An Efficient and Robust Adaptive Deinterlacing Technique

Qian Huang, *Student Member*, IEEE, Wen Gao, *Member*, IEEE, Debin Zhao, and Huifang Sun, *Fellow*, IEEE

Abstract — *This paper presents an efficient and robust adaptive deinterlacing technique that consists of three main steps, i.e. motion estimation, motion detection (MD), and adaptive decision-making. Firstly, five successive interlaced fields are used and four unidirectional motion vectors are calculated. Secondly, a probabilistic criterion is introduced to select the most reliable motion vector for each block before MD, which has four possible output motion states. Finally, an adaptive deinterlacing scheme is proposed based on the motion and texture information of each block, as well as the output of MD. Experimental results show that the proposed deinterlacing technique outperforms previous methods both objectively and subjectively. The analysis also shows that the proposed technique is robust to both unconventional motions and the video codec*[1].

Index Terms — Deinterlacing, motion estimation, motion vector reliability, motion detection.

I. INTRODUCTION

Due to the adoption of interlaced scan, current TV systems suffer from well-known visual artifacts such as interline twitter, line crawling, and field aliasing. Deinterlacing, which is the conversion from interlaced videos to progressive ones, aims at reducing these artifacts. Over the last three decades, various deinterlacing techniques have been proposed [1]-[4]. These techniques can be roughly classified into motion compensated (MC) ones and non-MC ones. Although MC methods generate better results in general, they are threatened by latent error propagations, and it is not always the case that we have the necessity to afford their space-time cost. Therefore, we need to introduce non-MC approaches sometimes, especially in real-time applications such as video compression and transcoding [5]. Despite the existence for years of the idea of combining MC and non-MC deinterlacing algorithms, there are no more than two hybrid strategies, i.e. weighted averaging [6], [7] and adaptive switching [8]-[10]. Since weighted averaging tends to introduce blurring, the latter one is preferable for robust and high-quality deinterlacing.

[1] This work is supported in part by the National Natural Science Foundation of China under Grant No. 60333020 and the Natural Science Foundation of Beijing under Grant No. 4041003.

Q. Huang, W. Gao, and D. Zhao are with the Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080 China, and Graduate School of the Chinese Academy of Sciences, Beijing 100039 China (e-mail: qhuang@jdl.ac.cn; wgao@jdl.ac.cn; dbzhao@jdl.ac.cn).

H. Sun is with the Mitsubishi Electric Research Laboratories, Cambridge, MA 02139 USA (e-mail: hsun@merl.com).

A well-performed adaptive deinterlacing scheme usually consists of the following three steps: motion estimation (ME), motion detection (MD), and adaptive MC/non-MC decision-making. The ME can be performed between either opposite-parity fields [11], [12] or same-parity fields [13]-[15]. Same-parity ME works better when the vertical component of motion vector (MV) is even (e.g. stationary areas), whereas opposite-parity ME is more efficient when the vertical component of MV is odd. Therefore, both kinds of ME methods should be considered to get the most reliable MV for each region. And a criterion is consequently needed to evaluate the MV reliability (MVR).

The task of the second step, MD, is to determine the motion state for each region. Jung [2] and Tai [9] defined two output states, i.e. stationary and moving. Lee [3] subdivided the input image into moving, background, and boundary regions. Huang [10] proposed a block-based directional adaptive MD (DAMD) method, but it did not work well to detect unconventional motions. Chang [4] tried to utilize global motion information—rotation, zooming, and panning, but the objective results seemed to be worse than those of traditional local motion compensation. As the detection of unconventional motions [16], [17] is rather time-consuming, more historical and future information can be used to alleviate the problem of such motions.

As for the final step, adaptive MC/non-MC decision-making, MC or non-MC candidates are chosen based on the motion and texture information of each region, as well as the results of MD. Different deinterlacing modes should be adaptively imposed on individual regions. Jung [2] and Tai [9] proposed to use intra-field interpolation for moving areas and inter-field interpolation for stationary ones. Kwon [8] used a finer-grained classification that defined three more moving modes, namely forward prediction, backward prediction, and bidirectional prediction. Huang [10] proposed to perform unidirectional interpolation for top fields and bidirectional interpolation for bottom fields. Note that in each case two MC candidates and two non-MC candidates were chosen based on the experimental results of [1].

This paper presents an adaptive and robust deinterlacing technique that selectively uses two kinds of ME. A probabilistic MVR criterion is used to reserve the most reliable MV for each block. The rest of this paper is organized as follows. Section II describes the two kinds of ME, as well as the subsequent MVR criterion and MD step. The adaptive decision-making step is presented in section III. Section IV shows the experimental results and compares our deinterlacing technique with existing algorithms. Finally, the conclusions are drawn in section V.

Contributed Paper
Manuscript received April 20, 2006

II. Motion Estimation and Motion Detection

A. Two Kinds of Motion Estimation

Traditional ME methods [10]-[15] used no more than three fields, thus are not robust to unconventional motions such as rotation and zoom. Chang [4] proposed a same-parity 4-field ME that had benefits in stationary areas. However, opposite-parity ME is also indispensable as stated above. In this subsection, we propose to take both kinds of ME into account, using five successive fields, as illustrated in Fig. 1.

