Proxy-based Error Tracking for H.264 Based Real-time Video Transmission in Mobile Environments

Wei Tu and Eckehard Steinbach
Institute of Communication Networks, Media Technology Group
Technische Universität München
{wei.tu, eckehard.steinbach}@tum.de

Abstract

Error tracking (ET) is an error resilience technique for real-time video transmission over error-prone communication channels. In this paper we propose proxy-based ET for communication scenarios where the sender is in the wired Internet and the receiver is connected via a wireless link. Our main assumptions in this work are that there is a strong imbalance between the transmission rates available in the wired and the wireless Internet and that the round-trip delay is mainly caused by the wired part of the connection. We also assume that the majority of the packet loss is caused by the wireless link. In order to allow the proxy server to perform error tracking, an additional update stream is sent through the wired network. This additional information is used by the proxy to improve the performance on the wireless link. We show that under these assumptions proxy-based error tracking leads to significantly improved performance for H.264 based real-time video communication in comparison to traditional end-to-end error tracking.

I. Introduction

In 3G networks, video services are expected to be the most popular ones and maybe the key factor for success. Wireless video applications without real-time constraints (e.g. Multimedia Message Service) have been introduced to the market. However, real-time video communication over wireless networks remains challenging. Decoding of erroneous or incomplete video bit-streams leads to severe quality degradations. Because of motion compensated prediction, these impairments also propagate in space and time and therefore stay visible for a significant amount of time. Hence, an error resilient transmission scheme is essential to achieve good quality in a wireless multimedia communication system. A recent overview of approaches for error resilient video transmission can be found in [4],[9].

Error Tracking (ET) [2],[5],[7] is an error resilience technique taking advantage of a back-channel to report corrupted image areas. The encoder reacts to this feedback by tracking the spatio-temporal error propagation. Those frame areas that have been identified to be corrupted are then updated using INTRA coding. Because the update happens in future frames to be encoded, ET does not introduce additional delay.

ET is suitable for real-time applications, but the performance is closely related to round-trip delay. In the Internet, a video sender may be located far away from the receiver and the long trip delay leads to serious error propagation in case of packet loss. Larger image areas are affected and need to be refreshed, which is critical when using a low bit-rate wireless channel, as INTRA coding leads to a bit-rate increase.

If we assume that the video sender is located in the wired Internet and the receiver is a wireless client, the round-trip time for ET is determined by the end-to-end delay between sender and receiver. One solution to cope with long delays between sender and receiver would be to use the base station that serves the mobile client as a gateway and to separate the video transmission into two separate parts: a real-time video communication between the sender and the base station and a video communication between the base station and the mobile client. This approach would lead to a small round-trip time for the second part and therefore good results for error tracking. The disadvantage of this approach, however, is that the base station would have to decode and re-encode (transcode) the video for feedback-triggered INTRA updates. In [3] a proxy-based approach for error robust video transmission has been proposed that transcodes only those MBs that are found to be distorted by transmission errors. The error signal is determined by the proxy and the DCT coefficients of the outgoing bitstream are updated accordingly. In [3] the complexity is still significant as the incoming bitstream has to be decoded and affected MBs have to be transcoded. In order to reduce computational complexity on the base station we propose in this paper to send side information along with the normal video stream that can be used by the base station to perform INTRA updates of affected image regions. In our approach, the base station tracks the error propagation by parsing the bit-stream, i.e., extracting the coding mode and motion vector information, and then uses the side information to perform INTRA refresh of corrupted image areas. This is achieved by replacing parts of the original bitstream with corresponding parts from the side information bitstream.

This paper is organized as follows. The proposed architecture of proxy-based ET is described in Section II. In Section III, implementation details of ET for H.264 encoded video are given. In Section IV experimental results for our proxy-based approach are presented and a comparison with end-to-end error tracking is made. Section V gives concluding remarks.

