International Journal of Computer Sciences and Engineering Systems
Vol. 11 No. 2 (December, 2017)

Improved Method for Construction of Side-Information for Distributed Video Coding

Shanshan OU and Fan LIANG
Electronic and Communication Department, Sun Yat-sen UniversityGuangzhou, 510275, China
E-mail: isslf@mail.sysu.edu.cn

Abstract: *This paper concerns the issue of construction of side-information for distributed video coding (DVC). A new Motion Compensated Interpolation algorithm is introduced. In the proposed scheme, both forward and backward motion estimation (ME) is performed to construct the motion vector (MV). In addition, vector median filter is used to smooth the motion vector field in motion vectors refinement step. The algorithm has been compared with three other common MCI methods in the aspect of PSNR performance and perceptual quality. Simulation results demonstrate a better performance of the proposed method.*

Keywords: *Distributed Video Coding, Forward Motion Estimation, Backward Motion Estimation, Vector Median Filter*

1. INTRODUCTION

With the development of network and multimedia, many new applications arise, such as wireless PC cameras, mobile camera phones and multimedia sensor networks. They require simple encoder. However, the traditional video coding standards, such as MPEG and H.26x which exploit complex motion estimation at the encoder, cannot meet this need. As a new video coding paradigm, distributed video coding (DVC) is presented. Distributed video coding is based on Slepian-Wolf Theorem for lossless distributed coding and Wyner-Ziv Theorem for lossy source coding which have been proposed in 1970s [1] [2]. These theories suggest that using dependence only at decoder can get almost the same compression performance as conventional dependent source coding. This gives theory base to simple the encoder. There are two types of video streams in DVC architecture: key frames and Wyner-Ziv frames [3]. The key frames are encoded by the traditional intra-frame codec, while Wyner-Ziv frames are Wyner-Ziv encoded. At the Wyner-Ziv decoder, after the key frames decoding, the side-information is got by interpolating the adjacent decoded key frames and the Wyner-Ziv bits are recovered after turbo decoder. Then on the basis of side-information, the Wyner-Ziv frames are reconstructed. Side-information is very important in DVC. The more accurate the side-information is, the less Wyner-Ziv bits are required to recover a reliable decoding of the Wyner-Ziv frames. The simplest method is frame repeating or simple Average Interpolation [4]. A. Aaron, E. Setton, and B. Girod have proposed motion compensation interpolation (MC-I) and motion compensation extrapolation (MC-E) in [5] and [6]. In [8], Chi Wah Tang and Oscar C. Au proposed a scheme called Motion Compensated Temporal Interpolation (MCTI) to perform block based motion compensated temporal interpolation. In MCTI scheme, the $(k-1)^{th}$ frame, the k^{th} frame and the inserted frame are divided into blocks of size NxN. Forward and backward block-based motion estimation are performed to find the motion vector. Then the appropriate motion vectors are added to the candidate motion vector lists of the blocks in the inserted frame. For each block, the candidate motion vector with the largest overlapping area is chosen and the inserted frame is then generated by motion compensation. But due to the block-based processing, the interpolated frames of MCTI tend to be blocky. In [9], Byung Tae Choi, Sung-Hee Lee and Sung-Jea Ko proposed propose a new frame rate up-conversion algorithm for high quality video. In this scheme, bi-directional motion estimation (ME) is performed using the existing previous and current frames to construct the motion vector (MV) field for the frame to be interpolated. The proposed scheme is composed of three functional units: bi-directional motion estimation, spatial-temporal smoothing, and overlapped block motion compensated interpolation. The main feature of this scheme is that, unlike conventional MCI algorithms, it does not produce any overlapped pixel and hole region in the interpolated frame. Moreover, by using the overlapped block motion compensation technique, the blocking artifact of the block-based motion estimation is effectively eliminated. In [10], Mohammed E. Al-Mualla proposed a method called motion field interpolation (MFI) for converting a block-level motion field to a pixel-level motion field. First, a block-matching algorithm is used to get the block-level motion vector field. Then pixel-level motion vectors can be attained according to the block-level motion vector field. The main advantage of this method is that it provides a smoothly varying motion field. But the disadvantage of these two methods is that the interpolated frame is blurry when the motion is severe.

