BIT-STREAM SWITCHING IN MULTIPLE BIT-RATE VIDEO STREAMING USING WYNER-ZIV CODING

Wen Sun[1,*], *Yan Lu*[2], *Feng Wu*[2]

[1]University of Science and Technology of China, Hefei, Anhui, 230027, China
[2]Microsoft Research Asia, Beijing, 100080, China

ABSTRACT

It has been commonly recognized that multiple bit-rate (MBR) encoding provides a concise method for video streaming over bandwidth-fluctuant networks. The key problem of the MBR technique lies in how to seamlessly switch one bit-stream to another one. To tackle this problem, we propose a bit-stream switching framework based on the Wyner-Ziv coding. Within the propose framework, the multiple bit-streams can be individually encoded without data exchange, which also supports the random switching at any desired frame without affecting the original coding efficiency of the regular bit-stream. In particular, two different implementation schemes under the same framework are presented. Different from the traditional switching schemes, the proposed method can use the same switching frame for the switching from any other bit-stream to the current one, which means less storage and less encoding efforts. Simulation results and comparison between the proposed method and the traditional switching method in H.264 are also presented.

1. INTRODUCTION

While the video streaming is becoming more and more popular, the bandwidth fluctuation and the requirement of constant visual quality are always two conflicting aspects in video streaming. To cope with the network bandwidth variations, multiple bit-rates (MBR) video streaming is often used in the video streaming services. In the MBR-based video streaming, the same video source is encoded multiple times, each at a chosen bit-rate. During the streaming, the server or receiver can dynamically select the proper bit-stream to transmit so as to accommodate the variations of the bandwidth available at the client. MBR technique has been extensively supported by the commercial systems such as Windows Media Services, RealSystem and QuickTime [1]~[3].

However, the MBR technique also suffers from some drawback due to the inherent feature of hybrid video coding. In the traditional way, the perfect switching between bit-streams is possible only at the Intra-coded frames (I-frames). As we know, the more I-frames the worse coding efficiency. Therefore, I-frames are usually coded far apart from each other. As an improvement, a new frame type referred to as SP frame [4][5] is defined in the MPEG-4 AVC/H.264 standard. In terms of the SP method, both primary and secondary SP frames are generated at some candidate switching points. The primary and secondary SP frames are alternatively transmitted according to whether or not the switching occurs.

Although the SP method can provide the drift-free bit-stream switching, it also suffers from the following drawbacks. Firstly, the primary SP frames embedded in the regular bit-stream require much more coding bits than the traditional Inter-coded frames (P-frames). Secondly, the multiple bit-streams have to be jointly encoded. Suppose the number of bit-streams is N. Each candidate switching point at each bit-stream corresponds to one primary SP frame and N-1 secondary SP frames, which means huge complexity and storage capability at the encoder. Therefore, it is very desirable to have an MBR technology that can tackle all the above problems.

Accordingly, we propose a new MBR-based video streaming framework, in which the Wyner-Ziv video coding is utilized. The Wyner-Ziv video coding is the extension of the distributed source coding (DSC), which has attracted a lot of research attentions recently. The theme of DSC lies in that two correlated signals can be separately encoded with the same coding efficiency as being jointly encoded, which perfectly meets the above requirements of the bit-stream switching in MBR-based video streaming. Based on the general framework, we propose two practical implementation schemes.

The rest of this paper is organized as follows. Section 2 presents the employed Wyner-Ziv coding scheme. Section 3 presents the proposed MBR framework as well as two practical implementations. The experiment results compared to the SP-based switching method are presented in Section 4. Finally, Section 5 concludes this paper.

* This work has been done while the author was with Microsoft Research Asia as an intern.

2. WYNER-ZIV VIDEO CODING

Typically, Wyner-Ziv coding indicates the lossy DSC with side information. Let *X* and *Y* be statistically dependent Gaussian random process. Let *Y* be known as side information for the encoding of *X*. Wyner and Ziv have proven that the conditional rate-mean squared error distortion function for *X* is the same whether the side information *Y* is available only at the decoder or both at the encoder and the decoder [7]. The theory has shown the potential of developing an intra-frame encoder – inter-frame decoder system that can come close to the efficiency of an inter-frame encoder –decoder system.

