Offset Distortion Traces for Trace-Based Evaluation of Video Quality after Network Transport (Extended Version)

Patrick Seeling and Martin Reisslein
Dept. of Electrical Engineering
Arizona State University
Tempe, AZ 85287-5706
{patrick.seeling, reisslein}@asu.edu
http://trace.eas.asu.edu

Frank H.P. Fitzek
Dept. of Communication Technology
Aalborg University
Aalborg, Denmark, DK-9220
ff@kom.aau.dk

Abstract

Video traces containing the sizes and (PSNR) qualities of the individual frames of a video encoding are a convenient video representation for the evaluation of video networking mechanisms. These video traces can be used to find the frame loss probabilities of a lossy networking mechanism, but can not give the PSNR video quality after lossy network transport. To date the video quality after lossy network transport could only be determined through experiments with actual video or by approximating the quality of the frames affected by a loss with some low PSNR quality.

In this paper we introduce and evaluate *offset distortion traces* with which the video quality after lossy network transport can be accurately determined without requiring experiments with actual video. We explain how the offset distortion traces can be used by networking researchers without equipment or experience in video signal processing to accurately evaluate video networking mechanisms in terms of the PSNR video quality.

Index Terms

Offset distortion, video traces, video quality estimation, PSNR, RMSE

I. Introduction

With the increasing popularity of multimedia applications and services, large portions of the traffic in the Internet are expected to be encoded video data. For networking research in the area of video transmission, the encoded video can be represented by (i) the encoded bit stream, (ii) video traces, or (iii) a model. The encoded video has the drawbacks that it is typically large in size and copyright protected, limiting exchange among researchers. Also, experiments with actual video bit streams require typically specialized equipment and experience in video signal processing. Accurate and parsimonious video traffic models, on the other hand, are still an ongoing research issue. Video traces provide an appealing approach for conducting research on the transmission of video. Video traces are typically in simple text format and carry only the video frame sizes and the video frame qualities. In contrast to encoded video data, video traces do not carry the actual video information and are therefore exchangeable among researchers without copyright issues. Also, no special equipment is needed, video traces can be employed in standard discrete event simulation, widely used in networking research.

Video traces have been used in networking research since the mid 1990s (e.g., [1], [2], [3], [4], [5], [6], [7], [8], [9]) and have evolved from simple video frame size traces to traces that additionally carry information about the video frame quality [10]. To determine video quality, subjective tests or objective metrics can be applied to video bit streams. Determining the video quality through subjective tests resulting in mean opinion scores (MOS) [11] requires test subjects and is therefore typically impractical for utilization in networking research. The objective video quality is typically measured in terms of the root mean square error (RMSE) and the peak signal to noise ratio (PSNR), which is computed from the RMSE and it is widely accepted that these metrics give a reasonably accurate measure of the perceived video quality. (Throughout this paper we refer to the RMSE as *distortion* and to the PSNR as *quality*.)

For networking research, the frame loss probability, which is defined as the long run fraction of frames that miss their playout deadline at the receiver, can be easily determined. This networking metric, however, can not be directly translated into the video quality perceived by the user. Video traces that contain the video frame qualities in addition to the video frame sizes allow to determine the video quality perceived by the recipient(s) as long as there are no losses, i.e., all frames arrive in time [10]. Most video transport mechanisms, however, exploit some form of statistical multiplexing to accommodate the highly variable video traffic and thus incur some loss of video frames.

The most basic and common approach for overcoming a lost video frame is to re-display the last successfully decoded video frame until a new video frame is successfully decoded and displayed at the client. The current video traces, however, contain only encoded video frame qualities. Hence, network researchers using the currently available traces can make only a very rough qualitative approximation of the PSNR (quality) of the re-displayed frames. For a quantitative account of the frame loss one would either need to experiment with actual video [12], [13], [14] or approximate the quality of the frames affected by a loss by a low PSNR value (depending on the quality of the encoded video frame before transmission), e.g., less than 20 dB [10], although the meaningfulness of such a quantitative approximation may be questionable. The perceived quality for a video stream can then be estimated using elementary statistics such as mean and variation of the video frame qualities, whereby the video stream quality is generally maximized if the quality of individual frames is maximized and the variability of the quality among the frames of a video is minimized [15].

In this paper, we introduce video frame offset distortion traces in addition to the currently available traces to allow networking researchers to meaningfully assess the perceived video quality using only video traces. The offset distortion traces contain the qualities of re-displayed video frames. By jointly using the currently available video traces for successfully delivered video frames and the offset distortion traces for frames that are re-displayed, networking researchers are able to accurately determine the impact of lost video frames on the video stream quality as it is perceived by the receiving client.

II. Video Frame Quality

In this section we introduce the video frame distortion and quality metrics and how lost frames impact these metrics. We additionally introduce the elementary statistics that are used to evaluate the video stream quality.