In this paper, a new interpolation strategy which based on motion estimation (ME) and motion compensation (MC) has been proposed. The proposed interpolation strategy can construct more accurate side-information for distributed video coding. According to our experiments, the performance of the proposed method is much better than motion compensated temporal interpolation (MCTI), frame up-conversion bi-directional motion estimation (FUBME) and motion field interpolation (MFI).

This paper is organized as follows. In Section 2, we describe the proposed interpolation strategy. The steps to get side-information are described in detail. Section 3 gives the experimental results are present. In Section 4, the paper is finally concluded.

2. PROPOSED METHOD

In the proposed method, we first perform forward and backward motion estimation on the key frames to get motion vectors. After that, we get the initial motion vector field. Then we perform a motion vector refinement algorithm to smooth the motion vector field, and with the refined motion vectors, adaptive frame interpolation is executed. Fig. 1 shows the block diagram of the overall scheme of proposed method.

The following subsections introduce the detailed steps of the proposed method

2.1 Motion Estimation

When we depose a key frame, we divide it into 8x8 blocks and operate on the block data. When perform forward and backward motion estimation on blocks, we extend the block size to 12×12. A 12×12 block size is used because it provides a good tradeoff between accuracy and complexity [7]. The criterion for block motion estimation is chosen to be mean absolute difference (MAD). For the present n^{th} frame, we denote the intensity value of the pixel with coordinates (i, j) by $F_n(i, j)$. The MAD between the block at (x, y) of the present frame and the block at $(x + dx, y + dy)$ of the previous frame can be calculated as

$$MAD_{(x,y)}(dx,dy) = \frac{1}{12^2}\sum_{i=-2}^{9}\sum_{j=-2}^{9} | F_n(x+i, y+j) - F_{n-1}(x+dx+i, y+dy+j) | \quad (1)$$

Where (x, y) is the coordinate of top left pixel of the 8×8 block.

The best matching block is defined as the block with minimum MAD value among all the locations within the search area.

2.2 Motion Vector Refinement

After motion estimation, motion vector field is constructed. However, there are still a few bad motion vectors. Fig. 2 shows an example of a single bad motion vector. It is often observed that inconsistencies or not-smoothness in the estimated vector field decreases the interpolated frame quality severely. In consistencies of the motion vector can be corrected by constraining the spatial-temporal smoothness on the motion vector field. Here, we used vector median filter in [8] on the motion vector filed to get more accurate motion vectors. Vector median filter finds the one motion vector among the eight neighbouring motion vectors that minimize

$$DFD(\vec{D}, B) = | \vec{D}(B) - \vec{D}(N_i) | \quad (2)$$

Where B and N_i are the motion vector of current block and the eight neighboring blocks respectively.

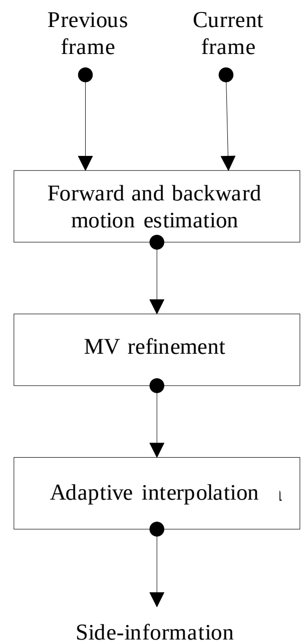

Figure 1: Proposed Method for Side-information Construction

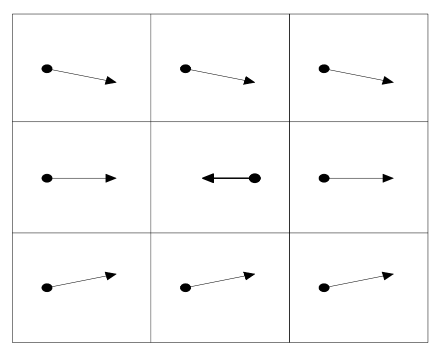

Figure 2: Motion Vector Discontinuity

2.3 Adaptive Interpolation

After the motion vector refinement, we interpolate the frame with the refined motion vector. The frame F_n to be inserted is halfway between frames F_{n-1} and F_{n+1}. Since a block $B_{n+1}(x,y)$ moves a distance V from frame F_{n+1} to frame

F_{n-1}. Assuming linear motion, a block $B_n(x, y)$ will move a distance V/2 from frame F_{n-1} to F_n and a similar distance from frame F_n to F_{n+1}. Fig. 3 shows the process of interpolating a frame F_n halfway between frames F_{n-1} and F_{n+1}.