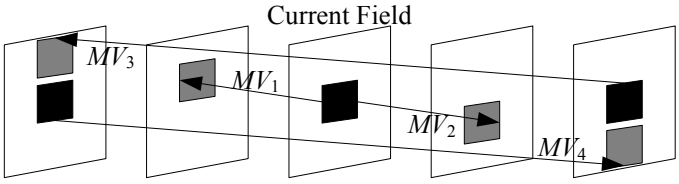

Fig. 1. Two kinds of ME. The three black blocks are co-located, whereas the gray blocks are referenced.

As can be seen from Fig. 1, the two kinds of ME are both unidirectional. MV_1 and MV_2 are calculated between opposite-parity fields, whereas MV_3 and MV_4 are calculated between same-parity fields. Note that the first two fields have already been deinterlaced.

B. Probabilistic MVR Criterion

Previously MVR was measured for adaptive switching between different deinterlacing techniques after the unique MV of each block had been determined [18], [19]. In this subsection, we measure the MVR to determine the final MV for each block after the above-mentioned two kinds of ME.

Before measuring the MVR, each of the four MV candidates MV_i (i=1, 2, 3, 4) is mapped as an opposite-parity MV candidate MVC_i of the current block:

$$MVC_i = (-1)^{i+1} \cdot MV_i, \tag{1}$$

where MVC_i is the forward MV candidate between opposite-parity fields, as depicted in Fig. 2.

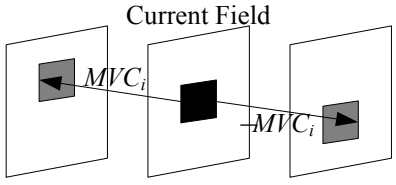

Fig. 2. Forward MV candidate between opposite-parity fields. The MV consistency is assumed.

Now we need a MVR factor to decide which MV candidate is the most reliable one. This can be achieved by considering the additive noise that is signal-independent. Assume that the interlaced video sequence is interrupted by additive white Gaussian noise. We can get the following:

$$f_n(x, y) = g_n(x, y) + v_n(x, y), \tag{2}$$

where (x, y) designates the spatial position. $f_n(x, y)$, $g_n(x, y)$ and $v_n(x, y)$ are the obtained image, original image and noise of the current nth field, respectively.

Suppose that the forward true MV $d_n(x, y)$ has already been successfully estimated between $f_n(x, y)$ and $f_{n-1}(x, y)$. We have:

$$f_{n-1}(x', y') = g_n(x, y) + v_{n-1}(x', y'), \tag{3}$$

where $x' = x + d_{nx}(x, y)$, $y' = y + d_{ny}(x, y)$. $d_{nx}(x, y)$ and $d_{ny}(x, y)$ are horizontal and vertical components of $d_n(x, y)$, respectively.

From (2) and (3) we can obtain:

$$f_n(x, y) - f_{n-1}(x', y') = v_n(x, y) - v_{n-1}(x', y'). \tag{4}$$

Since a two-dimensional matrix can be considered as a one-dimensional array by appending successive lines, we can assume that $v_n(x, y) \sim N(\mu_1, \sigma_1^2)$, $v_{n-1}(x', y') \sim N(\mu_2, \sigma_2^2)$. Let $Z_n(x, y) = f_n(x, y) - f_{n-1}(x', y')$. Due to the signal-independent characteristic of noise, we have:

$$Z_n(x, y) \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2). \tag{5}$$

Let $P_n(x, y)$ denote the conditional probability that measures how well f_n can be described by $d_n(x, y)$ and f_{n-1}. Then the following is approximately correct based on [19]:

$$Const + \ln P_n(x, y) \propto \frac{delta + SD^2}{delta + \sum_{(x,y)\in Bi} Z_n^2(x, y)}, \tag{6}$$

with *Const* a constant, *Bi* the current interlaced MC block and *delta* a very small positive constant used to avoid division by zero. SD^2 is defined as:

$$SD^2 = \frac{1}{N}\sum_{(x,y)\in Bi} Z_n^2(x, y) - \left(\frac{1}{N}\sum_{(x,y)\in Bi} Z_n(x, y)\right)^2, \tag{7}$$

where N is a constant indicating the number of pixels in *Bi*.

Given the current field numbered n, the following holds when substituting (7) into (6):

$$Const + \ln P_n(x, y) \propto \left[N - \frac{delta + \left(\sum_{(x,y)\in Bi} Z_n(x, y)\right)^2}{delta + \sum_{(x,y)\in Bi} Z_n^2(x, y)} \right]. \tag{8}$$

Since logarithm is a monotonically increasing function, the MV candidate that makes the right hand side of (8) largest will be chosen as the unique MV for the current block.

C. Motion Detection

At this point, the final MV for each block has already been determined. Note that the final MV is between successive opposite-parity fields, as shown in Fig. 3, where SAD_i (i=1, 2, 3, 4) can be formalized as follows. Hereinafter, each field will be considered as a full resolution frame that loses half the data.

$$SAD_i = \sum_x \sum_y \left\{ \left[0.5 - (-1)^n (y\%2 - 0.5) \right] \cdot \left| f_n(x,y) - f_{n+(-1)^i}(x+t_i, y+s_i) \right| \right\}, \quad (9)$$

where t_i and s_i designate the temporal and spatial offset from the current block, respectively. The positive directions are shown below.

