview sequences [21], [22], “Ballroom”, “Exit”, “Vassar”, “Flamenco2”, “Race1”, and “Rena”, with the image size of 640×480 were used for performance comparison. For simplicity, but without loss of generality, view 0 and view 1 were used as the left and right views, respectively. For the implementation, the proposed algorithm was built based on the Joint Multiview Video Model (JMVM version 8.0) [23]. The results of the proposed algorithm are compared with the three conventional

TABLE III
AVERAGE VALUES OF δ AND SEARCH RANGE AT DIFFERENT QPs

QP	Ballroom		Exit		Vassar		Flamenco2		Race1		Rena	
	δ	SR	δ	SR	δ	SR	δ	SR	δ	SR	δ	SR
22	4.53	6.26	5.38	5.77	0.87	2.05	1.63	2.75	2.56	4.94	3.13	3.37
27	3.87	5.71	2.76	3.46	0.70	2.05	1.61	2.77	2.53	4.74	3.12	3.29
32	3.52	5.06	2.14	3.19	0.47	2.03	1.55	2.75	2.43	4.51	2.83	3.11
37	3.48	4.64	1.97	3.10	0.38	2.03	1.58	2.77	2.42	4.18	2.40	2.64

methods, which are the full search algorithm (FSA), the fast algorithm proposed in [17], and the well-known fast TZ search algorithm in JMVM [23]. Basically, the bitstreams were encoded by different algorithms according to the test condition in [24]. Four different QPs (i.e. QP = 22, 27, 32, and 37) with IPPP... structure, quarter-pel motion and disparity estimation with ± 96 -pel search range, and CABAC were used. For the proposed algorithm, T_1 and T_2 , were selected by considering the tradeoff between the computational requirement and the prediction accuracy for most sequences. In this paper, T_1 and T_2 were experimentally set to 5 and 20, respectively. Besides, RSR_{MIN} and RSR_{MAX} were set to 2 and 96, respectively. It is interesting to point out that the full search algorithm is used to perform the exhaustive search with the largest search range if $\delta > 20$, which brings a huge amount of computational complexity. From Table I, it can be found that δ mainly distributes from 0 to 20 (more than 87% of MBs). It implies that the chance of using RSR_{MAX} to perform refinement is very low. All the simulations were carried out on an Intel(R) Xeon(R) X5550 2.67GHz computer CPU with 12GB RAM.

To evaluate the computational efficiency of various algorithms, the average numbers of search points per MB and the motion and disparity estimation time per sequence for different algorithms are shown in Table II, and all of these search points and estimation time were counted in runtime during the experiments. From this table, it shows that our algorithm can successfully improve the computational efficiency of FSA and either the number of search points or the estimation time is fewer than those of the algorithms in [17] and TZ search for all sequences. For all QPs, our algorithm requires the least number of search points and the least motion and disparity estimation time. The speed-up ratios compared to FSA of the tested algorithms with different QPs are also provided in the brackets of Table II. The speed-up ratio is defined as the number of search points (or the motion and disparity time) of FSA divided by those of various algorithms. As expected, the proposed algorithm gives the largest speed-up ratio among all algorithms, and the speed-up gain is up to 359 times of FSA in the ‘‘Vassar’’ sequence.

From Table II, it is interesting to note that the speed-up ratio of the proposed algorithm increases when QP increases. Generally speaking, the number of search points of the proposed algorithm depends greatly on an average value of δ . At high QP, the use of the adaptive refinement search range adjustment technique in the proposed algorithm always adopts a smaller search range for MBs due to the diminishing loop difference (δ). It results in lowering the complexity of the search. The results in Table III justify our analysis. This table shows the average

TABLE IV
STATISTICS OF AVERAGE NUMBERS OF ITERATIONS FOR DIFFERENT SEQUENCES

Step	equal to 1	be smaller or equal to 3	average no. of iterations	
Ballroom	QP22	81.65%	96.83%	1.33
	QP27	83.00%	96.90%	1.31
	QP32	84.18%	97.11%	1.29
	QP37	84.95%	97.35%	1.27
Exit	QP22	68.68%	95.68%	1.51
	QP27	75.09%	95.80%	1.43
	QP32	81.16%	96.67%	1.33
	QP37	85.02%	97.25%	1.27
Vassar	QP22	86.07%	98.60%	1.20
	QP27	88.65%	98.89%	1.16
	QP32	92.98%	99.18%	1.11
	QP37	95.70%	99.44%	1.07
Flamenco2	QP22	80.05%	97.45%	1.31
	QP27	82.75%	97.59%	1.28
	QP32	84.52%	97.86%	1.25
	QP37	85.26%	98.02%	1.23
Race1	QP22	81.03%	97.32%	1.31
	QP27	79.73%	97.06%	1.34
	QP32	79.50%	96.98%	1.34
	QP37	80.38%	97.15%	1.32
Rena	QP22	82.71%	96.84%	1.31
	QP27	83.39%	96.95%	1.29
	QP32	84.91%	97.28%	1.26
	QP37	87.89%	97.92%	1.21