Recently, a number of practical Wyner-Ziv coding schemes have been developed. For example, a Wyner-Ziv video codec that uses intra-frame encoding but inter-frame decoding has been proposed by Girod *et al* in [8]. In our previous work, we have further extended the method in [8] to the wavelet transform domain processing [9]. Fig. 1 shows the block diagram of the proposed algorithm. Briefly, the encoder includes wavelet transform, scalar quantization, zero-tree organization of wavelet coefficients and turbo encoder. Only Wyner-Ziv bits (parity bits) are transmitted to the turbo decoder until the Wyner-Ziv frame can be decoded correctly with side information. Note that the efficiency of the kind of Wyner-Ziv coding schemes greatly depends on the utilization of correlations between side information and coded frame. The proposed Wyner-Ziv coding method is used for bit-stream switching later.

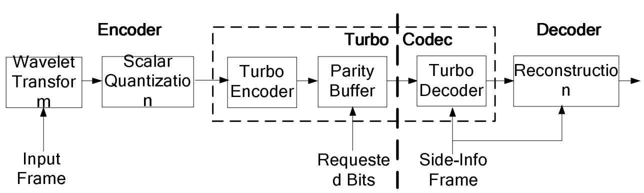


Fig. 1. Block diagram of the employed Wyner-Ziv codec.

3. PROPOSED BIT-STREAM SWITCHING

In this section, we first propose a general framework for bit-stream switching using Wyner-Ziv coding, and then turn to the discussion of two practical implementation schemes. For simplicity, we only discuss the case of the switching between two bit-streams. Let bit-stream 1 indicate the one with low bit rate, and bit-stream 2 indicate the one with higher bit-rate. Suppose $P_{k,n}$ (k = 1, 2) represents the coded bits of frame n in bit-stream k. Further, suppose $F_{k,n}$ (k = 1, 2) indicates the reconstructed frame in bit-stream k at the decoder. Since the down-switching (from higher bit-rate to lower bit-rate) is much easier to be handled rather than the up-switching we only discuss the case of up-switching, i.e., switching from bit-stream 1 to bit-stream 2, in this paper.

Assume that the server initially sends bit-stream 1 up to time n, and then turns to bit-stream 2 due to the variance of bandwidth. Thus, the decoder would have received {…, $P_{1,n-2}$, $P_{1,n-1}$, $P_{2,n}$, $P_{2,n+1}$, $P_{2,n+2}$, …} if the server switches directly at the switching point n. Obviously, in this case, the frame at time n (in terms of $P_{2,n}$) cannot be properly decoded due to the mismatch between $F_{1,n-1}$ and $F_{2,n-1}$, which may also lead to drifting errors. To handle this mismatch and avoid drifting errors, a straightforward method is to transmit the intra-coded $F_{2,n-1}$ rather than $P_{1,n-1}$. However, since there is great similarity between $F_{1,n-1}$ and $F_{2,n-1}$, it is more desirable to code $F_{2,n-1}$ using Wyner-Ziv method rather than intra coding based on the results on Wyner-Ziv coding.

In particular, we can produce the Wyner-Ziv bits of each frame at the encoder, and store them aside. When the switching occurs, we only need transmit part of the Wyner-Ziv bits. The number of bits sent to the decoder depends on the gap between the two streams. Moreover, instead of completely removing the mismatch, we propose to limit the drifting errors to a very small scale to avoid the lossless coding that is inefficient. Based on this basic framework, we propose two practical implementation schemes for bit-stream switching using Wyner-Ziv coding, as discussed below.

