

CODING WITH TEMPORAL LAYERS OR MULTIPLE DESCRIPTIONS FOR LOSSY VIDEO TRANSMISSION

*Sila Ekmekci**, *Thomas Sikora*⁺, *Pascal Frossard*^{*}

Ecole Polytechnique Fédérale de Lausanne (EPFL)^{*}
Signal Processing Institute
CH-1015 Lausanne / Switzerland

Technical University Berlin⁺
Institute for Telecommunications
D-10587 Berlin / Germany

ABSTRACT

In this paper, we compare temporal layered coding (TLC), as well as single-state coding (SSC), to multi-state video coding (MSVC) in the context of lossy video communications. MSVC is a MDC scheme where the video is coded into multiple independently decodable streams each with its own prediction process and state. The performance of these three coding schemes are analyzed at different loss rates and coding options, under the assumption that each packet contains the complete coded data for a frame, and the total bit rate is kept constant. To substitute the lost frames, MSVC employs state recovery based on motion compensated frame interpolation, whereas SSC and TLC repeat the last received frame. Results show that MSVC outperforms SSC and TLC for high motion sequences, and also for low motion sequences at high loss probabilities, due to increased state recovery ability of the system. Additionally, if one of the parallel channels of MSVC is in bad condition, unbalanced MSVC that allocates less bit rate to this channel, becomes favorable. Finally, increased error resilience with intra-GOB or frame update improves the system performance for high motion sequences at high loss rates, whereas for low motion sequences, intra updates are disadvantageous due to the penalty on the source coding quality.

1. INTRODUCTION

Video Communication over wireless networks and Internet is still a demanding issue due to long delays and packet losses which cause quality degradation. Multiple Description Coding [6] is a source coding technique used for transmission over error-prone channels. Two or more descriptions of the same source are generated which are mutually refining. If only one description is received the reconstruction distortion is D_1 or D_2 . If both descriptions are received, however, a lower distortion D_0 is achieved. Multi-state video coding (MSVC) is a particular multiple description scheme where the video frames are split into two subsequences constituted of even and odd frames. Each subsequence can be encoded and decoded independently from each other. The advantage is twofold: 1- Even if one of the streams is lost the other one can still be decoded. 2- The lost

frames can be reconstructed by interpolation of their previous and next neighbors from the other subsequence (state recovery). Block diagram of the MSVC system is given in Figure 1. Reference [1] shows that if each frame is transmitted in a separate packet, MSVC outperforms SSC in recovering from single as well as burst errors. In this work, we compare the average performance of MSVC to SSC and also to Temporal Layered Coding (TLC) at the same total bitrate and at various channel loss rates (independent and uniformly distributed losses). Similar to MSVC, multiple bitstreams are generated in TLC [5]. Even if some portion of the bitstream is dropped due to channel problems, a reconstruction may still be possible with the received rest of the bitstream. However, in layered coding the reception of the base layer is mandatory for the decoding of the enhancement layer, contrarily to multiple description coding that enables independent decoding of the descriptions. Descriptions are in general mutually refining, while layers are hierarchically ordered, and thus natural candidates for differentiated protection.

In the literature there are several comparison studies of specific layered and multiple description schemes for special transmission scenarios. A summary is given in [7] which shows that MDC has advantages over Layered Coding for networks where feedback is not available due to lack of back channel or due to long round trip times or low delay constraints as for streaming video applications. The results vary when the source coding techniques are combined with FEC or ARQ.

In this paper, we consider streaming scenarios where multiple channels are available between the server and the client. MSVC uses both transmission paths, as well as TLC that separates the base and enhancement layers, where the enhancement layer contains each second frame coded as a B-frame. In SSC, the complete bitstream is sent over the same path. We investigate two cases: In the first case, the loss probabilities of the two paths are the same, i.e. $p_1 = p_2$, but the losses are independent from each other. In the second case, SSC is compared to MSVC, where one of the paths used for MSVC is lossless. We assume that each packet contains a frame and when a packet is lost, all information about the frame including the motion information is lost. In case of loss, SSC uses the last received frame to replace lost frames and MSVC implements state recovery based on motion compensated frame interpolation [1]. For all the comparisons, we target the same total bitrate R_T for the

This work was supported in part by Technical University Berlin and the Swiss National Science Foundation.

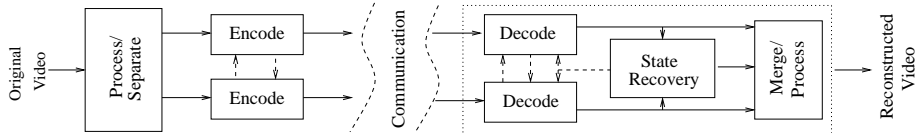


Fig. 1. Block Diagram of the MSVC System

three coding methods. For MSVC, we investigate both balanced as well as unbalanced quantized MSVC ([4] and [3]). In the sequel, $MSVC_b$ denotes balanced quantized MSVC where the total bitrate R_T is allocated equally between the two streams considered, whereas $MSVC_u$ is the unbalanced MSVC where more bitrate is allocated to the more reliable channel. Additionally, we investigate the effect of GOB and frame intra updates on the three coding techniques. This way, we increase the number of resources for optimal rate allocation.

