EFFICIENT VIDEO STREAM SWITCHING WITH PROGRESSIVE S-FRAMES

Byeong-Doo Choi, Ju-Hun Nam, Jin-Hyung Kim, Sung-Hoon Yun, and Sung-Jea Ko, Senior Member, IEEE

Department of Electronics Engineering, Korea University
5-1 Anam-Dong, Sungbuk-Ku, Seoul 136-701, Korea

ABSTRACT

In this paper, an efficient bitstream switching using the progressive S-frames is presented. The progressive S-frame approach can effectively switch two pre-encoded streams with different quantization steps by reducing progressively the quantization mismatch. The progressive S-frames can provide better coding efficiency than a single S-frame. The simulation results show that the proposed method is useful for the bit rate adaptation in heterogeneous networks.

Index Terms— Video streaming, dynamic switching, S-frame, progressive S-frame

1. INTRODUCTION

The fourth generation (4G) network will integrate heterogeneous networks since no single wireless network technology simultaneously can provide a low latency, high bandwidth, and wide area data service to a large number of mobile users [1]. Among different types of networks, the vertical mobility and vertical handover mean the movement of a user and the process of maintaining a user's active connection by changing its point of attachment [2]. To enable seamless and robust video streaming in vertical handover, one of the major challenges is the bit rate adaptation when the gap of bandwidth between two different networks is large. However, a scalable bitstream alone may not provide a large enough bit rate range to address large bandwidth variation in heterogeneous networks without sacrificing the coding efficiency. Dynamic bitstream switching among a set of bitstreams encoded at different bit rates may be a good way to effectively deal with bandwidth variation. The goals of dynamic bitstream switching are to achieve high utilization of the available bandwidth, and to avoid player re-buffering due to congestion to produce the best content quality.

Recently, S-frame has been proposed to serve as special bridging frames for switching from one bitstream to another [3]. When the transmitted video stream switches from high (low) quality to low (high) quality, an S-frame is generated by encoding the prediction error between the currently reconstructed frame in the following stream and the reference frame in the previous stream. However, the S-frame introduce a little drift error during bitstream switching. Using small quantization step size for each S-frame can reduce or eliminate the quantization mismatch error, but it increases drastically the bit rate of S-frames. As a more recent work, SP-frame was proposed and accepted to H.264/AVC, which is a new type of a predictive frame to be used in bitstream switching [4, 5]. Unlike the S-frame, It can completely eliminate the drift error. However, SP-frames also should be periodically inserted. It increases considerably the total bit rate and the complexity of an en/decoder.

In this paper, we propose the progressive S-frame approach for efficient bitstream switching. The progressive S-frame approach utilizes multiple S-frames with optimal quantization steps. The S-frame results in the drift errors caused by the mismatch between reference frames. However, multiple consecutive S-frames with optimal quantization steps minimizing the R-D cost function can reduce progressively the mismatch. It has better coding efficiency than the periodical I-frame and S-frame approaches with keeping away from the buffer overflow by smoothing the abrupt increment of bit rates caused by a single S-frame.

The rest of the paper is organized as follows. Section 2 presents the dynamic bitstream switching using the progressive S-frames. Section 3 compares the performances of the proposed algorithm with the conventional methods. Finally, Section 4 concludes this paper.

2. DYNAMIC BITSTREAM SWITCHING WITH PROGRESSIVE S-FRAMES

Dynamic bitstream switching is a rate control technique in which the server has access to multiple copies of the same content encoded at different rates. The server switches between these streams in order to adapt to changes in the server-player network bandwidth. The common approach of the bitstream switching is to encode periodically the input frame in intra mode. The encoding strategy of periodic I-frames is given by

$$\begin{cases} I_{k,1} = \mathrm{Intra}(F_k) & \text{if } k = Nt, \text{ for } t = 0, 1, 2, \ldots \\ P_{k,1} = \mathrm{Inter}(F_k, \widehat{F}_{k-1,1}) & \text{otherwise,} \end{cases} \tag{1}$$

1-4244-1437-7/07/$20.00 ©2007 IEEE

where F_k and $\widehat{F}_{k,1}$ denote the kth original input frame and the reconstructed frame of the first encoded stream, respectively. $I_{k,1}$ and $P_{k,1}$ denote the kth encoded frames of the first encoded stream in intra and inter mode, and N is the number of frames in one group of picture (GOP). The function $\text{Intra}(F_k)$ indicates the intra mode encoding of the input frame F_k while $\text{Inter}(F_k, \widehat{F}_{k-1})$ denotes the inter mode process using $\widehat{F}_{k-1}$ as the reference frame. We can successfully switch the bitstreams by firstly sending the I-frame of the following stream after transmitting all P-frames in a GOP of the previous stream. However, since I-frames typically require 5∼10 times as many bits as P-frames, a lot of overhead bits are required to transmit periodically I-frames. If I-frames are sparsely inserted to reduce the total bit rate, the bitstream switching mechanism can not correspond immediately to a sudden vertical handover.

