A Review of Multiple Description Coding Techniques for Error-Resilient Video Delivery

Mohammad Kazemi, Khosrow Haj Sadeghi, Shervin Shirmohammadi

Abstract- **Multiple Description Coding (MDC) is one of the promising solutions for live video delivery over lossy networks. In this paper, we present a review of MDC techniques based on their application domain and we explain their functionality, with the objective of giving enough insight to designers to decide which MDC scheme is best suited for their specific application based on requirements such as standard compatibility, redundancy tunability, complexity, and extendibility to n-description coding. The focus is mainly on video sources but image based algorithms applicable to video are considered as well. We also cover the well-known and important problem of drift and solutions to avoid it.**

Keywords Multiple Description Coding (MDC), Video Streaming, Best Effort Network

1 Introduction

Video transmission over noisy channels has been a challenging problem for more than two decades. Transmission of raw video is not feasible due to the very large bandwidth required and so video compression is inevitable. On the other hand, compressed video is very sensitive to data loss which happens in best-effort networks such as the Internet. To counter the effect of data loss for video transmission over noisy networks, there are three categories of approaches: a reliable Automatic Repeat Request (ARQ) based transport layer protocol such as TCP, Forward Error Correction (FEC), and Error Resilient Coding (ERC). ARQ and FEC are channel level protections, while ERC can be used as either source level protection, such as Multiple Description Coding (MDC), or as both source and channel level protection known as Joint Source Channel Coding (JSCC) such as Layered Coding (LC). In this paper, we specifically focus on MDC as a method to counter video packet loss. We review existing MDC schemes, and we provide a taxonomy and analysis to aid practitioners and researchers for better understanding and selecting the most suitable MDC scheme for their specific application. But before doing so, we need to have a high level understanding of loss resilient methods and where MDC fits in that group. In the following two subsections, we present this overview.

1.1 Overview of Packet Loss Resilient Methods for Video

ARQ: In ARQ, the receiver asks, through a back channel, for the retransmission of a lost packet. Lost packets are the packets not received at all or received with bit error. By checking the parity symbols, the erroneous packets are detected and requested for retransmission.

FEC: In this method, some correcting data is added to the main message in a redundant manner. If the lost data are within the correcting capability of the FEC codes, the whole message can be recovered. The correcting capability depends on the number of added parity symbols. For example, in Reed Solomon FEC, by adding $2t$ parities, t erroneous symbols can be corrected.

ERC options of H.264/AVC: These coding options such as MB Intra-refreshment, B frame coding, Slicing, and Flexible Macroblock Ordering (FMO) enable the decoder to conceal and regenerate the lost data from the received data, exploiting the correlation existed among blocks of the images.

Layered Coding: In this method, which is also called Scalable Video Coding (SVC), the source is coded into a Base Layer and one or more Enhancement Layers. At the receiver, the layers are superimposed on each other hierarchically. The quality of the video is enhanced by the number of received enhancement layers. The base layer is usually protected using FEC codes and hence LC categorizes as a JSCC method.

MDC: In MDC, independently-decodable and mutually-refinable streams of a video source are generated. The streams, also called descriptions, are then transmitted separately, possibly through different network paths. In MDC, as long as one or more descriptions arrive at the receiver, some video with certain quality can be displayed. If a packet is lost, the corresponding packets of the other descriptions, containing a different representation of the data in the lost packet, may be available and the video is decoded successfully but with a lower fidelity. This is depicted in Fig. 1.

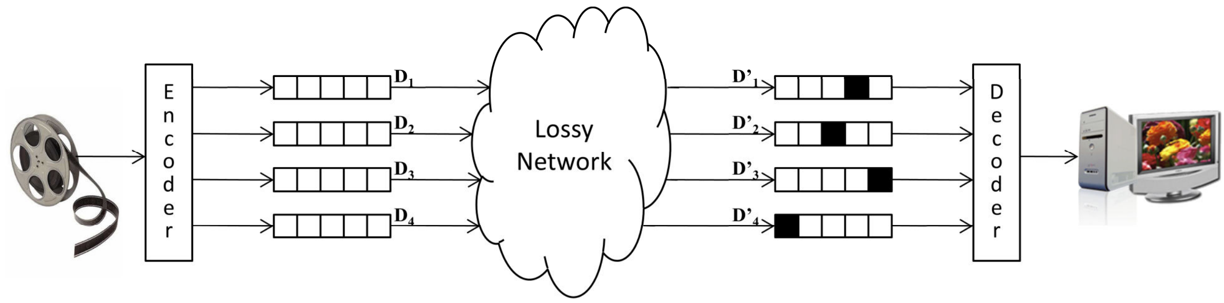

Fig. 1 The main rationale behind MDC video. Black boxes indicate lost information

1.2 MDC in the "Big Picture"

It should be noted that the aim of our paper is not to argue for or against any of the above schemes of video transmission over lossy networks. With this in mind, in order to show the application domain of MDC, we compare MDC to the other methods as follows:

ARQ versus MDC: ARQ's applications are mostly presentational ones such as YouTube type of video distribution and on-demand services, as opposed to conversational ones where two or more people interact in a live session, such as video conferencing. It is clear that in live communications and in channels with long Round Trip Time (RTT) this method is not suitable due to the time it takes to ask for retransmission and then to receive it. In a presentational application the video can be paused while the receiver waits for the retransmitted packet, but a conversational application cannot pause due to its live nature. Also, in multicast communication where we are dealing with a potentially large number of receivers, responding to the receivers' request is not possible. MDC on the other hand does not have these limitations, at the cost of producing a higher video bit rate.

FEC versus MDC: In FEC, if the loss rate is beyond the value based on which the FEC scheme is designed, no lost data can be recovered, unless the codes are over-designed which in turn reduces the efficiency of such schemes. Since channel condition is dynamic and loss rate is usually time variant, FEC protection is not always a good approach. Experiment results presented in [1-4] compare the performance of MDC and FEC and show the advantages of MDC in such cases.

ERC options of H.264/AVC versus MDC: These options work only if the loss rate is very low or at most low. In moderate or high loss rate environments, MDC is more beneficial.

SVC versus MDC: Layers in SVC are hierarchical which means that a given layer cannot be decoded unless all of its lower layers have been received correctly. This limits the error resiliency of SVC, while MDC descriptions are not hierarchical and can be decoded independently and hence are more resilient against loss than SVC layers. In other words, once the loss rate grows beyond a certain threshold, MDC outperforms SVC as shown in [5-8]. On the other hand, SVC has less overhead than MDC and is more suitable for low loss rate situations. Similar to MDC, SVC also supports heterogeneous receivers with varying bandwidth requirements.

From the above, we can identify situations where MDC has advantages over other approaches and should be seriously considered:

1- <u>In real-time and/or live applications, where retransmission of a lost packet will miss the deadline and is not acceptable.</u> A popular example of this, in the context of today's modern applications, is video conferencing, specially High Definition Video Conferencing (HDVC). HDVC's stringent requirements such as maximum delay threshold and minimum video quality threshold make it a challenging application for video streaming [9]. In HDVC, a lost packet cannot be retransmitted, and will adversely affect the quality of the video, rendering useless the very purpose of HDVC [9].This is why most HDVC systems today recommend the usage of a dedicated network with guaranteed services. But such networks are expensive and with recurring costs. In this context, MDC can help HDVC run over best-effort networks, in a manner that is not possible with ARQ.

2- In situations with higher than usual network loss. As mentioned above, it has been shown that MDC outperforms SVC, FEC, and the built-in loss resiliency options of video codecs when the loss rate goes above a certain threshold. A possible usage of MDC would therefore be in mobile video streaming, another modern application in terms of today's context, since wireless networks are more lossy than wired networks. Indeed, in the literature we can see a recent tendency in using MDC for mobile video streaming applications [10-13]. Another usage would be in video applications where packet loss is unacceptable, such as our HDVC example. Even if in an HDVC session the loss rate occasionally or rarely goes above a certain threshold, those rare moments could be the exact times that an important event is happening in the meeting (signing of a contract, waiting for an important answer, trying to see your counterpart's facial expression to see if s/he is bluffing, etc.) So for those types of applications, MDC could be a better option.

3- Supporting heterogeneous receivers. To cope with bandwidth heterogeneity and make its functionality resilient to peer joining and disjoining, MDC in conjunction with multiple-tree structure provides a promising solution for P2P video streaming [14], as used in Hotstreaming [15], Splitstreaming [16] , CoopNet [17] and TURINstream [18]. MDC as a robust stream delivery method for Application Layer Multicasting in heterogeneous networks has been presented in [19].

1.3 Contributions

Previous survey papers on MDC have been published by Goyal [20] and Wang *et al.*[21]. Goyal's paper has three parts. In the first part, the historical development of MDC is presented. In the second part, rate-distortion behavior of MDC with an information theoretic view is studied and in the last part, the applications scenarios of MDC are introduced. In other words, [20] focuses on introducing MDC for applications such as multimedia transmission over lossy channels. Even though some examples of early MDC methods such as splitting, MDSQ (Multiple Description Scalar Quantizer), MDTC (Multiple Description Transform Coding) are presented, by now it is somewhat outdated and lacks not only newer methods but also the domain based categorization presented in our paper. Wang's paper is not actually a review of MDC methods, as it mainly discusses mitigation methods to counter the error propagation incurred by reference mismatch. Based on how the inter frame prediction is performed, it defines three classes A, B, and C, and categorizes the existing works in these three classes. In other words, the predictor types used in MDC papers are reviewed in [21], and the MDC techniques themselves and their strengths and weaknesses are not discussed there. For example, for spatial domain MDC, [21] refers to only some sample papers and their predictor type, while in our paper several modification techniques to the basic spatial domain MDC are explained and reviewed in detail. The same is true for all MDC domains. In addition, in our paper we examine each MDC scheme in details through four characteristics of standard compatibility, redundancy tunability, complexity of the algorithm, and the possibility to increase the number of descriptions. We also recommend a 5-question process to help practitioners select an appropriate MDC scheme for a given application. Therefore, our paper complements and extends [20] and [21] without much overlap, except for basic MDC concepts explained in all three.

