



Computational Intelligence in Electrical Engineering
Vol. 12, No. 4, 2022
Research Paper

End-to-end Distortion Modeling and Channel Adaptive Optimization of Mixed Layer Multiple Description Coding Scheme

Mohammad Kazemi

Assistant Professor, Department of Electrical Engineering, University of Isfahan, Isfahan, Iran
m.kazemi@eng.ui.ac.ir

Abstract :

Multiple Description Coding (MDC) is an error resilient coding technique in which multiple streams from a video source are generated, each is individually and mutually decodable. This error resiliency is obtained at the cost of redundancy, and the amount of redundancy depends on the channel loss rate as well as the frame position in the sequence. Due to the nature of video codecs, an erroneous frame will infect the successive frames and causes an unacceptable quality. MDC redundancy allocation more than required leads to compression inefficiency. Therefore, a channel adaptive optimization for frame-wise redundancy allocation is inevitable. In this paper, the MDC scheme known as Mixed Layer Multiple Description Coding (MLMDC) is under consideration for end-to-end distortion modeling and channel adaptive optimization. The model works based on the side and central decoder outputs mismatch. The performance of the model as well as the optimizer are verified by experimental results measured from JM19.0, H.264/AVC reference software. The experiments also show that the optimal MLMDC outperforms the conventional methods for high enough loss rates.

Keywords: Video Transmission, Multiple Description Coding, End-to-end Distortion Optimization, Error Propagation.

1. Introduction

Multiple Description Coding (MDC) is an error resilient coding approach to combat channel loss in multimedia transmission [1]. MDC is a promising solution that provides loss resiliency for video communication over error prone environments through multiple streams sent simultaneously. In MDC, multiple descriptions of the video source are generated, each individually decodable and mutually refinable. The descriptions are sent, possibly over multiple separate channels, and the reception of at least one description allows the video to be decoded. In the case of receiving all descriptions, they are

mutually reconstructed. This leads to a higher quality. So, in lossy packet networks, the video is delivered to the receiver unless all channels fail simultaneously, which is less probable than single channel failure. In practice, the total number of descriptions is usually two, but higher number of descriptions is also possible and has been used for certain applications. However, two-description coding is most commonly used. The reason is that the most gain of MDC is achieved in two-description coding; most research papers are for two-description coding [1] [3] [3].

The descriptions, in order to be individually



2252-083X/ © 2022 The Authors. Published by University of Isfahan
This is an open access article under the CC BY-NC-ND/4.0/ License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).



<http://dx.doi.org/10.22108/isee.2019.118706.1266>

decodable with an acceptable quality, must have some information in common. In the central decoder, only one copy of this common information is useful and the other one is redundant. In other words, the property of independently decodable descriptions is achieved at the cost of *redundancy*. The higher the needed side quality is, the higher this redundancy will be. Of course, the desired side quality depends on the channel condition: in channels with high loss rates, the probability of description missing is high and hence a higher side quality is required, whereas in low loss rate channels, most of the time all descriptions are available and side quality becomes less important. As a result, channel adaptive optimization to achieve the best performance of MDCs is inevitable.

As already mentioned, in the case of missing one description, the quality is acceptable, but it is not ideal with respect to frame quality. As a result, the reference reconstructed at the decoder has mismatch with what has been used at the encoder. Since this erroneous reference is used for inter prediction for the next frames, the error will be propagated. The error propagation continues until an I-frame is correctly received.

Error propagation sometimes is not considered in MDC optimization but it has very important role in MDC performance. In [4], the end-to-end distortion is calculated, but the redundancy is equally allocated to the frames; that is, the frame positions are ignored which leads to non-optimal solution. The effect of error propagation in the design of quantization-based MDCs has been minimized in [7] and [8], too. In [7], based on an MB's contribution to motion compensation, the importance of each MB is determined and then its redundancy is determined accordingly. However, for motion path analysis, one needs to have the motion vectors (MVs) for one Group of Pictures (GOP) which sometimes cause about 1 sec delay. The algorithm used in [8] provides optimized per frame redundancy allocation for an MDC method (the same MDC used in [7]), but it is specific for this MDC scheme. Furthermore, it is based on error propagation model of [9], which is an empirical model but only valid for low loss rates. The recursive approaches such as [10] and [11] introduced for modeling of error

propagation for SDC are more accurate than [9].

MDC optimization has been also presented in [12]-[14]. But, the problem is that they are much complex algorithm which might not be feasible or recommended for mobile devices. Furthermore, they are proposed for a specific MDC algorithm. Finally, drift avoiding in MDCs is studied in [15][16][17]; to compensate the difference of the descriptions, some data are sent as side information, causing additional redundancy. However, these approaches offer no end-side distortion optimization. MDC with intra coding and MDC with multi-view coding has been presented in [5] and [6], respectively. The work of [28] proposes a MDC scheme based on the visual saliency map. Color contrast statistics are used to define saliency value of each pixel in that work, then when residual encoding, more bits are allocated to the salient region as compared to the non-salient region. Redundancy allocation can be performed with a complex method based on slopes of the redundancy-bitrate curves as presented in [29].

It can be seen that end-to-end distortion modeling as optimization in MDC has been considered for a few cases, and these cases have studied it for a specific MDC scheme. In [19], an MDC scheme is introduced which works based on the mixing layers (MLMDC). This MDC provides some favorable aspects compared to the some DCT-domain MDC techniques as presented there. In this paper, the MDC under study for end-to-end distortion modeling and optimization is MLMDC. Mixing the base and enhancement layers has been proposed in [27], too. In that work, the base layers are generated by spatially subsampling to provide lower resolution videos each sent over a different channel. The enhancement layers for even higher quality are transmitted with the base layers. With the assist of data partitioning, the not received blocks in the base layer are error concealed.

The rest of this paper is organized as follows: Section 2 provides an overview of the MLMDC scheme. In Section 3, the end-to-end distortion model is derived for MLMDC. The objective function is formulated in Section 4. The experimental results showing the performance of the model, and the optimal Rate-Distortion (RD) curves of the MLMDC

algorithms as well as the comparison with other methods are given in Section 5, and finally the paper is concluded in Section 6.