II. Proxy-based Error Tracking

ET does not introduce additional delay, but the quality of error recovery depends on the round trip delay. To minimize this delay, we perform ET on the base station as it is the interface between the wired and wireless network. The base station is also the nearest point to the mobile terminal. A low delay feed-back channel exists between the mobile terminal and the base station, which allows fast error recovery when using ET.

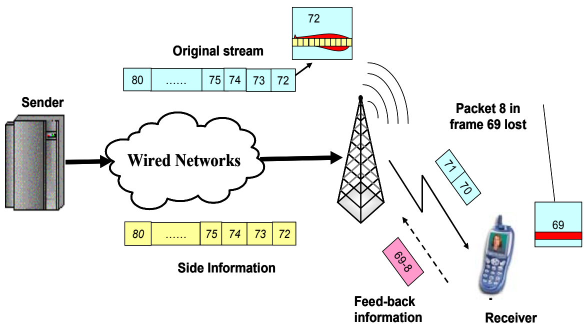

Fig.1 Proxy-based error tracking scenario

Fig.1 illustrates the architecture of our proposal as well as the procedure of ET. Two video streams are transmitted simultaneously to the base station, as shown in Fig.1, one P-frame stream (original stream) and an additional I-frame stream (side information). The downlink channel between the base station and mobile station is used for the video transmission and the uplink channel is used to transmit the error messages (feed-back information). In traditional ET, these error messages are forwarded back to the video sender. In our approach, these messages are processed by the proxy. In Fig. 1, we assume that the end-to-end delay corresponds to about 10 frames while the delay between the base station and the client is only about 2 video frames. From the figure we can see with end-to-end ET, the error can be recovered around frame 80. On the contrary, with the proxy at the base station, the error recovery will start from frame 72.

III. Implementation

The dependencies of blocks in successive frames are the essential information for ET. The motion vectors of the blocks determined during motion estimation provide adequate information for accurately tracing error propagation. MB-based and pixel-based ET have been compared in [1]. In this paper, pixel-based ET is used. Fig. 2 illustrates the pixel wise tracking procedure. Any pixel's motion dependency can be found by tracing back the motion vector of the block it belongs to. In Fig. 2, the proxy receives a NAK that indicates an error occurred at MB 30 of frame 4 while the next frame to be sent is 7. Thus, we can first trace every pixel in frame 5 with its motion vectors to see whether it refers to the erroneous area in frame 4. In this step, all affected pixels in frame 5 can be found and we can go on for the next frame with its own motion vectors. For this example, after 3 iterations, the distorted pixels in frame 7 are worked out.

Here, forward ET is used instead of the backward tracking proposed in [1]. In the configure file of the H.264 encoder, the maximum motion vector length allows us to determine the possibly affected area in the next frame, and only pixels in this area need to be checked.

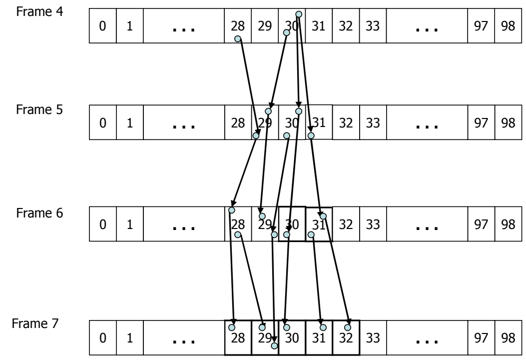

Fig. 2 Pixel wise forward error tracking procedure

The selection of side information is also important to the proxy-based ET. INTRA encoded P-frames of the original bitstream are used as the side information so that the mismatch obtained when replacing parts of the original video bitstream with INTRA side information is minimized. A similar approach has been proposed in [10] in the context of bitstream switching for Internet Media Streaming. To reduce the overhead of the side information, a subset of INTRA frames or a subset of INTRA slices can be transmitted.