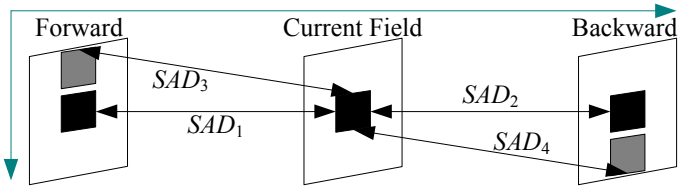

Fig. 3. Blocks used for MD. The three black blocks are co-located, whereas the gray blocks are referenced.

We define four possible output states of MD (*stationary*, *undirectionally moving*, *horizontally moving* and *vertically moving*) which will be detailed in the rest of this subsection.

If the average of SAD_3 and SAD_4 is not less than that of SAD_1 and SAD_2, which indicates that the similarity along motion trajectory is smaller than temporal similarity, we have good reason to believe that the current block is not moving. However, out of regard for the fact that large SAD does not necessarily lead to small similarity due to brightness change as shown in Fig. 4, the standard deviation σ of the current block should also be taken into account to rationalize our criterion.

255	255	255	255
255	255	255	255
255	255	255	255
255	255	255	255

253	253	253	253
253	253	253	253
253	253	253	253
253	253	253	253

Fig. 4. Small brightness change can also lead to large SAD.

The current block is considered *stationary* if:

$$SAD_1 + SAD_2 \le \min\{SAD_3 + SAD_4, 2\cdot\sigma\}, \quad (10)$$

where σ is a nonnegative value whose square satisfies:

$$\sigma^2 = \left[0.5 - (-1)^n (y\%2 - 0.5) \right] \cdot \left[N \sum_x \sum_y f_n^2(x,y) - \left(\sum_x \sum_y f_n(x,y) \right)^2 \right]. \quad (11)$$

As for nonstationary blocks, the output state depends on the ratio of (MV_x + *delta*) and (MV_y + *delta*). MV_x and MV_y are horizontal and vertical components of the final MV obtained according to the last subsection, respectively. If the ratio is less than or equal to 0.2, the output state is *vertically moving*. If the ratio is larger than or equal to 5, the output state is *horizontally moving*. Otherwise the output state of MD will be *undirectionally moving* except when $SAD_3 + SAD_4 \ge 2\cdot\sigma$, in which case the current block is still thought to be *stationary*.

III. Adaptive Deinterlacing

For clarity, the deinterlacing scheme for stationary blocks and that for moving blocks are discussed separately.

A. Deinterlacing for Stationary Blocks

Although field insertion is theoretically the best method for stationary regions, undesirable serrations will occur if we mistake a moving region as stationary. For safety, the median filtering algorithm shown below is used for stationary blocks.

$$MF_n(x,y) = \begin{cases} f_n(x,y), & if\ y\%2 = n\%2 \\ Med(up, down, C, D, \dfrac{up+down}{2}), & o.w. \end{cases}, \quad (12)$$

where *up*, *down*, *C*, and *D* represent the values of the spatially above, spatially below, forward referenced and backward referenced pixels, respectively.

B. Deinterlacing for Undirectionally Moving Blocks

The adaptive recursive (AR) method proposed in [20] is one of the best deinterlacing algorithms nowadays. However, its high performance is based on accurate ME. Thus the threat of error propagations is more pressing for AR than that for many other MC methods, because AR is also a time-recursive approach [21]. Therefore, an edge-based median filtering (EMF) algorithm is introduced below to alleviate the influence of error propagations on AR.

As shown in Fig. 5, we use directional correlations in sub-pixel precision. For each interpolated pixel in the current field f_n, nine possible edge directions are considered. The sub-pixel values are calculated using a linear interpolation, e.g.:

$$A = (f_n(x-1, y-1) + f_n(x, y-1))/2. \quad (13)$$

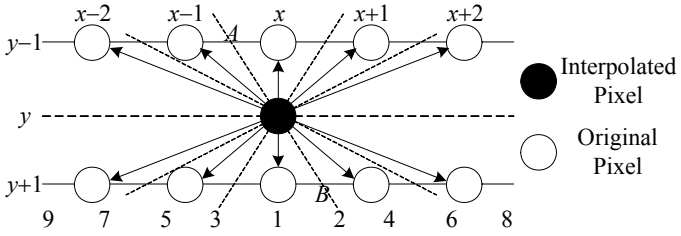

Fig. 5. EMF in sub-pixel precision. *A* and *B* are sub-pixel values.

To detect the edge, we employ a zigzag search heuristic whose search order is numbered above. The direction with minimal absolute difference will be reported as the dominant edge direction at the end of the search. Since the sub-pixel values are not original, we consider the first horizontal differences to exclude impossible sub-pixel directions so as to minimize mistakes. For example, the first horizontal difference $\Delta_x A$ is defined as:

$$\Delta_x A = |f_n(x-1, y-1) - f_n(x, y-1)|. \quad (14)$$

We assume that there is no edge in position A if $\Delta_x A$ is less than 15. If there is no edge in position B either, we exclude the possibility that AB is the dominant edge. The other sub-pixel directions can be dealt with by the same token.