In contrast to I-frame, the S-frame can provide seamless switching points without periodical insertion. The S-frame is generated and sent only when switching from one bitstream to another. The encoding mechanism of S-frame is as follows;

$$S_{k,12} = \text{Inter}(\widehat{F}_{k,2}, \widehat{F}_{k-1,1}), \quad (2)$$

where $S_{k,12}$ indicates the kth frame in S-mode which switches from the first stream to the second one. In detail, in order to enable to switch at kth frame, $S_{k,12}$ is generated by encoding $\widehat{F}_{k,2}$ as a P-type frame predicted from $\widehat{F}_{k-1,1}$.

However, S-frame introduces some drift errors caused by the quantization mismatch among reference frames. Using small quantization step for S-frame could reduce the quantization mismatch, but it would increase drastically the packet size of S-frame. Moreover, the abrupt increment of the bit rate caused by S-frame would occur the buffer overflow in an encoder or decoder side. Especially for vertical handover, it can produce severe traffic congestion or packet loss. To complement these problems, we propose the progressive S-frame approach. The progressive S-frame approach indicates that multiple frames are consecutively encoded in S-frame mode. By sending multiple progressive S-frames instead of a large size S-frame, we can reduce the total bit rate and smooth the impulse-like explosion of bit rate caused by a huge S-frame. The encoding mechanism of progressive S-frames is given by

$$S_{k,12} = \text{Inter}(\widehat{F}_{k,2}, \text{Inter}^{-1}(S_{k-1,12}, \widehat{F}_{k-2,1})), \quad (3)$$

where $\text{Inter}^{-1}(S_{k-1,12}, \widehat{F}_{k-2,1})$ is the decoding process of $S_{k-1,12}$ using $\widehat{F}_{k-2,1}$ as the reference frame. Table 1 compares the sequence structures of periodic I-frames, S-frame, and progressive S-frames.

The computational complexity of the progressive S-frame approach is heavier than that of single S-frame since several frames are encoded in S-frame mode. However, it can provide totally low bit rate by reducing progressively the drift errors of the decoded frames caused by the differences between reference frames. Let the rate and distortion function: $R_k(Q_k)$ and $D_k(Q_k)$. To minimize the overall distortion of progressive S-frames subject to a bandwidth constraint, the quantization parameter, Q_k, of each S-frame, $S_{k,12}$, is determined by

$$Q_k = \min \sum_{i=k-N_{ps}}^{k} D_i(Q_i) \text{ subject to } \sum_{i=k-N_{ps}}^{k} R_i(Q_i) \leq R_T, \quad (4)$$

where N_{ps} is the number of progressive S frames, Q_i denotes the quantization step-size of the ith progressive S-frame, R_T represents the target bit rate.

The distortion of encoding a S-frame can be divided into the quantization distortion, $D_k^q(Q_k)$, and the propagated distortion from its reference frame, D_k^{ref}. $D_k^q(Q_k)$ is only determined by the quantization step size, Q_k; whereas D_k^{ref} is affected by the errors propagated from the previous frames. We use the following function to represent the combined distortion:

$$D_k(Q_k) = D_k^q(Q_k) + D_k^{ref}. \quad (5)$$

We employ the quadratic distortion model in [6] and [7] for $D_k^q(Q_k)$ as:

$$D_k^q(Q_k) = a_1 Q_k^2 + a_2, \quad (6)$$

where a_1 and a_2 are constants. The propagation distortion function, D_k^{ref}, can be expanded as follows:

$$D_k^{ref} = \mu_{prop}(D_{k-1}^{ref} + D_{k-1}^q(Q_{k-1})), \quad (7)$$

where μ_{prop} is the propagation factor. Let Q_H and Q_L indicate the quantization step-sizes of the high and low bit rate streams, respectively. From [8], the initial distortion, $D_{k-N_{ps}}$, of two different bitstreams caused by the quantization mismatch is given by

$$D_{k-N_{ps}} = a_1(Q_H - Q_L)^2 + a_2. \quad (8)$$

If we assume that the quantization ,Q_k, for each S-frame is identical, the total distortion is approximated to

$$\sum_{i=k-N_{ps}}^{k} D_i(Q_i) \simeq N_{ps} \cdot D_k(Q_k) + \frac{1-\mu_{prop}^{N_{ps}}}{1-\mu_{prop}} D_{k-N_{ps}}. \quad (9)$$