In the next section, we present an overview of MDC and it's basic operations. Then, in Section 3 an overview of existing MDC techniques is presented. The solutions for avoiding drift are studied in Section 4. A comparison of MDC schemes with respect to their functionalities such as standard compatibility, redundancy adaptation, complexity and capability to n-description extension are presented in Section 5 and a discussion on the best suited MDC for a specific application is provided in Section 6. Finally the paper is concluded in Section 7.

2 MDC Basics

Multiple Description Coding was originally used for speech communicating over circuit-switched network in the 1970's. Traditionally, to avoid communication interruption, an additional transmission link was on standby and would be activated in the case of the outage of the main link. This approach however was not cost efficient and therefore the idea of splitting the information over two channels; i.e. MDC, was proposed. At the 1979 IEEE Information Theory Workshop, the MD problem was posed by Gersho, Witsenhausen, Wolf, Wyner, Ziv, and Ozarow. Suppose a source is described by two descriptions each coded at rate R_1 and R_2. Each description can be individually decodable with distortion D_1 and D_2, respectively, while decoding the two descriptions together leads to distortion D_0; the MD problem is to characterize the achievable quintuples $\{D_0, D_1, D_2, R_1, R_2\}$. The initial papers discussing this problem were [22-24], with more attention devoted to the topic in succeeding years first from a rate-distortion perspective and then in other engineering applications. Interested readers are referred to [20] for a deeper historical view of MDC.

Fig. 2 shows the basic block diagram of MDC. This figure shows a two-description case but a higher number of descriptions is possible. In the figure, a source is coded such that multiple complementary descriptions that are individually decodable are generated. After the descriptions are built, they can be transmitted separately, possibly through different network paths. At the receiver side, if only one description is available, it is decoded by the *side decoder* and the resulting quality (distortion) is called *side quality (distortion)*. When both descriptions are available, they are decoded by the *central decoder* and the resulting quality (distortion) is called *central distortion (quality)*. In central decoder the descriptions are merged and hence a video with higher quality is achieved. In other words, there exist two types of decoding at the receiver, when all descriptions are received the central decoding is used, and if one or more descriptions are not received the side decoder is used for the received description(s). Obviously, quality is enhanced by the number of received descriptions.

Since predictive coding is used in all modern video codecs, the quality of a predicted frame will depend on its reference frame. When MDC's side decoder is active; i.e., when some descriptions are lost, a reference frame may not be reconstructed correctly due to this loss, leading to noisy reconstruction of all other frames which are predicted from it. Subsequently, some of the erroneous frames could in turn be used as reference for other frames and so error propagation occurs. This phenomenon is known as *drift* and will be studies in details in section 4

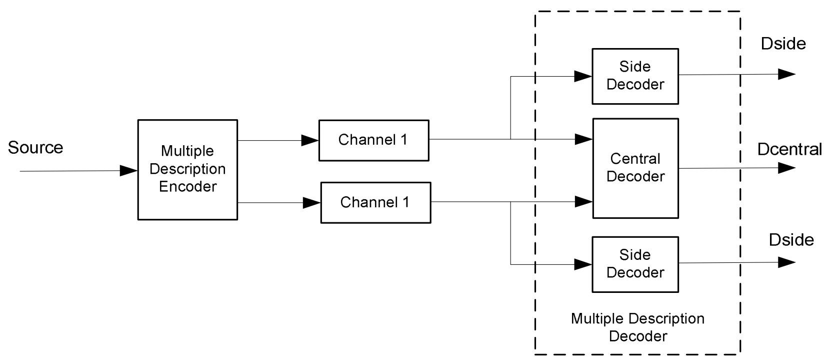

Fig. 2 Block diagram of MDC

In MDC, each description in order to be individually decodable must have some basic information from the source. At the central decoder, only one copy of that basic information is utilized and the others are redundant. Therefore, MDC's bit rate is larger than that of Single Description Coding (SDC) and they provide the same quality over a lossless channel. This excess rate is called redundancy. In other words, the descriptions, since they present the same source (but with distortion), are correlated and coding two correlated descriptions separately, produces redundancy. For higher side quality, since the descriptions are closer to the source and hence are more correlated, redundancy increases.

The cost of redundancy is accepted because of the error resiliency achieved by MDC. For channels with high packet loss rate (PLR), the probability of side decoding is more and hence higher side quality is of interest and therefore more redundancy is needed. However, for cases where all descriptions are received, this redundancy must be minimized. So, redundancy tunability is one of the main features of MDC schemes, and optimizing the amount of redundancy is the main challenge for MDC designers.

How to add the redundancy or equivalently enhance the side quality depends on the MDC method. In section 3, we categorize and explain each MDC method as well as ways to add redundancy. But before categorizing MDCs, we need to first define the characteristics used to categorize each MDC scheme. This is discussed next.

2.1 MDC characteristics

There are four qualitative factors which are important for selecting an appropriate MDC for a specific application: standard compatibility, redundancy tunability, complexity of the algorithm, and the possibility to increase the number of descriptions.

2.1.1 Standard compatibility

Standard compatibility allows the receivers to use their standard SDC decoder module without any change. Since currently there is no standard for MDC, this feature is of much interest. A single decoder can be used for SDC as well as for MDC. For merging the descriptions in the central decoder, some post-processing is needed, but the post-processing tasks are performed outside of the decoder and hence do not alter the standard compatibility of the MDC. This is the case for some side decoders, too. That is, as long as the MDC related processes are outside of the decompressor engine, the algorithm is standard compatible.

2.1.2 Redundancy tunability

As was explained, in MDC some common information must be copied into all descriptions in addition to the header information. This information is crucial to decode and use each description individually. In low loss-rate cases, mostly all of the descriptions are available and decoding one individual description rarely occurs; therefore there is no need for high redundancy. On the other hand, in high loss-rate cases, the probability of side decoding increases and side quality becomes more important. Therefore, controlling the amount of redundancy continuously and as wide as possible is a major feature which provides adaptive tunability, from very low to very high redundancy depending on the channel condition.

2.1.3 Complexity

Complexity is another important issue, especially for real-time applications. Even though the computation power of processors increases year by year, computation consumption also increases correspondingly. For example, the current video coding standard (H.264/AVC) is about eight times more complex than the preceding ones, and also video frame rate and resolution are at least doubled compared to the past. On the other hand, due to the rapid growth of battery powered devices such as smart phones and tablets, less power consumption or less computational complexity are still of high interest.

By complexity, we mean the volume of computations needed for generating the descriptions and not the computation usually done in standard coding routine. In other words, the additional complexity compared to SDC is referred to as complexity here.

2.1.4 Capability to increase the number of descriptions

The capability to increase the number of descriptions is the other feature of an MDC scheme. Even though two-description coding is suitable for most cases and the channels are not usually lossy enough to warrant usage of four-description coding, there are reasons other than channel loss for using four descriptions or even higher number of descriptions. For example, in some cases such as P2P video streaming, in order to provide bandwidth and/or spatial/temporal resolution heterogeneity, we are interested in, say, eight or sixteen descriptions because an MDC with a higher number of descriptions provides better bit rate scalability. Furthermore, MDC is sometimes paired with Multipath Transmission (MPT) where each description is sent over a separate path. Some features of MPT such as total aggregated bandwidth, probability of outage, and delay variability are improved with higher number of paths and hence with higher number of descriptions, as much as system complexity allows. [25].

3 Review of video MDC schemes

In this paper, MDC schemes are grouped and analyzed based on their application domain. We present here the details of these MDC groups and their specific MDC methods. Our focus is on video, but the algorithms proposed and applied to images which can also be used for videos are considered as well.

The MDC process, i.e. generating and splitting the information into the descriptions, can be carried out in several domains. The domains are spatial, temporal, frequency, and compressed, in which the descriptions are generated by partitioning the pixels, the frames, the transformed data, and compressed data, respectively. Under each domain, we present a number of categories of MDCs. For each of these categories, we describe their objective and functionality, and we evaluate them based on the characteristics described in 2.1. There exist also MDC approaches working in multiple domains. The MDCs which are not based on splitting, namely unpartitioning methods, are discussed as a new category. All of these will be explained in this section.

3.1 Spatial-domain MDC

In this category, the MDC process is carried out in the pixel domain. The simplest approach is to divide the image/frame into multiple subimages and encode each one independently. Fig. 3 shows Polyphase Spatial Subsampling (PSS) of a frame to generate four subimages to achieve four-description coding. At the decoder side, if all descriptions are received, the subimages are merged and the image in full resolution is reconstructed. Otherwise, any missed description must be recovered using interpolation or similar techniques. However, in this basic approach there is no way to add redundancy and hence no provision for improving the side quality. The following solutions were therefore proposed to address this problem. Note that subsampling in spatial or temporal domain reduces the correlation among the data which in turn reduces the compressibility of the source. So subsampling indirectly imposes some amount of redundancy compared to SDC, but this type of redundancy is not much advantageous in side quality.