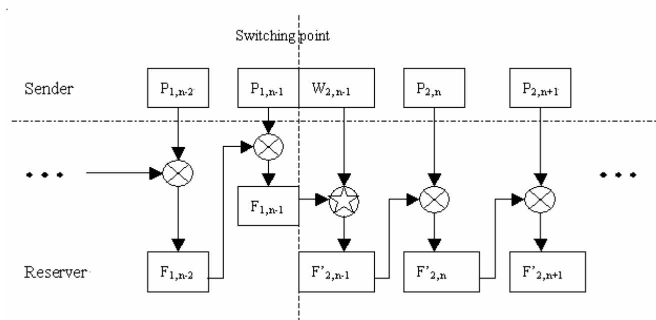


Fig. 2. Bit-stream switching *Scheme One*. "☆" denotes a Wyner-Ziv decoder that takes $F_{1,n-1}$ as side information, and "×" denotes the motion-compensated prediction.

3.1. Scheme One

In this scheme, the reconstructed frame $F_{2,i}$ is coded using the Wyner-Ziv method. The coded bits are denoted as $W_{2,i}$. Wyner-Ziv bits of bit-stream 1 are not needed in the up-switching case. Fig. 2 depicts how *Scheme One* works. When the streaming comes to the switching point, $P_{1,n-1}$ is transmitted. Then $F_{1,n-1}$ is reconstructed using $F_{1,n-2}$ as the reference frame. Subsequently, $W_{2,n-1}$ are transmitted to inform the receiver for the purpose of switching. At the receiver, a Wyer-Ziv decoder is employed to decode $W_{2,n-1}$

using $F_{1,n-1}$ as the side information. The decoded frame $F'_{2,n-1}$ should be identical or very close to $F_{2,n-1}$ depending on the quantization step-size in the Wynzer-Ziv coding. In the following, the subsequent frames $P_{2,n}$, $P_{2,n+1}$, … are transmitted to the receiver. In summary, based on *Scheme One*, the server transmits $W_{2,n-1}$ and $P_{2,n}$ in total at the switching point.

3.2. Scheme Two

In *Scheme One*, the whole $P_{2,n}$ is transmitted when switching occurs at time n. Instead, in *Scheme Two*, only part of $P_{2,n}$ is transmitted. Fig. 3 depicts how *Scheme Two* works. In this scheme, the motion information (i.e. $MV_{2,n}$) and the intra-coded blocks in $P_{2,n}$ are transmitted to the receiver. Thus, the prediction of $F_{2,n}$, namely $PF_{2,n}$, can be generated using the decoded motion vectors, the decoded intra blocks and the reconstructed previous frame $F_{1,n-1}$. Then, $W_{2,n}$ is transmitted. At the receiver, $W_{2,n}$ is decoded using the $PF_{2,n}$ as side information. In this way, the decoded frame $F'_{2,n}$ can be very close to $F_{2,n}$, depending on quantization step-size in Wyner-Ziv coding. The following process is the same as that in *Scheme One*. In summary, based on *Scheme Two*, the server should transmit $W_{2,n-1}$ and part of $P_{2,n}$ at the switching point.

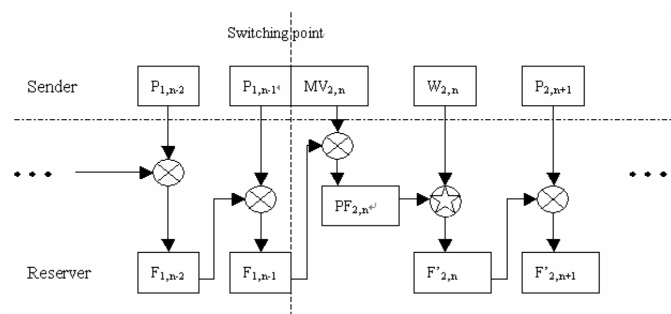


Fig. 3. Bit-stream switching *Scheme Two*. "☆" denotes a Wyner-Ziv decoder that takes $PF_{2,n}$ as side information. "×" denotes the motion-compensated prediction.

