2011 3rd International Conference on Signal Processing Systems (ICSPS 2011)
IPCSIT vol. 48 (2012) © (2012) IACSIT Press, Singapore
DOI: 10.7763/IPCSIT.2012.V48.9

Bi-directional Motion Compensated Reconstruction for Pixel Domain Distributed Video Coding

Chunling Yang+, Zhuohan Su, Lin Lee
School of Electronic and Information Engineering
South China University of Technology
Guangzhou, China

Abstract. Distributed Video Coding (DVC) is a new video coding paradigm that attracts more and more research interest these years. In DVC, reconstruction is a key module closely related to the quality of the decoded video. This paper proposes a reconstruction scheme of Bi-directional Motion Compensated Reconstruction (BMCR) for pixel domain DVC, in which bi-directional motion compensation is performed to find reconstruction values for the pixels whose corresponding side information values are outside the decoded quantization bins. Experimental results illustrate that the proposed BMCR scheme improves both the objective and subjective reconstruction quality without increasing the bitrate, compared with the commonly used reconstruction schemes.

Keywords: Distributed Video Coding; Pixel Domain Wyner-Ziv video code; Reconstruction; Bi-directional motion compensation

1. Introduction

Different from traditional video coding, for example, H.264/AVC, DVC is a new video coding paradigm that exploits video correlation partially or wholly at the decoder rather than encoder. Such video coding solution can greatly reduce the encoding complexity, thus is more suitable for the wireless video applications where computing power, storage capacity and power consumption are all limited. DVC is based on two theorems in Information Theory: Slepian-Wolf theorem [1] for lossless compression and Wyner-Ziv theorem [2] for lossy compression. These two theorems state that separate encoding and joint decoding could be as efficient as joint encoding and decoding. In recent years, DVC has attracted much research interest and some practical schemes have been proposed. One of the most interesting approaches is the turbo code based Pixel Domain Wyner-Ziv (PDWZ) video codec [3].

In pixel domain DVC, reconstruction is a key module which directly affects the quality of decoded video. In [3], a simple reconstruction scheme was first proposed. In this scheme, if the side information value falls in the decoded quantization bin, the pixel will be reconstructed as the side information value directly; otherwise it will be reconstructed as the quantization bin boundaries. This method is widely used for its low complexity and acceptable performance. However, annoying errors are still possible to be found in pixels that are reconstructed as the boundaries. Thus some improved algorithms [4-6] were then presented. In [4], D. Kubasov et al. developed Minimum Mean Squared Error (MMSE) reconstruction scheme, which exploits the virtual correlation model between source and side information to minimize the mean squared error. In this scheme, the conditional expectation of original pixel is calculated as its reconstruction value, which is optimal statistically but not for each pixel. Moreover, when side information value falls in the decoded quantization bin, reconstructing the pixel by the expectation may result in visual distortion.

In DVC, the task of the Wyner-Ziv decoder is to construct the Wyner-Ziv frames by exploiting video correlation and the turbo decoder. Producing side information by motion compensation is the first step to get

+ Corresponding author. *E-mail address*: eeclyang@scut.edu.cn.

an initial estimation of the original frame. In this step, the information of past and future Key frames is not completely used. That's because no information about the original Wyner-Ziv frame is available here. This paper proposes a reconstruction scheme of Bi-directional Motion Compensated Reconstruction (BMCR), which makes further use of the Key frames information to improve reconstruction. In this scheme, reconstruction is considered as a continuation of side information generation to construct the Wyner-Ziv frames. If the side information value falls in the turbo decoded quantization bin, it's considered to be a good estimation of the original pixel that doesn't need revision. Otherwise, bi-directional motion compensation is employed to find a more precise reconstruction value. Since the turbo decoded symbols, denoting the true quantization bins of the original frame, have been obtained now, the Key frames can be better exploited to find out more accurate compensation values to reconstruct the out-of-bin pixels.

This paper is organized as follows: section 2 describes the architecture of the pixel domain DVC. Proposed BMCR scheme is presented in section 3. In section 4, some simulation results are given. Finally, the conclusion of this paper is shown in section 5.