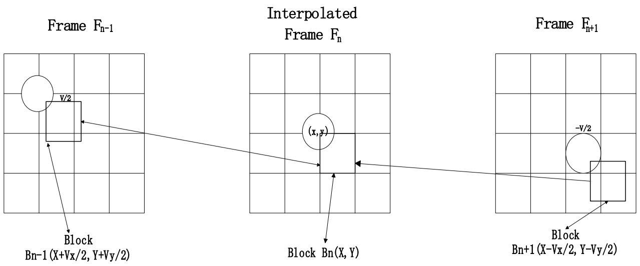

Figure 3: Motion-Compensated Interpolation

The interpolation process applies with the use of both forward and backward motion estimates. As a result, this yields two interpolated frames $F_{forward}$ and $F_{backward}$, respectively. These two frames then go through the following step to form final side information.

We use $F_{forward}$ and $F_{backward}$ to form a better quality frame.

First, a mean absolute difference (MAD) is calculated between each pair of blocks at the same location in frames $F_{forward}$ and $F_{backward}$. We set a threshold T, if MAD value dose not exceed a threshold T, the two are regarded as well-reconstructed in both frames. Between these two blocks, the one with a smaller extended block MAD in the motion estimation is chosen and copied to the final side-information frame F_{side}. If the MAD value is greater than the threshold, both the blocks in $F_{forward}$ and $F_{backward}$ are averaged and copied to the final side information frame F_{side}. This process is called adaptive interpolation.

3. EXPERIMENTAL RESULTS

In this section, we present the simulation results of the proposed method for the first 50 frames of Foreman and Paris sequences (QCIF), and we also choose Coastguard and Table tennis sequences (CIF) as the test sequences; only luminance component of each frame is accounted in.

Fig. 4 and Fig. 5 shows the PSNR comparisons for the Foreman and Paris sequence respectively. And the PSNR comparisons for CIF sequences Coastguard and Table tennis are present in Fig. 6 and Fig. 7. We can see the PSNR performance is improved when using the proposed method.

Table 1 and Table 2 summarize the simulation results of each method for Foreman and Paris sequence respectively. Table 3 and Table 4 summarize the simulation results of each method for CIF sequence Coastguard and Table tennis respectively. From the results, we can see the performance enhancement when the proposed method is used.

For comparatively evaluating the perceptual quality of the interpolated frames by different methods, Fig. 8 shows the perceptual results of the rebuilt frames for Foreman sequence. Fig. 9 presents the rebuilt frames for CIF sequence Table tennis. It is obvious that the proposed method produces a better perceptual quality with much less noticeable reconstruction artifacts and blurriness.

Table 1
PSNR Comparisons (Foreman)

Method	*Frame Up-Conversion BME*	*MCTI*	*MFI*	*Proposed Method*
Averaged PSNR(dB)	34.366	35.071	35.133	36.348

Table 2
PSNR Comparisons (Paris)

Method	*Frame Up-Conversion BME*	*MCTI*	*MFI*	*Proposed Method*
Averaged PSNR(dB)	33.86	35.071	35.133	36.068

Table 3
PSNR Comparisons (Coastguard)

Method	*Frame Up-Conversion BME*	*MCTI*	*MFI*	*Proposed Method*
Averaged PSNR(dB)	29.423	30.671	30.603	31.148

Table 4
PSNR Comparisons (Table tennis)

Method	*Frame Up-Conversion BME*	*MCTI*	*MFI*	*Proposed Method*
Averaged PSNR(dB)	22.285	27.601	27.378	29.21