A. Video Quality Definition

The objective video quality is typically calculated as peak signal to noise ratio (PSNR) between the unencoded original video data and the encoded and subsequently decoded video data. The PSNR is calculated using the root mean squared error (RMSE) between the pixels of the unencoded and the encoded and subsequently decoded video frame. Each individual pixel is represented by an 8-bit value for the luminance (Y) component, and a sub-sampled version of the image is used to store the two 8-bit values for the chrominance components hue (U) and intensity (V). Typically only the luminance component is taken into consideration for the calculation of the RMSE and PSNR, as the human eye is most sensitive to this component [16]. Let q denote the quantization scale (which relates inversely to quality) for an arbitrary video encoding and let N denote the total number of video frames in the video stream. We denote an individual pixel's luminance value in the nth original video frame at position (x, y) as $F_n^q(x, y)$ and its encoded and subsequently decoded counterpart by $f_n^q(x, y)$. Let X and Y denote the resolution in pixels of the source video. We calculate the video frame distortion as RMSE for all the luminance differences of an individual frame n encoded with the quantization scale q as

$$RMSE_n^q = \sqrt{\frac{1}{XY} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} [F_n^q(x, y) - f_n^q(x, y)]^2}. \tag{1}$$

The video frame quality as PSNR can be calculated from the RMSE as

$$Q_n^q = 20 \log_{10} \frac{255}{RMSE_n^q}. \tag{2}$$

With the N frames in a given video stream, we calculate the average video quality or video stream quality as

$$\overline{Q}^q = \frac{1}{N} \cdot \sum_{n=1}^{N} PSNR_n^q \tag{3}$$

and the variability of the video frame qualities measured as standard deviation as

$$\sigma^q = \sqrt{\frac{1}{(N-1)} \sum_{n=1}^{N} (Q_n^q - \overline{Q}^q)^2}. \tag{4}$$

To obtain a more useful variability metric taking the average video frame quality into account, we additionally calculate the coefficient of variation of the video frame qualities as

$$CoV^q = \frac{\sigma^q}{\overline{Q}^q}. \tag{5}$$

The video stream quality is generally maximized if the quality of individual frames is maximized and the variability of the quality among the frames of a video stream is minimized [15].

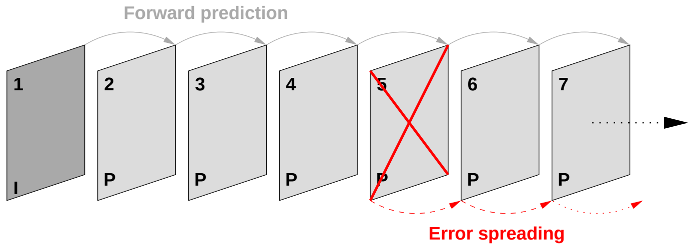

Fig. 1. Popular video coding scheme with inter-frame dependencies.

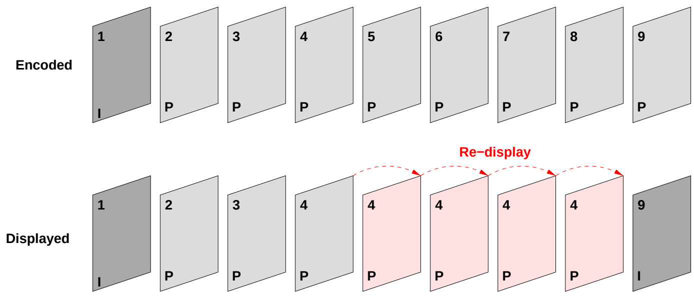

Fig. 2. Displayed video frames at the receiver with re-display of the last successfully received frame and I frame update $d = 4$ frames later.

B. Assessing Impact of Lost Frames with Video Bit Stream or through Approximation

To assess the impact of a lost video frame on the video quality, we consider without loss of generality a video sequence encoded with the *IPPP...* encoding pattern as illustrated in Figure 1. The *I* frames are intra-coded and rely on no other frame, whereas the forward predicted *P* frames rely on the previous I or P frames. We note that in addition to I and P frames, bidirectionally predicted (B) frames can be used as well. Frames of the B type rely on the previous and following I or P frames and can be accommodated in analogous fashion in the offset distortion traces.

For a general introduction, let us assume that frame number 5 is not received correctly as illustrated in Figure 1. Given the inter-frame dependencies resulting from the forward prediction used in the video encoding, the frames following the erroneous frame are considered to be not decodeable at the receiver. Let us furthermore assume that after d frames, the sender can update the reference at the receiver, e.g., by sending an I frame. In general, if frame $n+1$ is lost, frame n is re-displayed for frames $n+1, n+2, \ldots, n+d$ until the decoder receives a new I frame as reference for frame $n+d+1$. The resulting video that is displayed at the receiver with re-display of the frame before the lost frame for our example is illustrated in Figure 2. If the source video data (bit stream) would be available, the video frame distortion and qualities could be calculated as outlined in Section II-A for the correctly received frames 1–4. For frames 5–8, the video quality could be determined by calculation of the video frame quality metrics as outlined in II-A for the original unencoded frames 5–8 compared with the encoded and re-displayed frame 4. This quality assessment, however, requires the actual video bit stream [12], [13], [14]. With the current video frame size and quality traces [10], the qualities for frames 1–4 can be simply taken from the qualities recorded in the traces. For determination of the qualities of frames 5–8 (for which frame 4 is re-displayed), however, no specific information is available in the current traces. Only a rough approximation, e.g., saying that the quality for frames 5–8 is very low, e.g., 20 dB, can be made.