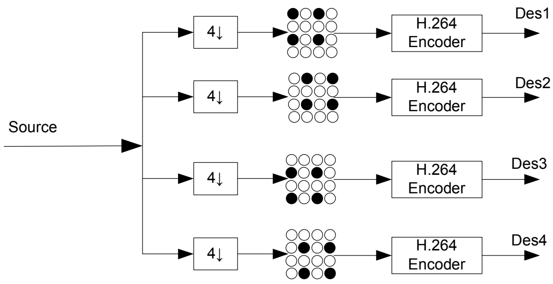

Fig. 3 Spatial polyphase subsampling MDC

3.1.1 Zero padding

Shirani *et al.* proposed zero padding in the DCT domain to control the redundancy for images [26] and for videos [27]. Zero padding in the DCT domain means oversampling in the spatial domain and this way an amount of redundancy is added to the image. At the preprocessing stage, the image is transformed by DCT and then a number of zeros are padded to the transformed image which is followed by inverse DCT (IDCT). The new larger size image is then partitioned and subimages are generated applying PSS. The direction and dimension of zero padding are discussed in [28] which shows that one-dimension (1D) zero padding (either vertical or horizontal direction in which the image is less predictable) is more efficient than 2D zero padding. Fig. 4 shows both of these options for zero padding.

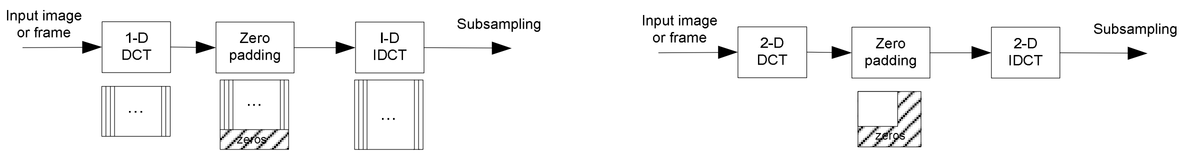

Fig. 4 Zero padding before subsampling, left: 1-D zero padding, right: 2-D zero padding

Standard compatibility: Zero padding is actually a way for subsampling and does not change the coding rule. Therefore, the streams can be decoded by any standard decoder.

Redundancy tunability: The amount of redundancy is determined by the number of padded zeros and, due to aliasing, it has an upper limit.

Complexity: Compared to SDC, some additional processes are needed for DCT transform, padding zeros and then inverse DCT and subsampling. Furthermore, due to the redundant samples added to the image, more samples must be coded which leads to additional complexity. This complexity is increased by the number of padded zeros.

Capability to increase the number of descriptions: This depends on the original image size and the display size. However, as long as the image and display sizes correspond, producing more than four number of descriptions using this algorithm leads to the subimages much less detailed than the original image, which might not be acceptable.

3.1.2 Duplicating least predictable data

A method called Least Predictable Vector MDC has been proposed in [28] that duplicates the data which is difficult to be estimated when it is lost. The image is divided into two subimages composed of even rows and odd rows; the even rows that cannot be estimated in the odd subimage are duplicated in both descriptions, the same concept is applied to the odd rows as well. The decoder is informed about the duplicated lines through side information.

Standard compatibility: Like zero padding, this method is standard compatible.

Redundancy tunability: The amount of redundancy is controlled by the number of duplicated lines.

Complexity: Due to the processing needed to find the lines with higher priority for duplication, this algorithm is complex.

Capability to increase the number of descriptions: The duplicated lines are totally redundant (unused) in the central decoder. In a high number of descriptions, the unused lines increase even more which reduces the MDC efficiency.

3.1.3 Filtering at the encoder and/or decoder

There are some methods in which advanced filtering has improved the quality of interpolation when some of the subimages are not available. That is, instead of simple bilinear or bicubic interpolation filters, a more complex filter is used to generate the full resolution picture, as shown in Fig. 5. The filters are optimally designed at the encoder and the filter coefficients are embedded into the descriptions. Based on the experiments performed in [29] for both high texture and smooth area pictures, this method outperforms the tested approaches: namely, bilinear interpolator, bicubic interpolator and zero padding. Except for some images and one-description decoding for which zero-padding is the best, this method gives the highest quality. This method can be combined with zero-padding to achieve even better results [30].

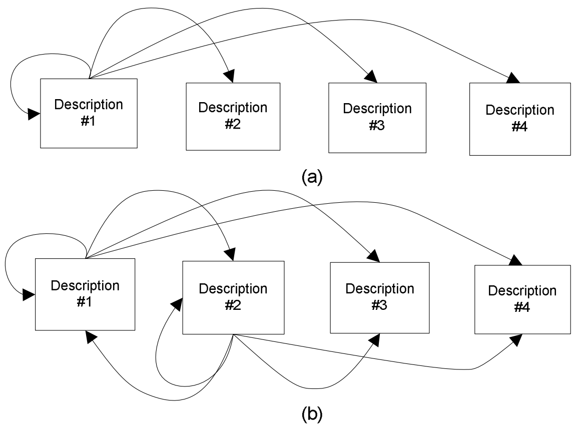

Fig. 5 Approximating the lost descriptions when (a) only Description#1 is received and (b) the first and second descriptions are received.

Standard compatibility: Since the filtering process is done as post processing, this algorithm is standard compatible.

Redundancy tunability: The first constraint on filter coefficients is estimation error, and considering the amount of redundancy as the second constraint, we would need complicated optimization algorithms to tune redundancy. Therefore, the redundancy is not usually tuned in this algorithm.

Complexity: Finding the filter coefficients needs computation and it increases by the number of descriptions.

Capability to increase the number of descriptions: Finding the filter coefficients and filtering itself is not an easy task to for a high number of descriptions. The filters become more complex (higher order) with the number of descriptions.

3.1.4 Sending the difference of descriptions

In this approach, the subimage of description2 is predicted from the subimage of description1 and the residual data is quantized and sent in description1, as shown in Fig. 6. This way, at the side decoder we have one subimage and the prediction signal of the other subimage which will help achieving a higher side quality; this was proposed for Intra frames and images [31-32].

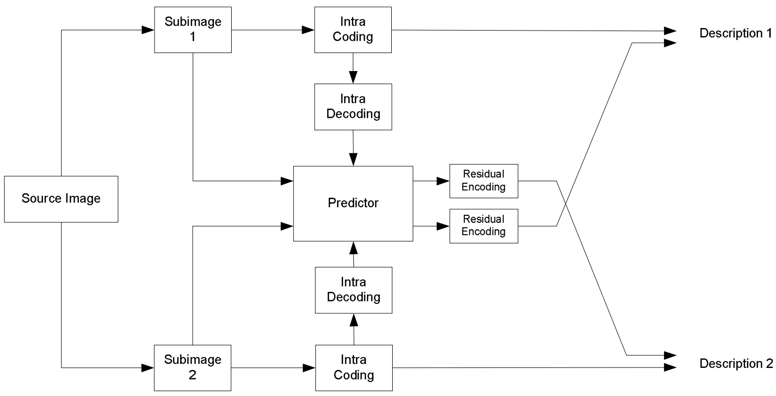

Fig. 6 Prediction of one subimage from the other subimage

Standard compatibility: Each description is composed of two streams, the subimage and the residual image. The decoder must be customized to separate and decode the two streams.

Redundancy tunability: By adjusting the quantization parameter used for residual encoding, redundancy can be tuned well.

Complexity: The residual encoding is an additional image encoding pass which leads to additional computational complexity. Note that compressing the residual image is much easier than compressing the image itself, since the residual image has much lower content.

Capability to increase the number of descriptions: In two-description-coding, there is one image and one residual image in each description. For four-description-coding, there is one image while we have (at least) three residual images. This makes the algorithm complicated and hence having four and higher number of descriptions using this algorithm is not recommended.

3.1.5 Nonlinear polyphase subsampling

The method uses linear ploNyphase subsampling help to split the image along rows and columns with the same sampling rate over the whole image. However, using nonlinear polyphase subsampling helps us to insert more samples of the region of interest (ROI) in each description and hence a higher subjective side quality is achieved [33], the same goal is achieved if first the image is nonlinearly transformed and then linear PSS is applied, as shown in Fig. 7 [34].

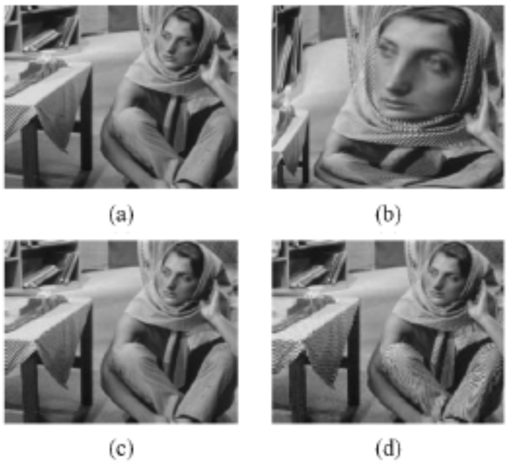

Fig. 7 MDC using nonlinear polyphase subsampling on BARBARA image (a) original image, (b) the image after nonlinear transformation, (c) multiple description, decoded from all four descriptions, (d) multiple description decoded from one description [34]

Standard compatibility: As discussed, this is a way for subsampling and does not change the decoding routine. Therefore, it is standard compatible.

Redundancy tunability: Using the transform, we can tune the redundancy to some extent, but we do not have much flexibility. Redundancy is allocated to the ROI.

Complexity: For this MDC, some image processing tasks must be performed on each frame. The tasks are finding ROI, finding a suitable transform and then applying it. So complexity is high.

Capability to increase the number of descriptions: Finding an efficient transform for a high number of descriptions is not an easy task. Even by nonlinear transform, more than four subimages leads to low quality descriptions.

3.1.6 Multi-rate MDC

In this method, a copy of description2 will be embedded into description1, but with a coarser quantization. In the spatial polyphase subsampling approach, this method is introduced in [35]. In this structure, as is shown in Fig. 8, after partitioning into subimages, the subimage of one description is encoded by Q_2 and multiplexed by the subimage of the other description. This algorithm is useful for coding of images and I-frames.