2. Mixed Layer MDC (MLMDC)

As already described, two-description coding is the focus of this paper. When both descriptions are available, they are decoded by the *central decoder* and the distortion (quality) achieved is called *central distortion (quality)*. In the case that one description is received, it is decoded by the *side decoder* and the resulting distortion (quality) is called *side distortion (quality)*.

A preliminary version of the MLMDC method was presented in [18], with detailed descriptions and performance evaluations presented later in [19]. But, in order to make this paper readable and self-contained, in this section we repeat some basic descriptions of MLMDC from those papers.

The MLMDC technique is inspired from [20], but instead of combining different-frequency DCT coefficients, the base layer coefficients are combined with the *enhancement* layer coefficients. The base and enhancement layers are produced such as those of Coarse-Grain Scalable coding [21]. Fig. 1 shows the block diagram of the encoder, the central decoder, and the side decoder of MLMDC. Signals x_i and \hat{x}_i are DCT coefficients at the i^{th} position before quantization and after dequantization by Q , respectively. Subtracting \hat{x}_i from x_i results in a coarse quantization error which clearly cannot be quantized again by Q , and so a smaller quantization step size is required. The second quantization is carried out with Q/c , $c > 1$, to produce the enhancement coefficients x_{Qe_i} .

The enhancement coefficients are then added to the base coefficients to produce the combined coefficients z_{Q_i} . As shown in Fig. 1(a), the combined and base coefficients are alternated between the descriptions. At the central decoder, as shown by Fig. 1 (b), the base and enhancement coefficients are separated and after dequantization, a two-layer decoding is performed, as follows:

$$\begin{aligned} x_{Qe} &= z_Q - x_Q \\ x_{cen} &= Qx_Q + (Q/c)x_{Qe} \\ D_{cen} &= E[(x - x_{cen})^2] \end{aligned} \quad (1)$$

At the side decoder, the mentioned separation is not possible, since one description is not available. Here, the base coefficients are estimated from the combined coefficients, as given by

$$x_{Q_{est}} = \begin{cases} z_Q - \text{sign}(z_Q)Z_0 & |z_Q| \geq 2 \\ N_0 z_Q & |z_Q| < 2 \end{cases} \quad (2)$$

where Z_0 and N_0 are functions of λ , Q and c ; their functionality and the derivation details can be found in [19]. Consequently the side distortion is computed by

$$D_{side} = E[(x - Qx_{Q_{est}})^2] \quad (3)$$

MLMDC with the above structure is an MDC scheme shows promising performance especially at high Packet Loss Rates (PLRs).

3. End-to-end Distortion Model for MLMDC

Based on general model presented in [12], the end-to-end distortion of frame n is obtained as:

$$D_{e2e}^n = D_c^n + D_Q^n = \left(\sum_{i=0}^n \alpha^{i,n} \Delta_X^i \right) + E[(\delta_Q^n)^2] \quad (4)$$

where D_c^n and D_Q^n are channel and quantization distortions, respectively; Δ_X^i is the mismatch distortion and is defined as:

$$\Delta_X^i = E[(\delta_X^i)^2] \quad (5)$$

in which δ_X^i is the mismatch between the transmitted signal and received signal for the i^{th} frame; and δ_Q^n is the difference signal caused by quantization at the encoder. For the channel distortion we can write that:

$$\begin{aligned} D_c^n &= \sum_{i=0}^n \alpha^{i,n} \Delta_X^i \\ \alpha^{i,n} &= \prod_{j=i+1}^n (1 - \beta^j) \quad \text{for } i < n \\ \alpha^{i,n} &= 1 \quad \text{for } i = n \end{aligned} \quad (6)$$

where β^j is the intra rates of the j^{th} frame. Equation (6) says that in order to find the channel distortion at each frame, one must calculate the mismatch distortion associated with that frame as well as the mismatch

distortion of all previously coded frames.

The propagated error will be mitigated by two factors. First, by intra coding of some blocks (nonzero β) as shown in (6). And the second one is spatial filtering which results from sub-pel motion compensation and de-blocking filtering, acts as an averaging filter and

suppresses the power of the error propagated from the previous frames.

The core of the end-to-end distortion model is the mismatch signal used in (5). The MDC parameter controls side and decoder outputs mismatch, and for MLMDC this parameter is denoted as C :

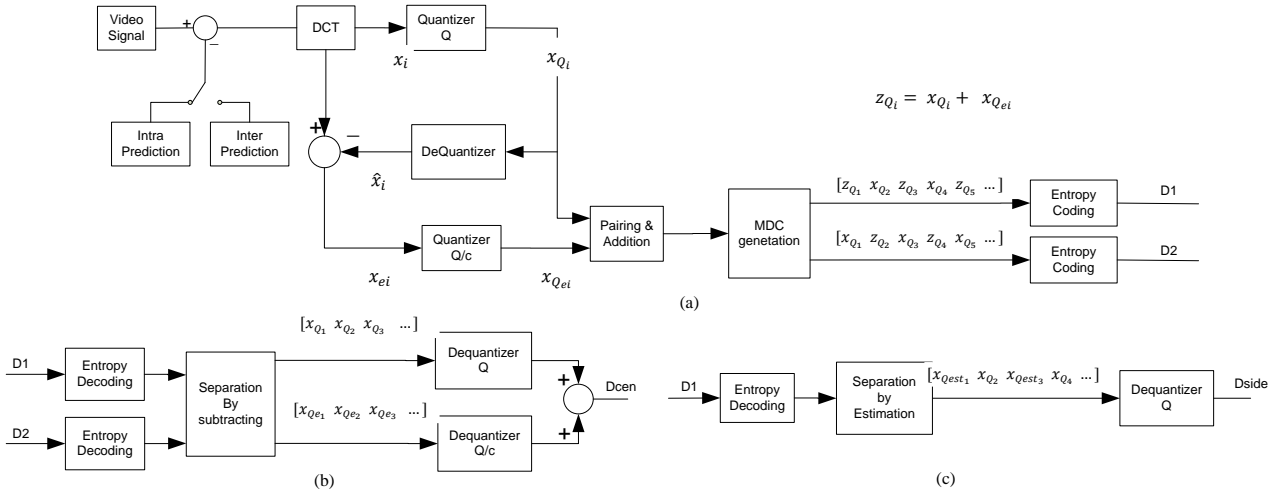


Fig. 1. The Block Diagram of MLMDC (a) Encoder (b) Central Decoder (c) Side Decoder

$$C = \{c^{(0)}, c^{(1)}, \dots, c^{(N-1)}\} \quad (7)$$

which $c^{(i)}$ is the parameter associated with i^{th} frame. This parameter is used in (1), the basic equation of MDC generation in MLMDC.