III. Offset Distortion Traces

In order to allow for accurate trace-based assessment of the impact of lost frames on the video quality, we introduce the video frame offset distortion as outlined in the following. Let d denote the offset in frames between the last successfully received frame n and the frame to which the offset distortion should be calculated. Assuming that in general, transmission errors can be healed (e.g., by sending an I-frame) after a certain number of video frames, we denote the maximum offset for which we calculate the offset distortion as $d_{\max}$. The calculation of the RMSE for the re-display of frame n is given as function of the frame offset d similar to Equation (1) as

$$RMSE_n^q(d) = \sqrt{\frac{1}{XY} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} [F_n^q(x,y) - f_{(n+d)}^q(x,y)]^2}. \quad (6)$$

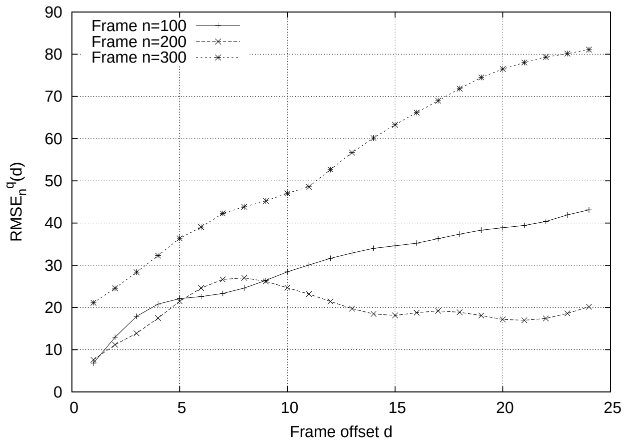

Fig. 3. Video frame offset distortion for frames $n = 100, 200$, and 300 from the *Foreman* video sequence encoded with quantization scale $q = 3$.

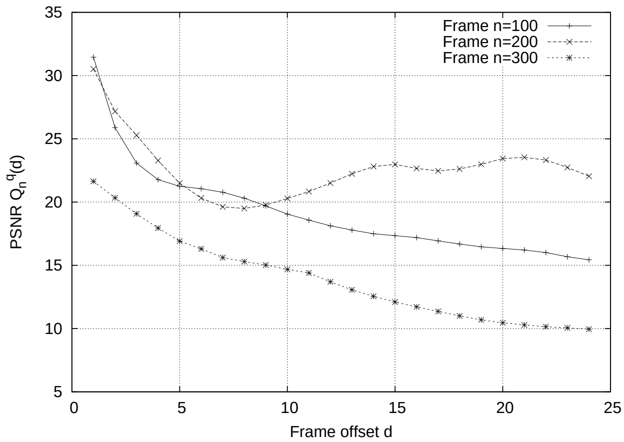

Fig. 4. Video frame offset quality for frames $n = 100, 200$, and 300 from the *Foreman* video sequence encoded with quantization scale $q = 3$.

The corresponding video frame quality can be calculated similar to Equation (3) as

$$Q_n^q(d) = 20 \log_{10} \frac{255}{RMSE_n^q(d)}. \quad (7)$$

The thus calculated values can be stored in a video frame offset distortion trace file, whereby for each video frame n, each video frame offset distortion from $d = 1$ to $d = d_{\max}$ is stored in column d of a row indexed with n. Note that $Q_n^q(0)$ corresponds to the encoded quality of frame n, which is stored in the current video traces.

We illustrate the video frame offset distortion $RMSE_n^q(d)$ in Figure 3 for frames $n = 100, 200$, and 300 from the *Foreman* video sequence. We observe that for frames at different positions in the *Foreman* video sequence, the behavior of the video frame offset distortions is quite different. We also note that the video frame offset distortion first increases for all video frames. For some video frames, the offset distortion afterwards decreases again. This is because the frames that are further away from the frame that the offset distortion is calculated for can have quite different content and thus a comparison can result in a decrease of the distortion due to the low correlation in content. (This effect can be adjusted with a perceptually adjusted PSNR metric, which we can not include here due to space constraints, and for which we refer to [17].)

We illustrate the video frame offset quality $Q_n^q(d)$ in Figure 4 for frames $n = 100, 200$, and 300 from the *Foreman* video sequence encoded with quantization scale $q = 3$. We observe the typical inverse relationship between the video distortion and quality values. Importantly, we also observe that the approximation of the video frame offset quality with a low value, e.g., $Q(d) = 20$ dB, is a very rough approximation, as the values for the offset quality vary between approx. 31 dB and 10 dB. The approximation with a low fixed value therefore captures the real behavior of the offset qualities only very roughly.