values of δ for different QPs. It is known that $RDCost$ used in (4) and (6) is to optimize the amount of distortion (D) against the amount of data (R) required to encode the video using Lagrangian optimization. This cost can be defined as

$$RDCost = D + \lambda R \quad (9)$$

The amount of distortion, D , is defined as the sum of absolute differences between the current block (B_c) and reference block (B_r), and can be written as,

$$D = \sum_{(k,k') \in (B_c, B_r)} |B_c(k) - B_r(k')| \quad (10)$$

where $B_c(k)$ denotes the pixel value of the k th pixel in B_c , $B_r(k')$ denotes the pixel value of the k' th pixel in B_r .

In (9), R is the number of coded bits for each MB which includes the required bits to encode the header information, quantized residual block and motion/disparity vector difference between the optimal motion/disparity vector and the predictors, and λ is the Lagrange multiplier which is equal to $0.85 \times 2^{\min\{52, QP\}/3-4}$. From (9), a high QP value increases λ . In this case, R gets more important in the calculation of the RDCost where smooth and regular motion/disparity vectors are always obtained at high QP. As a consequence, the probability of getting a smaller value of δ is higher, which reduces the size of the search range, as demonstrated in Table III. Therefore, in the case of high QP, motion and disparity estimation has a good chance to be performed with a smaller search range in the proposed algorithm, which results in higher speed-up ratio, as shown in Table II. In contrast, the resulting motion/disparity vectors tend

Fig. 8. Rate-Distortion performances of the tested algorithms for (a) “Ballroom”, (b) “Exit”, (c) “Vassar”, (d) “Flamenco2”, (e) “Race1”, and (f) “Rena” sequences.

to be chaotic at low QP. Hence, δ becomes larger as shown in Table III which increases the complexity of the proposed algorithm, but it is still better than the TZ search. The statistics of the average number of iterations required by the proposed algorithm are also shown in Table IV. It is seen that over 95% of macroblocks requires less than 3 iterations, and the average numbers of iterations keep a small value.

Fig. 8 shows the rate-distortion (R-D) curves of all algorithms for different sequences. It is shown that there is nearly no difference between the proposed algorithm and the full search, but the proposed algorithm provides higher coding efficiency as compared with [17]. Results for different sequences using the Bjontegaard delta PSNR (BD-PSNR) [25] and Bjontegaard delta bitrate (BD-Bitrate) compared to the full search are summarized in Table V. Compared with the full search, the PSNR degradation

TABLE V
THE RESULTS OF BD-PSNR AND BD-BITRATE FOR DIFFERENT ALGORITHMS

Sequence	Algorithm in [17]		TZ search		Proposed	
	BD-PSNR (dB)	BD-Bitrate (%)	BD-PSNR (dB)	BD-Bitrate (%)	BD-PSNR (dB)	BD-Bitrate (%)
Ballroom	-0.21	+6.60	+0.00	-0.04	+0.00	-0.07
Exit	-0.07	+3.42	-0.02	+0.62	-0.04	+1.64
Vassar	-0.02	+1.04	-0.01	+0.39	-0.02	+0.51
Flamenco2	+0.01	-0.25	-0.02	+0.33	-0.02	+0.33
Race1	-0.14	+3.46	-0.05	+1.20	-0.01	+0.33
Rena	-0.03	+0.71	-0.01	+0.12	+0.04	-0.81

of the proposed algorithm is less than 0.04 dB. However, [17] has a degradation from -0.21 dB to -0.02 dB. It demonstrates

that the computational redundancy of motion and disparity estimation can be effectively removed by the proposed algorithm with nearly the same RD performance.