In MLMDC, there are two types of mismatch: mismatch between the estimated signal and the signal decoded from the two layers, δ_{2L-est} , and the mismatch between the estimated signal and the signal decoded from one layer, δ_{1L-est} . In other words, δ_{2L-est} is originated from the difference of the x_{cen} and x_{Qest} and δ_{1L-est} is originated from the difference of the x_Q and x_{Qest} . Signal δ_{1L-est} is the case of reference mismatch and the source of error propagation; since, similar to the previous MDCs, the references are reconstructed by only the base layer coefficients. It is worth noting that in one-layer decoding, for the separation of the base and enhancement coefficients, both descriptions are needed but only the base coefficients are used. Considering these two types of mismatch, channel distortion for this MDC scheme is modified as:

$$D_c^n = \sum_{i=0}^{n-1} \alpha^{i,n} \Delta_{X_{1L-est}}^i + \Delta_{X_{2L-est}}^n \quad (8)$$

in which the mismatch distortion associated with the n^{th} frame, $\Delta_{X_{2L-est}}^n$, which is due to the current frame data loss and not due to the error propagation, has been separated. Correspondingly, (4) becomes:

$$\begin{aligned} D_{e2e}^n &= \sum_{i=0}^{n-1} (\alpha^{i,n} \Delta_{X_{1L-est}}^i) + \Delta_{X_{2L-est}}^n \\ &+ E [(\delta_Q^n)^2] \\ &= \sum_{i=0}^{n-1} (\alpha^{i,n} E [(\delta_{X_{1L-est}}^i)^2]) \\ &+ E [(\delta_{X_{2L-est}}^n)^2] + E [(\delta_Q^n)^2] \end{aligned} \quad (9)$$

For the last two terms of (9), we can write:

$$\begin{aligned}
 E[(\delta_{x_{2L-est}}^n)^2] + E[(\delta_Q^n)^2] &= 2P_{01}E[(\delta_{01}^i)^2] \\
 &+ E[(\delta_Q^n)^2] \quad (10) \\
 &= 2P_{01}E[(\hat{X}_0^n - \hat{X}_1^n)^2] \\
 &+ E[(X^n - \hat{X}_0^n)^2]
 \end{aligned}$$

then the right hand side of (10) is:

$$\begin{aligned}
 2P_{01}E[(\hat{X}_0^n - \hat{X}_1^n)^2] + E[(X^n - \hat{X}_0^n)^2] &= (2P_{01})\left(E[(\hat{X}_0^n - \hat{X}_1^n)^2]\right. \\
 &+ E[(X^n - \hat{X}_0^n)^2]) \quad (11) \\
 &+ (1 - 2P_{01})E[(X^n - \hat{X}_0^n)^2]
 \end{aligned}$$

With the assumption of independency between channel and quantization distortions [10], (10) becomes:

$$\begin{aligned}
 E[(\delta_{x_{2L-est}}^n)^2] + E[(\delta_Q^n)^2] &= 2P_{01}E[(X^n - \hat{X}_1^n)^2] \\
 &+ (1 - 2P_{01})E[(X^n - \hat{X}_0^n)^2] \quad (12) \\
 &= 2P_{01}D_{side}^n \\
 &+ (1 - 2P_{01})D_{cen}^n
 \end{aligned}$$

where D_{side}^n and D_{cen}^n are given by (3) and (1), respectively. Using (12), the distortion of (9) becomes:

$$\begin{aligned}
 D_{e2e}^n = \sum_{i=0}^{n-1} (\alpha^{i,n} E[(\delta_{x_{1L-est}}^i)^2]) &+ 2P_{01}D_{side}^n \\
 &+ (1 - 2P_{01})D_{cen}^n \quad (13)
 \end{aligned}$$

In equation above only $E[(\delta_{x_{1L-est}}^i)^2]$ is unknown which can be computed as follows:

$$\begin{aligned}
 E[(\delta_{x_{1L-est}}^i)^2] &= 2P_{01} \left(\frac{1}{2}\right) \left(\frac{1}{16}\right) \sum_{k=0}^{15} E[(\delta_{01k}^M)^2] \quad (14)
 \end{aligned}$$

in which

$$\delta_{01k}^M = Qx_Q - Qx_{Qest} \quad (15)$$

where superscript M signifies MLMDC, and x_{Qest} is calculated by (2). The factor (1/2) in (14) is due to this fact that the estimation is performed for the combined coefficients; that is for half of the coefficients in each description.

4. Objective Function

With the aim of minimum distortion at the

receiver, the end-to-end distortion derived before is used for objective function formulation. The error propagation continues until the end of the GOP; thus, the objective function is defined as the summation of individual frames distortion in the GOP, as follows:

$$\begin{aligned}
 \min_{\{C, QP\}} \{D_{GOP}\} &= \min_{\{P, QP\}} \left\{ \sum_{n=0}^{N-1} D_{e2e}^n \right\} \quad (16) \\
 s. t. \quad \sum_{n=0}^{N-1} R_1^n + R_2^n &\leq R_t
 \end{aligned}$$

where C is the vector of MLMDC parameters, QP is the quantization parameter, D_{GOP} is the total distortion over a GOP, N is the size of the GOP, R_1^n , R_2^n are the rates of frame n in description 1 and description 2, respectively, and R_t is the total allocated rate for the GOP under consideration.