Continuing the example outlined in II-B, the video stream quality can now be determined as illustrated in Figure 5 with the combination of the currently available and the offset distortion traces. For every correctly received frame, say up to frame n, the traditional quality metrics as in Equations (1) and (3) can be taken from the currently available video traces (as for these frames $d = 0$). In our example this constitutes the frames 1–4. For frames that were not correctly received, say $n+1, \ldots, n+d_{\max}$,

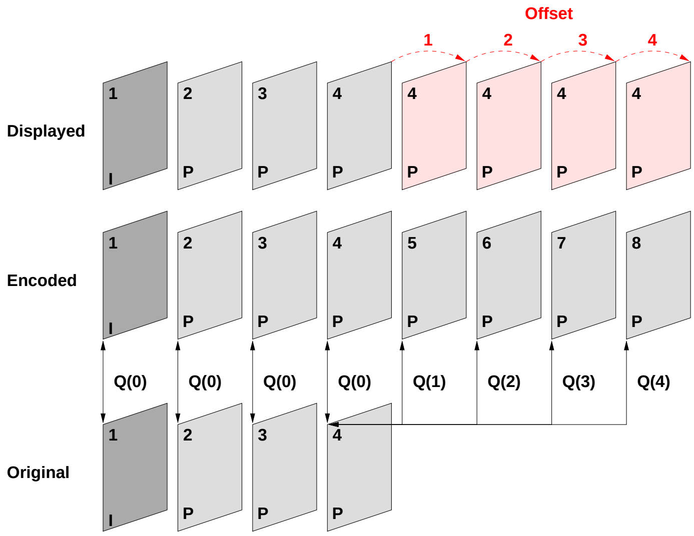

Fig. 5. Offset quality $Q(d)$ calculation for a sequence of video frames.

the video quality metrics calculated for the offset distortion can be taken from the offset distortion trace at position $(n; d)$ with $d = 1, \ldots, d_{\max}$. In our example we would thus retrieve the offset distortion values $(4; 1)$ to $(4; 4)$ for frames 5–8 from the offset distortion trace.

IV. Perceptual Adjustments for Utilization in Trace-Based Approaches

The RMSE and the RMSE-based PSNR metrics are based on the comparison of individual video frames. They therefore do not take the flow of the video frames into account. This can lead to a decrease in the offset distortion when consecutive frames have only little correlation as illustrated in Figure 3 for the RMSE and Figure 4 for the PSNR. In order to derive a more suitable metric for comparing the source video and the encoded video with errors and offset distortions, it is necessary to take the impact of re-displaying the current frame multiple times on the perceived video quality into account. Several complex metrics, e.g., the VQM [18], have been developed to study the impact of losses on the perceived video quality on a fine granularity (and generally require the actual decoded video data). These metrics are not suitable to be included in video traces due to their complexity and the unavailability of the source video data. In the following, we therefore present an approximation of perceptual considerations based on the RMSE and PSNR values.

We propose an adjustment of the RMSE and PSNR metrics that takes the number of consecutive displays of a video frame into account yet does not require more information than storable in video traces. Specifically, we use the sum of distortions that were seen by a client as basis to calculate the *perceptually adjusted RMSE* ($pRMSE$). In particular, we define

$$pRMSE_n^q(d) \quad = \quad \sum_{d=0}^{d} RMSE_n^q(d) \cdot \log_{\kappa_n}(d + \kappa_n), \tag{8}$$

where κ_n is a perceptual factor adjusting the impact of the frame offset distortion on the perceptual quality with respect to the video's content dynamics. In other words, the function $\log_{\kappa_n}(d + \kappa_n)$ is applied to determine "how much less worth is the frame n at offset d for the perceived quality?"

We determine the perceptually adjusted quality pQ for each frame and offset as

$$pQ_n^q(d) = 20 \cdot \log_{10} \frac{(d+1) \cdot 255}{pRMSE_n^q}. \tag{9}$$

We illustrate the traditional versus the perceptually adjusted video frame offset quality in Figure 6 for the frames 100 and 200 from the *Foreman* video sequence encoded at quantization scale $q = 1$. We observe that the perceptual adjustment results in a smooth decline of the video frame offset quality for all frames and offsets. Comparing the originally obtained video frame qualities Q_n^1 and the perceptually adjusted video frame qualities pQ_n^1, we observe that the perceptual quality values obtained at small offsets are generally higher than their traditionally calculated counterparts. This reflects the generally higher correlation for these frames, which in case of re-display would not result in a large degradation of perceived quality. For frames that are further away, the quality can be further reduced compared to the traditional video quality. The lower correlation of the more distant frames reduces the perceived video quality.

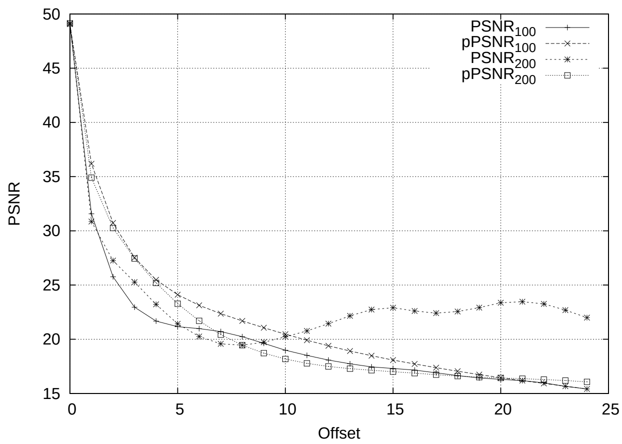

Fig. 6. Video frame offset quality (original and perceptually adjusted) for frames 100 and 200 from the *Foreman* video sequence encoded with quantization scale $q = 1$.

