Multiple description video coding for real-time applications using HEVC

1 Introduction

Remote control vehicles require the transmission of large amounts of data, and video is one of the most important sources for the driver. To ensure reliable video transmission, the encoded video stream is transmitted simultaneously over multiple channels. However, this solution incurs a high transmission cost. To address this issue, it is necessary to use more efficient video encoding methods that can make the video stream robust to noise. Moreover, it should have a less complexity to adapt to the real-time requirement. In this paper, we propose a low-complexity, low-latency 2-channel Multiple Description Coding (MDC) solution with an adaptive Instantaneous Decoder Refresh (IDR) frame period, which is compatible with the HEVC standard which adaptive redundancy adjustment. This method shows a better resistance to high packet loss rates with lower complexity.

Résumé

Remote control vehicles require the transmission of large amounts of data, and video is one of the most important sources for the driver. To ensure reliable video transmission, the encoded video stream is transmitted simultaneously over multiple channels. However, this solution incurs a high transmission cost. To address this issue, it is necessary to use more efficient video encoding methods that can make the video stream robust to noise. Moreover, it should have a less complexity to adapt to the real-time requirement. In this paper, we propose a low-complexity, low-latency 2-channel Multiple Description Coding (MDC) solution with an adaptive Instantaneous Decoder Refresh (IDR) frame period, which is compatible with the HEVC standard which adaptive redundancy adjustment. This method shows a better resistance to high packet loss rates with lower complexity.

Mots clefs

Multiple Description coding, HEVC, noisy channel, Error Correction, low latency

2 Proposed Method

We propose a balanced MDC with two descriptions as described in Figure 1. Assume that the packet error distribution of both channels is independently and identically distributed (i.i.d). Therefore, for each frame of the sequence, the expected distortion at the decoder is expressed as:
The first-order conditions, leading to:

$$D_e = (1-p_e) \sum_{j=1}^{2} D_{p,j} + p_e (1-p_e) \sum_{j=1}^{2} D_{r,j} + p_e^2 D_{error} \tag{1}$$

Here, $D_{p,j}$ and $D_{r,j}$ are the total distortion of the principal CTUs and the redundant CTUs, respectively, in the same description $j$ of the same frame. We define the following relationship:

$$D_{p,j} + D_{r,j} = \sum_{i=1}^{N} d_{i,j} \quad \forall j \in \{1, 2\} \tag{2}$$

In this equation, $d_{i,j}$ is the quantization distortion of a CTU$_{i,j}$, where $i$ is the CTU index in a frame, $j$ is the description and $N$ is the total number of CTUs in each frame. $p_e$ is the probability of packet error. The term $D_{error}$ is the distortion when the two descriptions are lost simultaneously. Thus $D_{error}$ is a constant and can be omitted from the cost function. The problem is to find the set of $QP_{i,j}$ to use for each CTU$_{i,j}$ in a frame which minimizes the expected distortion $D_e$ under a frame target bit rate $R_t$. The optimal MDC rate-constrained optimization problem is then given by:

$$\min_{QP_{i,j}} D_e \quad \forall j \in \{1, 2\} \quad \forall i \in \{1, ..., N\}$$

subject to:

$$R_j = \frac{R_t}{2}, \quad QP_{min} \leq QP_{i,j} \leq QP_{max}$$

This problem can be solved using the standard Lagrangian approach and minimizing the following cost function:

$$J_{\lambda_1, \lambda_2}(R_1, R_2) = D_e + \sum_{j=1}^{2} \lambda_j (R_j - R_t/2) \tag{3}$$

As the two descriptions are independent from each other, we can therefore establish:

$$J_{\lambda_1, \lambda_2}(R_1, R_2) = J_{\lambda_1}(R_1) + J_{\lambda_2}(R_2) \tag{4}$$

where $J_{\lambda_j}(R_j)$ contains only the terms of the corresponding description $j$ and is given by:

$$J_{\lambda_j}(R_j) = (1-p_e)D_{p,j} + p_e (1-p_e)D_{r,j} + \lambda_j (R_j - R_t/2) \tag{5}$$

The solution to the optimization-constrained problem is given by the first-order conditions, leading to:

$$\partial J_{\lambda_j}(R_j) / \partial R_j = 0 \quad \forall j \in \{1, 2\} \tag{6}$$

$$\partial J_{\lambda_i}(R_i) / \partial R_{i,j} = 0 \quad \forall j \in \{1, 2\} \quad \forall i \in \{1, ..., N\} \tag{7}$$

Due to the nonlinearity of the scalar quantizer in HEVC, minimizing the cost function is not straightforward. As a result, the rate-distortion function $d_{i,j}(R_{i,j})$ is not continuous and therefore not differentiable. To address this challenge, we approximate the rate-distortion relationship using the exponential function, given by:

$$d_{i,j}(R_{i,j}) = a_i e^{b_i R_{i,j}} \quad \forall i, j \tag{8}$$

where the parameters $a_i$ and $b_i$ are estimated using linear regression in each CTU of the residual frame produced by a pre-encoding process before encoding the two descriptions. Then, we can solve the problem using Algorithm 1 where $R_{min}$ and $R_{max}$ are given by $QP_{min}$ and $QP_{max}$ respectively for each CTU$_{i,j}$. To prevent error propagation, the Instantaneous Decoder Refresh (IDR) frame allows the encoder to send an intra-frame signal to the decoder, clearing the Central Reconstructed Buffer. All frames can then be decoded from this IDR frame. Therefore, the encoder needs to select the appropriate amount of IDR frames to achieve the best coding quality concerning the channel noise. The study [9] has shown that the optimal IDR frame period under i.i.d. packet error distribution is given by:

$$T_{IDR} = \frac{1}{p_e} \tag{9}$$
To decode the erroneous bitstream, if the principal CTU is lost, its redundant version will replace it. If two versions of the CTU are lost, the basic error concealment, which consists of replacing the block with the previous one, is applied.

3 Experimental Result

In this section, we evaluate the performance of our framework under a packet erasure channel. As mentioned earlier, the solution is implemented inside the HM codec [10]. To simulate the transmission, we use HEVC compressed streams with varying packet loss rates. Figure 2 shows that our proposed solution, which employs the LD-P configuration, outperforms the SF-PMDVC method with Random Access (RA) configuration for high packet error rates and high-motion sequences like BQMall, RaceHorses, and BlowingBubbles, while having a lower complexity encoding profile. Therefore, our proposed method is better suited for real-time applications.

4 Conclusion

In this study, we proposed a spatial-based multiple description encoding bit allocation and decoding solution that is adapted to HEVC standard. Our proposed MD coding scheme includes a bit allocation that distributes the redundancy between descriptions by adjusting the QP value for each CTU within a frame based on the channel characteristics and an IDR adaptation. This solution meets the requirements of low latency and good compression performance, making it suitable for use in remote control vehicles.

In our perspective, a more robust error handling mechanism with finer grain error detection at the CU level will enhance the performance of the system. Various methods as those discussed in [11] [12], could be employed to improve the decoding performance. Additionally, a scheme of optimization with error mismatch propagation model should improve the performance of the system.

Appendix

This is the summary of the article that has been submitted for the IEEE ICIP 2023 conference and is currently awaiting the review process.

Références


[10] High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Encoder Description Update 10 | MPEG.