2. Pixel Domain DVC

The pixel domain DVC codec that this paper uses is based on the PDWZ codec. The architecture is shown in figure 1.

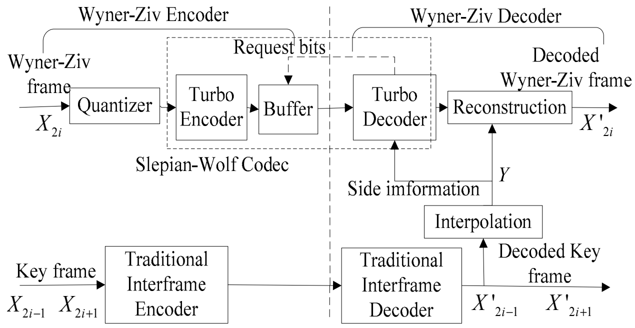


Figure 1. Architecture of pixel domain DVC

At the encoder, the frames are organized into two groups: Key frames (odd frames) and Wyner-Ziv frames (even frames). Key frames are encoded by traditional I-frame encoder, while Wyner-Ziv frames are encoded by Wyner-Ziv encoder. Each Wyner-Ziv frame is first quantized by a uniform scalar quantizer with 2^M levels and the quantized symbols are regrouped based on bitplanes. Then each bitplane is encoded by the turbo encoder. The output parity bits are stored in the buffer waiting to be transmitted whereas systematic bits discarded.

The decoder obtains side information by interpolating the decoded past and future Key frames. Then the side information and the parity bits are fed to the turbo decoder, which runs the log-MAP decoding algorithm and generate the information bits of the Wyner-Ziv frame. A feedback channel is used here to request more parity bits from the encoder until the bit error rate is acceptable. A Laplacian model with offline-estimated parameter is used for the virtual correlation channel between side information and original Wyner-Ziv frame.

After these, reconstruction is performed by exploiting the side information and turbo decoded symbols (assumed error-free). Given that x denotes the original pixel value of the Wyner-Ziv frame and y denotes the side information value, the simple reconstruction solution in [3] can be expressed as follow:

$$\hat{x}=\begin{cases} z_i & \text{if}\, y<z_i \\ y & \text{if}\ y\in[z_i,z_{i+1}) \\ z_{i+1} & \text{if}\, y\geq z_{i+1} \end{cases}, \tag{1}$$

where $\hat{x}$ denotes the reconstructed value, and $[z_i, z_{i+1})$ denotes the quantization bin of x, which is determined by the turbo decoded indices in the decoder. This scheme has a low computational complexity but fails to consider the correlation between x and y. So in MMSE reconstruction scheme, the conditional expectation $E[x \mid x \in [z_i, z_{i+1}), y]$ is calculated to reconstruct the pixel.

$$\hat{x} = E[x \mid x \in [z_i, z_{i+1}), y] = \frac{\int_{z_i}^{z_{i+1}} x f_{x|y}(x) dx}{\int_{z_i}^{z_{i+1}} f_{x|y}(x) dx}, \tag{2}$$

where $f_{x|y}(x)$ is the conditional p.d.f of x given y. Though the result of this scheme is statistically optimal, a better reconstruction value for each pixel can be found with the help of past and future Key frames. So this paper proposed BMCR scheme, in which the past and future decoded Key frames are utilized to assist reconstructing the Wyner-Ziv frames. Details of the algorithm will be presented in the next section.

3. Bi-directional Motion Compensated Reconstruction

In proposed BMCR scheme, the side information value y that falls in $[z_i, z_{i+1})$ is considered to be a perfect estimation of x, thus it's used as the reconstruction value directly. In other cases, forward and backward compensation values are first found out by bi-directional motion compensation and then are rectified if necessary. The pixels are reconstructed as the mean of rectified forward and backward compensation values.

The proposed algorithm provides a reconstruction function as follow:

$$\hat{x} = \begin{cases} y & \text{if } y \in [z_i, z_{i+1}) \\ \dfrac{x_f + x_b}{2} & \text{others} \end{cases}, \tag{3}$$

where x_f denotes the rectified forward compensation value and x_b denotes the backward case.