Finally the dominant edge direction will be determined for each interpolated pixel. Let *ae* denote the average of the above and below pixel values in the dominant edge direction. Then the output of our EMF algorithm is defined as:

$$EMF_n(x,y)=\begin{cases} f_n(x,y), \ if \ \ y\%2=n\%2 \\ Med(ae,\dfrac{up+down}{2},C,D,\dfrac{C+D}{2}), \ o.w. \end{cases}. \quad (15)$$

Let $Edge_n(x, y)$ denote the binary value returned by Canny edge detection [22], indicating if there is any edge information in the current block. To speed up the implementation, the edge information from the forward deinterlaced field can be used if the motion is not acute. Thus the deinterlacing scheme for undirectionally moving blocks can be defined as follows.

$$\begin{aligned} UF_n(x,y) = {} & Edge_n(x,y)\cdot AMC_n(x,y) \\ & +[1-Edge_n(x,y)]\cdot TAR_n(x,y) \end{aligned}. \quad (16)$$

If there is no edge information in the current block, we adaptively use two MC approaches, i.e. AR and EMF. The adaptive MC (AMC) scheme is defined as:

$$AMC_n(x,y)=\alpha\cdot AR_n(x,y)+(1-\alpha)\cdot EMF_n(x,y), \quad (17)$$

where $AR_n(x, y)$ is the deinterlacing output of AR. α can be formulated as:

$$\alpha = clip(0,1,clip(0,1,MV_n(x,y)-4)), \quad (18)$$

with $MV_n(x, y)$ the absolute value of the current MV. The *clip* function is given as follows.

$$clip\ (a,b,c)=\begin{cases} a, \ if \ \ c<a \\ b, \ else \ if \ \ c>b \\ c, \ o.w. \end{cases}. \quad (19)$$

If there exists edge information in the current block, we take the texture information into account and use the following texture-based adaptive recursive (TAR) scheme:

$$\begin{aligned} TAR_n(x,y) = {} & Smooth_n(x,y)\cdot LA_n(x,y) \\ & +[1-Smooth_n(x,y)]\cdot AMC_n(x,y) \end{aligned}, \quad (20)$$

where $Smooth_n(x, y)$ and $LA_n(x, y)$ are defined in (21) and (22), respectively.

$$Smooth_n(x,y)=floor(clip(0,1,\frac{SAD_3+SAD_4+delta}{\sigma+delta})), \quad (21)$$

with *floor*(*x*) a function that returns the largest integer less than or equal to *x*, for any real number *x*.

$$LA_n(x,y)=\begin{cases} f_n(x,y), \ if \ \ y\%2=n\%2 \\ \dfrac{f_n(x,y-1)+f_n(x,y+1)}{2}, \ o.w. \end{cases}. \quad (22)$$

C. Deinterlacing for Horizontally Moving Blocks

As shown in [1], the linear spatio-temporal filtering (LSTF) algorithm is the best non-MC approach for horizontal motion. In this subsection, the LSTF proposed in [20] is adopted as the only non-MC candidate for horizontally moving blocks. However, as shown in Fig. 6, LSTF is typically weak on vertically moving regions and introduces strong noise. The median filter in (23) can effectively eliminate such noise, as shown in Fig. 9(e).

Fig. 6. Strong noise introduced by LSTF in the *Basketball* sequence with SD resolution.

The deinterlacing scheme for horizontally moving blocks is formulated as:

$$\begin{aligned} HF_n(x,y) = {} & \beta\cdot UF_n(x,y) \\ & +(1-\beta)\cdot Med(Ae,Be,C,D,LSTF_n(x,y)) \end{aligned}, \quad (23)$$

where *Ae* and *Be* are values of the above and below pixels in the dominant edge direction, respectively. $LSTF_n(x, y)$ is the deinterlacing output of LSTF. β is defined as:

$$\beta = floor(clip(0,1,\frac{SAD_1+SAD_2+delta}{SAD_3+SAD_4+delta})). \quad (24)$$

D. Deinterlacing for Vertically Moving Blocks

Experiments suggest that line averaging (LA) [23] or its improved version edge-based LA (ELA) [24] is generally the best non-MC approach. However, the performance of ELA depends on the implementation to a great extent. Thus in this subsection we simply choose the LA algorithm as the only non-MC candidate.

The deinterlacing scheme for vertically moving blocks is:

$$VF_n(x,y)=\gamma\cdot UF_n(x,y)+(1-\gamma)\cdot LA_n(x,y), \quad (25)$$

where γ is defined as:

$$\gamma = floor(clip(0,1,\frac{2\cdot\sigma + delta}{SAD_3 + SAD_4 + delta})) . \quad (26)$$

IV. EXPERIMENTAL RESULTS

In this section, the proposed deinterlacing technique is evaluated both objectively and subjectively. Several other methods are also implemented for comparison.