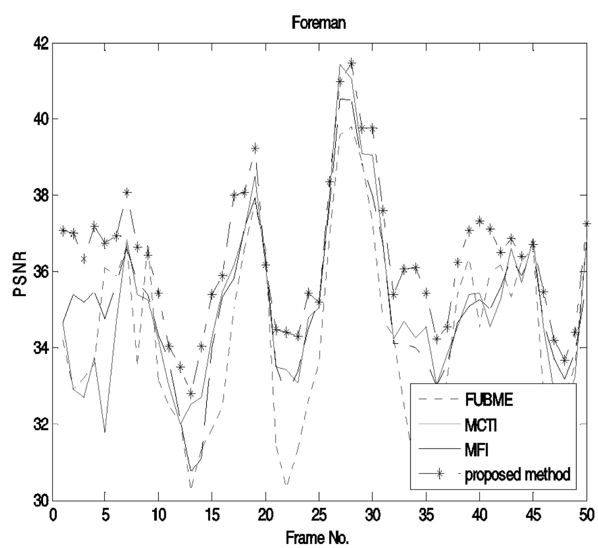

Figure 4: PSNR Comparisons of Four Methods for Foreman (QCIF)

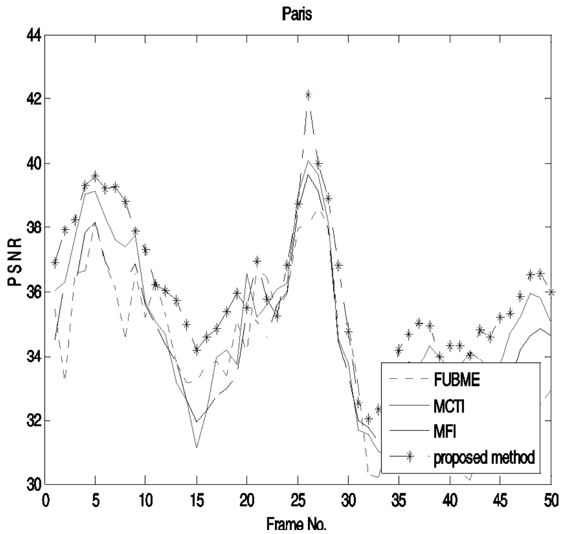

Figure 5: PSNR Comparisons of Four Methods for Paris (QCIF)

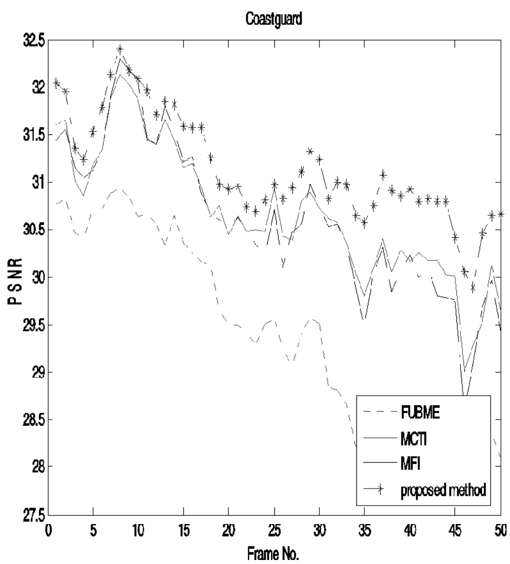

Figure 6: PSNR Comparisons of Four Methods for Coastguard (CIF)

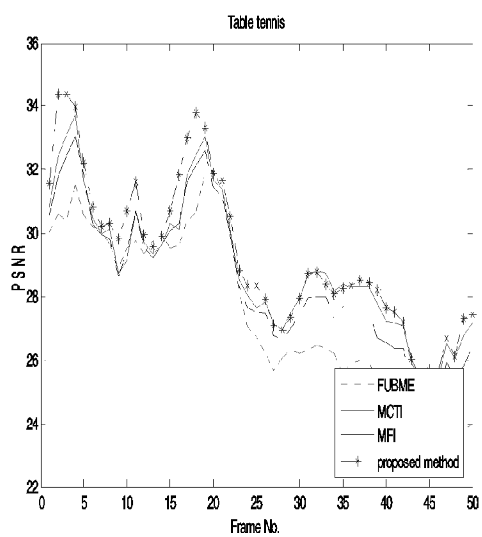

Figure 7: PSNR Comparisons of Four Methods for Table Tennis (CIF)