The constrained problem of (16) can be solved by Lagrange method. The Lagrangian function of the problem (16) is:

$$J = \sum_{n=0}^{N-1} (D_{e2e}^n) + \lambda \left(\sum_{n=0}^{N-1} (R_1^n + R_2^n) - R_t \right) \quad (17)$$

where λ is the lagrange multiplier. Then the objective optimization function becomes

$$\begin{aligned}
 \min_{\{C, QP\}} J &= \min_{\{C, QP\}} \left\{ \sum_{n=0}^{N-1} (D_{e2e}^n) \right. \\
 &+ \lambda \left(\sum_{n=0}^{N-1} (R_1^n + R_2^n) - R_t \right) \left. \right\} \quad (18)
 \end{aligned}$$

It shows that the solution of equation (18), $\{C_\lambda^*, QP_\lambda^*\}$, is the solution of the constrained problem of (16) but with the constraint $\sum_{n=0}^{N-1} (R_1^n + R_2^n) < R(\lambda)$, as shown in [22]. In other words, if λ^* is found such that $R(\lambda^*) = R_t$, the solutions of (18) and the solutions of (16) are identical.

Some parameters used in equations (13)-(15) must be also known. For estimating the quantization and mismatch distortions, the distribution of the DCT coefficients is extracted by fitting a Laplacian distribution function [23]. With the distribution parameters, the distortions and bitrates can be computed [19]. Note that

since MLMDC parameter (C) affects the quantization and mismatch distortions and also the bitrates, these parameters must be estimated, and cannot be measured directly.

The steps of solving the objective function are as follows:

1- The distribution parameter of DCT coefficients of the GOP under optimization is calculated.

2- By selecting QP_0 from set $\{20, 24, 28, 32\}$, the rate budget (R_t) defined in (16) is determined.

3- According to the distributions parameters, the quantization distortion of each frame is computed as a function of QP .

4- Δ_X associated with each frame is obtained as a function of MLMDC parameter and QP .

5- Using Entropy function, the total rate of each description as a function of QP and MLMDC parameter is obtained.

6- The objective function of (18) is solved iteratively with the algorithm given in [12].

It is worth mentioning that, this approach, similar to the other approaches dealing with error propagation, needs to gather the information of the future frames. This causes delay in this approach. However, this issue is solved by predicting the DCT coefficients distributions. It is possible and practical, since DCT coefficients distributions, due to the motion compensation, are very close for the successive frames.

5. Experimental Results

In order to validate the end-to-end distortion model and also measure the performance of the optimizer, we implemented MLMDC schemes in the H.264/AVC reference software, JM 19.0.

Foreman and *Mobile* CIF videos are used for the tests. The state of the art codec is HEVC and the test videos resolutions are HD and beyond. However, HEVC is several times more complex than H.264/AVC [24]; furthermore, it has been shown in [25] that HEVC is less error resilient than H.264/AVC. For these reasons and due to the fact that there are many conditions (e.g. display size or limited bandwidth) that small pictures are more favorable, our implementations are carried out in H.264/AVC

standard codec. In order to provide competitive curves and fair comparison, the encoder configuration settings is as set in [19], therefore one may see the bitrate values somewhat more than with default settings. But this configuration is common for all schemes and conditions.

No B-frame is used, Context-Adaptive Binary Arithmetic Coding (CABAC) is used for the entropy encoder and rate control is off.

Each packet, in order to be independently decodable, must contain an integer number of slices and be smaller than the Maximum Transmission Unit (MTU) of the network (to avoid fragmentation). For MTU size of 1500 bytes, in order to maximize channel throughput, the payload size is set to be 1460 bytes where 40 bytes are reserved for RTP/UDP/IP header information. Each slice contains a group of MBs of the frame in raster scan order and also the header information. In the experiments, given the PLR, 40 random packet loss patterns with Bernoulli distribution are generated and applied on each channel. The channels are assumed to be independent.

5.1 End-to-End Distortion

In this subsection, the derived end-to-end distortion model is validated experimentally by measuring the frames' distortion when sending the MDC streams of the test videos over lossy channels. The MLMDC parameter range is as follows:

$$\begin{aligned} C_1 &= \{1, 1.3, 1.3, \dots, 1.3\} \\ C_2 &= \{1, 2, 2, \dots, 2\} \end{aligned} \tag{19}$$

The parameters are chosen such that the first frame of the GOP (I-frame) is duplicated in both descriptions and for the next P-frames, a constant MDC parameter is used.

All of our analytical derivations assumed laplacian distributed DCT coefficients. However, the distribution of the coefficients cannot be exactly modeled by a specific function. Therefore, the distortion models based on the laplacian distribution of coefficients are not necessarily very accurate in reality. Our simulations show that this error leads to a shift in quality when working in PSNR. In other words, the distortion curves predicted by the model must be added by a constant shift to match the actual distortion curves. The measured frame-wise quality as well as what is predicted by our end-to-end distortion model are

illustrated in Fig. 2 for *Foreman* and *Mobile* video sequences, $QP = 20$ and $QP = 28$, $PLR = 0.05$ and $PLR = 0.20$ and two set of parameters C_1 and C_2 . The required shifts are also reported in the figures. One can see that with only a shift calibration, the derived model can predict the actual distortion accurately most of the times. The Cauchy distribution model provides better distribution fitting than laplacian [26], however, working with Cauchy model is rather difficult, and also some amount of residual error would persist. Using the

models instead of the actual data, although very beneficial for analytical derivation, always produces some amount of inaccuracy.

It is possible that due to the occlusion, a significant number of MBs are intra coded in a specific frame. In Fig. 3, 50% of MBs in the fourth frame of the second GOP are forcedly intra coded. This way, a jump in PSNR occurs since intra coding clears the propagated error; however, we can see that the derived model follows the experiments closely.