A. Objective Performance

In the literature on deinterlacing, the mean square error (MSE) or the peak signal to noise ratio (PSNR) in decibel (dB) is most frequently used as an objective performance criterion [1]-[4], [8]-[12]. Both criteria require progressively scanned original video sequences. In this subsection, six progressive sequences in CIF format (*Foreman*, *News*, *Bus*, *Mobile Calendar*, *Paris*, and *Tempete*) and nine progressive sequences in HD format (*City*, *Crew*, *Cyclists*, *Jets*, *Night*, *Optis*, *Sailormen*, *Sheriff*, and *Spin Calendar*) are selected to generate interlaced videos by dropping one field in each frame, as illustrated in Fig. 7.

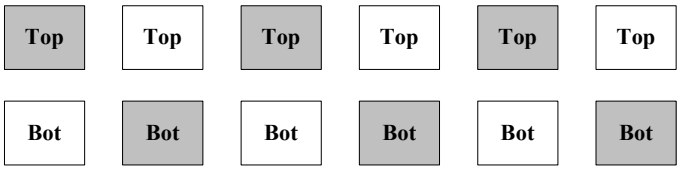

Fig. 7. Extract interlaced videos by dropping one field in each frame. Fields in white will be removed, whereas fields in gray will be directly merged.

PSNR comparison between original progressive sequences and deinterlaced ones is shown in Fig. 8, where eight deinterlacing methods are compared: LA, LSTF [20], conventional EMF [10], the sub-pixel EMF [25], AR [20], adaptive mechanism based on DAMD (Adaptive DAMD) [10], adaptive deinterlacing using MVR criterion proposed in [25] (Adaptive MVR), and the proposed adaptive deinterlacing selectively using two kinds of ME (Adaptive Two ME). Note that each progressive CIF sequence contains 150 frames, whereas each progressive HD sequence contains 50 frames.

The results show that the proposed adaptive deinterlacing technique that selectively using two kinds of ME generally has better objective performance than other methods. Though we obtain only the second highest PSNRs on sequences *Optis*, *Sheriff*, *City*, and *Spin Calendar*, we get the highest PSNRs in all the other cases. Note that on sequences *Optis*, *Sheriff*, *City*, and *Spin Calendar*, the highest PSNRs are also obtained using our methods, e.g. the sub-pixel EMF method, whose subjective performance is not as good as that of the proposed scheme, as shown in Fig. 9. The average PSNRs over all the fifteen sequences are shown in Table I.

From Fig. 6 and Fig. 9(e), we can see that the noise caused by LSTF can be effectively removed by the EMF method used in (23). Table II shows that significantly better objective results can also be obtained. The seven CIF sequences tested are the same as those stated above.

B. Subjective Performance

Fig. 9 depicts the subjective performance on the *Basketball* sequence in SD format. Fig. 9(a) is obtained by directly merging two interlaced fields together and shows the worst subjective effect. Fig. 9(b) suggests that LA is generally a very good non-MC deinterlacing algorithm, but it cannot work well on edges that are not horizontal. Fig. 9(c) shows that there are indeed error propagations for MC methods such as AR. Thus the sub-pixel EMF algorithm shown in Fig. 9(d) is indispensable. Note that there are problems with the diagonal lines in the first three figures, whereas we successfully remove such artifacts in the next three figures.

TABLE I
AVERAGE PSNR OVER ALL THE FIFTEEN SEQUENCES

Deinterlacing Method	Average PSNR (dB)
LA	32.76
LSTF	32.12
Conventional EMF	34.71
Sub-pixel EMF	35.11
AR	35.40
Adaptive DAMD	35.78
Adaptive MVR	36.11
Adaptive Two ME	36.30

From Fig. 8, Table I, and Fig. 9(d) we can see that the sub-pixel EMF outperforms conventional EMF methods. We do not need to worry about the artifacts introduced by sub-pixel EMF, since they are guaranteed to be eliminated by the proposed adaptive scheme.

The proposed adaptive deinterlacing technique effectively combines the advantages of LA, LSTF, AR, and sub-pixel EMF while reducing the artifacts caused by each of them. Thus we generate perfect performance, especially in the areas with continuous motions where most of flickers and serrations have been eliminated.

The robustness of the proposed adaptive technique can be interpreted in three ways. Firstly, we use two more fields than traditional ME methods. Hence the proposed deinterlacing is more robust to unconventional motions, since more historical and future information is utilized. Secondly, the adaptive scheme makes our deinterlacing technique more robust to inaccurate MVs than conventional MC approaches. For example, if we implement both AR and the proposed technique in the MPEG2 decoder, we can find that the average PSNR gap is much more obvious due to the inaccuracy of MVs. Thirdly, the proposed adaptive deinterlacing technique is more suitable for video coding than conventional MC methods, as can be seen in Fig. 10, where four SD sequences (*Basketball*, *Flower Garden*, *Horse Riding*, and *Mobile Calendar*) and three D1 sequences (*Football*, *News Football*, and *Interview*) are coded after deinterlacing. The improved LSTF algorithm is used as anchor, and the average PSNR gain of each method over all the seven sequences is shown in Table III.