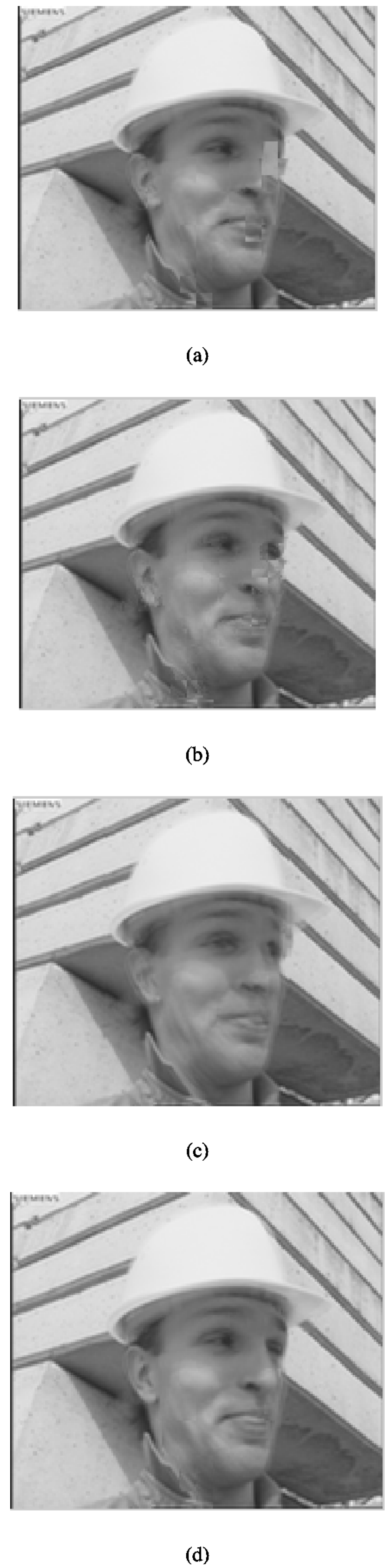

Figure 8: Test Sequence Foreman (QCIF): (a) MCTI, (b) MFI, (c) Frame up-conversion BME, (d) Proposed Method

