Parameter Selection for a Video Communication System based on HEVC and Channel Coding

Y. Samayoa, J. Ostermann
Institut für Informationsverarbeitung
Gottfried Wilhelm Leibniz Universität Hannover
30167 Hannover, Germany
Email: {samayoa, office}@tnt.uni-hannover.de

***Abstract*—A low delay video transmission over error prone channels with limited bandwidth requires both video and channel coding. A reduction of the distortions added to the video sequence by the video coding and through the channel can be achieved by selecting appropriate configuration parameters for both codecs. In this paper we address this problem from a bit-allocation perspective under the joint source-channel coding approach. We propose an evaluation methodology with a new metric that allows the measurement of the overall distortion of a video communication system. Experimental results, for a system conformed by High Efficiency Video Coding (HEVC) and a channel coding with variable coding rate, demonstrate the procedure to evaluate the video communication system and to find a suitable parameter set to reduce the overall distortion of the video at the receiver. For the evaluated system, it was also found a bit allocation strategy between the video and channel codec that reduces the overall distortion. It gives less protection against errors from the channel while reduces the distortion introduced by the video encoder.**

I. Introduction

Low delay video communication systems usually involve separate source coding and channel coding components. This division comes from Shannon's separation principle [1], which allows the design of the source and channel coding separately without loss in performance. However, this separation principle relies on some assumptions, i.e., an arbitrary long or infinity length of blocks for both source and channel coding, infinity computational resources and an exact knowledge of the statistics of the channel. For systems with delay constraints, and in general for real systems, on the one hand, these assumptions may not hold. On the other hand, academic and industry research has been inspired on this separation principle to develop state-of-the-art systems. When keeping the separation principle, the remaining tasks are to select a suitable application-dependent video and channel codec, and search their parameters jointly to minimize the distortion of the transmitted video.

In the literature, several solutions for this problem are found under the umbrella of joint source-channel coding (JSCC) [2]. Most of them propose a modification of existent video and channel codecs, which may not be desirable in practical systems. In [3], the authors take another approach. They developed a theoretical model for a system that combines H.263 standard and Reed Solomon (RS) codes. Moreover, a review of the H.264/AVC standard capability for wireless environments is given in [4]. Nonetheless, these last results are not applicable for newer video coding standards such as High Efficiency Video Coding (HEVC) that achieves a higher coding efficiency in comparison with its preceding standards. However, a new standard brings also new challenges for video transmission over error prone channels. For example, the data partitioning and the flexible macro block ordering (FMO), which are options in H.264/AVC to enhance its error resilience capability, have been removed in HEVC. Little research regarding HEVC under error prone channels can be found in the literature. For example, in [5] and [6] the coding performance of HEVC under packet transmission errors is investigated. The former compares the performance between HEVC and H.264/AVC. The latter searches for an optimal configuration of the HEVC encoder under energy constraints. Nevertheless, both ignore the need of channel coding which is an essential component for any communication system without an ideal channel. This indicates that further research is necessary for video transmission systems with HEVC codecs.

978-1-7281-8903-1/20/$31.00 ©2020 European Union

In this paper, we focus on a very low delay, point-to-point video communication system over an error prone channel. The main goal is to find the set of parameters that minimize its overall distortion while making use of off-the-shelf systems for video and channel coding without any modification. For this purpose, we develop an evaluation system capable of estimating the end-to-end distortion, which otherwise would be difficult to accomplish with common methods. This evaluation system synchronizes the recovered video at the receiver with the original video at the transmitter in case some frames have been lost, thus, finding the correspondence between transmitted and available frames at the receiver. Moreover, a new metric that allows the evaluation of the received video is introduced. With it, a suitable measurement of the end-to-end distortion can be performed even if there are errors after the video decoder. We investigate the performance of a system conformed with HEVC for the video coding and with forward error correction (FEC) by means of serial concatenated codes that pair a Reed Solomon (RS) with convolutional codes for the channel coding.

The remainder of this paper is organized as follows. In Section II, we present the video communication system for our model, the parameters selection and the bit-allocation dilemma between the video and channel codec. In Section III we introduce the evaluation methodology and the performance comparisons of different configuration parameters are presented as well. Section IV provides a conclusion for this paper.