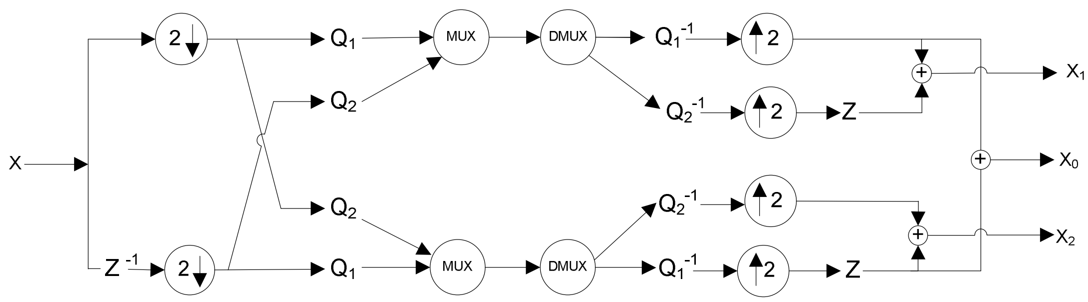

Fig. 8 Spatial Multi-rate MDC

Standard compatibility: Each description is composed of subimage1 and a lower quality of subimage2, therefore the decoder must be modified to handle the received data.

Redundancy tunability: Redundancy is controlled by the quantization parameter used for the low quality subimage.

Complexity: Additional complexity for coarse encoding is needed.

Capability to increase the number of descriptions: Due to the additional subimages having to be sent and processed at the receiver, having a high number of descriptions is difficult. Furthermore, the higher the number of descriptions, the less useful the added redundancy becomes and the lower the efficiency of this algorithm. The situation here is the same as the method explained in 3.1.4 .

3.2 Temporal domain MDC

In temporal domain MDC, the descriptions are generated by a process performed at the frame level; i.e., the granularity of this category of methods is a frame. A simple case is frame splitting between descriptions: odd frames in one description and even frames in the other description, as shown in Fig. 9. At the side decoder, the lost frames are substituted by frame freezing or estimated by concealment methods. Motion estimation/compensation is performed intra description, meaning even (odd) frames are predicted from the even (odd) frames [36].

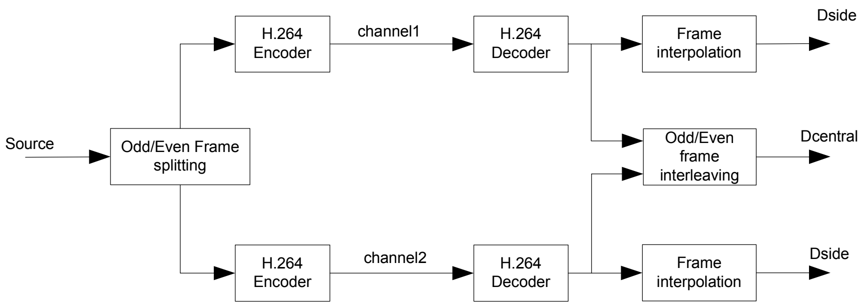

Fig. 9 Block diagram of temporal domain MDC

However, the efficiency of this approach depends on the inter-frame or temporal correlation. In other words, if the inter frame motion is not slow enough, the interpolated version of the lost frame in the side decoder is quite different than the original. Some solutions for this problem are proposed as follows. As we will see, except for the first solution, the others lead to additional frame insertions into the descriptions; these new frames can be exactly the same in all descriptions, be lower rate versions, or be intermediate virtual frames.

3.2.1 Duplicating the motion information in both descriptions

Motion information of the absent, say even, frames can be inserted in the description containing the odd frames. This way, for estimating the lost frames in the side decoder, the Motion Vectors (MV) are available and hence concealment can be performed more efficiently [37].

In [38], for each frame of description1, Motion Estimation (ME) against both neighbor frames in description2 (as references) is performed and the MVs are embedded in the corresponding reference. This approach is helpful in the reconstruction of the unavailable frames as mentioned earlier. In the test results, the proposed method, denoted as "MVpred", is studied against some other approaches, namely:

- "average", in which the frame is recovered by averaging the two neighbor frames in the received description;
- "inplaceMC", in which the decoder recover the lost frame using the MVs estimated as the 1/2 of those between two neighbor (previous and next) frames; and
- "MCinterp", which presents a motion-compensated interpolation between the previous and future correctly received frames in the received description; the motion information of the lost frame is estimated using a phase-correlation ME algorithm presented in [39].

Simulations carried out for sequences "garden" and "football" show that performance-wise "MVpred" is the first, "MCinterp" is the second, "inplaceMC" is the third and "average" is the worst.

Standard compatibility: The decoder must separate the second description's MVs and use them appropriately. The standard decoder does not do this.

Redundancy tunability: The only redundancy is due to the copied MVs, therefore the algorithm is not very flexible in this regard.

Complexity: When, as in [37], MVs of the descriptions are simply copied, it has no additional complexity. But when, as in [38], additional ME is performed, the algorithm becomes complex; since ME needs a lot of computations.

Capability to increase the number of descriptions: In four-description-coding, the MVs of three descriptions are copied into one description; this produces a high amount of redundancy which cannot be used efficiently. In other words, the efficiency of the algorithm degrades by the number of descriptions.

3.2.2 Duplicating the less recoverable data

Extra frame insertion is discussed in [40]; the frames which are hard to be estimated from the other frames, are duplicated in all descriptions. These frames are found by some pre-processing tasks at the encoder. The pattern of frame insertion is sent to the receiver as side information.

Standard compatibility: Since the sent and dropped frames are specified in the headers, the standard decoder can be used.

Redundancy tunability: Each frame can be either copied (full redundancy) or dropped (no redundancy); therefore we cannot precisely tune the amount of redundancy.

Complexity: The process needed to find the frame being duplicated or dropped makes this algorithm complex.

Capability to increase the number of descriptions: Producing four-description-coding by this algorithm means that some frames exist in all four descriptions with no usage when at least two descriptions are received. So, unused redundancy increases and makes the algorithm inefficient for a high number of descriptions.

3.2.3 Multi-rate MDC

In multi-rate MDC, the redundant data is the coarse quantized (lower-rate) representation of the primary data. The lower rate part might be Group of Pictures (GOP) [41], or one frame [42]. In [41], a GOP is encoded multiple times but in different rates, one for each description and hence the generated descriptions are unbalanced for each GOP. The lower rate descriptions are discarded in the side and central decoders. In each description, the prediction references are chosen from the frames of that description itself and for this reason, switching between descriptions is carried out when receiving an Intra frame. In this approach, the side quality may be low for one or multiple GOPs. In [42], each description contains the frames that alternatively are high quality (fine quantization) and low quality (coarse quantization).

Standard compatibility: Obviously this algorithm does not change the standard decoding routine.

Redundancy tunability: Since the quantization parameter is used for low quality frames, redundancy can be adjusted.

Complexity: The process needed to code odd (even) frames in even (odd) description, makes this algorithm complex to some extent. Note that the ME and mode decision of low quality frames are less complex than those of high quality frames.

Capability to increase the number of descriptions: For four-description coding, each side decoder has one high quality frame and three low quality frames which leads to flicker in the video. Furthermore, in high number of descriptions, we have multiple low quality frames are not useful and this reduces the efficiency of the algorithm. However, since the redundant frames are coarsely quantized, the situation is better compared to the temporal domain methods presented above.

3.2.4 Changing the local frame rate

A method was proposed in [43] which makes the inter frame motion smooth before splitting frames into odd and even frames by adding some new frames. This way, the local frame rate of high activity intervals is increased. The motion between the subsequent frames in the new sequence has been reduced which will be exploited in the side decoder for concealment of the lost frames. The pattern of frame adding will be sent to the decoder. The same idea but with the option of removing the frames of slow intervals is proposed in [44]. In other words, here, based on the motion activity of the sequence, some new frames might be inserted or some of the frames might be removed. This method provides better rate-distortion performance than what is presented in [36]. Fig. 10 shows the encoder and decoder structure of this method.

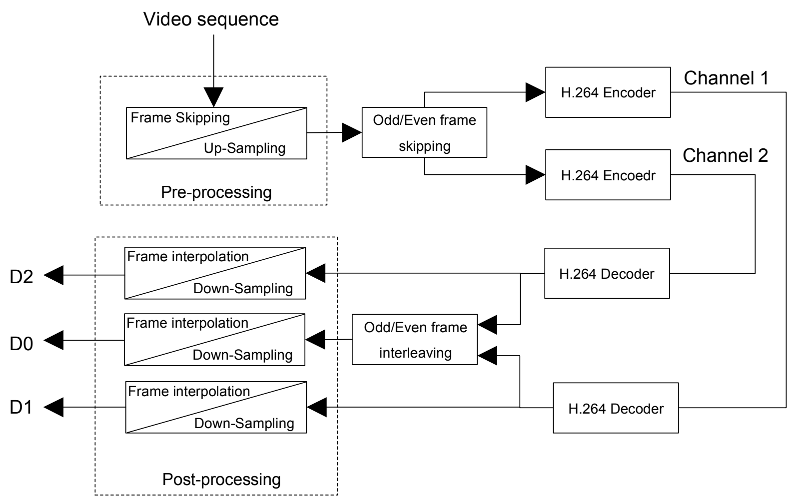

Fig. 10 Block diagram of adaptive fame skipping/Up-sampling MDC

Standard compatibility: The variable frame rate cannot be handled by the standard decoder and it must be customized for this purpose.

Redundancy tunability: Redundancy is controlled by the frame rate, though is not very flexible.

Complexity: The process needed to find whether to add or remove the frames makes this algorithm complex.