The process of BMCR scheme includes 4 main steps, which are presented as follows.

3.1. Initial Reconstruction

In this step, each pixel with side information in is reconstructed as directly. As to the others, reasonable reconstruction values should be taken in order to apply block matching motion estimation technology. Hence these pixels are initially constructed using MMSE solution due to its favorable performance.

3.2. Bi-directional Motion Compensation

For each pixel whose side information is outside the quantization bin, a 5×5 block with current pixel as central point is employed to run block matching motion estimation. The matching criterion used here is Weighted Sum of Absolute Difference (WSAD), of which the weighting coefficients are defined based on quantization bins [7]. Assume (x_0, y_0) to be the coordinate of current pixel,

$$D(px, py) = | X'_{2i}(px, py) - X'_k(px + dx, py + dy) | \tag{4}$$

$$Q(X'_{2i}(px, py)) = Q(X'_k(px + dx, py + dy)) \tag{5}$$

$$WSAD = \sum_{\substack{px \in [x_0 - 2, x_0 + 2] \\ py \in [y_0 - 2, y_0 + 2]}} \begin{cases} \dfrac{D(px, py)}{2^M} & \text{if (5) is satisfied} \\ D(px, py) & \text{others} \end{cases} \tag{6}$$

where X'_{2i} stands for the initial reconstructed frame, X'_k stands for the decoded past ($k = 2i - 1$) or future ($k = 2i + 1$) Key frame, (px, py) is the coordinate of the pixel in X'_{2i}, (dx, dy) is the motion vector, and Q represents the quantization index of the pixel value in the bracket.

In motion estimation, the forward and backward motion vectors with corresponding minimum WSAD are found out respectively. Then they are used as the motion vectors of current pixel to get its forward compensation value x_{fmc} and backward compensation value x_{bmc}.

$$x_{fmc} = X'_{2i+1}(x_0 + dx_f, y_0 + dy_f) \tag{7}$$

$$x_{bmc} = X'_{2i-1}(x_0 + dx_b, y_0 + dy_b) \tag{8}$$

with the motion vector (dx_f, dy_f) pointing to the future key frame and (dx_b, dy_b) pointing to the past key frame.

3.3. Rectifying the Compensation Values

During motion compensation process, more accurate estimation values that fall in the true quantization bins can be found for some pixels. But for the others, the compensation values are still outside the right quantization bins. Thus rectification is performed here. Assume that x_{mc} represents x_{fmc} (or x_{bmc}),

$$x'_{mc} = \begin{cases} z_i & \text{if } x_{mc} < z_i \\ z_{i+1} - 1 & \text{if } x_{mc} \ge z_{i+1} \\ x_{mc} & \text{others} \end{cases} \tag{9}$$

where x'_{mc} denotes the rectified compensation value. After rectification, x'_{mc} is set within $[z_i, z_{i+1})$. And here x'_{mc} is x_f (or x_b) in equation (3).

3.4. Final Reconstruction

Finally, the mean of the rectified forward and backward compensation values is calculated as the reconstruction value of each pixel whose corresponding side information is outside the decoded quantization bin. This was presented in case 2 of equation (3). Since x_f and x_b are both within $[z_i, z_{i+1})$, the reconstruction value must fall in the decoded quantization bin, too.

4. Experimental Result

In our experiments, the first 201 frames of two video sequences (QCIF@30Hz, 100 Wyner-Ziv frames) are used: Foreman with medium motion complexity and Soccer with high motion complexity. The Key frames are encoded by H.264 Intra-coding. Different QP parameters for each rate point are selected so as to keep the quality of decoded Key frames and Wyner-Ziv frames at the same level, The side information is generated by algorithm in [8].

Figure 2 shows the average PSNR vs. bitrate of the luminance component for the Wyner-Ziv frames. As clearly illustrated, reconstruction quality of proposed BMCR scheme is much better than that of the simple reconstruction in [3]. Compared with MMSE reconstruction, proposed scheme still gains up to 0.2 dB for Foreman sequence while 0.9 dB for Soccer sequence. The reason is that Soccer sequence has higher motion complexity and the proportion of pixels with side information values outside the right quantization bins is larger. Since the reconstruction quality of these pixels has been improved, the gain for Soccer sequence is larger than Foreman. The worse the side information is, the more increase BMCR scheme will gain.