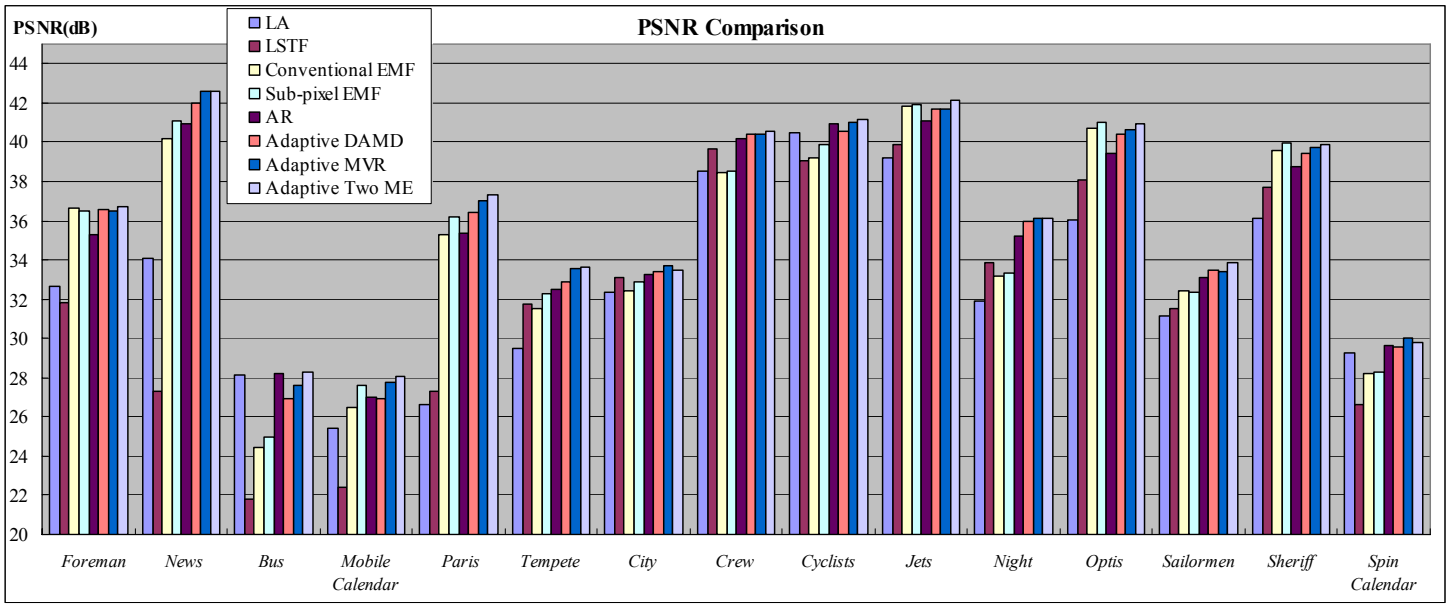

Fig. 8. PSNR comparison between original progressive sequences and deinterlaced ones. Eight deinterlacing methods are compared.

Fig. 9. Subjective performance on the *Basketball* sequence with SD resolution. (a) "Weave" [23], (b) LA, (c) AR, (d) sub-pixel EMF, (e) LSTF + median filtering, (f) proposed adaptive deinterlacing technique selectively using two kinds of motion estimation.

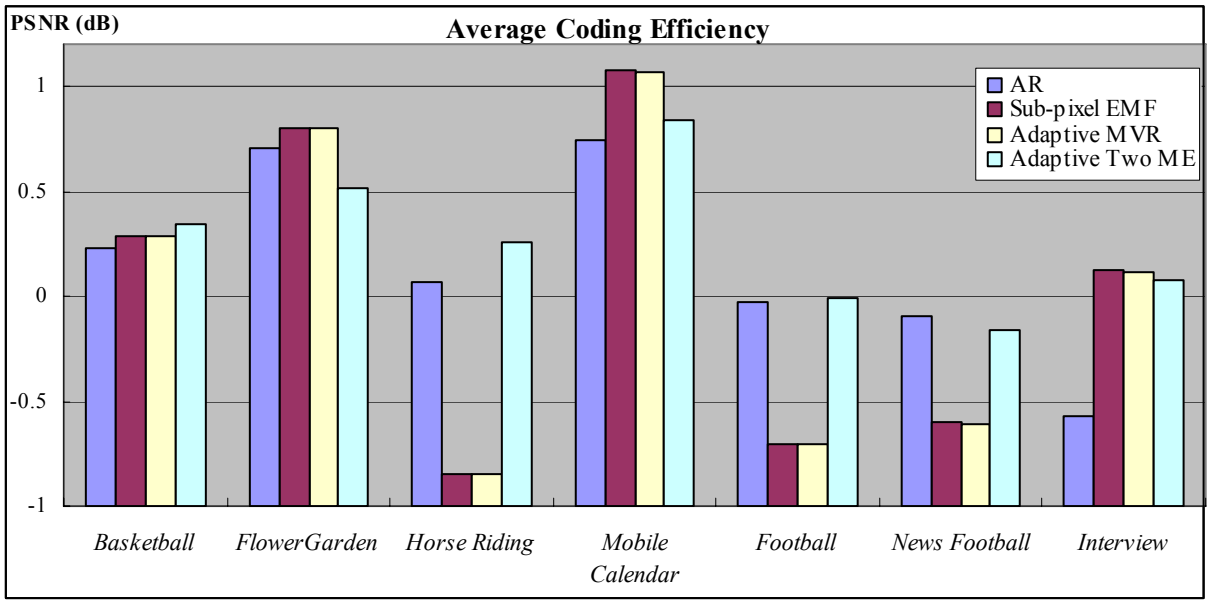

Fig. 10. Coding efficiency comparison after deinterlacing. The improved LSTF algorithm is used as anchor. An encoder of the Audio Video coding Standard (AVS) of China, which targets at higher coding efficiency and lower complexity than existing standards for HD video coding, is employed. For each deinterlacing method, the first 100 frames of the seven deinterlaced sequences are coded at fixed QPs 28, 32, 36, 40, using IBBP mode as the basic coding structure. Note that all the seven interlaced sequences used here are decoded from MPEG2 streams.