For the rate function, the quadratic rate model is employed. Let $\text{MAD}(\widehat{F}_{k-1,1}, \widehat{F}_{k,2})$ denote the mean of absolute difference between $\widehat{F}_{k-1,1}$ and $\widehat{F}_{k,2}$. The rate function is given by

$$R_k(Q_k) = (b_1 Q_k^{-1} + b_2 Q_k^{-2}) \text{MAD}(\widehat{F}_{k-1,1}, \widehat{F}_{k,2}), \quad (10)$$

where b_1 and b_2 are constants. From (5) and (10), we can find the optimal quantization parameters minimizing (4) with Lagrangian or dynamic optimization method.

Table 1. Frame structures of transmitted streams

Periodic I-frames	$I_{1,1}P_{2,1}P_{3,1}\cdots I_{k-3,1}\,P_{k-2,1}\quad P_{k-1,1}\quad \mathbf{I_{k,2}}\quad P_{k+1,2}P_{k+2,2}I_{K+3,2}\cdots$
S-frame	$I_{1,1}P_{2,1}P_{3,1}\cdots P_{k-3,1}P_{k-2,1}\quad P_{k-1,1}\quad \mathbf{S_{k,12}}P_{k+1,2}P_{k+2,2}P_{K+3,2}\cdots$
Progressive S-frames	$I_{1,1}P_{2,1}P_{3,1}\cdots P_{k-3,1}\mathbf{S_{k-2,12}}\mathbf{S_{k-1,12}}\mathbf{S_{k,12}}P_{k+1,2}P_{k+2,2}P_{K+3,2}\cdots$

3. EXPERIMENTAL RESULTS

The performance of the proposed algorithm is evaluated using the standard H.264 coder [9]. In order to evaluate the coding efficiencies of periodic I-, S-, and progressive S- frames, we measure the peak signal-to-noise ratios (PSNRs) and the bit rates. Each one is switched 10 times within 300 frame between two different quality bitstreams representing the same sequence. A GOP consists of 10 frames for periodical I-frame approaches while IPPP structure is used for others. Five S-frames are employed for one switching in progressive S-frame approach. From the results in Table 2, it can be noted that the total number of bits when using progressive S-frames is considerably smaller than for S-frame approaches. It shows that progressive S-frames can switch bitstreams successfully with small bit rate increments while preserving same PSNR performance.

In Table 3, we illustrate the performance of each scheme when there is no switching between bitstreams. It can be seen that in this case the performance of the periodic I-frame approach is significantly lower than that of the S-frame and progressive S-frame approaches. Without switching, the performance of S-frame or progressive S-frame approach is equal to that of the IPPP structure.

In Figs. 1 and 2, we illustrate switching between bitstreams using S-frame and progressive S-frames in terms of bit rate and PSNR. In Fig. 1, the bit rate increases abruptly in S-frame. It can yield buffer overflow or congestion. In contrast to that of S-frame, the bit rate of progressive S-frames gradually increases. It brings bit rate smoothing effect and prevents the buffer overflow. Fig. 2 illustrates the PSNR performance. We can see that the PSNR of progressive S-frames changes step by step while keeping the same quality after switching.

4. CONCLUSION

In this paper, we proposed an efficient bitstream switching method using the progressive S-frame approach. The progressive S-frame approach using multiple S-frames with optimal quantization steps can reduce progressively the mismatch between two pre-encoded streams using different quantization steps. The progressive S-frame approach can provide better coding efficiency than the periodical I-frame and S-frame approaches, and smooth the abrupt increment of bit rate. The simulation results demonstrated that the proposed method provides much better performance than other existing methods. Based on the simulation results, we can conclude that the proposed method is useful for video stream switching in heterogeneous networks.

5. REFERENCES

[1] M. Stemm and R.-H. Katz, "Vertical handoffs in wireless overlay network," *ACM Trans. Networking and Applications*, vol. 3, pp. 335–350, June 1998.

[2] J. McNair and F. Zhu, "Vertical handoffs in fourth-generation multinetwork environments," *ACM Trans. Networking and Applications*, vol. 35, no. 3, pp. 272–278, Aug. 1989.