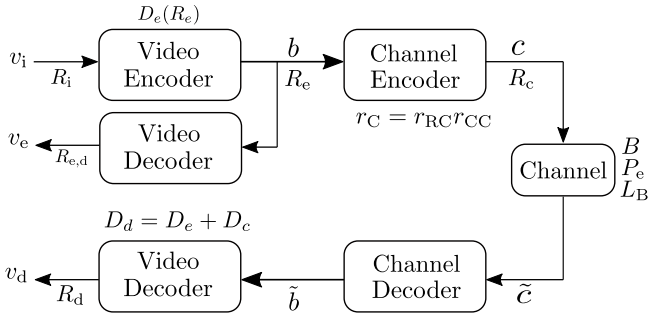

Figure 1: Video communication system block

II. System Overview

The video communication system that is considered in this paper is shown in Figure 1. The camera captures and generates a sequence of raw digital frames or pictures v_i at some bit-rate R_i in bits per second (bps), e.g., a sequence of image in YUV color space. This discrete-time and discrete-space signal is compressed by the video encoder, thus, transforming it in a bit-stream b with a bit-rate $R_e \leq R_i$ according to the rate-distortion $D_e(R_e)$ function that characterizes HEVC. The channel encoder block maps b to c with rate $R_c \geq R_e$. We assume that the digital bandwidth B in bps and the statistical characteristics of the channel are known parameters. Moreover, c is transmitted over the channel with a rate $R_c = R_e r_C^{-1} \leq B$. At the receiver, the processes are inverted by the channel decoder and video decoder blocks in order to recover the sequence of raw digital frames and if all works well $v_\mathrm{d} = v_\mathrm{e}$.

There are two sources of uncorrelated distortions D_e and D_c that are added to the video signal throughout its transmission such that the overall distortion $D_\mathrm{d} = D_\mathrm{e} + D_\mathrm{c}$. The first one is introduced systematically by the quantization process in the video encoder and can be determined after the decoder block at the transmitter: $D_\mathrm{e} = f(v_\mathrm{i}, v_\mathrm{e})$. The second source of distortion is introduced randomly by the channel; it can be, to some extent, reduced by the error concealment (EC) block at the video decoder and by the channel decoder. It can be measured indirectly at the output of the video decoder, i.e., $D_\mathrm{c} = D_\mathrm{d} - D_\mathrm{e}$, where $D_\mathrm{d} = f(v_\mathrm{i}, v_\mathrm{d})$.

The goal of the video communication system, as described in Figure 1, is to minimize the overall distortion D_d for a given B. If most of the bit-budget is spent for R_e, there will be less bit-budget available for R_c and the channel coding will have less capability to correct the errors introduced by the channel, which in turn increases D_d. If most of the bit-budget is spent for R_c, less bits-budget is spent for R_e, which in turn increases D_d as well. Consequently, the main goal may be reached by establishing a suitable strategy for the bit allocation between the two given codecs.

A. HEVC Video Encoder/Decoder

In this section, we briefly describe the parameters for the video coding that brings the most impact in the overall performance of the video communication system. The control of the rest of parameters are left to the *coding control* block. An overview of the HEVC standard and its parameters can be found in [7]. Among the vast number of parameters required to be set for a video compression, we consider the following:

- *Quantization Parameters (QP)*: it controls the quantizer step size. Higher QPs corresponds to higher quantizer steps. This quantizer is a non-linear lossy procedure and the only source of distortion at the encoder, due to the fact that, it removes signal information of the video sequence.
- *Structure of Pictures (SOP):* it is also refereed as group of pictures (GOP) in prior standards which has been also adopted informally in the context of HEVC. We will make use of the term GOP in this paper. As its name indicates, this parameter controls the number of consecutive pictures that comprises a coding video sequence (CVS), which can be encoded/decoded independently of other CVS. This may be seen as the temporal division of the video sequences that HEVC makes.
- *Slices per frame*: this parameter indicate a spacial division of each picture. A slice is a group of Coding Tree Unit (CTU) and every CTU is included in exactly one slice, therefore, the entire picture can have one or more slices. Slices gives more flexibility for recovery and synchronization in case of data loss. This means that each slice can be encoded/decoded independently from other slices in the same picture. Therefore, if an error appears in any slice, due to the temporal prediction process of the codec, the error could propagate in the same slice position in subsequent frames until the next intra-predicted slice. Nevertheless, the error will not spread to other slice positions.
- *The spatial resolution:* is the total number of pixels per frame. By reducing the video resolution (or just by low-pass filtering it), a new source of distortion is introduced. We found in our experiments that for a given output rate R_e, the total distortion after the video encoding is more likely to be greater when the resolution is reduced than with the original resolution. For this reason, we do not include spatial resolution in our set of parameters to investigate.