TABLE II
PSNR COMPARISON BETWEEN ORIGINAL LSTF AND IMPROVED LSTF

Sequence	Original LSTF PSNR (dB)	LSTF + EMF in (23) PSNR (dB)
Foreman	31.79	36.60
News	27.30	38.12
Bus	21.81	27.47
Football	30.80	32.53
Mobile Calendar	22.41	26.48
Paris	27.33	33.75
Tempete	31.76	32.46

TABLE III
AVERAGE PSNR GAIN OVER ALL THE SEVEN SEQUENCES

Deinterlacing Method	Average Gain (dB)
AR	0.149246
Sub-pixel EMF	0.020508
Adaptive MVR	0.014648
Proposed Adaptive Two ME	0.264949

V. CONCLUSION

In this paper, we have proposed an efficient and robust adaptive deinterlacing technique, which consists of three steps—ME, MD, and adaptive decision-making. In the ME step, two kinds of unidirectional ME are selectively used and only the most reliable MV for each block is reserved based on a probabilistic MVR criterion. Then four possible motion states are defined in the MD step. In the decision-making step we adaptively switch among four deinterlacing algorithms based on the motion and texture information of each block.

Experimental results show that our objective and subjective performance are both better than conventional approaches. Since only one deinterlacing algorithm will be performed in each region, we also have lower implementation complexity than weighted averaging algorithms such as [6], [7], and [19]. Furthermore, the proposed adaptive technique is robust to unconventional motions, inaccurate MVs, and the video encoder.

ACKNOWLEDGMENT

The authors would like to thank the editor and the anonymous reviewers for their invaluable comments and suggestions.

REFERENCES

[1] G. de Haan and E. B. Bellers, "Deinterlacing—an overview," *Proc. IEEE*, vol. 86, no. 9, pp. 1839-1857, Sept.1998.

[2] Y. Y. Jung, S. Yang, and P. Yu, "An effective de-interlacing technique using two types of motion information," *IEEE Trans. Consum. Electron.*, vol. 49, no. 3, pp. 493-498, Aug. 2003.

[3] S. G. Lee and D. H. Lee, "A motion-adaptive de-interlacing method using an efficient spatial and temporal interpolation," *IEEE Trans. Consum. Electron.*, vol. 49, no. 4, pp. 1266-1271, Nov. 2003.

[4] Y. L. Chang, S. F. Lin, C. Y. Chen, and L. G. Chen, "Video de-interlacing by adaptive 4-field global/local motion compensated approach," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 12, pp. 1569-1582, Dec. 2005.

[5] A. Vetro, C. Christopoulos, and H. Sun, "Video transcoding architectures and techniques: an overview," *IEEE Signal Process. Mag.*, vol. 20, no. 2, pp. 18-29, Mar. 2003.

[6] A. Nguyen and E. Dubois, "Spatio-temporal adaptive interlaced-to-progressive conversion," in *Signal Process. HDTV, IV*, E. Dubois and L. Chiariglione, Eds. The Netherlands: Elsevier, 1993, pp. 749-756.

[7] J. Kovacevic, R. J. Safranek, and E. M. Yeh, "Deinterlacing by successive approximation," *IEEE Trans. Image Process.*, vol. 6, no. 2, pp. 339-344, Feb. 1997.