This paper is organized as follows: Section 2.1 presents the experimental setup, whereas several streaming scenarios are analyzed in section 2.2. Section 3 discusses the experimental results, and presents a series of heuristics particularly useful in the choice of an efficient coding strategy. Section 4 concludes the paper.

2. COMPARATIVE ANALYSIS

2.1. Experimental Setup

We consider two types of sequences: Foreman as a high motion sequence and Akiyo as a low motion sequence. The coding parameters (quantization stepsizes of intra and remaining frames, periods of GOB and frame updates and total bitrate R_T) are given in Tables 1, 2 and 3 for MSVC, SSVC and TLC respectively, “A.” denotes Akiyo and “F.” stands for Foreman. The different rate allocations under consideration for $MSVC_u$ are given in Table 4. For all the comparisons, we target the same total bitrate R_T for all three coding methods (140 kbit/s for Foreman and 19 kbit/s for Akiyo). We considered the first 200 frames from each sequence.

The lossy transmission is simulated using random loss patterns. The average PSNR over all frames in each run is averaged over all loss patterns. 100 randomly generated loss patterns are used for each loss rate. MSVC uses Approach 2 from [1], that aims at maximizing the average frame PSNR by using interpolation from the past and future frames not only for lost frames, but also if the current frame PSNR can be increased through interpolation instead of using the received packet [4].

| | QP I/P | i.-GOB per. | i.-fr. per. | R_T [kbit/s] |
|------------------|------------|----------------|----------------|-------------------|
| F. | 17/17 | | | 158.21 |
| F. i.-GOB | 17/(20/21) | 1 | | 139.31 |
| F. i.-fr. | 17/23 | | 9 | 140.82 |
| A. | 21/21 | | | 18.68 |

Table 1. MSVC+intra-updates, Coding Parameters

| | QP I/P | i.-GOB per. | i.-fr. per. | R_T [kbit/s] |
|------------------|-----------|----------------|----------------|-------------------|
| F. | 16/16 | | | 137.28 |
| F. i.-GOB | 16/17 | 3 | | 136.51 |
| F. i.-fr. | 16/17 | | 30 | 133.88 |
| A. | 18/17 | | | 20.80 |

Table 2. SSVC Coding Parameters

| | QP I/P,B | i.-GOB per. | i.-fr. per. | R_T [kbit/s] |
|------------------|-------------|----------------|----------------|-------------------|
| F. | 14/14 | | | 147.95 |
| F. i.-GOB | 15/15 | 2 | | 142.86 |
| F. i.-fr. | 15/15 | | 15 | 143.60 |
| A. | 17/17 | | | 18.62 |

Table 3. TLC Coding Parameters

| | QP | QP |
|------------------|--------|-------|
| F. | 111.88 | 27.48 |
| F. i.-GOB | 106.38 | 34.02 |
| F. i.-fr. | 83.90 | 55.68 |
| A. | 13.92 | 5.03 |

Table 4. Unbalanced rate allocation

2.2. Results and Observations

Figure 2 gives a comparison of the coding methods SSC, TLC and MSVC. MSVC outperforms SSVC by 5 to 7 dB over the loss rate range. This is a huge gain although we assumed that both channels are error prone. Moreover SSC outperforms TLC as the loss rate increases: at 20% loss rate, the gap between the two methods is about 0.8 dB. Figure 3 shows the case when the first channel used for MSVC is lossless whereas the second one has the same loss rate as the channel used by SSC. The probability that we catch a second channel with a better transmission condition is the main idea behind path diversity. At 20% loss rate, $MSVC_u$ outperforms SSC by 14 dB when $p_1 = 0\%$. The PSNR gap between $MSVC_u$ and $MSVC_b$ is about 1 dB at 20% loss rate, i.e.: unbalanced channels call for unbalanced rate allocations.

Figures 4 and 5 show the same comparisons for the low motion sequence Akiyo. Error concealment is easier due to low motion. Therefore SSC with repetition of the last received frame as concealment technique gives good performance in lossy environment and outperforms MSVC when

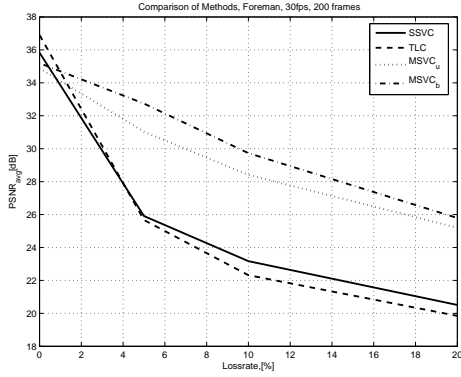


Fig. 2. Comparison of SSVC, SSVC-TC, TLC, TLC-TC and MSVC, all channels have the same loss rate, Foreman.