A. Determination of Perceptual Factor κ_n

Without loss of generality, we determine κ_n assuming that at the first lowest $PSNR$ value obtained, the correlation between successive frames is too low to give meaningful results for larger offsets. In particular, let $d_{\max}$ denote the offset for which the distortion is first locally maximal, either before the distortion degrades again or in case of no degradations for $d_{\max}$, i.e., $d_{\max} = \min\{\min[d : RMSE_n^q(d+1) < RMSE_n^q(d)], d_{\max}\}$. Using the identities of $RMSE_n^q(d=0) = RMSE_n^q$, we derive

$$Q_n^q(d_{\max}) \quad = \quad pQ_n^q(d_{\max}) \tag{10}$$

which results in

$$\log_{\kappa_n}(d_{\max} + \kappa_n) = \frac{(d_{\max}+1) \cdot RMSE_n^q(d_{\max})}{\sum_{d=0}^{d_{\max}} RMSE_n^q(d)}. \tag{11}$$

B. Further Adjustments for Trace-Based Video Quality Estimation

Typically, the video quality at the receiving client is maximized if the distortion per frame is minimized and the variation of the distortion is minimized as well [15]. As the distortion caused by each frame's encoding alone is different, we introduce the *relative quality* of an encoded and subsequently decoded video frame as

$$rQ_n^q(d=0) = \frac{Q_n^q(d=0)}{Q_n^1(d=0)}, \tag{12}$$

where Q_n^1 denotes the PSNR value for encoding frame n with quantization scale $q = 1$. We thus exploit the fact that the encoding at the smallest quantization scale factor is the best achievable quality given the compression scheme under consideration. Using the relative distortion, the previously determined approximation for the perceived quality can be further adjusted to

$$rpQ_n^q(d) = \frac{20 \cdot \log_{10} \frac{(d+1)\cdot 255}{pRMSE_n^q}}{Q_{n+d}^1(0)} \tag{13}$$

This results in a relative and perceptually adjusted PSNR value that allows for comparison between different quantization levels and between different coding modes as well.

We illustrate the relative perceptually adjusted PSNR in Figures 7, 8, and 9 for frame 100 from the *Foreman* sequence, encoded at several different quantization scale settings q.

We observe that the relativization of the offset quality shows that the quality degradation is similar for different higher offsets, resulting in a drop to approximately 30% of the original frame quality at the highest calculated offset. We also observe that the perceptual adjustment results in a larger relative value for most of the offsets for this particular frame.

V. Offset Distortion Influence on Simulation Results

In this section, we compare the actual video quality obtained with the offset distortion traces introduced in this paper with the previously outlined approximation approach in [10]. We consider sending the *Foreman* video sequence encoded with a quantization scale $q = 3$ over an error-prone link. We utilize the MPEG-4 reference software and encode the video sequence in simple profile, single layer mode. The link is modeled using uncorrelated bit-errors with different error probabilities. We

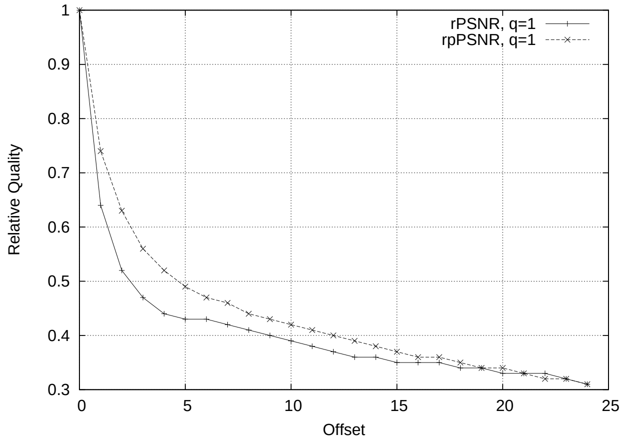

Fig. 7. Relative and relative perceptually adjusted video frame offset quality for frame 100 from the *Foreman* video sequence encoded with quantization scale $q = 1$.

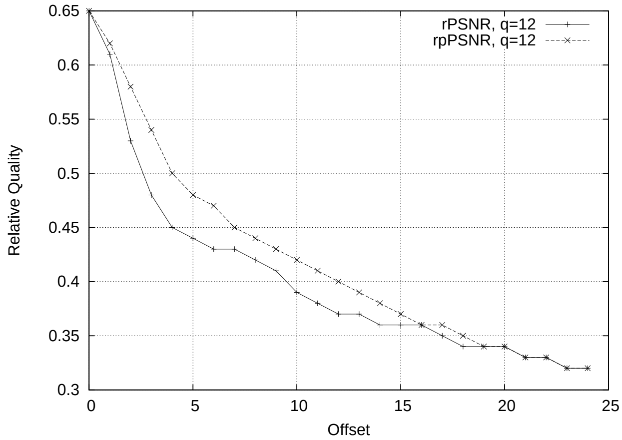

Fig. 8. Relative and relative perceptually adjusted video frame offset quality for frame 100 from the *Foreman* video sequence encoded with quantization scale $q = 12$.