HEVC is entropy encoded by means of the Context-Based Adaptive Binary Arithmetic Coding (CABAC) method. This means that any error that goes unnoticed through the channel decoder may have devastating effects in v_d. Moreover, spatial and temporal prediction processes in the encoder create high dependencies in the compressed video, therefore, uncorrected errors may also lead to error propagation in both domains that in the worst case will remain until the next intra-coded frame. The error concealment (EC) component is responsible for minimizing the impact of errors in $\tilde{b}$. It may be seen as a filter of errors prior v_out.

B. Channel

To simulate the channel, we use a Gilbert-Elliott channel model with a 2-state Markov model [8] [9]. This model can be used to simulate a more realistic errors in a bit level. It is simple and it requires only two parameters to be described. Commonly, the states are denoted good (G) and bad (B). The

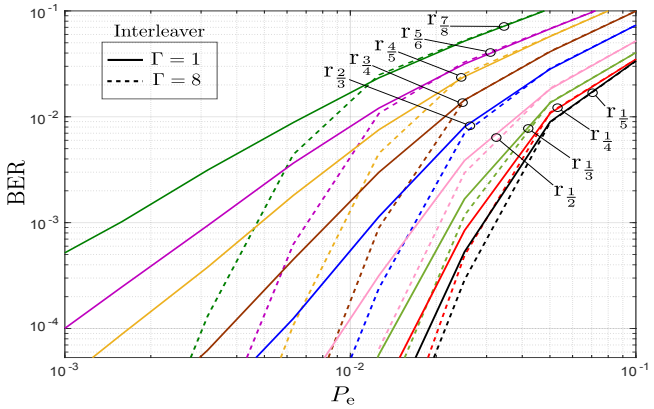

Figure 2: BER of the channel coding for $\mathrm{r}_{k_{\mathrm{CC}}/n_{\mathrm{CC}}} = r_{\mathrm{RC}} r_{\mathrm{CC}}$ and interleaver deep $\Gamma = \{1, 8\}$ over a Gilbert-Elliott channel model with P_{e} and $L_{\mathrm{B}} = 8$.

former implies a reception of an error-free bit and the letter a bit with error. We define a burst error as a consecutive number of received errors. The model is described by the transition probabilities p_{GB} and p_{BG} denoting the probabilities from the G to the B state and the probability from the B to the G state respectively. These probabilities are not intuitive, therefore, a useful correspondence between channels realizations and its statistics is to compute the equivalent error probability and the average burst error length as

$$P_{\mathrm{e}} = \frac{p_{\mathrm{GB}}}{p_{\mathrm{GB}} + p_{\mathrm{BG}}} \tag{1}$$

and

$$L_{\mathrm{B}} = \frac{1}{p_{\mathrm{GB}}}\ , \tag{2}$$

respectively.

C. Channel Encoder/Decoder

Due to delay constraints, a forward error correction (FEC) solution is selected for the channel coding block. In the context of Figure 1 a FEC could be also considered as a method to increase the error resilience capability of the video codec. It corrects errors introduced by the channel, therefore, D_c is decreased. In this paper we use a very popular serial concatenated code, i.e., a code pair consisting of a Reed Solomon (RS) code, as the outer code, with a convolutional code (CC), as the inner code. A block interleaver with deep Γ connects both codecs. A hard decision Viterby algorithm is used as the decoder of the convolutional code.

We use a (204,188) RS shortened code derived from the original (255, 239) RS code. With $m = 8$ bits per information symbol, the input of the RS consist of a block of $k_{\mathrm{RS}} = 188$ information symbols being encoded (or mk_{RS} bits), it appends $n - k = 2t = 16$ parity bits to it, resulting in a code word of $n_{\mathrm{RS}} = 204$ encoded symbols at the output. Hence, the code rate is $r_{\mathrm{RS}} = 188/204$. Its maximum symbol-error correcting capability is $t = \lfloor (n-k)/2 \rfloor = 8$, meaning that this specific RS can correct up to 8 symbols per code word.