Capability to increase the number of descriptions: Adding or removing frames would be efficient for enough non-uniform video content such that the ensuing variable frame rate is justified. For four-description coding, the required non-uniformity is increased even more. So this algorithm is not recommended for a high number of descriptions.

3.3 Frequency Domain

In frequency domain MDC, several approaches exist. They are MDSQ, coefficients partitioning, MDTC, Multi-Rate MDC, and Mixed Layer MDC (MLMDC).

3.3.1 MDSQ

The concept of MDSQ is to use different quantization methods such that they refine each other at the central decoder. The simplest way of MDSQ is shifting the quantization intervals of each side encoder by half; as shown in Fig. 11.

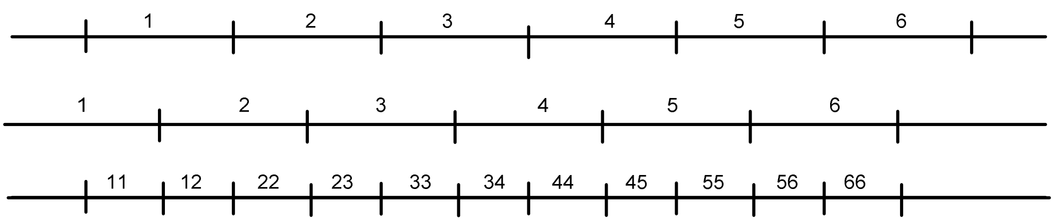

Fig. 11 The simplest MDSQ, upper two lines: side encoder (decoder) quantization (dequantization), lowest line: central decoder dequantization

Quantization can be performed for descriptions with an offset not necessarily ½ of the quantization interval. In H.264/AVC, there is an option (adaptive quantization offset) that can be used for this purpose. This idea has been presented in [45] where the descriptions are generated using different quantization offsets (Δ_{e1} and Δ_{e2}) as follows:

$$Z_{1,i,j} = \left\lfloor \frac{|W_{i,j}| + \beta + \Delta_{e1}}{\Delta} \right\rfloor . sign(W_{i,j})$$
$$Z_{2,i,j} = \left\lfloor \frac{|W_{i,j}| + \beta + \Delta_{e2}}{\Delta} \right\rfloor . sign(W_{i,j})$$

At the central decoder, the two reconstructed values are averaged, and at the side decoder simple dequantization is performed but with a shift as follows:

$$W'_{1,i,j} = (|Z_{1,i,j}|.\Delta - \Delta_{d1}).sign(Z_{1,i,j})$$

The offset can be fixed or changed adaptively. The adaptive offset is optimally obtained given coefficients' distribution and channel conditions. The simulations show better performance of this approach compared with fixed offset and also with the temporal subsampling method, particularly for high loss rates. The redundancy is decreased by increasing the difference between the offsets. The offset has an upper limit and thus, the redundancy is always beyond a certain value. For this reason, this method is not efficient for low enough loss rates.

In more complicated MDSQ, for each quantization level two indices are assigned, one for each description. Fig. 12 shows an example with 21 levels which are mapped to 8 indices in each description. When both descriptions are available, both indices are available and hence the quantization level is uniquely determined according to the index assignment table; otherwise, the side decoder must choose the level from the received single index and the possible values from the table. The index assignment table comes from an optimization where given the maximum rate and side distortion, the central distortion is obtained optimally. The first and most famous paper in this category is [46], which solves the problem for uniform sources. This work was modified for video sources and applied in H.264 encoder in [47].

Description 2 indices

Description 1 indices	1	2	3	4	5	6	7	8
1	1	2						
2	3	4	6					
3		5	7	8				
4			9	10	11			
5					12	14		
6					13	15	16	
7						17	18	20
8							19	21

Fig. 12 An example of MDSQ index assignment table.

Modified MDSQ (MMDSQ) [48], uses the two-stage quantization of Fig. 13. In the second stage, the quantization bins of the first stage are finely quantized again and the corresponding levels are inserted in each description, alternatively. At the side decoder, the second stage quantization data are not complete and discarded; but at the central decoder both stages are used and a high quality reconstruction is possible. The same approach is used and applied for predictive coding in [49].

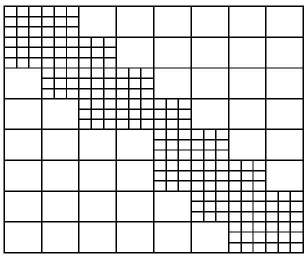

Fig. 13 Staggered index assignment, MMDSQ

MDSQ is optimized by information theorists even more [50-52]; however, those optimizations have not been applied in practice for image/video sources. For this reason, we have not covered them in this review.

Standard compatibility: Due to the special procedure of side and central decoding, the standard decoder cannot be used.

Redundancy tunability: As explained, redundancy can be tuned in each method of MDSQ. For maximum redundancy, the indices of both descriptions are the same and only the diagonal of the index assignment table are filled. For lower redundancy, the number of central levels is increased and more table cells are used. Minimum redundancy is achieved when all cells correspond to specific central levels.

Complexity: Having AN index assignment table, allows for the generation of descriptions without complicated processes.

Capability to increase the number of descriptions: The descriptions are generated by differently quantizing the DCT coefficients; this does not have enough degree of freedom for a high number of descriptions.

3.3.2 Coefficients partitioning

In this group, the transformed coefficients are partitioned or divided between descriptions. By the transformed coefficients or simply coefficients, we mean the outputs of the DCT transform. In this paper, we consider only DCT transform, and not wavelet transform, since all standard video codecs are DCT based.

In this method, some of the coefficients are duplicated in both descriptions and the others are split between them; splitting a coefficient means that it is unchanged in one description and it is made to be zero in the other description. In other words, the coefficients are treated as two groups, the coefficients of the first group are inserted in both descriptions and the coefficients of the other group are alternatively inserted. In the central decoder, all of the coefficients are available and the video is reconstructed with a high quality. In the side decoder, all of the coefficients of the first group and half of the coefficients of the second group (for two-description coding) are available, and a lower quality is achieved. The remaining problem is how to categorize the coefficients into these two groups. There are three approaches for this purpose:

In the first approach, coefficients larger than a threshold are duplicated in all descriptions and the others are alternated among descriptions. However, copies of the duplicated coefficients are totally redundant in central decoder. This redundancy is determined by the value of the distinction threshold

[53-54]. In [54], the threshold is obtained such that the descriptions are balanced with respect to rate and distortion. In another work presented in [55], the position and the sign of the coefficients of the description2 are inserted into description1, and vice versa. This way the unreceived coefficients can be approximated by interpolation exploiting the exponential decaying behavior of DCT coefficients of a block. This capability achieved at the cost of additional rate for sending the sign and position and also the syntax modification.

In the second approach, some of the early first coefficients (low frequency coefficients) are duplicated and the rest are forced to be zero in each DCT block of one description while they are remained unchanged in the other description [56]. The main advantage of this approach compared to the first one is the lower bit rate, at the same side distortion. The reduced bit rate is the result of the procedure used for entropy coding in image/video coding standards. In this approach, there exists a run of zeroes, while in the first approach the zeros are not consecutive. From the view point of bit rate saving, a run of zeros is more efficient than the same number of zeros distributed randomly.

As the third approach, coefficients partitioning can be carried out block wise; however, block splitting has poor side quality since a lost block must be estimated from its neighboring blocks with which it might have no correlation. One solution is overlapping block transform, such as Lapped Orthogonal Transform (LOT), in which the blocks have for example 50% overlapping, as is shown in Fig. 14. Due to the overlaps, some common data exist in the descriptions so that can then be used for estimation of the lost blocks. The results are much better than conventional non-overlapping block coding [57]. The same idea has been used in [58] with this difference that the blocks of description2 are predicted from those of description1 and the residuals are coded and sent in description1 as an additional part .This way, the side decoder1 has more information about description2, thus it provides higher side quality but at the cost of increased redundancy.

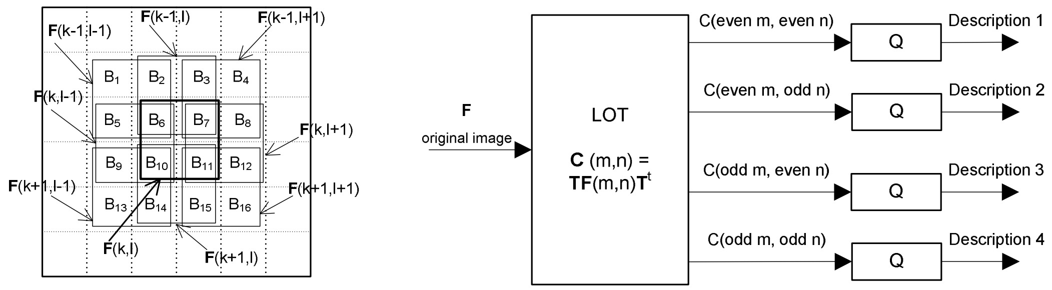

Fig. 14 (a) The overlapping blocks (b) decomposition into four descriptions by splitting the coefficient blocks

Standard compatibility: There is no need to change the decoding routine for side decoders, except for [55], [57] and [58]. However, merging the descriptions must be done inside the decoder (and not as post processing) and hence the central decoder must be modified appropriately.

Redundancy tunability: As discussed, some of the coefficients are duplicated and the others are split. By adjusting the number of duplicated coefficients, redundancy can be tuned well. The higher the number of duplicated coefficients, the higher the amount of added redundancy.

Complexity: The only work is to split the blocks or split the coefficients based on a predetermined pattern, so it does not need a lot of computations.

Capability to increase the number of descriptions: Due to low quality of each description even for two-description-coding, this algorithm is not used for a high number of descriptions. For higher number of descriptions, the number of nonzero coefficients in each description is further reduced.