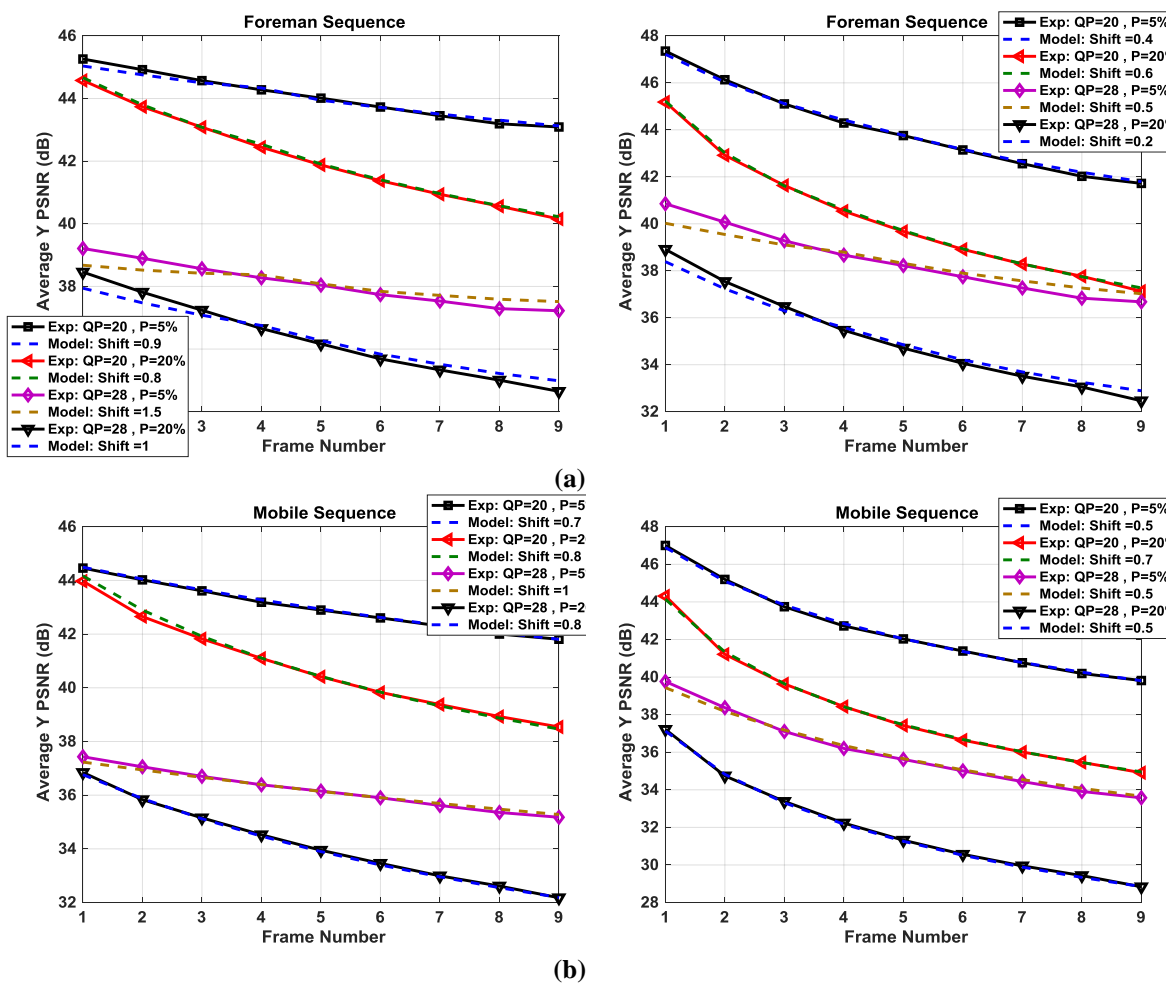


Fig. 2. The receiver side frames quality in PSNR, the experiments and the model curves when MLMDC parameter is C_1 and C_2 for the figures on the left and right respectively and (a) *Foreman* and (b) *Mobile* sequence

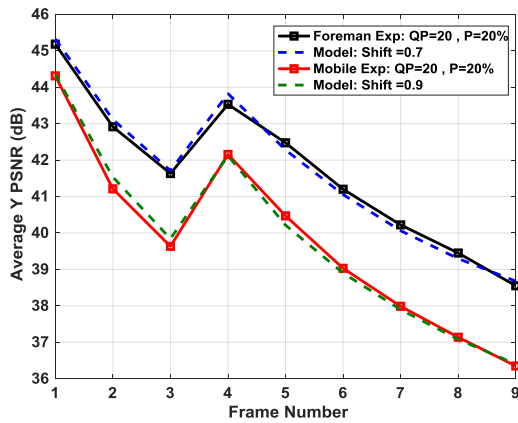


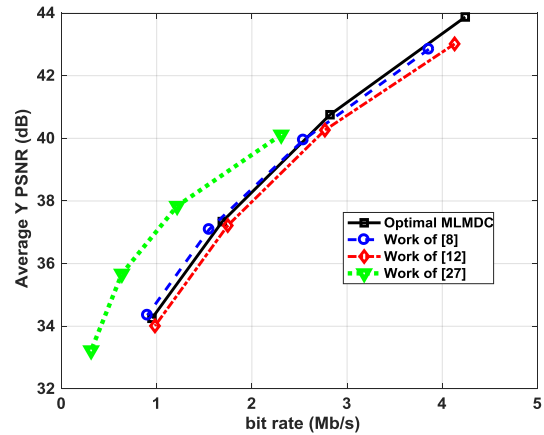
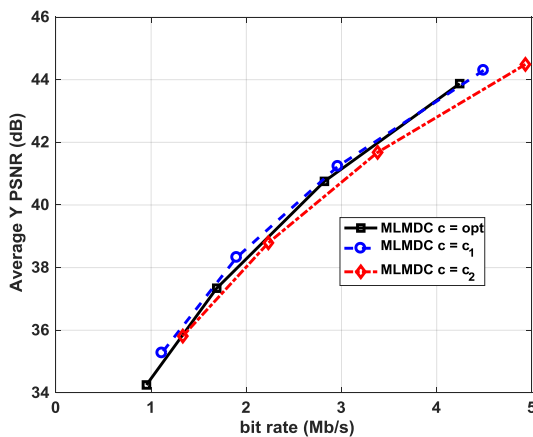
Fig. 3. The receiver side frames quality in PSNR in the presences of occlusion which leads to 50% intra MBs, the experiments and the model curves for both Foreman and Mobile sequences

It is worth mentioning that our model predicts the average distortion taken up over a sufficient number of packet loss patterns. The exact value of end-to-end distortion in the transmission of the video over lossy channels