IV. Experimental Results

A wireless channel is always subject to various kinds of errors. The bursty packet loss behavior of a wireless channel is modeled using the well known Gilbert two-state Markov Model [8]. Foreman in QCIF resolution is used as the test sequence and 10 fps is selected as the frame rate. The H.264 test software version JM 6.1d [6] is employed as the video codec. When we perform ET and INTRA-refresh, some INTER- MBs will be replaced by INTRA-MBs, which means that additional bits are needed for the error control. When using ET in our experiments, we reserve a fixed part of the total transmission rate for error control. In the following, when using ET the target bitrate of the video encoder is 80% of the total transmission rate.

A) Performance of traditional ET at 64kbit/s

In Fig.3 four experiments show the improvement obtained when using ET at a burst loss rate of 1%. The upmost curve shows the PSNR for the 150 frame sequence at 64kbps in an error free environment, 38.40dB PSNR on average. The bottom curve stands for reconstructed video suffering from 1% packet loss. Simple previous frame error concealment is used to conceal lost packets [9]. The average PSNR is 31.03dB. In this experiment, two round trip delays, 200ms and 1000ms are used to show the performance of ET as a function of feed-back delay. When using ET the target bitrate at the encoder is 52kbps. This leaves about 12 kbps for selective INTRA updates trigged by ET. ET with short delay allows the stream to recover to a good quality in a much shorter time than ET with longer delay. The average PSNR is 35.72 dB and 34.59 dB, respectively. Compared with the concealment-only case, ET can achieve about 4 dB gain and another 1 dB gain is obtained if the short delay applies.

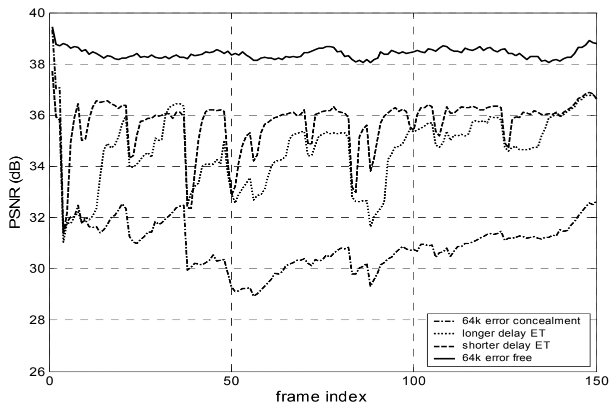

Fig.3 Error tracking improvement curves
at a mean packet loss rate of 1%

The improvements obtainable with ET increase with the packet loss rate, as shown in Table 1. Here the round trip time is 2 video frames (200ms).

	0.5%	1%	5%	10%	15%
52kbps with ET	36.27	35.72	33.65	31.58	30.05
64kbps w/o ET	33.51	31.03	26.51	23.83	21.98

Table 1 Performance of ET as a function of loss rate

B) Best QP for proxy-based ET at different channel loss rates

We now investigate how the improvement achieved by ET changes depending on the QP of the INTRA side information. In Fig.4, the solid curves illustrate the average PSNR as a function of the QP of the INTRA side information at different error rates. The larger the error rate, the more we can gain by using the proposed proxy-based ET approach. When QP is small, very few MBs can be updated every frame. At large QP, many MBs can be updated but the update quality is poor.

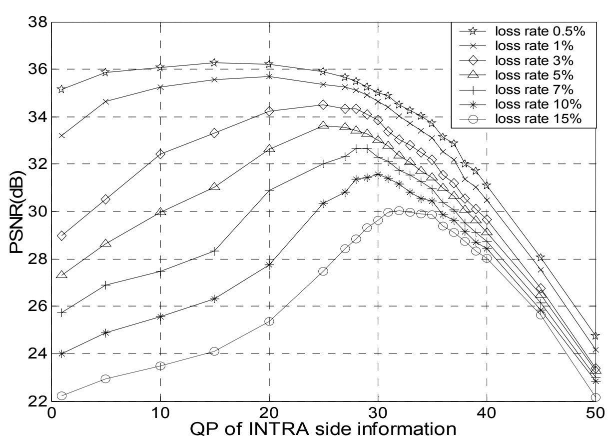

Fig. 4 Best QP for different loss rates in a 64kbps channel, QP of the original bit-stream is 27

Fig.4 shows that this best QP increases with increasing channel loss rate. Knowing this relationship, we can select QP according to the observed channel loss rate.