3.3.3 MDTC

The main property of the DCT transform is decorrelation of the coefficients, so a coefficient cannot be estimated from the others in the same DCT block. Instead, the same-position DCT coefficient of neighboring blocks can be used. In cases of images with uniform texture, this approach may be helpful; however, this is not beneficial for video. The reason is that, in video coding, the DCT transform is applied on the residual signal and the data of residual signals are less correlated than the data of an image. MDTC provides a solution for this problem, as follows.

In MDTC, after the DCT transform, another transform called Pair-wise Correlating Transform (PCT) is applied to a pair of data, and the two correlated pieces of data are generated. The correlating transform would be helpful in estimation of the lost data from the received one and hence better side quality is achieved. Correlating, on the other hand, acts in opposition to DCT and leads to increased bit rate for transmission. This additional bit rate is the redundancy rate of this MDC method. As is shown in Fig. 15, the transform can be applied only to some of the coefficients to avoid unusable redundancy. In fact, which coefficients are suitable for pairing is also important. For example, pairing the coefficients with the same distribution parameter is not beneficial at all, this problem is addressed in [59].

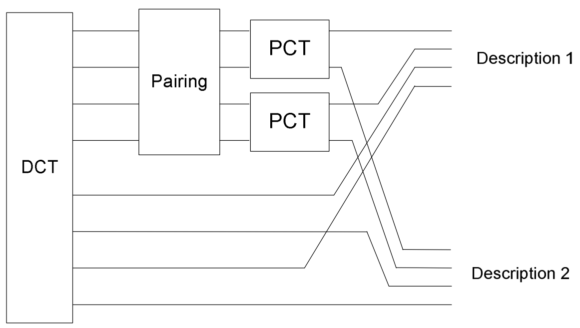

Fig. 15 Block diagram of MDTC

The progress track of this algorithm which is done by two research groups: Reibman *et al.* and Goyal *et al.* is explained as follows: the basic framework of the algorithm is introduced with a fixed transform matrix [60]; rate-distortion analysis and discussing nonorthogonal/orthogonal transform as the transform matrix are presented in [61], the algorithm was generalized for more than two descriptions with not necessarily balanced or independent channels in [62] for general Gaussian sources and for image sources in [63]. An arbitrary transform is also considered and optimally obtained with Redundancy-Rate-Distortion sense in [64]. Deep analysis and experimental results are concluded in [59] where it is shown that the MDTC in basic format has a good performance specially at low redundancy but its performance becomes increasingly poor at high loss rates. This is due to the fact that we always send one of two paired coefficients in each description and the side decoder cannot completely recover the original coefficient from the paired coefficient even at high redundancy. This problem is addressed in [65] in which the difference between the value of paired coefficient and the estimated coefficient is coded and incorporated into the descriptions, at the cost of even more redundancy rate. And finally, this method was used for video coding in [66]. A comprehensive study addressing Goyal's work on MDTC can be found in [67].

Standard compatibility: Due to the tasks needed for separating the coefficients and estimating the lost coefficients, the standard decoder cannot decode MDTC descriptions.

Redundancy tunability: Due to the splitting nature of this algorithm, the amount of redundancy is limited. However, as mentioned earlier, the algorithm presented in [65] does not have such a limitation.

Complexity: Pairing of the coefficients and generating the correlated coefficients make this algorithm complex to some extent.

Capability to increase the number of descriptions: How to extend this algorithm for a high number of descriptions is presented in Goyal's works. However, due to limited number of coefficients on which the algorithm is applied, we cannot generate efficiently say eight descriptions or higher for each video.

3.3.4 Multi-rate MDC

Multi-rate coding in the transform domain can be performed coefficient-wise or block-wise. As a case of the former, DCT coefficients are coded using different quantization parameters, where the redundant coefficients are quantized coarsely instead of being set to zero as in the DCT coefficient partitioning method. An example is [68] where in each description $1/n$ (n: number of descriptions) of coefficients are coded with fine quantization parameter and the rest are coded by a coarse quantizer as shown in Fig. 16.

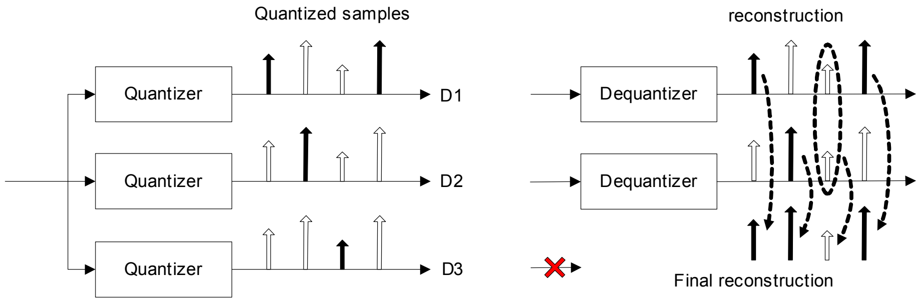

Fig. 16 Multi-Rate MDC, coefficient-wise[68] (a) encoder (b) decoder

In block-wise multi-rate MDC, when two-description coding, for a pair of blocks, each description contains a low rate version of one block and a high rate version of the other block. This idea has been proposed in [69] where two slice groups are defined: Slice Group A (SGA) and Slice Group B (SGB) which are selected based on Dispersed Slice Group shown in Fig. 17. SGA is finely quantized and SGB is coarsely quantized in description1, and the opposite is used for description2. Furthermore, the authors propose not to send any MV for low rate blocks and instead estimate the MVs of them from those of the high rate group and save the bit rate. The same idea with some minor modifications has been presented in [70].

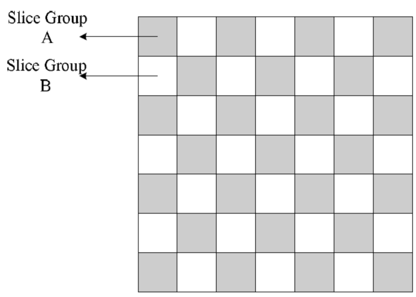

Fig. 17 Macroblock map for Dispersed Slice Group [69]

The low-rate MBs can be coded as redundant slices as in [71] as shown in Fig. 18. Redundant slice is another tool of the H.264/AVC standard, beneficial for video transmission in lossy conditions. For each of the primary slices, the standard allows inserting the redundant slices in the bitstream representing a lower-rate version of the primary slices which are used when the primary data are lost. For MDC, the redundant slices with the same role can be used. The work in [71] discusses the quantization parameter of the redundant slices in the rate-distortion optimization sense, taking into account the network condition and effect of drift. The quality is improved even more considering the role of each slice in error propagation within a GOP and optimizing redundancy allocation at slice level as in [72] which is published by the same research group.

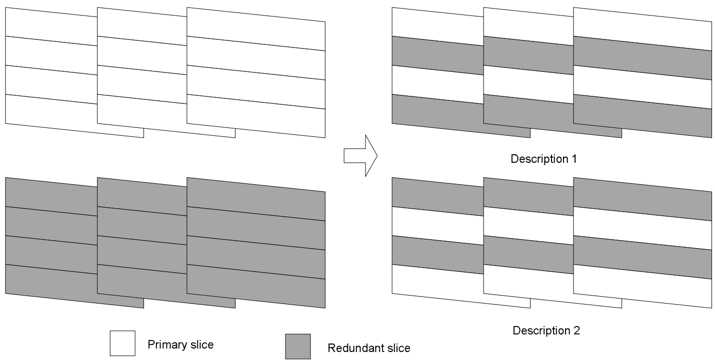

Fig. 18 Multiple description scheme using redundant slice

Standard compatibility: Each description is composed of finely and coarsely quantized DCT coefficients, so the standard decoder can be used when only one description is received. But the coarse quantized coefficients are discarded in cases that at least two descriptions are received, so, the decoder must be modified appropriately in such cases. However, when coding low rate blocks as redundant slices, an H.264/AVC decoder can be used as side and central decoder.

Redundancy tunability: By adjusting the quantization parameter, redundancy can be tuned easily.

Complexity: The additional process of encoding the low quality coefficients/blocks makes the algorithm complex to some extent. Note that MVs and mode types of fine data are usually used for coarse data; this in turn saves computational complexity.

Capability to increase the number of descriptions: As discussed, low quality data are discarded when high quality ones are available. The data discarded increases by the number of descriptions; so this algorithm loses its efficiency proportional to the number of descriptions.

3.3.5 MLMDC

In MLMDC, proposed by us in [73-74], each description is composed of the combination of the base layer and the enhancement layer of coarse granular SNR scalable coding. The coefficients of the base and enhancement layers are combined and the resulting combined coefficients and base coefficients are alternatively placed in the descriptions, as shown in Fig. 19(a). At the central decoder, the coefficients are decombined and similar to SVC layers are superimposed to achieve a higher two-layer central quality, as shown in Fig. 19(b). At the side decoder, the coefficient for which the base coefficient is not available is estimated from the received combined coefficient, as shown in Fig. 19(c).