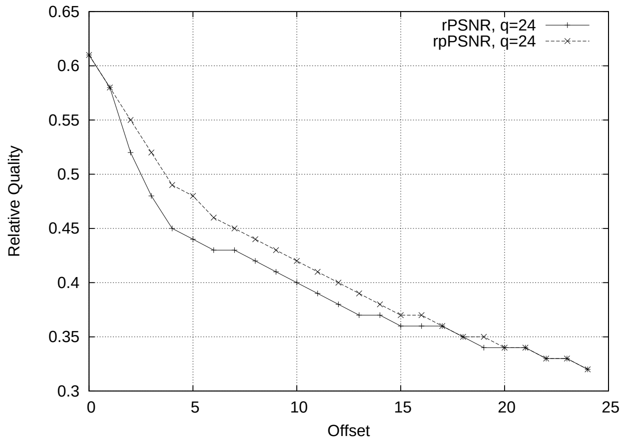

Fig. 9. Relative and relative perceptually adjusted video frame offset quality for frame 100 from the *Foreman* video sequence encoded with quantization scale $q = 24$.

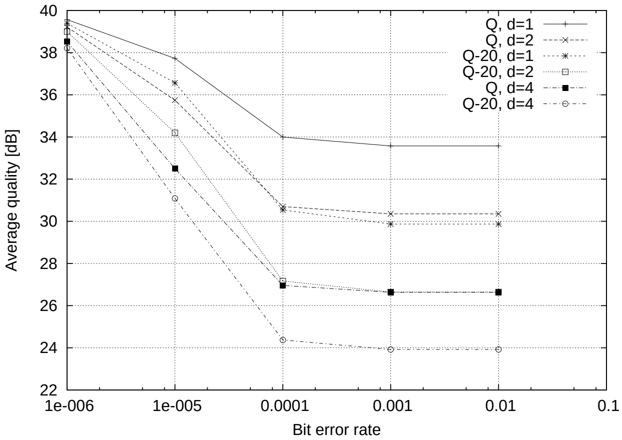

Fig. 10. Comparison of approximation of quality of loss-affected frames (`Q-20`) with actual quality obtained from offset distortion trace (`Q`): Average video stream qualities for the *Foreman* video sequence encoded with quantization scale $q = 3$ as function of bit error rate for different offsets d.

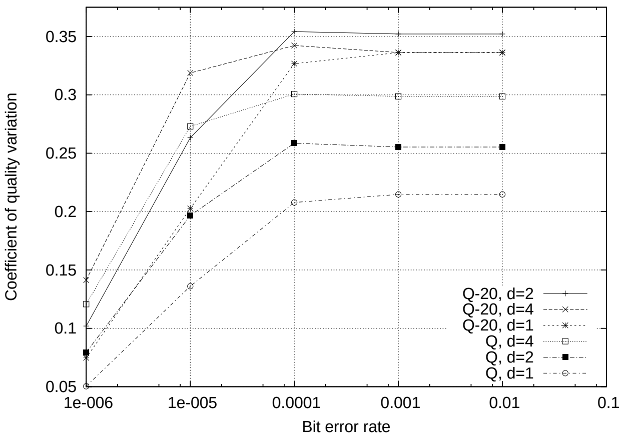

Fig. 11. Comparison of approximation of quality of loss-affected frames (`Q-20`) with actual quality obtained from offset distortion trace (`Q`): Coefficient of video frame quality variation for the *Foreman* video sequence encoded with quantization scale $q = 3$ as function of bit error rate for different offsets d.

consider the error probability for the size of the video frames (in bits) only and include no protocol overhead for comparison. We utilize the elementary *IPPP...* GoP pattern and assume that after each erroneous frame $d - 1$ more frames are lost before the sender can update the receiver by sending an I-frame. Without loss of generality, we assume that the I frame has the same distortion values as the P frame otherwise sent at that position. We report our results for a 99% confidence interval of the mean quality.

We illustrate the effect of different bit error rates on the video quality $\overline{Q}^q$ for different offsets d in Figure 10. We observe that only for very low bit error rates the approximation with $Q = 20$ dB (Q-20) results in a close fit of the value obtained by the framewise exact PSNR calculation (Q) using the offset distortion traces. As the bit-error rate increases, the difference between the approximation and the actual quality obtained with the offset distortion traces becomes larger, reaching differences between 2 dB and 4 dB, which are quite significant. We conclude that the approximation does not capture the effect of different offsets d well and in turn results in a large deviation from the correct simulation outcomes that can be obtained using the offset distortion traces.