VI. CONCLUSIONS

In this paper, an iterative search for motion and disparity estimation in stereoscopic video coding has been proposed to reduce the computational complexity. The strategy of the iterative search utilizes a stereo-motion consistency constraint such that motion and disparity vectors can be determined in a recursive manner by updating the local optimal vectors. A new confidence measure based on this constraint has also been designed to adjust the search range for each iteration. This process can further reduce the errors inherent to the iterative search. As compared with the full search algorithm, the proposed algorithm achieves an insignificant quality drop with over 95% computational savings.

REFERENCES

- [1] A. Smolic, K. Mueller, P. Merkle, P. Kauff, and T. Wiegand, "An overview of available and emerging 3D video formats and depth enhanced stereo as efficient generic solution," in *Proc. Picture Coding Symp. (PCS)*, Chicago, IL, May 2009, 10.1109/PCS.2009.5167358.
- [2] S. Y. Chien, S. H. Yu, L. F. Ding, Y. N. Huang, and L. G. Chen, "Efficient stereo video coding system for immersive teleconference with two-stage hybrid disparity estimation algorithm," in *Proc. Int. Conf. Image Process. (ICIP)*, Barcelona, Spain, Sep. 2003, pp. 749–752.
- [3] D. Kim, D. Min, and K. Sohn, "A stereoscopic video generation method using stereoscopic display characterization and motion analysis," *IEEE Trans. Broadcast.*, vol. 54, no. 2, pp. 188–197, Jun. 2008.
- [4] N. Dodgson, "Autostereoscopic 3D displays," *Computer*, vol. 38, no. 8, pp. 31–36, Aug. 2005.
- [5] Y. Luo, Z. Y. Zhang, and A. Ping, "Stereo video coding based on frame estimation and interpolation," *IEEE Trans. Broadcast.*, vol. 49, no. 1, pp. 14–21, Mar. 2003.
- [6] A. Smolic, K. Mueller, N. Stefanoski, J. Ostermann, A. Gotchev, G. B. Akar, G. Triantafyllidis, and A. Koz, "Coding algorithms for 3DTV – A survey," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 11, pp. 1606–1621, Nov. 2007.
- [7] A. Aksay and G. B. Akar, "Evaluation of stereo video coding schemes for mobile devices," in *Proc. True Vision Capture, Transmiss. Display of 3D Video (3DTV)*, Potsdam, Germany, May 2009, 10.1109/3DTV.2009.5069664.
- [8] S. Malassiotis and M. G. Strintzis, "Joint motion/disparity MAP estimation for stereo image sequences," *IEEE Vision, Image, Signal Process.*, vol. 143, no. 2, p. 101, Apr. 1996.
- [9] L. F. Ding, S. Y. Chien, and L. G. Chen, "Joint prediction algorithm and architecture for stereo video hybrid coding systems," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 11, pp. 1324–1337, Nov. 2006.
- [10] W. Yang, K. N. Ngan, and J. Cai, "An MPEG-4-compatible stereoscopic/multiview video coding scheme," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 2, pp. 286–290, Feb. 2006.
- [11] A. Tamtaoui and C. Labit, "Coherent disparity and motion compensation in 3DTV image sequence coding schemes," in *Proc. Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, Toronto, Canada, Apr. 1991, pp. 2845–2848.
- [12] W. Miled, B. Pesquet-Popescu, and W. Chérif, "A variational framework for simultaneous motion and disparity estimation in a sequence of stereo images," in *Proc. Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, Taipei, Taiwan, Apr. 2009, pp. 741–744.
- [13] I. Patras, N. Alvertos, and G. Tziritis, "Joint disparity and motion field estimation in stereoscopic image sequences," in *Proc. Int. Conf. Pattern Recogn. (ICPR)*, Vienna, Austria, Aug. 1996, pp. 359–363.
- [14] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 17, pp. 560–576, Jul. 2003.
- [15] Y. Kim, J. Lee, C. Park, and K. Sohn, "MPEG-4 compatible stereoscopic sequence codec for stereo broadcasting," *IEEE Trans. Consum. Electron.*, vol. 51, no. 4, pp. 1227–1236, Nov. 2005.
- [16] L. F. Ding, P. K. Tsung, W. Y. Chen, S. Y. Chien, and L. G. Chen, "Fast motion estimation with inter-view motion vector prediction for stereo and multiview video coding," in *Proc. Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, Las Vegas, NV, Mar. 2008, pp. 1373–1376.
- [17] P. Lai and A. Ortega, "Predictive fast motion/disparity search for multiview video coding," in *Proc. SPIE Vis. Commun. Image Process.*, San Jose, CA, Jan. 2006, pp. 6077091–60770912.
- [18] S.-H. Lee, S. H. Lee, N. I. Cho, and J.-H. Yang, "Disparity vector prediction methods in MVC," Hangzhou, China, Joint Video Team, Doc. JVT-U040, Oct. 2006.
- [19] H. Yang, J. Huo, Y. Chang, S. Lin, P. Zeng, and L. Xiong, "Regional disparity based motion and disparity prediction for MVC," Marrakech, Morocco, Joint Video Team, Doc. JVT-V071, Jan. 2007.
- [20] Y. L. Chan and W. C. Siu, "Reliable block motion estimation through the confidence measure of error surface," *Signal Process.*, vol. 76, no. 2, pp. 135–146, Jul. 1999.
- [21] A. Vetro, M. McGuire, W. Matusik, A. Behrens, J. Lee, and H. Pfister, "Multiview video test sequences from MERL," Busan, Korea, Doc. MPEG-M12077, Apr. 2005.
- [22] ISO/IEC JTC1/SC29/WG11, "Call for proposals on multi-view video coding," Poznan, Poland, Doc. MPEG-N7327, Jul. 2005.
- [23] A. Vetro, P. Pandit, H. Kimata, A. Smolic, and Y. K. Wang, "Joint multiview video model (JMVM) 8.0," Geneva, CH, Joint Video Team, Doc. JVT-AA207, Apr. 2008.
- [24] A. Vetro and A. Smolic, "Common test conditions for multiview video coding," Hangzhou, China, Joint Video Team, Doc. JVT-U211, Oct. 2006.
- [25] G. Bjontegaard, "Calculation of average PSNR differences between RDcurves," Austin, TX, Doc. VCEG-M33, Apr. 2001.