C) Best QP as a function of channel capacity

The relation between the best QP of the INTRA-stream and the channel loss rate has been investigated in B). Now we investigate how it changes with the available transmission rate as different rates will result in a different number of MBs that can be refreshed. In this simulation, 32kbps, 64kbps, 128kbps, and 256kbps are used as target transmission rates. From the curves in Fig.5, a similar relationship as in Fig. 4 can be observed between the best QP and the transmission rate. As the available transmission rate decreases, the optimal QP of the INTRA side information increases. So here we can make the conclusion that the best QP of the I-stream should be a joint function of channel rate and loss rate. These are two important characteristics of the wireless channel. If we know the channel condition, we can select the most suitable QP to obtain the highest gain by ET.

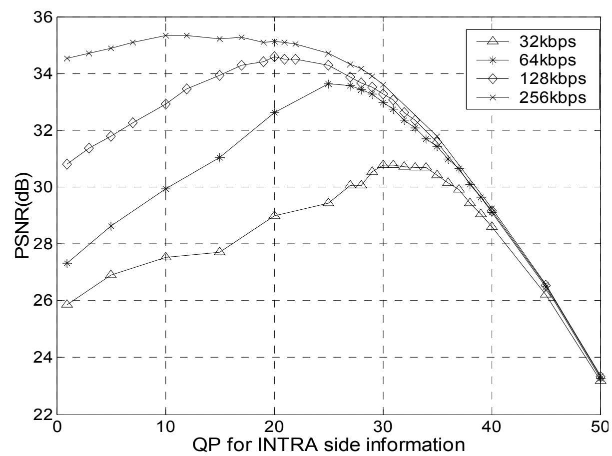

Fig.5 Best QP for different transmission rates at a loss rate of 5%

D) MB refresh versus slice refresh

Because of the differential motion vectors in the INTER-Macroblocks, INTRA update can only be done at the slice level. When one slice includes only one MB, we can update selected MBs. However, this would lead to significant overhead, which is not feasible in a wireless environment. As a compromise between overhead and ET performance, we put one row of MBs into one slice and the INTRA update is also done at the slice level. The overhead cost for 9 slices per frame for foreman at 64kbps is about 10kbps, so we set the QP of the P-frame to be 30 instead of 27 to save the bits for this overhead. For end-to-end ET one frame corresponds to one slice and INTRA refresh is performed for individual MBs. For proxy-based ET one frame corresponds to 9 slices and entire slices are INTRA-refreshed. Although the PSNR decreases compared to the MB level update, we can still get significant improvements for larger loss rates compared to the case that end-to-end ET is used. This is shown in Table2.

Loss rate		0.5%	1%	3%	5%	7%	10%	15%
End-to-end QP_P=27 1s, RTT	PSNR	35.71	34.59	31.96	30.80	30.04	29.13	24.43
	Best QP	15	20	24	25	28	30	32
Proxy QP_P=30 0.2s RTT	PSNR	34.60	34.52	33.90	33.25	32.74	32.22	31.51
	Best QP	17	21	25	28	29	31	34

Table 2 Comparison of end-to-end and proxy based error tracking

E) Overhead in the wired network

As shown in the above experiments, the proxy-based ET leads to improved quality compared with the conventional ET at the encoder. However, an extra stream is needed to be sent at the same time, which will add overhead to the transmission in the core network.