Figure 11 shows the calculated coefficients of video quality variation CoV^q for different bit error rates. We note that the variability in the video frames' qualities increases for all different settings and metrics under consideration. We observe that in terms of capturing the variability of the quality of the video frames, the approximation approach results in an overly high estimate of the video quality variability.

We further illustrate the effect of different fixed frame intervals needed to update the receiver (offsets d) on the calculated video stream quality in Figure 12 for a fixed bit error rate of 10^{-4}. We observe that for a given bit error rate, the video stream quality as function of the offset (or frames needed to allow the receiver to receive an I frame) has an asymptotic behavior towards 20 dB for the approximation approach. The impact of increased offsets on the actual quality obtained with the offset distortion traces, however, continues to decline more rapidly as a function of the offset than the approximation. We also note

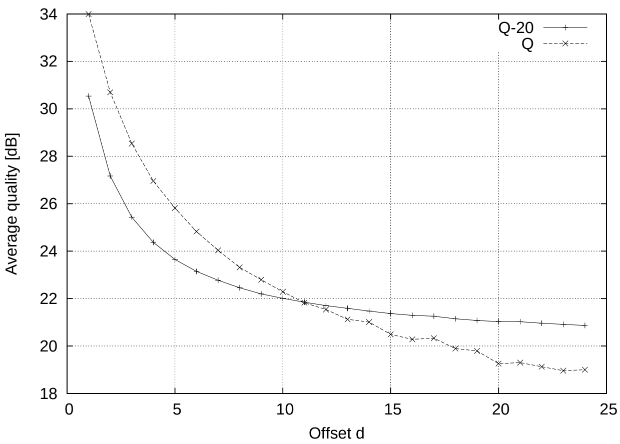

Fig. 12. Comparison of approximation of quality of loss-affected frames (`Q-20`) with actual quality obtained from offset distortion trace (`Q`): Average video stream qualities for the *Foreman* video sequence encoded with quantization scale $q = 3$ at different offsets for bit error rate 10^{-4}.

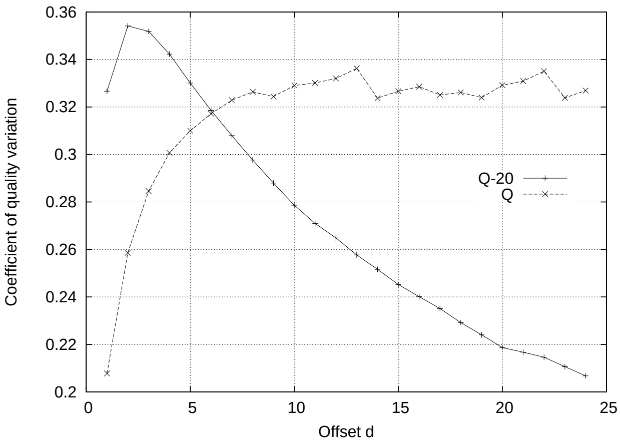

Fig. 13. Comparison of approximation of quality of loss-affected frames (`Q-20`) with actual quality obtained from offset distortion trace (`Q`): Coefficient of video frame quality variation for the *Foreman* video sequence encoded with quantization scale $q = 3$ at different offsets for bit error rate 10^{-4}.

a significant difference between the approximation and the actual quality values of close to 4 dB in case only the erroneous frame is affected ($d = 1$). The continuing decline with larger offsets and the higher quality at smaller offsets can only be obtained using the actual quality values obtained from the offset distortion traces.

Figure 13 illustrates the calculated coefficients of (relative) video quality variation CoV^q as a function of the offsets d. We observe that with larger offsets d the video quality variability increases for the actual quality. We also note that the approximative approach does not only fail to capture the video quality variability correctly, but also drops with increasing offsets. In the region of smaller offsets d, the approximation approach results in too high variation estimates. Only for an offset of $d = 6$ do the approximation approach and the offset distortion trace based approach result in approximately similar simulation results of the variability of the video frame qualities. For larger offsets, the approximation approach greatly underestimates the video frame variabilities.

VI. Estimation of the Video Frame Offset Distortion for Unknown Quantization Scales

The influence of the quantization scale parameter on the encoded video frame offset distortion is illustrated in Figure 14 for frame 100 of the *Foreman* test sequence. We observe that the video frame offset distortion as function of the quantization scale parameter q approximately resembles a linear function for each individual offset d. It is thus possible to approximate the offset distortion $RMSE_n^q(d)$ for a given d by a linear function once the offset for two different quantization scale settings q_1 and q_2 is known. (For the intricacies of scaling the video traces, we refer the reader to [10].)

VII. Conclusion

In this paper, we reviewed the shortcoming of the currently available video traces with respect to their ability to facilitate networking research with exact video stream quality evaluation. Currently available video traces only allow for very rough approximations of the distortion that is caused by re-displaying the last successfully received frame until a new video frame

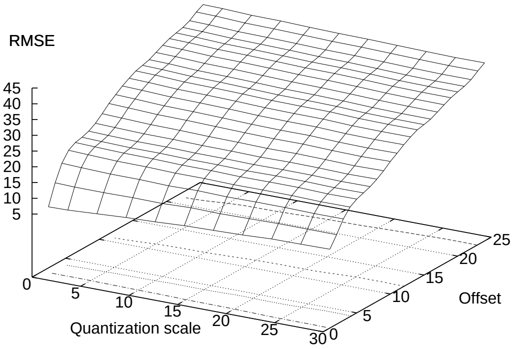

Fig. 14. Video frame offset distortion as function of the quantization scale and offset for frame 100 of the *Foreman* video sequence.

can be decoded at the receiver. We introduced offset distortion video traces that allow to accurately calculate the video frame distortion as RMSE and video frame quality as PSNR in case of re-displayed frames.