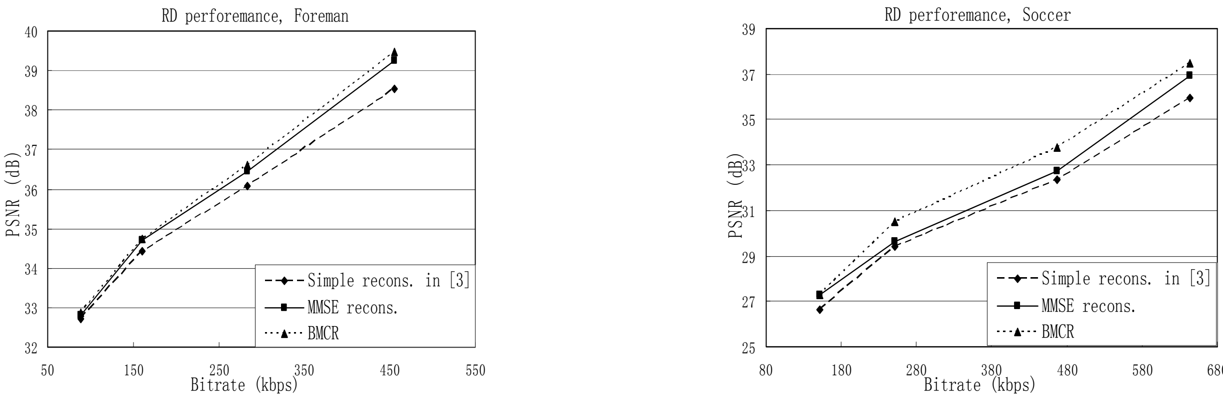


Figure 2. RD performance for foreman and soccer sequence

Figure 3 and figure 4 show the subjective quality of decoded frames. As stated above, MMSE reconstruction in pixel domain DVC may produce unexpected edges in some smooth areas, just as the hat of the man in foreman sequence and the football field in soccer sequence demonstrate. But BMCR does not have this kind of problem. As for the edge areas, BMCR provides clearer boundaries than both simple reconstruction and MMSE reconstruction. See the hats in row 2, figure 3 and feet in row 2, figure 4.

The simulation results show that the proposed BMCR scheme improves the objective and subjective quality of the reconstructed frames compared with the simple reconstruction in [3] and MMSE reconstruction.

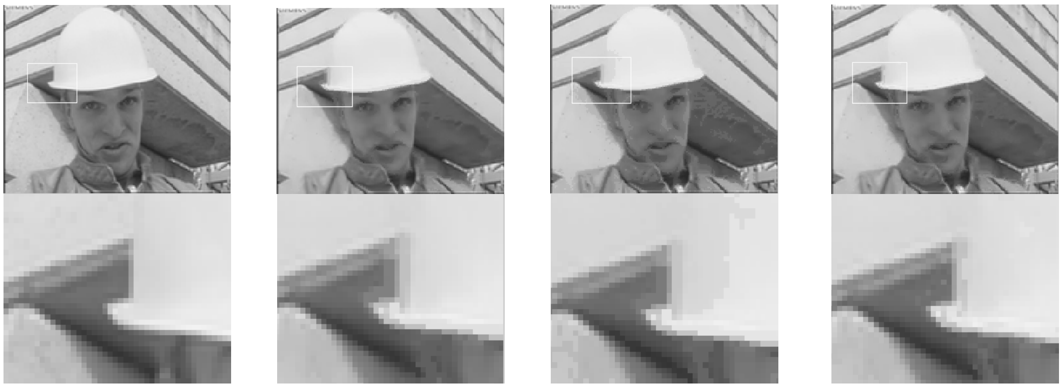
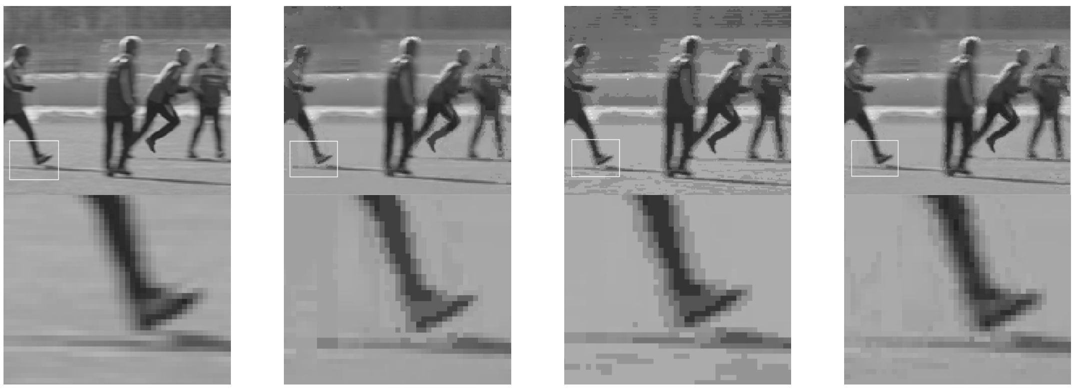