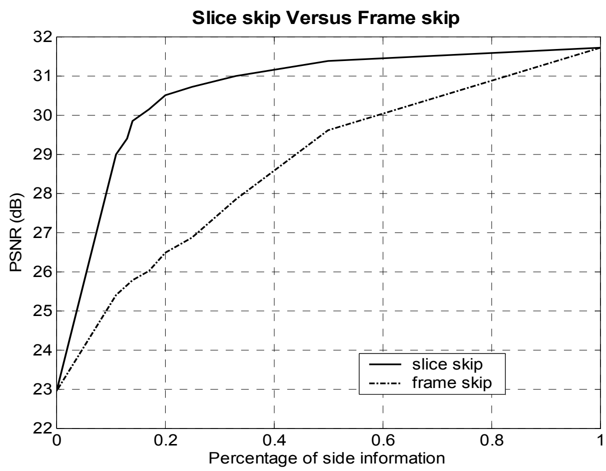

Fig.6 Performance of error tracking as a function of side information skipping rate at 10% packet loss rate

Two approaches can be used to control the bitrate of the additional I-stream. One is frame skip. In this case, I-frames are sent as side information every N frames. The other is slice skip. INTRA-slices are sent as side information every N slices.

Our experiments (Fig.6) show that the slice-based skipping approach achieves significantly better quality compared to the frame-based approach at the same bitrate. When side information is set to 20% of the full I-stream bitrate, 1 dB drop from the best quality is observed but still more than 7 dB improvement in comparison to the case when no ET is used and 2 dB in comparison to end-to-end ET.

V. Conclusion

Error Tracking is a video resilience technique suitable for real-time video transmission as it does not introduce any additional delay. In our investigation, a low bitrate last link transmission on a wireless link is added as an additional constraint. ET is performed on a proxy that is located at the base station.

In order to enable low complexity ET and INTRA refresh on the base station, we propose to send an INTRA side information bitstream in addition to the regular video stream from the sender to the base station. The side information is generated at the sender by INTRA re-encoding decoded P-frames from the original video stream.

We determine the best QP of the I-frame side information as a function of the channel condition. It is shown that the side information does not have to be available for every slice in the original stream. Very small degradations in quality are observed when only every 5th slice is sent as INTRA side information.

References

[1] P.C. Chang, T.H. Lee, “Precise and Fast Error Tracking for Error-Resilient Transmission of H.263 Video,” IEEE Transaction on Circuits and Systems for Video Technology, vol. 15, no. 4, Jun. 2000.

[2] B. Girod, N. Färber, “Feedback-Based Error Control for Mobile Video Transmission,” Proceedings of the IEEE, Special Issue on Video for Mobile Multi-media, vol. 97, no. 10, pp. 1707-1723, Oct. 1999.

[3] T.C. Wang, H.C. Fang, L.G. Chen, “Low Delay and Error Robust Wireless Video Transmission for Video Communications,” 2002 IEEE International Conference on Multimedia and Expo (ICME 2002), Lausanne, Switzerland, Aug. 2002.

[4] Y. Wang, S. Wenger, J.T. Wen, A.K. Katsaggelos, “Review of Error Resilient Coding Techniques for Real-Time Video Communications,” IEEE Signal Proc. Magazine, vol. 17, no. 4, pp. 61-82, Jul. 2000.

[5] E. Steinbach, N. Färber, B. Girod, “Standard Compatible Extension of H.263 for Robust Video Transmission in Mobile Environments,” IEEE Trans. Circuits and Sys. for Video Tech., vol. 7, no. 6, pp. 872-881, Dec. 1997.

[6] http://bs.hhi.de/~wiegand/JVT.html

[7] M. Wada, “Selective recovery of video packet loss using error concealment,” IEEE Journal on Selected Areas in Communications, vol. 7, pp. 807-814, 1989.

[8] E.N. Gilbert, "Capacity of a burst-noise channel," Bell Syst. Tech. J., vol. 39, pp. 1253--1265, Sep. 1960.

[9] Y. Wang and Q. Zhu, "Error control and concealment for video communications: A review," Proceedings of IEEE, Special Issue on Multimedia Signal Processing, pp. 974 - 997, May 1998.

[10] N. Färber and B. Girod, “Robust H.263 compatible video transmission for mobile access to video servers,” in Proc. IEEE Int. Conference to Image Processing (ICIP), vol. 2, pp. 73-76, Santa Barbara, CA, Oct. 1997.