RoSE Technical Report - I

—

Performance Evaluation of MDC Phase I

I. INTRODUCTION

This report presents the first phase of performance evaluation of MDC schemes in the Robust Streaming Environment (RoSE) project [1]. The spatial 2MD scheme is studied. The performance is evaluated using PSNR metric under perfect network conditions, i.e. transmission impairments are not considered. This baseline scenario provides the fundamental overhead of integrating MDC into a multimedia streaming system. Other issues such as increased complexity are also crucial but not considered in this phase.

II. SPATIAL 2MD SCHEME

A typical and low-cost way to produce multiple descriptions is to partition the source data into several sets and then compress independently to produce descriptions. The separation can be into odd- and even-numbered samples [2]. In the spatial dimension, this corresponds to spatial sampling of frames into N subsets. For $N = 2$, two balanced descriptions can be generated by separating odd/even lines [3]. This technique is denoted as *spatial 2MD* and illustrated in Figure 1 [4].

In spatial 2MD, each frame in the input video is separated into odd and even subframes at the Remux module of RoSE. The odd and the even subframes contain the odd and even lines, respectively. Therefore, the height of the frames are halved but the width does not change a shown in Figure 1. These two descriptions are encoded with half the bitrate of the original stream to keep the total bitrate constant. Then these descriptions are merged at the Postmux to reconstruct the received video. All these MDC-related operations are performed in the data plane. This property achieves compatibility with any incumbent codecs.

III. PERFORMANCE EVALUATION AND METHODOLOGY

The MDC-enabled system is evaluated using experiments with various encoding bitrates. The relevant system parameters are listed in Table I. In the experiments, *Foreman* test sequence is used as the input video. It has 300 frames and CIF size. The original raw video is encoded into MPEG4 with varying bitrates between 96 kbps and 2048 kbps. The frame type structure is IPPPPPPPPPPP (no B frames). Then the raw video (scenario 1) or this encoded

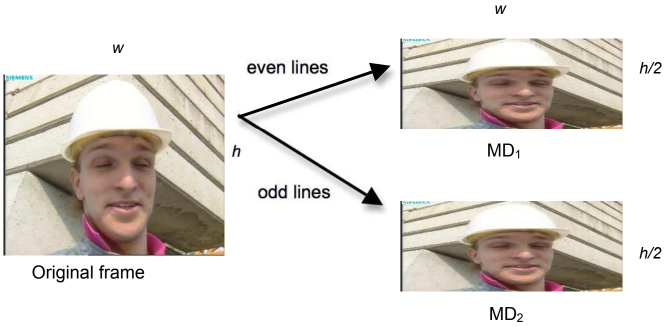


Fig. 1. Spatial 2MD coding scheme.

sequence (scenario 2) is fed into the Remux module and processed to generate descriptions in MPEG4 format. MPEG4 is preferred since it is one of the state-of-the-art coding formats. However, any codec can be employed as long as both Remux and Postmux modules support it. This multiple description coded video is encapsulated in a .*nut* file to be streamed. Subsequently, Postmux processes these descriptions to reconstruct the original video at the receiver side. This experimental setup is depicted in Figure 2.

Scenario 1 (the raw video input case) corresponds to streaming or content delivery systems where the multimedia can be processed beforehand in an offline fashion to take care of multiple description transmission. Moreover, it constitutes the baseline scenario for MDC based streaming. In scenario 2, the Remux module acts as a multiple description transcoding engine where the input is an already coded and compressed video. This may occur when MD is integrated integrated into a network as an plug-in edge device for interfacing heterogeneous systems such as wireless access networks. Obviously, this setup also entails some drawbacks, the most notable being that the MDC performance will be “constrained” by the quality degradation caused by the prior encoding. Additionally, real-time transcoding comes with processing and delay overhead. However, cache-based delivery systems such as personalized TV or radio can benefit from this configuration since the delay burden is not a major issue. The overhead is also acceptable when various quality levels (distortions) are acceptable and distinguishable and the reconstructions produced at side decoders should be more valuable than nothing [2].

The Peak-Signal-to-Noise-Ratio (PSNR) metric is used to assess the penalty of the method and the quality of the MDC video. PSNR is one of the simplest and most widely used quality metrics. PSNR is calculated with the mean squared error (MSE), computed by averaging the squared intensity differences of distorted and reference frame pixels, along with the related quantity of PSNR. These are appealing because they are simple to calculate, have clear physical meanings, and are mathematically convenient in the context of optimization. The simplest implementation of this concept is the MSE, which objectively quantifies the strength of the error signal. But two distorted images with the same MSE may have very different types of errors, some of which are much more visible than others.