3.3. Discussion

The difference between *Scheme One* and *Scheme Two* lies in the bits transmitted to the receiver at the switching point. *Scheme One* induces less Wyner-Ziv bits than *Scheme Two*, whereas it needs transmit the more original bits of the frame at the switching point. Undoubtedly, the essence of the two schemes is the Wyner-Ziv coding method described in Section 2. The performance greatly depends on the correlations between the two reconstructed frames at the two bit-streams. The stronger dependency the better performance and the less switching bits. Because the Wyner-Ziv frame can be perfectly reconstructed by the Wyner-Ziv decoder as long as sufficient parity bits are transmitted, the proposed switching method can be actually seamless. However, in practice, we often have to trade off between the drifting errors and the number of bits transmitted at the switching points.

The following summarizes the advantages of the propose method over the traditional bit-stream switching methods. Firstly, with the proposed method, each bit-stream can be encoded individually with whatever parameters, without considering the bit-stream streaming problems. Secondly, with the proposed method, the same switching frame can be used for the switching from any other bit-streams to the current one. The overall storage and encoding burden for the switching frames can be significantly saved. Thirdly, with the propose method, the switching points can be set at any positions, which has no effect when the bit-stream switching never occurs.

4. EXPERIMENTAL RESULTS

In this section, we present some simulation results of the two proposed bit-stream switching schemes, which are also compared with the SP frame-switching in H.264. The test sequences include Foreman (300 frames) and Football (200 frames) at QCIF format and 30 fps. Both are encoded with the first frame as I frame and the rest as P frames. We create two bit-streams for each sequence. For Foreman, the bit-rate of bit-stream 1 is 103 kbps and the bit-rate of bit-stream 2 is 236 kbps. For Football, the bit-rate of bit-stream 1 is 331 kbps and the bit-rate of bit-stream 2 is 576 kbps. In addition, for each sequence, we also create an alternative bit-stream 2 at the same bit-rate as the regular coding but with the SP frames inserted. Particularly, every second frame is encoded as an SP frame in order to support the almost equally flexible switching as the proposed method.

Suppose the switching occurs at the 150^{th} frame for Foreman and 100^{th} frame for Football. Fig. 4 shows the results in terms of the luminance PSNR in the sequences frame by frame. The bits transmitted at the switching points are listed in Table 1. The simulation results have shown that both proposed schemes works well for the up-switching cases. Compared with the switching using SP frames, our schemes make the encoder significantly simplified. Because all of the bits transmitted at the switching point are generated within the single bit-stream and without any information from the other bit-streams, no extra work is required when the number of bit-streams increases. The duty of the sender is just to transmit the proper bits of the proper bit-stream at the proper time. Another impressive advantage of our schemes is that the switching can be performed at each frame with little delay.

As for the comparison between the proposed two schemes, it is really hard to say which is better based on the current implementations. A number of experiments have shown that, when the motion in the switching frame is not very fast, *Scheme Two* almost has the identical performance to *Scheme One* in compensating the mismatch, but can save quite a few bits. However, when the video contains fast motion, *Scheme Two* may cost much more bits than *Scheme One* due to the poor side information frame generated from the previously reconstructed frame. In this case, *Scheme One* surely outperforms *Scheme Two*. Nevertheless, both schemes have the drawback. As shown in Table 1, the number of bits for the switching frame is still very large, although it is much less the intra coding at the same level of the compensation of mismatch.

5. CONCLUSIONS

A bit-stream switching method using Wynzer-Ziv coding has been proposed for the MBR-based video streaming. Based on a general framework, two practical switching schemes are realized, resepctively. By using the Wyner-Ziv coding, we can achieve low encoder complexity without extra bits added to the transmitted bitstream when no switching occurs. The capability of random switching at any desired frame can also be provided. Compared with traditional switching schemes, the larger the number of bit-streams is, the more benefits our method can achieve. Particularly, the proposed method has shown its own advantages over the switching method using SP frames in H.264. Nevertheless, it still needs the future work to refine the Wyner-Ziv coding to further limit the drifting errors.