[8] S. Kwon, K. Seo, J. Kim, and Y. Kim, "A motion-adaptive de-interlacing method," *IEEE Trans. Consum. Electron.*, vol. 38, no. 3, pp. 145-150, Aug. 1992.
[9] S. C. Tai, C. S. Yu, F. J. Chang, "A motion and edge adaptive deinterlacing algorithm," in *Proc. IEEE Int. Conf. Multimedia Expo*, Taipei, Taiwan, Jun. 2004, pp. 659-662.
[10] Q. Huang, W. Gao, D. Zhao, and H. Sun, "Adaptive deinterlacing for real-time applications," in *Proc. Pacific-Rim Conf. Multimedia*, *LNCS3768*, Y.-S. Ho and H. J. Kim, Eds. Heidelberg: Springer-Verlag, 2005, pp. 550-560.
[11] G. de Haan, P. W. A. C. Biezen, H. Huijgen, and O. A. Ojo, "True-motion estimation with 3-D recursive search block matching," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 3, no. 5, pp. 368-379, Oct. 1993.
[12] G. de Haan and P. W. A. C. Biezen, "Sub-pixel motion estimation with 3-D recursive search block-matching," *Signal Process.: Image Commun. 6*, pp. 229-239, 1994.
[13] P. Delogne, P. Francis, L. Vandendorpe, and F. Vermaut, "Conversion from interlaced to progressive formats by means of motion compensation based techniques," in *Proc. IEE Int. Conf. Image Process. Appl.*, Maastricht, The Netherlands, Apr. 1992, pp. 409-412.
[14] M. Biswas and T. Nguyen, "A novel de-interlacing technique based on phase plane correlation motion estimation," in *Proc. IEEE Int. Symp. Circuits Syst.*, Bangkok, Thailand, May 2003, pp. 604-607.
[15] S. Li, J. Du, D. Zhao, Q. Huang, and W. Gao, "An improved 3DRS algorithm for video de-interlacing," in *Proc. Picture Coding Symp.*, Beijing, China, Apr. 2006, to appear.
[16] H. Onishi and H. Suzuki, "Detection of rotation and parallel translation using Hough and Fourier transforms," in *Proc. IEEE Int. Conf. Image Process.*, Lausanne, Switzerland, Sept. 1996, pp. 827-830.
[17] R. Jin, Y. Qi, and A. Hauptmann, "A probabilistic model for camera zoom detection," in *Proc. IEEE Int. Conf. Pattern Recognit.*, Quebec, Canada, Aug. 2002, pp. 859-862.
[18] O. A. Ojo and G. de Haan, "Robust motion-compensated video upconversion," *IEEE Trans. Consum. Electron.*, vol. 43, no. 4, pp. 1045-1056, Nov. 1997.
[19] D. Wang, A. Vincent, and P. Blanchfield, "Hybrid de-interlacing algorithm based on motion vector reliability," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 8, pp. 1019-1025, Aug. 2005.
[20] G. de Haan and E. B. Bellers, "Deinterlacing of video data," *IEEE Trans. Consum. Electron.*, vol. 43, no. 3, pp. 819-825, Aug. 1997.
[21] F. M. Wang, D. Anastassiou, and A. N. Netravali, "Time-recursive deinterlacing for IDTV and pyramid coding," in *Proc. IEEE Int. Symp. Circuits Syst.*, New Orleans, USA, May 1990, pp. 1306-1309.
[22] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 6, pp. 679-698, Nov. 1986.
[23] Microsoft Corp., "Broadcast-enabled computer hardware requirements," in *WinHec'97*, *Broadcast Technologies White Paper*, 1997, pp. 11-12.
[24] C. J. Kuo, C. Liao, and C. C. Lin, "Adaptive interpolation technique for scanning rate conversion," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 3, pp. 317-321, Jun. 1996.
[25] Q. Huang, W. Gao, D. Zhao, and Q. M. Huang, "An edge-based median filtering algorithm with consideration of motion vector reliability for adaptive video deinterlacing," in *Proc. IEEE Int. Conf. Multimedia Expo*, Toronto, Canada, Jul. 2006, to appear.

Qian Huang (S'06) received the B.S. degree in computer science from Nanjing University, China, in 2003. He is currently working toward the Ph.D. degree at the Institute of Computing Technology, Chinese Academy of Sciences. In 2004, he joined the Joint Research & Development Laboratory as a research assistant.

His research interests include digital image/video processing, pattern recognition and artificial intelligence.

Wen Gao (M'99) received the Ph.D. degree in computer science from Harbin Institute of Technology, China, in 1988 and the Ph.D. degree in electronics engineering, University of Tokyo, Japan, in 1991. He was a research fellow at the Institute of Medical Electronics Engineering, University of Tokyo, in 1992; a visiting scientist at the Robotics Institute, Carnegie Mellon University, in 1993; a visiting scientist at the MIT AI Lab, from May 1994 to December 1995. Now he is a professor of Peking University and the Institute of Computing Technology, Chinese Academy of Sciences and an adjunct professor in Hong Kong University of Science and Technology, China. He is the head of the Chinese National Delegation to MPEG working group (ISO/SC29/WG11) and chairs the Audio Video coding Standard (AVS) workgroup of China.

His research interests include pattern recognition, artificial intelligence, image understanding, data compression, hand gesture recognition, multimodal interface, and computer vision. He has published 4 books and over 400 scientific papers.

Debin Zhao received the B.S., M.S., and Ph.D. degrees in computer science from the Harbin Institute of Technology, China, in 1985, 1988, and 1998, respectively. He was a research fellow in the Department of Computer Science, City University of Hong Kong, from 1989 to 1993. And now he is with the Institute of Computing Technology, Chinese Academy of Sciences.

His research interests include data compression, image processing, and human–machine interface. He has published 2 books and over 60 scientific papers.

Huifang Sun (S'83-M'85-SM'93-F'00) received the Ph.D. degree in electrical engineering from University of Ottawa, Canada. In 1995, he joined Mitsubishi Electric Research Laboratories, where he now serves as deputy director of the Technology Laboratory. Prior to joining Mitsubishi, he was with the Electrical Engineering Department of Fairleigh Dickinson University and later at Sarnoff Corporation, where he received the AD-HDTV Team Award in 1992 and Technical Achievement Award for optimization and specification of the Grand Alliance HDTV video compression algorithm in 1994.

His research interests include digital video/image compression and digital communication. He has published a number of papers in these areas.