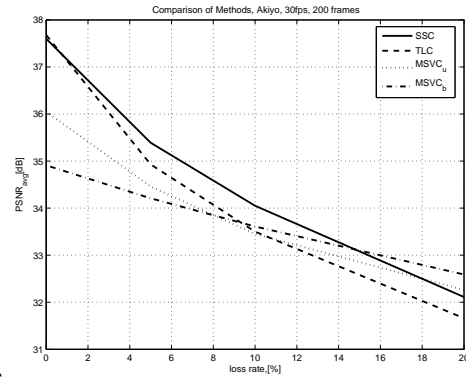


Fig. 4. Comparison of SSVC, SSVC-TC, TLC, TLC-TC and MSVC, all channels have the same loss rate, Akiyo

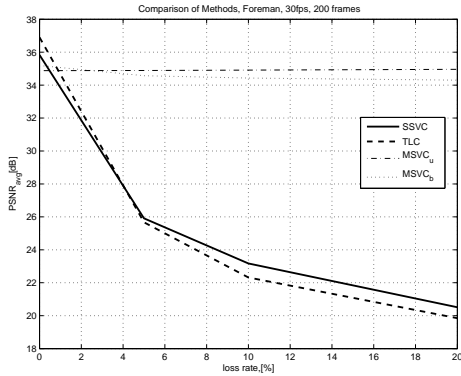


Fig. 3. Comparison of SSVC, SSVC-TC, TLC, TLC-TC and MSVC, one of the MSVC channels is lossless, Foreman.

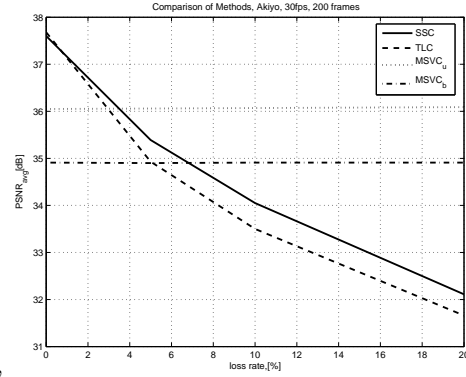


Fig. 5. Comparison of SSVC, SSVC-TC, TLC, TLC-TC and MSVC, one of the MSVC channels is lossless, Akiyo.

loss rate is smaller than about 15% as shown in figure 4. But when loss rate increases beyond this limit, it is better to employ MSVC. Although unbalanced rate allocations are better at smaller loss rates ($MSVC_u$), larger loss rates require balanced rate allocations ($MSVC_b$). When the first channel is lossless, $MSVC_u$ performs always better than $MSVC_b$. Moreover, $MSVC_u$ performs 4dB better than SSC at 20% loss rate.

In the next step, we compare the methods when intra-updates are used. Figures 6 and 7 show the cases with intra GOB- and frame updates for Foreman respectively. The threshold loss probability increases with the introduction of updates. All coding techniques profit from updates at high loss probabilities. The performance increase in SSC and TLC is larger than in MSVC. Using intra updates for Akiyo is not a good idea, since the gain of motion compensation is very high and a small rate is available for coding (the corresponding figures are omitted here due to limited space). TLC performs best, since enhancement layer uses no update. The differences between different methods are smaller for Akiyo. MSVC outperforms SSC at about 15% loss rate. Moreover, MSVC outperforms SSC when the first channel is lossless.

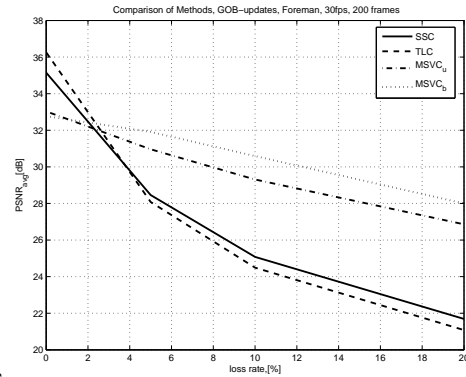


Fig. 6. Comparison of SSVC, SSVC-TC, TLC, TLC-TC and MSVC, all channels have the same loss rate, Foreman with GOB-intra-updates