The frame (image) quality is measured as the MSE value which is defined as

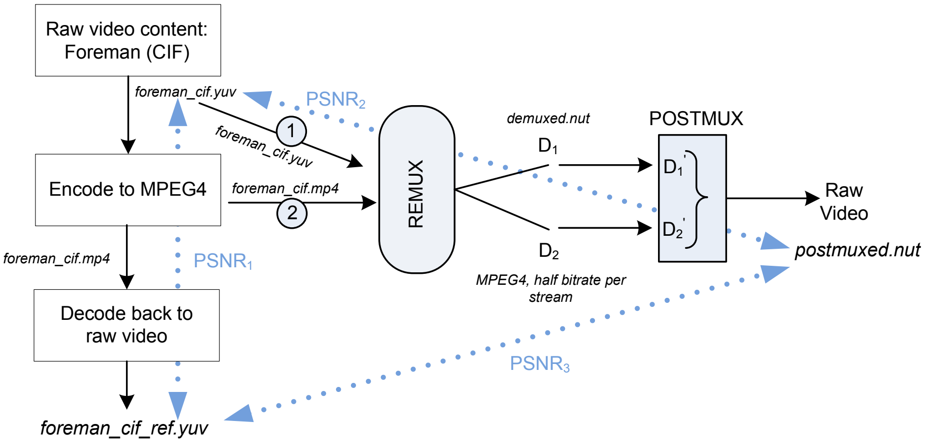


Fig. 2. Experimental setup and performance evaluation.

TABLE I

SYSTEM AND SIMULATION PARAMETERS

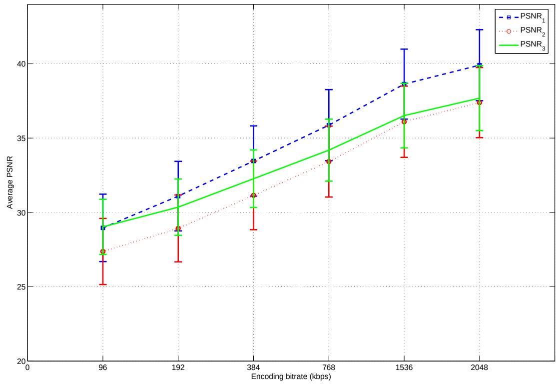
Parameter	**Value**
Channel Model	No impairments
Input raw video	*Foreman* YUV 420
Input codec	1: RAW, 2:MPEG4
Bitrate	96, 192, 384, 768, 1536, 2048 kbps
No. frames	300
Frame size	CIF
MPEG4 frame order	IPPPPPPPPPPP
MDC codec	MPEG4
MDC scheme	Spatial 2MD

$$MSE = \frac{1}{N^2}\sum_{i}^{N}\sum_{i}^{N}\left[X(i,j) - \widehat{X}(i,j)\right] \tag{1}$$

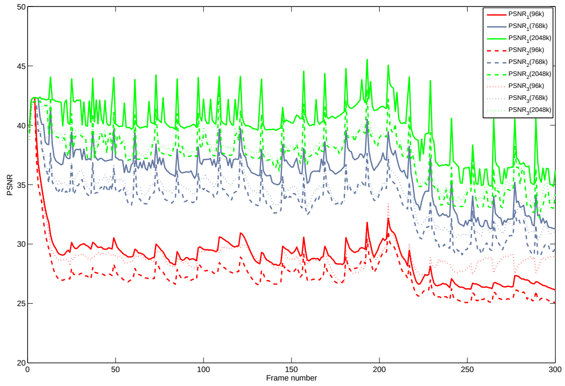
where, N^2 is number of pixels in image, $X(i,j)$ and $\widehat{X}(i,j)$ are the pixel values of the reference frame and of the final frame reconstructed from the received multiple descriptions, respectively. We use the following PSNR metric which is

$$PSNR(dB) = 20\log_{10}\frac{2^n - 1}{RMSE} \tag{2}$$

where p is the largest possible value of the signal (n = 8, i.e. $2^n - 1 = 255$ for grayscale images), and RMSE is the Root Mean Square Error between the two images given above, respectively.



(a) Average PSNR values for varying bitrates using raw video input and MPEG4-coded output at the Remux module (scenario 1).



(b) PSNR performance for the entire sequence with encoding bitrates 96 kbps (red), 768 kbps (blue), and 2048 kbps (green). Solid lines are for PSNR$_1$, whereas dashed and dotted lines represent PSNR$_2$ and PSNR$_3$, respectively.

Fig. 3. PSNR values using MPEG4 as the Remux input. The raw video is *Foreman* sequence with 300 frames and CIF size.

IV. EXPERIMENTAL RESULTS

Three cases are considered for PSNR evaluation. These measurements are marked as PSNR$_i$ on Figure 2. For the time being, scenario 2 is not considered. In the first metric, PSNR$_1$, the distortion due to encoding in MPEG4 is measured. This case is also equivalent to a single sender and receiver transmission without MDC by sending all the video packets using a single route under perfect channel conditions. This is plotted as the dashed line in Figure 3(a). In the second case, we measure the distortion cumulatively due to both MPEG4 and the subsequent MDC stage (MPEG4 decoding/encoding and MDC operations). This corresponds to PSNR$_2$ and shown as the solid line in Figure 3(a). And in the final case, PSNR$_3$ measures the distortion due to MDC. This is the most meaningful metric for focusing solely on the effect of MDC.