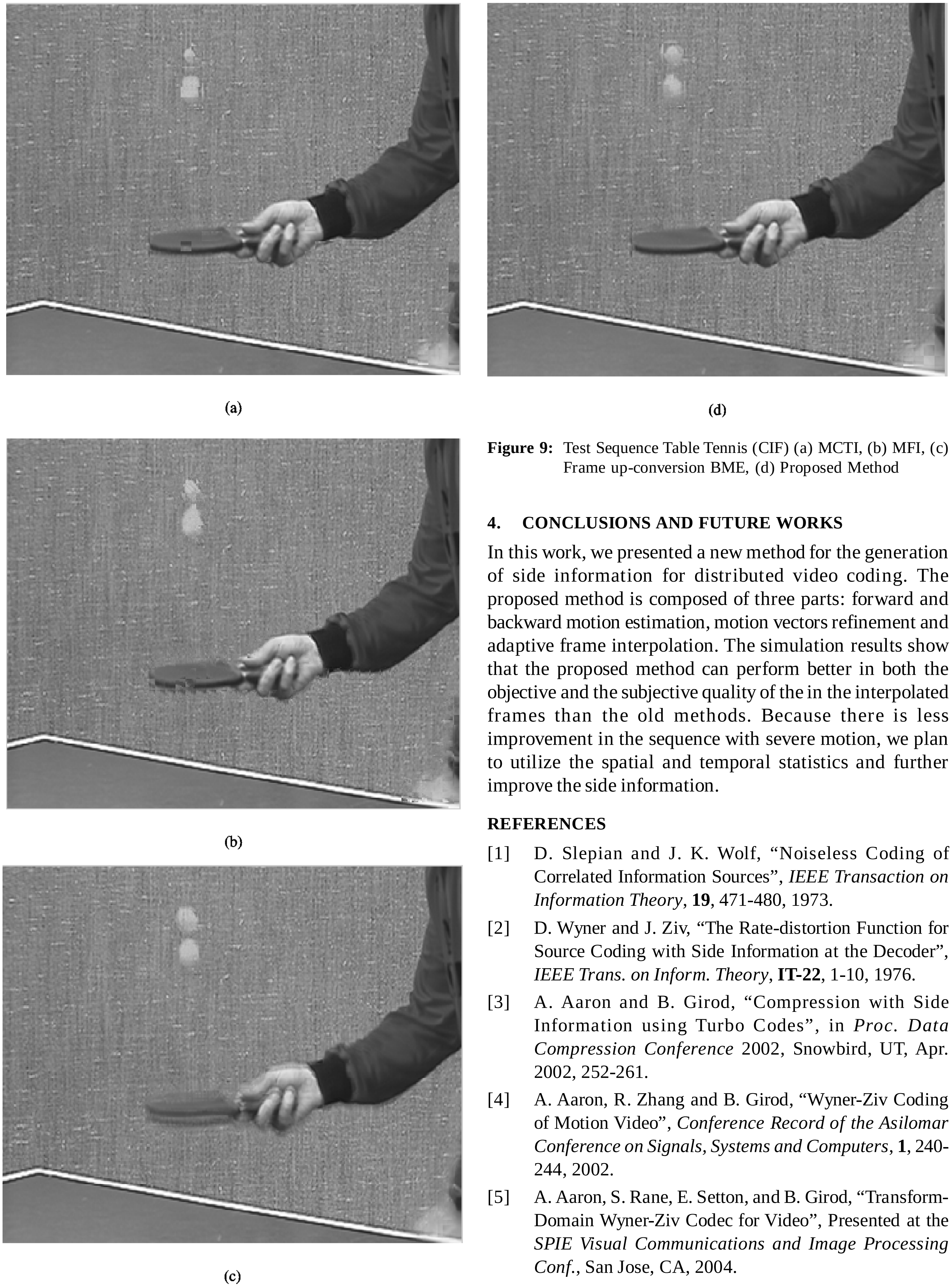

Figure 9: Test Sequence Table Tennis (CIF) (a) MCTI, (b) MFI, (c) Frame up-conversion BME, (d) Proposed Method

4. CONCLUSIONS AND FUTURE WORKS

In this work, we presented a new method for the generation of side information for distributed video coding. The proposed method is composed of three parts: forward and backward motion estimation, motion vectors refinement and adaptive frame interpolation. The simulation results show that the proposed method can perform better in both the objective and the subjective quality of the in the interpolated frames than the old methods. Because there is less improvement in the sequence with severe motion, we plan to utilize the spatial and temporal statistics and further improve the side information.

REFERENCES

[1] D. Slepian and J. K. Wolf, "Noiseless Coding of Correlated Information Sources", *IEEE Transaction on Information Theory*, **19**, 471-480, 1973.

[2] D. Wyner and J. Ziv, "The Rate-distortion Function for Source Coding with Side Information at the Decoder", *IEEE Trans. on Inform. Theory*, **IT-22**, 1-10, 1976.

[3] A. Aaron and B. Girod, "Compression with Side Information using Turbo Codes", in *Proc. Data Compression Conference* 2002, Snowbird, UT, Apr. 2002, 252-261.

[4] A. Aaron, R. Zhang and B. Girod, "Wyner-Ziv Coding of Motion Video", *Conference Record of the Asilomar Conference on Signals, Systems and Computers*, **1**, 240-244, 2002.

[5] A. Aaron, S. Rane, E. Setton, and B. Girod, "Transform-Domain Wyner-Ziv Codec for Video", Presented at the *SPIE Visual Communications and Image Processing Conf.*, San Jose, CA, 2004.

[6] A. Aaron, E. Setton, and B. Girod, "Towards Practical Wyner-Ziv Coding of Video", in *IEEE Int. Conf. Image Processing*, Barcelona, Spain, 2003.

[7] Jiefu Zhai, Keman Yu, Jiang Li and Shipeng Li, "A Low Complexity Motion Compensated Frame Interpolation Method", in *IEEE International Symposium on Circuits and Systems*, Kobe, Japan, 23-26, 2005.

[8] C. K. Wong and O. C. Au, "Fast Motion Compensated Temporal Interpolation for Video", *Proc. SPIE: Visual Communication and Image Processing* '95, 1108-1118.

[9] B. Choi, S. Lee and S. Ko, "New Frame Rate up-Conversion using Bidirectional Motion Estimation," *IEEE Transactions on Consumer Electronics*, **46**, 603-609, 2000.

[10] Mohammed E. Al-Mualla, "Motion Field Interpolation for Frame Rate Conversion", *IEEE International Symposium on Circuits and Systems* 2003, 652-655.