6. REFERENCES

[1] G. Conklin, G. Greenbaum, K. Lillevold, A. Lippman, Y. Reznik, "Video coding for streaming media delivery on the Internet," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 11, no. 3, pp. 269-281, Mar. 2001.

[2] D. Wu, Y. Hou, W. Zhu, Y. Zhang, J. Peha, "Streaming video over the Internet: approaches and directions," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 11, no. 3, pp. 282-300, Mar. 2001.

[3] J. Lu, "Signal processing for Internet video streaming: a review," *SPIE on Image and Video Communication and Processing 2000*, vol. 3974, pp. 246-258, Australia, 2000.

[4] M. Karczewicz and R. Kurceren, "The SP- and SI-frames design for H.264/AVC," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 637–644, July 2003.

[5] X. Sun, S. Li, F. Wu, J. Shen, and W. Gao, "The improved SP frame coding technique for the JVT standard," *Proc. IEEE ICIP 2003*, vol. 3, pp. 297–300, Barcelona, Spain, Sept. 2003.

[6] W. Zhang and B. Zeng, "An efficient bit-stream switching in multi-rate based video streaming systems," *Proc. SPIE*, vol. 5960, pp. 2220-2229, Bellingham, WA, 2005.

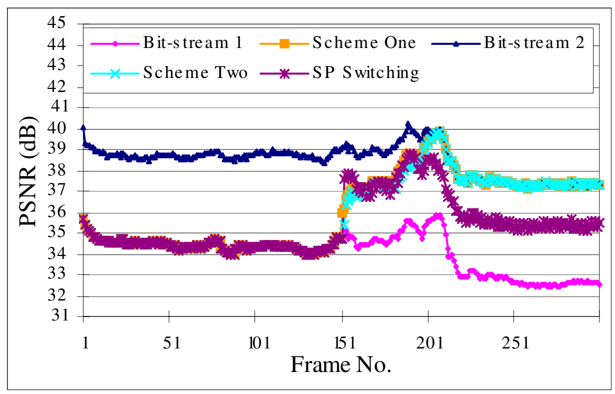
[7] A. D. Wyner, "Recent results in the Shannon theory," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 1, pp. 2–10, Jan. 1974.

[8] A. Aaron and B. Girod, "Wyner-Ziv video coding with low-encoder complexity," *Proc. Picture Coding Symposium*, PCS 2004, San Francisco, CA, December 2004.

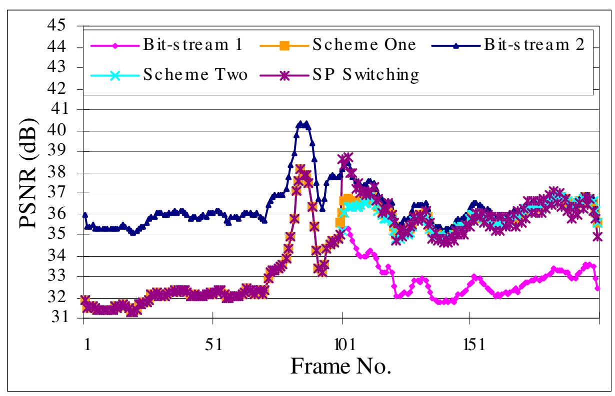
[9] X. Guo, Y. Lu, F. Wu and W. Gao, "Wyner-Ziv video coding with low-encoder complexity," *Proc. ISCAS 2006*, Greece, May 2006.

Table 1. Number of bits for the switching frame.

Scheme	Type	Number of Bits	
		Foreman	Football
Scheme One	W-Z bits	23760	22176
	P/MV bits	8464	8176
	Total	32224	30352
Scheme Two	W-Z bits	25344	22176
	P/MV bits	3952	4848
	Total	29296	30216
SP Switching	Total	31000	31000



(a) Foreman



(b) Football

Fig. 4. Switching between bit-stream 1 and bit-stream 2 using *Scheme One*, *Scheme Two* and SP Switching.