A set of nine convolutional codes is available, each one with a constraint length $K = 7$. Four codes have a code rate $r_{\mathrm{CC}} = 1/n_{\mathrm{CC}}$ and five codes are rate-compatible punctured convolutional codes (RCPC) [10] with $r_{\mathrm{CC}} = k_{\mathrm{CC}}/n_{\mathrm{CC}}$ and a mother code rate $r_{\mathrm{m,CC}}$, where every k_{CC} information bit produceds n_{CC} code bits at the output. Similarly to RS codes, convolutional codes can correct up to $t = \lfloor (d_{\mathrm{free}} - 1)/2 \rfloor$, where d_{free} is defined as the smallest Hamming distance between all possible code sequences of the code. This error capability assumes that the minimum separation among error is at least K. Table I summarizes the parameters of each convolutional codes.

The serial concatenation of RS with convolutional codes reaches an equivalent code rate $r_{\mathrm{c}} = r_{\mathrm{RS}} r_{\mathrm{CC}}$. Moreover, convolutional codes are very efficient, even more than RS for single errors, but due to its own memory they may tend to produce burst errors at the decoder. RS codes are very suitable to correct burst errors effectively. Hence, in this constellation, convolutional codes acts as a kind of filter for short error pattern. Figure 2 plots the performance of the channel codec in terms of bit error rate (BER), with and without an interleaver. Clearly, the use of interleaver divides the average burst length of the channel, therefore, it enhances the channel coding capability at expenses of an increment in delay.

III. Evaluation Methodology and Results

As mentioned earlier, the main goal of the video communication system depicted in Figure 1 is to make possible the transmission of a video while minimizing D_{d} or, in other words, maximizing the video quality of v_{out}. This implies that both distortion D_e and D_d must be measured. The effect of this two distortion in the output video are different. The first one is introduced by the quantizer, which, for example, affects evenly the encoded frames for video sequences with similar frequency content in every frame. The second one is introduced randomly by the channel and if it is not concealed correctly at the decoder, it may change very strongly the decoded video in time and space domain. A significant part of some frames may be lost, or even it can provoke some frame drops.

We introduce in this section the system that allows the measurement of D_e and D_{d}, often converted to the quality measure peak-signal-to-noise ratio (PSNR) in logarithmic domain.

1) Quality measure at the encoder: An easy method to measure the video distortion introduced by the encoder is

Table I: Convolutional codes configuration parameters

r_{CC}	$r_{\mathrm{m,CC}}$	Generator	Puncturing	d_{free}
1/5	*	[131 135 135 147 175]	*	25
1/4	*	[133 135 147 163]	*	20
1/3	*	[133 145 175]	*	15
1/2	*	[171 133]	*	10
2/3	1/2	[171 133]	[1 0; 1 1]	10
3/4	1/2	[171 133]	[1 0 1; 1 1 0]	10
4/5	1/2	[171 133]	[1 0 0 0; 1 1 1 1]	10
5/6	1/2	[171 133]	[1 0 1 0 1; 1 1 0 1 0]	10
7/8	1/2	[171 133]	[1 0 0 0 1 0 1; 1 1 1 1 0 1 0]	10

Figure 3: Metric comparison between conventional PSNR and $\text{PSNR}_{\text{d},\alpha}$ for a frame with equal amount of error bur different presentation. Left frame: PSNR = 17.69 dB and $\text{PSNR}_{\text{d},\alpha}$ = 34.86 dB. Rigth frame: PSNR = 13.90 dB and $\text{PSNR}_{\text{d},\alpha}$ = 34.86 dB

the mean-squared-error (MSE). Therefore, $D_{\text{e}} = \text{MSE}(v_{\text{i}}, v_{\text{e}})$. We are aware that normally the PSNR is computed after measuring the MSE of the entire video sequence, nevertheless, for convenience as a metric to evaluate the quality of v_{d} afterwards, we prefer an alternative definition for the PSNR at the encoder: $\text{PSNR}_{\text{e}} = \text{mean}(\text{PSNR}(n))$, where mean$(\cdot)$ computes the average, $n = 1, 2, ..., N$ is the frame number and N denotes the maximum number of frames.