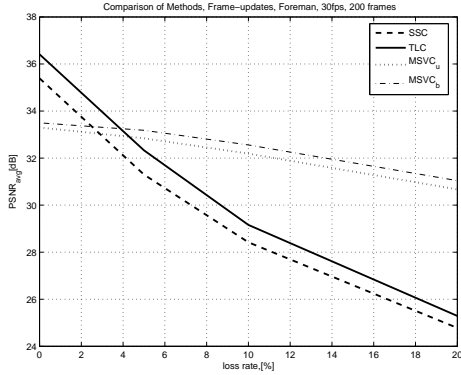


Fig. 7. Comparison of SSVC, SSVC-TC, TLC, TLC-TC and MSVC, all channels have the same loss rate, Foreman with frame-intra-updates

3. DISCUSSION

For Foreman, when both of the channels have the same loss rate, MSVC outperforms SSC and TLC. The difference increases with increasing loss rate. But at lossless transmission there is a penalty for MSVC due to sequence splitting, i.e. increased temporal distance which decreases the prediction gain. Moreover introducing intra-updates increases the performance of MSVC as well as of other coding methods for high loss rate. But for lossless transmission, the performance drops due to the wasted bitrate for intra coding. For the same total bitrate R_T , intra updates give better performance than GOB intra updates. Additionally, we see that balanced loss probabilities call for balanced rate allocations.

For Akiyo, however, repetition of the last received frame in case of losses gives good results due to low motion. MSVC outperforms SSC only at high loss probabilities. Frame splitting in MSVC is disadvantageous due to the high cost of intra frames (the first frame of each subsequence is coded intra).

The slopes of distortion-loss rate relations for MSVC_u and MSVC_b are very small when one of the channels is lossless, as shown in Figures 3 and 5. The reason is that the average PSNR for the lossless received stream does not change with the loss rate, and that the lost frames in the lossy stream are reconstructed through interpolations from the lossless stream. Moreover, the slope of MSVC_u is smaller than of MSVC_b. Since more bitrate is allocated to the reliable channel (smaller quantization distortion), interpolation errors in case of losses are smaller. For Akiyo, the slopes of MSVC_u and MSVC_b are nearly zero, since frame interpolation gives always good results due to low motion.

The experiments show that if one of the channels is lossless, MSVC outperforms both SSC and TLC for both sequences. Even if both of the channels have the same loss probability, at high loss probabilities MSVC gives the best performance. The threshold loss probability for MSVC is dependent on the motion content of the sequence, e.g. about 0.5% for Foreman and 15% for Akiyo. The performance gap between SSC and MSVC is larger for high motion sequences.

4. CONCLUSIONS

For deciding which coding technique and which coding options are to choose, the following factors are important: motion content of the sequence, and loss rate of the channels (obtained by methods like channel probing etc.). Generally, MSVC is to be preferred at high loss probabilities. The threshold loss probability is higher for low motion sequences. Moreover, introduction of intra-updates increases the threshold loss probability, i.e. SSVC and TLC profits more from intra updates than MSVC. For high motion sequences, MSVC combined with frame-intra-frames gives the best results at high loss probabilities. For low motion sequences, however, intra-updates decrease the system performance. Frame-intra-updates are more efficient than GOB-intra-updates in recovering from state errors.

In this paper, we compared MSVC to SSC and TLC at different loss rates and coding options. In each case, we targeted a constant total bitrate R_T to allow a fair comparison. Both for MSVC and TLC, we assumed that two independent channels are in use with independent loss patterns. We investigated both balanced and unbalanced operation for MSVC. In balanced operation half of the total bitrate is allocated to each stream, whereas in unbalanced case more bitrate is assigned to the first channel which is more reliable than the second one.

Further work will focus on joint optimization of redundancy, frame rate and also the quantization stepsize of the MSVC streams depending on the channel loss probabilities.

5. REFERENCES

- [1] J. Apostolopoulos. Reliable video communication over lossy packet networks using multiple state encoding and path diversity. *VCIP*, January 2001.
- [2] J. Apostolopoulos and S.J. Wee. Unbalanced multiple description video communication using path diversity. *ICIP*, October 2001.
- [3] S. Ekmekci and T. Sikora. Unbalanced quantized multi-state video coding: Potentials. *Picture Coding Symposium (PCS'04)*, December 2004.
- [4] S. Ekmekci and T. Sikora. Unbalanced quantized multiple description video transmission using path diversity. *Electronic Imaging 2003, SPIE*, January 2003.
- [5] M. Ghanbari. Two-layer coding of video signals for vbr networks. *IEEE JSAC*, vol. 7, June 1989.
- [6] V.K. Goyal. Multiple description coding: Compression meets the network. *IEEE Signal Processing Mag.*, vol. 18, no. 5, pp. 74-93, Sept. 2001.
- [7] Y. C. Lee, J. Kim, Y. Altunbasak, and R. M. Mersereau. Layered coded vs. multiple description coded video over error-prone networks. *Signal Processing: Image Communication*, May 2003.