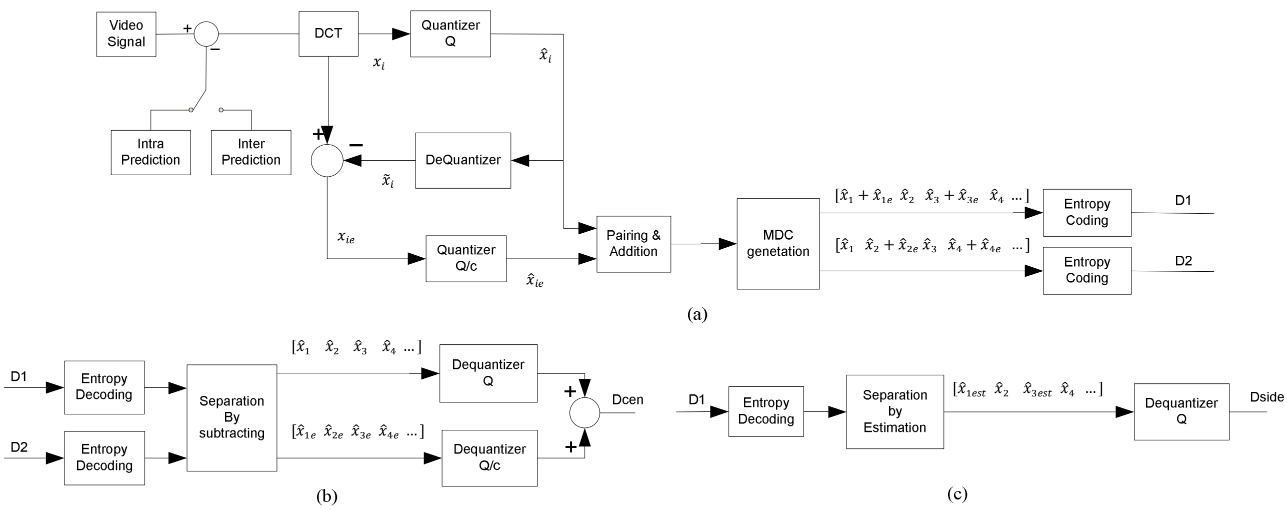

Fig. 19 MLMDC: (a) encoder, (b) central decoder and (c) side decoder

In other words, each description contains the base layer and one half of the enhancement layer, which are combined together. This combination of 1.5 layers is sent as one description. At the central decoder both descriptions are available and the base and enhancement layers are separated and superimposed on each other, thus we have two-layer video quality. At the side decoder, an estimator is used to separate the layers.

This method gives a side quality significantly higher than what is achieved by other tested DCT domain MDC approaches with almost the same central quality, leading to better side-central quality trade-off and hence higher average video quality when transmission in error prone environments. In addition, higher side quality reduces the descriptions mismatch and leads to less drift. For four descriptions, due to dividing the combined coefficients among the four descriptions, with at least two received descriptions, all base coefficients and hence the reference frame can be exactly reconstructed and thus MLMDC's performance is better for four number of descriptions.

Standard compatibility: Due to the estimation task for side decoding, this algorithm is not standard compatible, although for low values of c, the estimation can be bypassed and hence a standard decoder can be used as side decoder. Central decoding also needs some special tasks which do not exist in the standard decoder.

Redundancy tunability: The redundancy is tuned by c. However, the redundancy cannot be reduced arbitrarily since the base layer, mixed by or itself, is common in all descriptions.

Complexity: Generating the enhancement layer and mixing the layers make this algorithm complex.

Capability to increase the number of descriptions: Due to high amount of redundancy that exists in this algorithm, having a high number of descriptions is not efficient. In four-description-coding, the algorithm is capable to mitigate the drift and is useful, but higher number of descriptions is not recommended.

3.4 Compressed domain MDC

As the name implies, these MDC schemes are applied after the video has been encoded already. This can also be called packet domain MDC. The main idea here is to partition a layer into K segments which are then expanded to $N > K$ segments using FEC codes. Due to using channel codes, this scheme can be assumed to be a JSCC method. Then, these N segments are partitioned again into multiple descriptions which are sent independently. It is known that with any K out of these N segments, the K source segments are recoverable. The idea was inspired from [75] and then introduced and customized for multiple description coding in [76] and [77]. As is shown in Fig. 20, the video or image is encoded into a scalable bit stream where layer i is partitioned into i segments, and for N description coding, Reed-Solomon codes, $RS(N, i)$, are generated for these segments. The descriptions are composed by collecting one and only one segment (source or FEC code) from each layer. This configuration and the optimal bit rate for each layer are addressed in [77]. The algorithm proposed in [76] is basically the same but is performed byte-wise and is advantageous where we have fine granular layered coding.

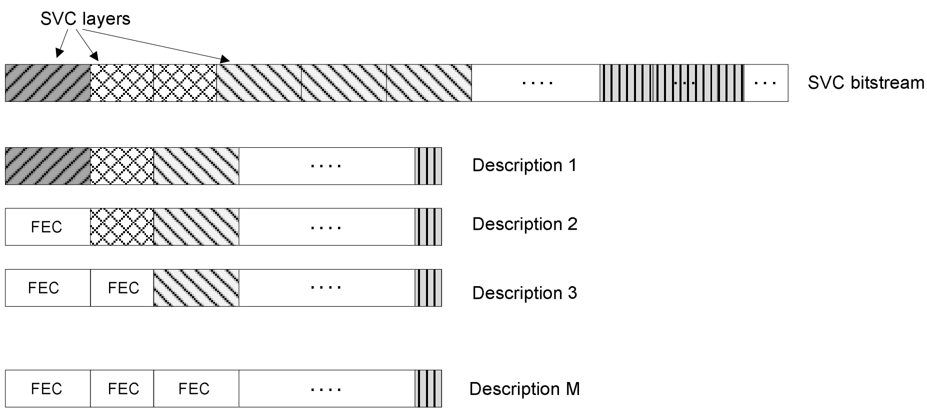

Fig. 20 MDC-FEC: with equal number of descriptions and SVC layers

Generally, the number of segments for each layer determines the trade-off between loss resiliency and redundancy. The lower the number of segments in a layer, the higher the redundancy of that layer and the more likely for it to be reconstructed at the receiver. This idea can be used for adapting MDC streams to channel condition as in [78]. This algorithm is essentially MDC transcoding and is useful when we do not have access to the original source and hence changing the rate of each layer is not possible. The number of segments in each layer is determined by an optimization approach while the original source is not changed or at most truncated in high packet loss conditions.

Standard compatibility: This algorithm is applied on the coded stream and does not affect the coding process. Therefore, it is standard compatible.

Redundancy tunability: The number of segments in each layer directly determines the amount of redundancy, so we can see that redundancy cannot continuously and precisely be tuned.

Complexity: Generation of FEC codes is a complex task.

Capability to increase the number of descriptions: Since the algorithm is applied on the coded stream, there is no limitation on the number of descriptions.

3.5 Hybrid domain MDC

In hybrid domain MDC, the MDC process is carried out in two of the above-mentioned domains. This is particularly used when we want higher number of descriptions where working in a single domain might not be efficient. For example, generating four descriptions using temporal domain MDC may not be reasonable, since each single description contains one quarter of the frames and the others must be estimated, leading to a poor side quality. As in [79], temporal and spatial partitioning can be used concurrently to generate the descriptions, or as in [80] and [81], a hybrid of spatial and frequency domain MDC (DCT partitioning and MDSQ, respectively) can be used, as shown in Fig. 21. The main advantage of working in two domains is preserving the inter description correlations for concealment of the lost information and hence creating a higher chance to have a higher side quality.

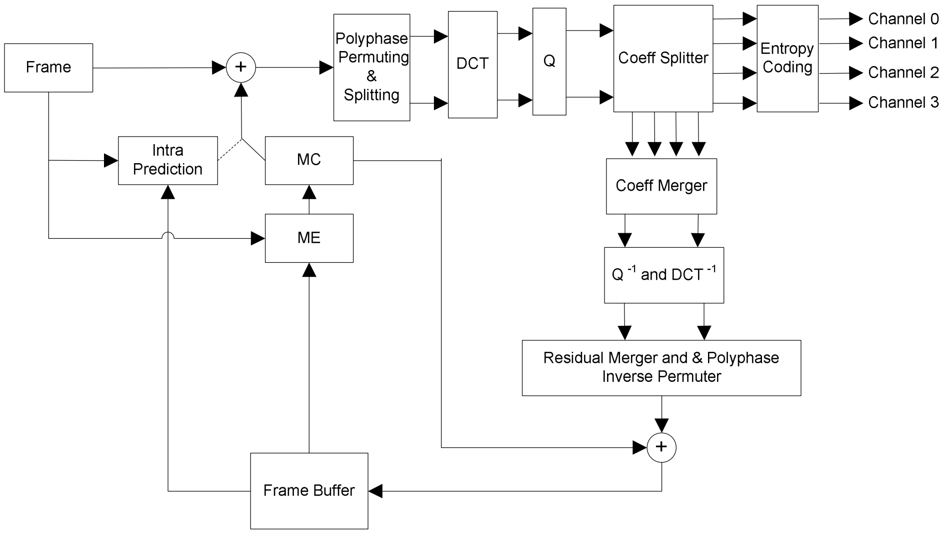

Fig. 21 Hybrid spatial-frequency MDC

Standard compatibility: The hybrid of spatial and temporal MDC is decodable by a standard decoder; since spatial and temporal partitioning/merging are carried out outside of the encoder/decoder engine. However, this is not the case for the spatial and frequency hybrid.

Redundancy tunability: In the hybrid of spatial and temporal MDC, redundancy cannot be controlled unless additional provisions are supplied. In frequency partitioning, some of the coefficients are duplicated which enables us to manage the amount of redundancy to some extent.

Complexity: Only partitioning is needed. The descriptions have lower spatial/temporal resolution and the total complexity is comparable to that of SDC.

Capability to increase the number of descriptions: Due to working in two domains, the algorithm can be used to generate a high number of descriptions easily.

3.6 Non-partitioning methods

In this category, two descriptions of the video are generated so that their reconstruction error is uncorrelated; these two noisy representations of the video are then merged to better reconstruct the original video. The more independent the distortions or errors can be made, the more accurately the original values can be estimated. This process is used for central decoding or when at least two descriptions are available. The inserted redundancy in this method is high but there is no provision to reduce it to the desired value, if the loss rate becomes low.