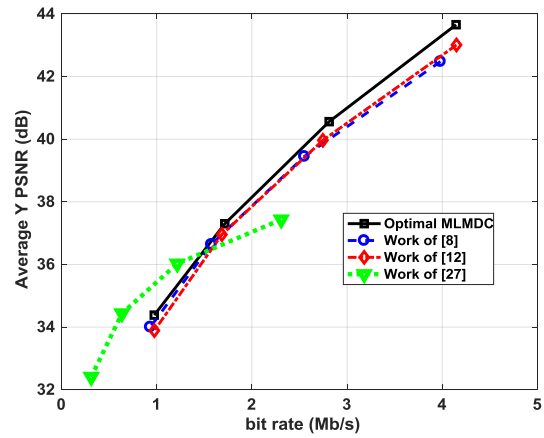
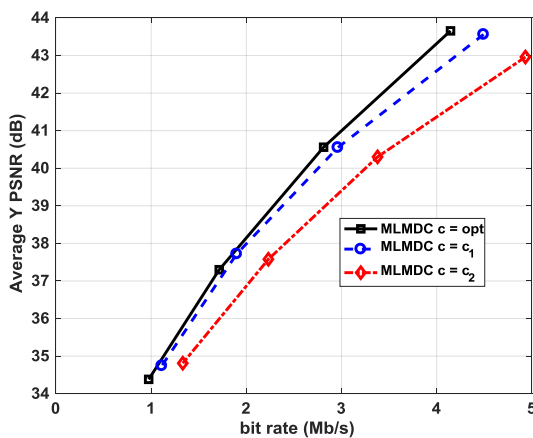
cannot be modeled neither by our model nor any other model. The reason is that the transmission distortion is highly dependent on not only the number of lost packets but also the pattern of their loss, and none of them are known *a priori*.

5.2 Optimization and RD Curves

As seen from the experimental results in Section 5.10, the analytically derived end-to-end distortion model needs calibration in the form of a shift to match the experimental results. In other words, the model predicts the slope of the quality decaying (in PSNR) fairly accurately but a constant shift must be added. However, as it is evident from Fig. 2, at a given PLR, while the MLMDC parameter varies from typically low to typically high values, the shift variation is mostly less than 0.5dB. The *QP* is the other optimization



(a)



(b)

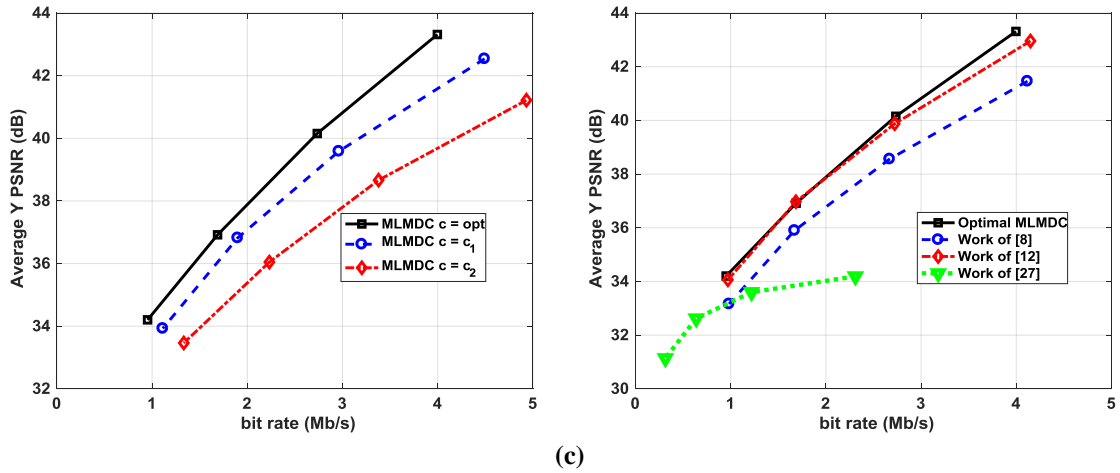
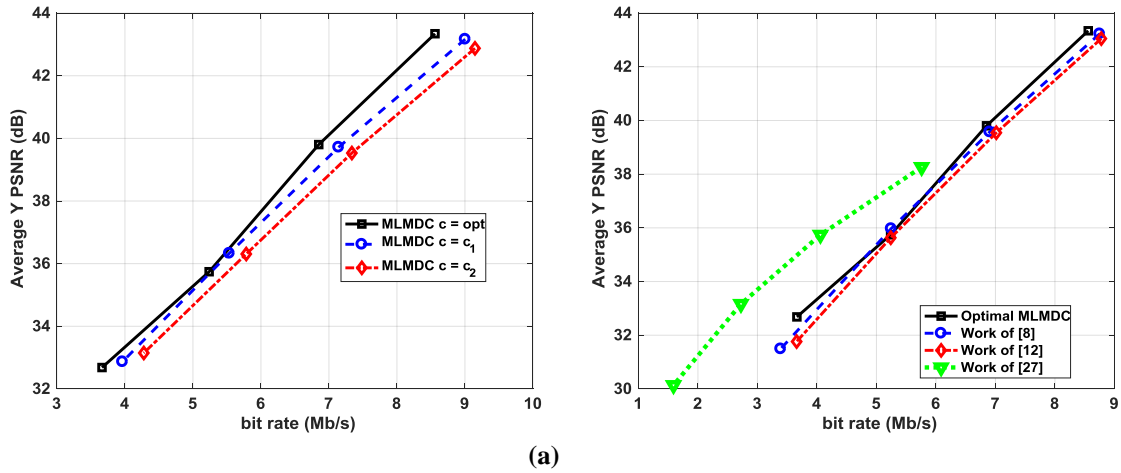


Fig. 4. Foreman sequence: The performance curves of the optimizer on the left and comparison of the MLMDC with the works presented in [8], [12] and [27] on the right for (a) PLR = 5% (b) PLR = 10% and (c) PLR =20%

variable over which the shift variation must be checked; the QP is almost constant and very close to QP_0 . Therefore, we can assume that the calibrating shift is well constant over the optimization variables and, hence, can be ignored for optimization.