In Figure 3(a), the average PSNR values for varying bitrates using raw video input and MPEG4-coded output at the Remux module are shown. Error bars in the figure show the standard deviation along the average PSNR curves. These standard deviation values range between 1.8 and 2.4 dB. The PSNR_1 value is about 28.96 dBs for 96 kbps compared to 39.90 dBs for 2048 kbps. The PSNR_2 has smaller values but follows a similar trend: about 27.37 dBs for 96 kbps compared to 37.39 dBs for 2048 kbps. The PSNR_3 value increases more sluggishly and is about 29.03 dBs for 96 kbps compared to 37.69 dBs for 2048 kbps. All PSNRs monotonically increase for increasing bitrates. The PSNR degrades substantially for low bitrates in all cases. As the bitrate increases, the difference between encoded-decoded MPEG4 output and original video fades since, for the high bitrates, single compression-decompression affects the quality of the video marginally. Therefore the gap between PSNR_3 and PSNR_2 closes and PSNR_3 converges to PSNR_2 (i.e., the MDC-free encoded-decoded sequence is almost identical to the original raw sequence in that case). These objective metrics match the actual quality of the videos because an obvious difference between qualities can be detected when the sequences are visually evaluated: the perceptual quality improves for increasing bitrates as seen in Figure 4 and 5.

We also investigate the difference between PSNR_1 and PSNR_2, denoted as δ and defined as

$$\delta = PSNR_1 - PNSR_2 \tag{3}$$

It represents the pure PSNR overhead or penalty for integrating MDC in the transmission chain. Because the experienced PSNR would be PSNR_1 without MDC whereas it is degraded to PSNR_2 with MDC. δ values start from 1.59 for 96 kbps and increases up to 2.53 and 2.51 for the last two bitrates. This is expected since both metrics suffer from extreme degradation for very low bitrate. The average value (μ_δ) is 2.26 and the standard deviation (σ_δ) is 0.36. Thus, it shows a consistent behaviour with a relatively small σ value.

The same trend is also apparent in Figure 3(b). In this figure, PSNR performance for the entire sequence with encoding bitrates 96 kbps (blue), 384 kbps (red), and 1536 kbps (green) are shown for representing the general behaviour. Other bitrates are omitted for the sake of brevity. Solid lines are for PSNR_1, whereas dashed and dotted lines represent PSNR_2 and PSNR_3, respectively. All PSNR values behave as expected and degrade for decreasing encoding bitrates. We also observe the diminishing returns for increasing the bitrate to very high bitrates. The PSNR gain becomes much less sensitive to increasing bitrate for high bitrates (the gap closes between consecutive bitrates.) As the limit case, when the compression is omitted (the bitrate passes the raw video bitrate), the gain from increasing the bitrate will be zero. Also, a periodicity is observed due to independent-coded I-frames in the sequence.

For visual evaluation, we provide sample frame captures from original, remuxed and postmuxed video sequences in Figure 4 and 5. These captures exhibit the positive correlation between the quality and encoding bitrate in the system. The performance loss due to MDC stage (remux+postmux) is hard to perceive with visual inspection. The significant quality degradation in all of the different cases for low bitrate encoding is to be noted.

(a) Raw video (b) 1536 kbps

(c) 384 kbps (d) 96 kbps

Fig. 4. Frame captures for the original and postmuxed video.

(a) 768 kbps (b) 192 kbps (c) 48 kbps

Fig. 5. Frame captures for the remuxed video. Please refer to Figure 4(a) to inspect the orginal frame.

V. CONCLUSION

In this report, we have presented the first phase evaluation of MD schemes in terms of PSNR performance in the baseline scenario. This scenario entails the bare PSNR penalty due to decoding into a new codec, separating into multiple descriptions at Remux module and then merging them in the Postmux module. In other words, the cost of MDC employment in video transmission is studied. The transmission impairments are not considered in this phase. As anticipated, the introduction of MDC incurs a PSNR penalty on the final performance. However, this penalty is relatively minor compared to the potential gains. As future work, more MDC schemes will be implemented. Additionally, the performance will be evaluated for additional sequences against transmission simulations and real network environments.

REFERENCES

[1] G. Gur and L. Abeni, “Bringing Multiple Description Coding and Scalable Video to Consumer Multimedia Communications,” submitted to *IEEE Communications Magazine*, 2009.

[2] V. K. Goyal, “Multiple description coding: Compression meets the network,” *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 74–93, September 2001.

[3] Y. Wang, A. Reibman, and S. Lin, “Multiple description coding for video delivery,” *Proceedings of the IEEE*, vol. 93, no. 1, pp. 57–70, Jan. 2005.

[4] A. Vitali and M. Fumagalli. (2005) Standard-compatible multiple-description coding (MDC) and layered coding (LC) of audio/video streams. Internet draft. [Online]. Available: http://tools.ietf.org/tools/rfcmarkup/rfcmarkup.cgi?draft=draft-vitali-ietf-avt-mdc-lc-00.txt