[3] N. Färber and B. Girod, "Robust H.263 compatible video transmission for mobile access to video servers," in *Proc. ICIP-97*, Oct. 1997, vol. 2, pp. 73–76.

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

[5] E. Setton and B. Girod, "Rate-distortion analysis and streaming of SP and SI frames," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, pp. 733–743, June 2006.

[6] T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 246–250, Feb. 1997.

[7] T. Wiegand and B. Girod, *Multi-frame Motion-compensated Prediction for Video Transmission*, Kluwer Academic, Boston, USA, 2001.

[8] C.-P. Chang and C.-W Lin, "R-D optimized quantization of H.264 SP-frames for bitstream switching under storage constraints," in *Proc. ISCAS 2005*, May 2005, vol. 2, pp. 1242–1245.

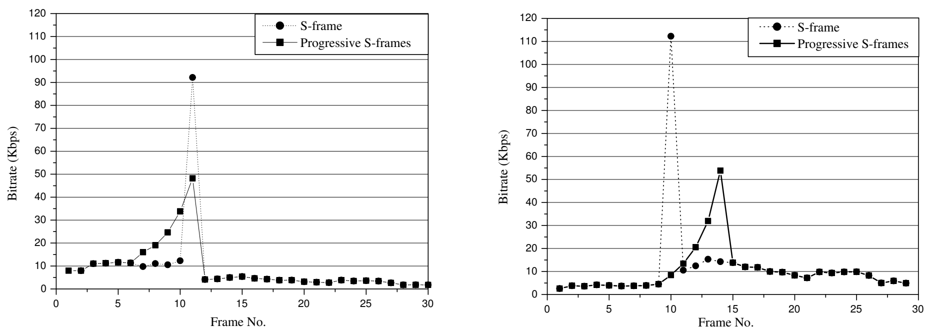
[9] *Joint Video Team Reference Software JM10.2*, Jan 2006. Available: http://iphome.hhi.de/suehring/tml/download/.

Table 2. Comparisons of the total bit rates and average PSNRs over 300 frames with 10 times switching

Sequences	Periodic I-frames		S-frames		Progressive S-frames	
$(QP_1 - QP_2)$	Total Bits	PSNRs	Total Bits	PSNRs	Total Bits	PSNRs
Foreman (37-29)	3082400	34.11	3645400	33.84	3273840	33.72
News (33-24)	3742560	38.05	4239640	37.84	3976840	37.76
Stefan (44-36)	3341280	27.44	4182720	27.04	3795280	26.98

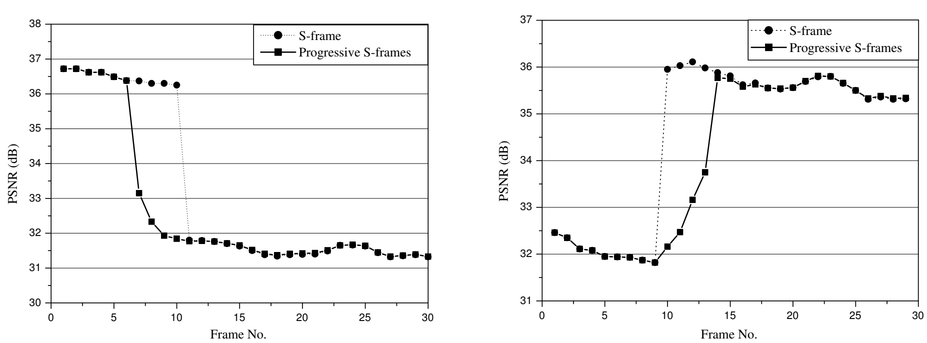
Table 3. Comparisons of the total bit rates and average PSNRs over 300 frames without switching

Sequences	Periodic I- frames		S-frames		Progressive S-frames	
(QP)	Total Bits	PSNRs	Total Bits	PSNRs	Total Bits	PSNRs
Foreman (37)	4426080	36.28	3508880	35.97	3508880	35.97
Foreman (29)	1738720	31.95	1295600	31.74	1295600	31.74
News (33)	5409840	40.98	3803680	40.80	3803680	40.80
News (24)	2075280	35.13	1289680	34.97	1289680	34.97
Stefan (44)	4791120	30.02	3549680	29.67	3549680	29.67
Stefan (36)	1891440	24.87	1314800	24.48	1314800	24.48



(a) (b)

Fig. 1. The changes of bit rates: (a) from high to low, (b) from low to high.



(a) (b)

Fig. 2. The changes of PSNRs: (a) from high to low, (b) from low to high.