For $QP_0 = \{20, 24, 28, 32\}$, the total allowable rate is computed, then for each GOP of the video source the optimal QP and the MDC parameters are calculated. The optimal values for the test sequences are applied and the average quality and the bit rate are measured.



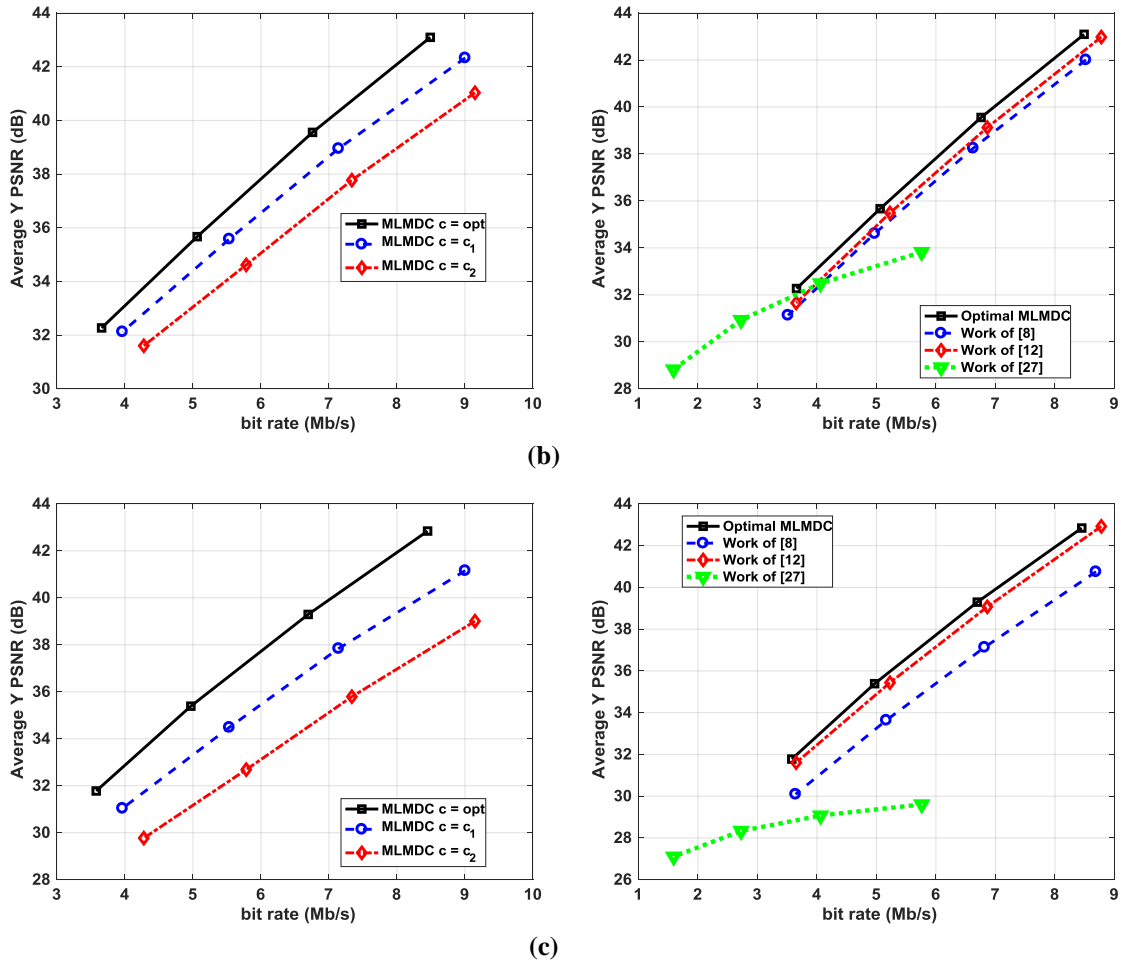


Fig. 5. Mobile sequence: The performance curves of the optimizer on the left and comparison of the MLMDC with the works presented in [8], [12] and [27] on the right for (a) PLR = 5% (b) PLR = 10% and (c) PLR = 20%

Furthermore, the curves for the case of typical constant parameters as defined in (19) are included. The results are shown in figures 4 and 5 for *Foreman* and *Mobile* videos. These results confirm the performance of the objective function and the optimally designed MLMDC. There are a few points that the optimal curves are slightly below than the curve of non-optimal parameters; the reason is that in our optimization, the Entropy Function is used for rate computation while the rate in the above figures is measured from the CABAC output. However, this difference is negligible.

In order to show the comparative performance of our optimizer, the optimum MLMDC is compared with the algorithms presented in [8], [12] and [27]. These MDC works try to provide an optimal video quality in the receiver side. The work of [27] has also some networking techniques which are not

applicable in our comparison scenario. One can see that for PLR of 5%, the output of [27] is significantly better than others, but it is not the case for PLRs of 10% and 20%. In other words, it can be said that for high enough PLRs, MLMDC outperforms them especially for *Mobile* sequence. The reason of success in [27] for low PLRs is the fact that this algorithm is actually greedy for redundancy allocation; therefore, it loses its performance for high enough loss rates. Note that with higher redundancies, the MDC becomes more robust against the packet losses.

6. Conclusion

In this paper, a fully analytical end-to-end distortion model was derived for MLMDC already introduced in [19] where MLMDC shows promising performance compared to the other DCT-domain MDCs. The model of end-

to-end distortion and its optimization is presented in this paper. The core of the model is the mismatch of the side and central decoder outputs. It was shown that the derived end-to-end distortion model, with a shift calibration, is fairly well matched to the actual distortion seen by the measurements. The model is then used for optimization of MLMDC. We also discussed that the shift calibration is not needed for finding the optimal MDC parameters. With a closer look at the results, it can be seen that at high enough loss rates (e.g. 10% and beyond), the optimal MLMDC outperforms the other examined MDCs to some extent.