Zhi-Pin Deng received the B.S. degree in electronic information engineering in 2006 from the Beijing University of Technology, Beijing, China, where she is currently working toward the Ph.D. degree, also in electronic information engineering.

She was a Research Assistant of the Hong Kong Polytechnic University (HKPOLYU) from 2008 to 2010. Her research interests include digital image/video signal processing, stereoscopic video coding, multiview video coding, video segmentation, video feature extraction, and video search.

Yui-Lam Chan (S'94–A'97–M'00) received the B.Eng. with a First Class Honours degree and the Ph.D. degree from the Hong Kong Polytechnic University in 1993 and 1997, respectively.

He joined the Hong Kong Polytechnic University in 1997, and is now an Associate Professor in the Department of Electronic and Information Engineering. He is also actively involved in professional activities. In particular, he serves as a reviewer and Session Chairman for many international journals/conferences. He was the Secretary of the 2010 IEEE

International Conference on Image Processing (ICIP'2010). He has published over 70 research papers in various international journals and conferences. His research and technical interests include multimedia technologies, signal processing, image and video compression, video streaming, video transcoding, video conferencing, digital TV/HDTV, 3DTV, multi-view video coding, future video coding standards, error-resilient coding, and digital VCR.

During his studies, Dr. Chan was the recipients of more than 10 famous prizes, scholarships and fellowships for his outstanding academic achievement, such as being the Champion in Varsity Competition in Electronic Design, the Sir Edward Youde Memorial Fellowship, and the Croucher Foundation Scholarships.

Ke-Bin Jia received the Bachelor degree in electronic engineering from the Lan Zhou University, China, in 1984, the M.Phil. and Ph.D. degrees from University of Science and Technology of China in 1990 and 1998, respectively.

He joined Beijing University of Technology as a Lecturer in 1998 and has become Chair Professor in the School of Electronic Information and Control Engineering since 2002. He is now the Dean of School of Electronic Information and Control Engineering of Beijing University of Technology. He was a visiting researcher at the Institute of International Information and Telecommunication of Waseda University, Tokyo, Japan during June 2001 and May 2002. He was a collaborative researcher at the Hong Kong Polytechnic University and the United State University of New York at Buffalo in 2008 and 2009, respectively. He is an expert in the Chinese Digital Audio Video Standardization Working Group, and has published 150 research papers, 1 monograph, and 9 National Invention Patents. He held over 25 significant projects, including the National 973 Project, 863 Project, the National Natural Science Foundation of China, the Natural Science Foundation of Beijing, and the Science Foundation of Education Committee of Beijing.