2) Quality measure at the decoder: PSNR is a per-pixel metric. It requires $R_{\text{d}} = R_{\text{i}}$, i.e., a perfect pixel alignment among the compared videos which at the decoder may not be fulfilled due to possible loss of frames. Therefore, a corresponding pixel alignment is compulsory prior the estimation of PSNR_{d} at the decoder.

We solve the pixel alignment problem by means of the well known algorithm dynamic time warping (DTW) [11]. First the number of frames in v_{d} and v_{i} are compared. If they are the same, the quality of v_{d} is measured. However, if at the receiver v_{d} is missing N_{loss} frames due to their lost as a result of the channel's errors, a correct synchronization for the $N - N_{\text{loss}}$ remaining pictures can be found by means of DTW. The synchronization means to find for each one of the $N - N_{\text{loss}}$ frames its correspondent frame in v_{i}. Afterwards, the gap of the lost frames in v_{d} is filled with a copy of the last decoded frame until that frame number. This replicates thus a frozen video. Hence, the quality of all N frames can be measured at the receiver.

Assuming a correct frame alignment, we still have to deal with a proper metric to measure the quality of v_{d}. For example, Figure 3 shows a frame of a transmitted video with two different presentations of the same unconcealed error. The frame lost an slice during the transmission leaving an empty gap. The first presentation fills the gap with a patch of color green while the second fills it with pink. The conventional PSNR of the frame with green and pink patch result in 17.69 dB and 13.9 dB respectively. We get two different estimations with this metric although both frames have the same amount distortion, i.e., the same slice is missing. In this example, a suitable metric should give same results. Hence, the normal PSNR may give a misleading interpretation as an end-to-end distortion metric.

We introduce in this paper a simple but effective procedure to measure the quality between v_{i} and v_{d}. The quality of v_{d} depends also indirectly on the performance of the channel coding which is commonly measured in terms of an error rate. In analogy, we first define the distortion at a pixel level:

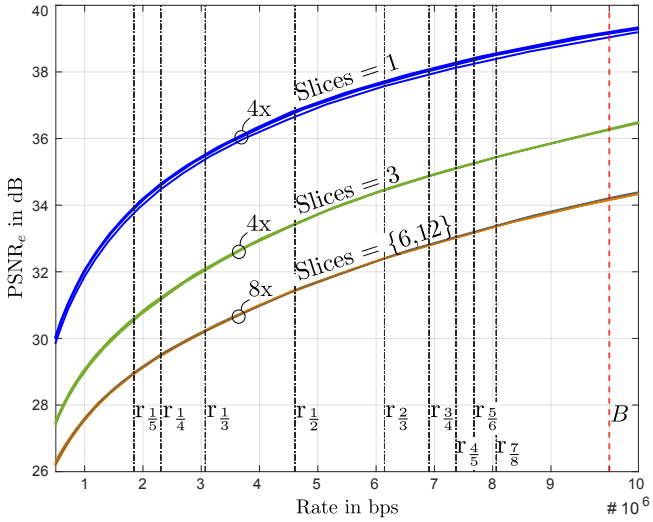

Figure 4: PSNR_e for different slices = {1,2,6,9}, QP = [22,50] and GOP = {30,60,90,120}. Coding rate denoted by $\text{r}_{k_{\text{CC}}/n_{\text{CC}}} = r_{\text{RC}} r_{\text{CC}}$ and channel bandwidth $B = 10$ Mbps

$$\gamma_i = \begin{cases} 1 & \text{if } |p_{\text{i},i} - p_{\text{d},i}| > \beta \\ 0 & \text{else} \end{cases} , \tag{3}$$

where $p_{\text{i},i}$ and $p_{\text{d},i}$ are the i-th pixel of the input and output video respectively. The parameter $\beta \geq 0$ is an arbitrary threshold. Moreover, from (3) the total pixel error rate (ε) can be computed as an average: $\varepsilon = \text{mean}(\gamma)$ with $\gamma = [\gamma_i]_{i=1}^{M}$, where M is the total number of pixels of the video. Finally, the quality metric for v_{d} is defined as