References

- [1] V. K. Goyal, "Multiple description coding: Compression meets the network," *IEEE Signal Processing Mag.*, Vol. 18, pp. 74–93, 2001.
- [2] M. Kazemi, S. Shirmohammadi, K. H. Sadeghi, "A Review of Multiple Description Coding techniques for Error-Resilient Video Delivery", *Multimedia Systems*, Springer, 2013.
- [3] Y. Wang, A. Reibman, and S. Lin, "Multiple description coding for video delivery," *Proc. IEEE*, vol. 93, no. 1, pp. 57–70, 2005.
- [4] N. Kamnoonwatana, D. Agrafiotis and N. Canagarajah, "Flexible Adaptive Multiple Description Coding for Video Transmission," *IEEE Trans. Circuits Syst. Video Technol*, Vol. 22, pp. 1-11, 2012.
- [5] M. Kazemi, R. Iqbal and S. Shirmohammadi, "Joint Intra and Multiple Description Coding for Packet Loss Resilient Video Transmission," *IEEE Transactions on Multimedia*, Vol 20, No. 4, pp. 781-795, 2018.
- [6] Jing Chen ; Jie Liao ; Huanqiang Zeng ; Canhui Cai, "Multiple description coding for multi-view video with adaptive redundancy allocation," *International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, 2018.
- [7] C. Lin, T. Tillo, Y. Zhao, and B. Jeon, "Multiple Description Coding for H.264/AVC with Redundancy Allocation at Macro Block Level," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 21, pp. 589-600, 2011.
- [8] Tammam Tillo, Marco Grangetto, and Gabriella Olmo "Redundant Slice Optimal Allocation for H.264 Multiple Description Coding," *IEEE Trans. Circuits Syst. Video Technol*, Vol. 18, pp. 59-70, 2008.
- [9] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE J. Select. Areas Commun.*, Vol. 18, pp. 1012–1032, 2000.
- [10] Z. He, J. Cai, and C. W. Chen, "Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, Special Issue on Wireless Video, No. 6, pp. 511–523, 2002.
- [11] Y. Wang, Z. Wu, and J. M. Boyce, "Modeling of transmission-loss induced distortion in decoded video," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 16, No. 6, pp. 716–732, 2006.
- [12] M. Kazemi, K. H. Sadeghi, S. Shirmohammadi, and P. Moallem, "Rate/distortion optimization in multiple description video coding," *Signal Processing: Image Communication*, Vol. 36, pp. 95-105, 2015
- [13] Y. Xu and C. Zhu, "End-to-End Rate-Distortion Optimized Description Generation for H.264 Multiple Description Video Coding," *IEEE Trans. Circuits Syst. Video Technol*, Vol. 23, issue 9, pp. 1523 – 1536, 2013
- [14] M. Kazemi, R. Iqbal, S. Shirmohammadi, "Redundancy allocation based on weighted mismatch-rate slope for multiple description video coding," *IEEE Trans. Multimedia*, Vol. 19, No. 1, pp. 54-66, 2017.
- [15] A. Reibman, H. Jafarkhani, Y. Wang, M. Orchard, and R. Puri, "Multiple-description video coding using motion-compensated temporal prediction", *IEEE Trans. Circuits Syst. Video Technol.* , Vol. 12, pp. 193 - 204, 2002.
- [16] P. Correia, P. A. Assuncao, and V. Silva, "Multiple Description of Coded Video for Path Diversity Streaming Adaptation," *IEEE Trans. Multimedia.*, Vol. 14, No. 3, pp. 935–375, 2012.
- [17] M. Ma, O. C. Au, L. Guo, S.-H. Gary Chan, and P. H. W. Wong, "Error Concealment for Frame Losses in MDC," *IEEE Trans. Multimedia.*, Vol. 14, No. 3, pp. 935–375, 2012.
- [18] M. Kazemi, K. Sadeghi, and S. Shirmohammadi, "A high video quality Multiple Description Coding scheme for lossy channels," *IEEE International Conference on Multimedia and Expo (ICME)*, 2011.
- [19] M. Kazemi, K. H. Sadeghi, and S. Shirmohammadi, "A Mixed Layer Multiple Description Video Coding Scheme," *IEEE Trans. Circuits Syst. Video Technol*, Vol. 22,

- pp. 202-215, 2012.
- [20] Y. Wang, M. T. Orchard, V. Vaishampayan, and A. R. Reibman, "Multiple Description Coding Using Pairwise Correlating Transforms" *IEEE TRANS. Image Processing*, Vol. 10, pp. 351-366, 2001.
- [21] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 17, No. 9, pp. 1103-1120, 2007.
- [22] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. 36, pp. 1445-1453, 1988.
- [23] X. Li, N. Oertel, A. Hutter, and A. Kaup, "Laplace distribution based Lagrangian rate distortion optimization for hybrid video coding," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 19, No. 2, pp. 193-205, 2009.
- [24] F. Bossen, B. Bross, K. Sühling, and D. Flynn, "HEVC complexity and implementation analysis," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 22, No. 12, pp. 1684-1695, 2012.
- [25] B. Oztas, M.T. Pourazad, P. Nasiopoulos, and V.C.M. Leung, "A study on the HEVC performance over lossy networks," 19th IEEE International Conference on Electronics, Circuits and Systems (ICECS), 785-788, 2012.
- [26] Y. Altunbasak and N. Kamaci, "An analysis of the DCT coefficient distribution with the H.264 video coder," in *Proc. IEEE ICASSP*, Vol. 3. 2004, pp. III-177-III-80.
- [27] Sh. Ghahremani, M. Ghanbari, "Error resilient video transmission in ad hoc networks using layered and multiple description coding," *Springer Multimed. Tools App.*, Vol. 76, No. 6, pp. 9033-9049, 2017.
- [28] M. Majid, Muhammad Owais, S. M. Anwar, "Visual saliency based redundancy allocation in HEVC compatible multiple description video coding," *Springer Multimed Tools Appl.*, Vol. 77, No. 16., pp. 20955-20977, 2018.
- [29] M. Kazemi, R. Iqbal, S. Shirmohammadi, "Redundancy allocation based on weighted mismatch-rate slope for multiple description video coding," *IEEE Trans. Multimedia*, Vol. 19, No. 1, pp. 54-66, 2017.