Dr. Jia won the National Excellent Teacher Award, the Beijing Outstand and Innovator Award, the Beijing Excellent Teacher and Innovator Award, etc. The group under his leadership won the Beijing Top-notch creative team Award in 2010. His research interests include Image/Video processing, Content-based Image/Video Retrieval and Multimedia Information Processing.

Chang-Hong Fu received the B.Eng. (with first class honours) degree and the Ph.D degree from the Hong Kong Polytechnic University in 2002 and 2008, respectively.

He entered South East University in 1998 and transferred to the Hong Kong Polytechnic University in 1999, by the support of Hong Kong Jockey Club Scholarship for outstanding Mainland students. During his studies, he was the recipients of several scholarships for his outstanding academic achievements. He joined Nanjing University of Science and

Technology in 2011, and is now an Associate Professor in School of Electronic and Optical Engineering. He has published over 20 research papers in various international journals and conferences. His research and technical interests include the area of multimedia technologies, signal processing, image and video compression, video transcoding, video streaming, bitstream switching, digital VCR (Video Cassette Recoding), multi-view/3D video coding, and future video coding standards.

Wan-Chi Siu (S'77–M'77–SM'90) received the M.Phil. degree from The Chinese University of Hong Kong in 1977, and the Ph.D. Degree from the Imperial College of Science, Technology & Medicine, University of London, U.K., in October 1984.

He joined The Hong Kong Polytechnic University as a Lecturer in 1980, and has become Chair Professor in the Department of Electronic and Information Engineering (EIE) since 1992. He was Head of EIE and subsequently Dean of Engineering Faculty

between 1994 and 2002. He is now Director of the Center for Signal Processing of the same university. He is an expert in Digital Signal Processing, specializing in fast algorithms and video coding, and has published 380 research papers, over 160 of which appeared in international journals, such as IEEE TRANSACTIONS ON IMAGE PROCESSING. His research interests include also transforms, image coding, transcoding, 3D videos, wavelets, and computational aspects of pattern recognition. His work on fast computational algorithms (such as DCT) and motion estimation algorithms have been well received by academic peers, with good citation records, and a number of which are now being used in hi-tech industrial applications, such as modern video surveillance and video codec design for HDTV systems of some million dollar contract consultancy works. Prof. Siu is now Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING, and member of editorial board of a number of other journals such as Journal of VLSI Signal Processing Systems for Signal, Image, Video Technology, etc. He was Guest Editor and Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS. He is a co-editor of the book, 'Multimedia Information Retrieval and Management', Springer Berlin Heidelberg 2003. He is a very popular lecturing staff member within the University, while outside the University he has been a keynote speaker of over 10 international/national conferences in the recent 10 years, and an invited speaker of numerous professional events, such as IEEE CPM'2002 (keynote speaker, Taipei, Taiwan), IEEE ISIMP'2004 (keynote speaker, Hong Kong), and IEEE ICICS'07 (invited speaker, Singapore) and IEEE ICNNSP'2008 (keynote speaker, Zhenjiang). He is the organizer of many international conferences, including the MMSP'08 (Australia) as General Co-Chair, and three IEEE Society sponsored flagship conferences: ISCAS'1997 as Technical Program Chair; ICASSP'2003 as the General Chair; and recently ICIP'2010 as the General Chair (2010 IEEE International Conference on Image Processing, which was held in Hong Kong, 26–29 September 2010). Professor Siu is also Vice President-Conference (Elect) of the IEEE Signal Processing Society (2012–2014). He is a member (2010–2012) of the Engineering Panel and also was a member of the Physical Sciences and Engineering Panel (1991–1995) of the Research Grant Council (RGC), Hong Kong Government. In 1993/4, he chaired the first Engineering and Information Technology Panel of the Research Assessment Exercise (RAE) to assess the research quality of 19 departments from all universities in Hong Kong, and initiated to set up objective indicators to assess the basic research quality of academia, which gives substantial impact to the research culture in Hong Kong for the recent 18 years.

Prof. Siu is the recipient of a number of awards, including the Distinguished Presenter Award (1997, HK), IEEE Third Millennium Medal (2000, USA), the Best Teacher Award (2003, HK), the Outstanding Award in Research (2003, HK), Plaque for Exceptional Leadership (2003, IEEE SPCB, USA), and Honorable Mention Winner Award (Pattern Recognition, 2004, USA). For more information, please see <http://www.eie.polyu.edu.hk/~wcsiu/mypage.htm>