We explained how networking researchers can use the offset distortion traces to accurately assess the received video quality achieved by arbitrary video transport mechanisms without requiring equipment for or experience in video signal processing.

Our simulations indicate that the approximation using a low PSNR value for re-displayed video frames can lead to rather large approximation errors. For networking researchers this implies that more accurate results in terms of estimating the video stream quality can only be achieved if the additional information in video frame offset distortion traces is available and used in conjunction with the currently available video traces. We are in the process of incorporating these new offset distortion traces into our existing library of video traces at [19].

Acknowledgement

Supported in part by the National Science Foundation under Grant No. ANI-0136774.

References

[1] M. Dai and D. Loguinov, "Analysis and modeling of MPEG-4 and H.264 multi-layer video traffic," in *Proceedings of IEEE INFOCOM*, Miami, FL, Mar. 2005.
[2] W.-C. Feng, *Buffering Techniques for Delivery of Compressed Video in Video–on–Demand Systems*. Kluwer Academic Publisher, 1997.
[3] P. Koutsakis and M. Paterakis, "Call-admission-control and traffic-policing mechanisms for the transmission of videoconference traffic from MPEG-4 and H.263 video coders in wireless ATM networks," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 5, pp. 1525–1530, Sept. 2004.
[4] M. Krunz and S. Tripathi, "Exploiting the temporal structure of MPEG video for the reduction of bandwidth requirements," in *Proc. of IEEE Infocom*, vol. 1, no. 1, Kobe, Japan, Apr. 1997, pp. 67–74.
[5] M. Krunz, R. Sass, and H. Hughes, "Statistical characteristics and multiplexing of MPEG streams," in *Proceedings of IEEE INFOCOM*, Boston, MA, Apr. 1995, pp. 455–462.
[6] J. Liebeherr and D. Wrege, "Traffic characterization algorithms for VBR video in multimedia networks," *Multimedia Systems*, vol. 6, no. 4, pp. 271–283, July 1998.
[7] J. W. Roberts, "Internet traffic, QoS, and pricing," *Proceedings of the IEEE*, vol. 92, no. 9, pp. 1389–1399, Sept. 2004.
[8] O. Rose, "Statistical properties of MPEG video traffic and their impact on traffic modelling in ATM systems," University of Wuerzburg, Institute of Computer Science, Tech. Rep. 101, Feb. 1995.
[9] U. Sarkar, S. Ramakrishnan, and D. Sarkar, "Study of long-duration MPEG-trace segmentation methods for developing frame-size-based traffic models," *Computer Networks*, vol. 44, no. 2, pp. 177–188, Feb. 2004.
[10] P. Seeling, M. Reisslein, and B. Kulapala, "Network performance evaluation with frame size and quality traces of single-layer and two-layer video: A tutorial," *IEEE Communications Surveys and Tutorials*, vol. 6, no. 3, pp. 58–78, Third Quarter 2004.
[11] ITU-T Recommendation P.800.1, "Mean opinion score (MOS) terminology," Mar. 2003.
[12] W. Luo and M. ElZarki, "Analysis of error concealment schemes for MPEG–2 video transmission over ATM based networks," in *Proceedings of SPIE Visual Communications and Image Processing 1995*, Taiwan, May 1995, pp. 102–108.
[13] ——, "MPEG2Tool: A toolkit for the study of MPEG–2 video transmission over ATM based networks," Department of Electrical Engineering, University of Pennsylvania, Tech. Rep., 1996.
[14] J. Shin, J. W. Kim, and C.-C. J. Kuo, "Quality–of–service mapping mechanism for packet video in differentiated services network," *IEEE Transactions on Multimedia*, vol. 3, no. 2, pp. 219–231, June 2001.
[15] T. Lakshman, A. Ortega, and A. Reibman, "VBR video: Tradeoffs and potentials," *Proceedings of the IEEE*, vol. 86, no. 5, pp. 952–973, May 1998.
[16] S. Winkler, "Vision models and quality metrics for image processing applications," Ph.D. dissertation, EPFL, Switzerland, 2000.
[17] P. Seeling and M. Reisslein, "Offset distortion traces for trace-based evaluation of video quality after network transport (extended version)," Apr. 2005, availabe from `http://trace.eas.asu.edu`.
[18] M. Pinson and S. Wolf, "Low bandwidth reduced reference video quality monitoring system," in *First International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, Scottsdale, AZ, Jan. 2005.
[19] "Video traces for network performance evaluation," Traces available at: http://trace.eas.asu.edu.