$$\text{PSNR}_{\text{d},\alpha} = \text{PSNR}_{\text{e}} - \alpha \text{PSNR}_{\text{e}} , \tag{4}$$

where $\alpha = (e^{\upsilon\varepsilon} - 1)/(e^{\upsilon} - 1)$ controls the subjective interpretation that penalizes ε by means of $\upsilon \geq 0$. If υ tends to zero, α tends to ε, otherwise it is lower until $\upsilon = 1$, then $\alpha = 1$. These nonlinear relations between α and $\{\varepsilon,\beta\}$ serves to adjust the metric subjectively if necessary. A per-frame PSNR can be calculated with its corresponding per-frame α following the same rationale of (4). In Figure 3, a comparison of the $\text{PSNR}_{\text{d},\alpha}$ with the conventional PSNR is given. As shown in this example, $\text{PSNR}_{\text{d},\alpha}$ in (4) gives a consistent estimation. Same amount of distortion gives same measurement independently of the presentation color of the EC. An analogous result could be obtained if under the same system configuration with same color patch a video with different frame content is transmitted. The color of the gap and the content of the frame affects the classical PSNR while not the results of $\text{PSNR}_{\text{d},\alpha}$.

A. System Parameters

For the evaluation of the video communication system, we use the ParkScene.yuv video sequence as v_{i}, with 1920x1080 pixel resolution and $fps = 24$. We use the popular x265 v.3.4 for the video encoder, with following parameters: slices $= \{1, 3, 6, 12\}$, GOP length or keyint $= \{30, 60, 90, 120\}$, $QP = [20, 50]$, only one reference

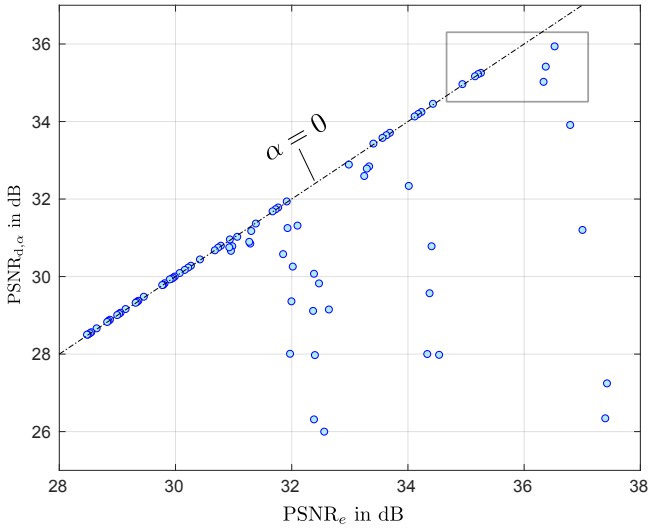

Figure 5: $\mathrm{PSNR}_{d,\alpha}$ vs PSNR_e. An average over 30 simulations are depicted for each combination of parameters and 100 simulations for the points inside the square.

frame and only P-Frames to minimize the delay, close GOP and *psnr* as the tuning criteria. The values of r_c for the channel coding are given in Section II-C and $\Gamma = \{1,8\}$. The parameters for our channel model are: $B = 10$ Mbps, $P_e = 0.001$ and $L_\mathrm{B} = 8$. For the video decoder, we use the popular FFMPEG v.3.4.5. The evaluation metric in (4) is configured with $\upsilon = 2$ and $\beta = 0$.