Figure 3. Foreman sequence that has been reconstructed with 2 bitplanes (frame NO. 146)



(a) Original frame (b) Simple Recons. (c) MMSE Recons. (d) BMCR

Figure 4. Soccer sequence that has been reconstructed with 3 bitplanes (frame NO. 112)

5. Conclusion

In the paper, a reconstruction scheme of BMCR was proposed to improve the reconstruction quality for pixel domain DVC. When the side information value falls in the decoded quantization bin, it is taken to reconstruct the pixel directly. Otherwise, the reconstruction value is determined by rectified forward and backward motion compensation. Simulation results illustrate that this proposed scheme increases the RD performance of the codec by 0.2 dB for foreman sequence and 0.9 dB for soccer sequence compared to MMSE reconstruction. What's more, the visual quality of decoded video was also improved evidently.

6. Acknowledgment

The work described in this paper was substantially supported by research project from National Natural Science Foundation of China (60972135), the Fundamental Research Funds for the Central Universities, SCUT (2009ZM0231), and Guangdong National Science Foundation (06025642)

7. References

[1] D.Slepian and J. Wolf, "Noiseless coding of correlated information sources," IEEE Transactions on Information Theory, vol. 19(4), pp. 471-480, July 1973.

[2] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," IEEE Transactions on Information Theory, vol. IT-22(1), pp. 1-10, 1976.

[3] A. Aaron, R. Zhang and B. Girod, "Wyner-Ziv coding of motion video," Thirty-Sixth Asilomar Conference on Signals, Systems & Computers - Conference Record, Vols 1 and 2, Conference Record, pp. 240-244, 2002.

[4] D. Kubasov, J. Nayak and C. Guillemot, "Optimal reconstruction in Wyner-Ziv video coding with multiple side information", 9th IEEE Workshop on Multimedia Signal Processing, 2007, pp. 183-186

[5] C.L. Yang, Y.G. Lin, L.M. Po and W.R. Gao, "Variable Reconstruction Scheme for Distributed Video Decoding," in 2008 Ieee Region 10 Conference: Tencon 2008, Vols 1-4, ed, 2008, pp. 1917-1921.

[6] S. Y. Shim, J.K. Han, J. Bae "Adaptive reconstruction scheme using neighbour pixels in PDWyner-Ziv coding," Electronics Letters, vol. 46, pp. 626-U43, Apr 2010.

[7] J. Ascenso, C. Brites and F. Pereir, "Motion compensated refinement for low complexity pixel based distributed video coding," Advanced Video and Signal Based Surveillance, 2005. AVSS 2005. IEEE Conference on pp. 593 - 598, 2005.

[8] J. Ascenso, "Hierarchical motion estimation for side information creation in Wyner-Ziv video coding," Proceedings of the 2nd International Conference on Ubiquitous Information Management and Communication. Suwon, Korea, pp. 347-352, 2008