B. Results

Figure 4 shows the PSNR_e for different parameters sets at the encoder; it depicts B and each r_c for which the QPs are selected to maximize R_e under the condition that $R_\mathrm{e} r_\mathrm{C}^{-1} \leq B$. There are 144 different combinations of parameters ($|\mathrm{slices}| \times |\mathrm{GOP}| \times |r_\mathrm{c}|$) to encode and transmit v_i. Each encoded video is sent over the channel to produce Figure 5 in which a comparison of the overall performance of the system for each parameter set is depicted. In our results, $\mathrm{PSNR}_{\mathrm{d},\alpha} = [2, \mathrm{PSNR_e}]$ dB. Points on the line $\alpha = 0$ indicates that $D_\mathrm{d} = 0$ for all simulations, i.e., channel codes with $r_\mathrm{c} \leq r_\mathrm{RS}/3$ were capable to correct all errors. Similar results are obtained with $\Gamma = 8$ and $r_\mathrm{c} \leq (3r_\mathrm{RS})/4$, which enhances $\mathrm{PSNR}_{\mathrm{d},\alpha}$ if an increment in delay can be tolerated. Moreover, unlike H.264/AVC, we found out that for HEVC, videos encoded with slices $= 1$ give the highest $\mathrm{PSNR_e}$ and $\mathrm{PSNR}_{\mathrm{d},\alpha}$; this can be verified with a closer inspection in Figures 4 and 5. Moreover, it can be observed in the small square in Figure 5 that the two highest $\mathrm{PSNR}_{\mathrm{d},\alpha}$ have $\alpha > 0$ and are encoded with $r_\mathrm{c} = r_\mathrm{RS}/2$. This means that despite sporadic uncorrected errors it may be worth it to assign more bits to reduce D_e instead of using a stronger channel codes but letting D_e increase. This trend was observed for $\upsilon \geq 0.5$.

In terms of GOP length, as expected, shorter GOPs reduces D_d. Conventional PSNR gives similar results to the Figure 5, nevertheless, the maximum $\mathrm{PSNR_d}$ is reached for parameters that gives $D_\mathrm{d} = 0$. We recall from the example in Figure 3 that conventional PSNR may be strongly biased with how the errors are presented in v_d in relation with the video content. It is convenient for estimating small changes on pixel values but not for error introduced by the channel. Our proposed metric $\mathrm{PSNR}_{\mathrm{d},\alpha}$ is based in both the same conventional $\mathrm{PSNR_e}$ and pixel errors rate.

IV. CONCLUSIONS

In this paper, we evaluate a low delay video communication system and introduce for its evaluation a methodology based on the dynamic time warping algorithm for the synchronization of the received video with the transmitted video if loss of frames occurs. Furthermore, a new metric based in classical PSNR measurements is proposed to measure the overall distortion of the system. We observed that the coding efficiency of HEVC permits to assign enough bits to the channel coding, hence, no more than one slice per frame is required to enhance the quality of the transmitted video. Moreover, the parameter set that minimizes the overall distortion can be found for a video even if it has few unconcealed errors at the decoder.

REFERENCES

[1] C. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, Jul. 1948.

[2] Robert E Van Dyck and David J Miller, "Transport of wireless video using separate, concatenated, and joint source-channel coding," *Proceedings of the IEEE*, vol. 87, no. 10, pp. 1734–1750, 1999.

[3] Klaus Stuhlmuller, Niko Farber, Michael Link, and Bernd Girod, "Analysis of video transmission over lossy channels," *IEEE Journal on selected areas in communications*, vol. 18, no. 6, pp. 1012–1032, 2000.

[4] Thomas Stockhammer, Miska M Hannuksela, and Thomas Wiegand, "H. 264/avc in wireless environments," *IEEE transactions on circuits and systems for video technology*, vol. 13, no. 7, pp. 657–673, 2003.

[5] Kostas P., "Hevc in wireless environments," *Journal of Real-Time Image Processing*, vol. 12, no. 2, pp. 509–516, 2016.

[6] M. Abdollahzadeh, H. Seyedarabi, J. M. Niya, and N. Cheung, "Optimal hevc configuration for wireless video communication under energy constraints," *IEEE Access*, vol. 6, pp. 72479–72493, 2018.

[7] G. J. Sullivan, J. Ohm, W. Han, and T. Wiegand, "Overview of the high efficiency video coding (hevc) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, 2012.

[8] E. N. Gilbert, "Capacity of a burst-noise channel," *The Bell System Technical Journal*, vol. 39, no. 5, pp. 1253–1265, 1960.

[9] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *The Bell System Technical Journal*, vol. 42, no. 5, pp. 1977–1997, 1963.

[10] J. Hagenauer, "Rate-compatible punctured convolutional codes (rcpc codes) and their applications," *IEEE Transactions on Communications*, vol. 36, no. 4, pp. 389–400, 1988.

[11] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 26, no. 1, pp. 43–49, 1978.