These different representation of the video source can be achieved using different coding parameters such as prediction reference, ME direction, and quantization parameters, as shown in [82]. Likewise in [83], all frames are included in both descriptions but frames in each description are coded with a different quantization parameter and different prediction references. The different representations of each frame in the central decoder are used for better reconstruction of the frame leading to a higher fidelity. The reconstructed frame is a weighted sum of the two frames in each description; the weights are obtained by estimation theory and sent to the decoder as side information. To reduce the bit-rate cost of these weights, they are quantized and entropy encoded.

To make the coding noise signals independent, one approach is to change the block boundaries in each frame: the video codecs are block based and changing the boundaries will lead to different position of pixels in the block and hence different coding representation. This idea was proposed in [84]. Different coding noise can be obtained using different transforms as shown in [85] where DCT is used for one description and no transform used for the second description; this is not necessarily optimum and one can find another pair of more efficient transforms.

In another method, the encoder configuration and settings are the same for both descriptions, but the input frames of the side decoders are different [86]. In this scheme, each original video frame is encoded with the encoder to get the first description. Then the error residuum of each coded frame of the first description is then added up to the corresponding original frame and the resulting new frame is encoded by the same encoder to generate the second description. At the central decoder, the two received descriptions of a frame are averaged which leads to a frame with lower quantization noise or higher quality. This is shown in Fig. 22.

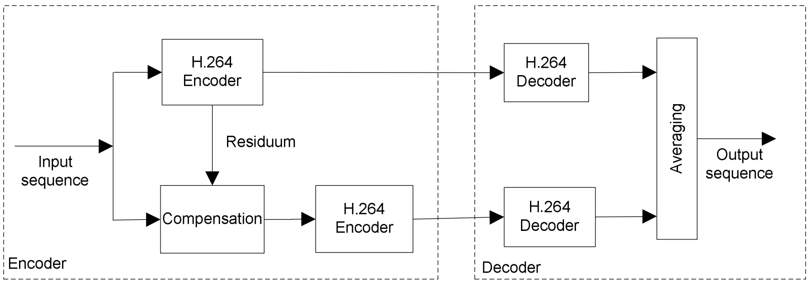

Fig. 22 The encoder and decoder of MDC presented in [86]

Standard compatibility: As discussed, except for [85] which uses different a transform in each description, other algorithms do not change the encoding process and are standard compatible.

Redundancy tunability: The difference of the descriptions is the quantization error and does not provide enough degree of freedom for redundancy tuning.

Complexity: Each description is coded as complex as SDC. Furthermore, some tasks such as finding estimation parameters are added. That is, for two-description-coding, the complexity is more than twice of SDC's complexity.

Capability to increase the number of descriptions: As mentioned, these MDCs are based on making different the quantization error in each description. The quantization error cannot be different as much as that required for four descriptions, for example.

4 Drift

In the previous sections, *drift* was brought up a number of times while discussing various MDC approaches. In this section, we will study this phenomenon and solutions to mitigate it. Drift originates from reference mismatch at the encoder and the decoder. Video compression standards make use of predictive coding for higher efficiency but this also makes a frame's quality dependent on its reference frame. When a reference frame is not correctly reconstructed at the decoder, it leads to the noisy reconstruction of all other frames which are predicted from it. Some of the erroneous frames are in turn used as the reference for other frames and so error propagation occurs. This is known as drift and results in quality degradation of frames until the end frame of the GOP.

As mentioned earlier, an excellent review of drift avoiding techniques are provided in [21], but since drift is an important issue in video MDC and in order to make our paper self-contained, the proposed solutions to mitigate or solve this problem in MDC are presented in this section, with our point of view. In the following, based on the number of prediction loops, the drift solutions are categorized and shortly explained.

4.1 Single-prediction-loop

A simple yet efficient method to mitigate drift is to formulate and incorporate the effect of error propagation when designing MDCs. The result is to make the first few frames in the GOP more similar in the descriptions, since the error in the earlier frames is more destructive than the later frames, and then to decrease the redundancy gradually as moving toward the last frames [71] [87]. The slice-wise or block-wise redundancy allocation is also proposed in [72], [88], respectively.

Another method is prediction based on a virtual frame, as proposed in [89]. In this paper; instead of previous frame itself, an intermediate frame based on the previous frame and leaky prediction is generated and used as prediction reference. The more different this virtual frame is from the current frame, the lower the sensitivity of the reconstruction quality of the current frame on the fidelity of the reference frame. This is achieved at the cost of higher bit rate; i.e., we move closer to intra coding. In this approach, the compression efficiency and error resiliency are something in between intra coding and inter coding.

4.2 Two-prediction-loop

One solution is to have a specific prediction loop for each description; e.g., for two-description coding, there exist two prediction loops at the encoder and the reference picture for each description is reconstructed separately at the corresponding side loop [90]. For the DCT domain MDCs, having different reference pictures leads to different DCT coefficient sets in each description which are hardly mutually refinable. Therefore, this algorithm is efficient for other MDCs such as spatial and temporal domain. Furthermore, in this approach, when a side reference, say description1, is corrupted, even though description2 with its own reference can be decoded, description1 still suffers from the drift until the end of the GOP, similar to single reference approaches. Therefore, this is suitable for the applications in which for a while one description is available and the other is not, for example when MDC is used for rate scalability. However, in a packet loss scenario where each frame might by decoded either by the side or the central decoder, two-prediction loop is not efficient.

4.3 Three-prediction-loop

In three-prediction-loop approach, corresponding to the possible decoder reconstructions, three prediction loops are used at the encoder; one for central decoding and two for side decoding (the case of missing both descriptions is not taken into account). The idea is to code and send a signal to compensate the mismatch between the side and central references. In the following, the existing methods are explained.

The idea of using three prediction loops is introduced in [66] where two three-loop algorithms are proposed, namely algorithm1, algorithm2. In these algorithms, the difference between the residual signals when using central reference and when using side references, is sent to decoder.

Fig. 23 shows the encoder structure of algorithm1, for the first loop only, the loop corresponding to the second description is not drawn due to space limitation.

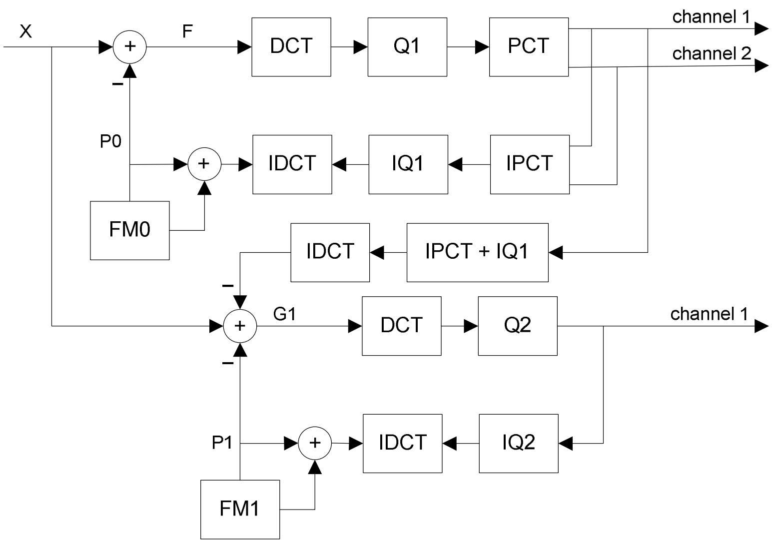

Fig. 23 Three-prediction-loop method, algorithm1 for drift avoiding

At the encoder, in addition to central reference (FM0), each side decoder is simulated and the side references are built (FM1 and FM2 for description1 and description2, respectively). The residual signal based on the side references is produced which is then subtracted by the residual signal of the main stream in description1 and the result (G_1) is sent as side information of description1. If description2 is lost, description1 with it's side information are able to reconstruct the frame, correctly and without error propagation. The central reference and the other side reference can be reconstructed, as well. In this procedure, there will be almost no drift, but at the cost of some redundancy rate. This redundancy is controlled by Q_2, the quantization parameter of the mismatch encoder. This algorithm is optimized for redundancy-rate-distortion in [91], and also is utilized with odd/even frame splitting MDC in [92] where in order to reduce the mismatch signal, the central predictor is made to be the average of the side predictors. This way, the side and central predictors are more similar and hence there is less mismatch between the residual signals created based on them, so the side information is reduced.

In algorithm2, the aim is to decrease the redundancy rate. The idea is the same as algorithm1 but instead of sending the whole side information, it is itself coded by MDC and only one of the descriptions is sent. Therefore, compared to algorithm1, we have only partial mismatch control and in return less redundancy rate. Algorithm2 provides better rate-distortion performance compared to the algorithm1, the results illustrated in [66] show.

Another method is an architecture which is basically similar to the algorithm1 but with the difference that instead of the original signal (signal X in Fig. 23), its reconstruction by the central decoder is used in the side prediction loops [93]. This way, in the case of packet loss, the decoder central references are made more similar to the encoder central reference. The achievement is a gain up to 0.4 dB in PSNR, the simulations provided in [94] show that. This algorithm with reduced computational complexity has been presented in [95] too.

In the algorithms explained above, the side information was totally redundant when both descriptions are available. The algorithms presented in [96] and [97] try to make them beneficial. The authors of [96] propose to combine the difference of the reference picture decoded by the central decoder and the corresponding side decoder with residual signal and send it as one description. This algorithm shows better performance compared to the structure of algorithm1, the rate-distortion curves show. However, due to the mentioned combination performed at the encoder, this algorithm has the limitation of two-prediction-loop method; that it, it is suitable for the applications in which for a while one description is available and the other is not.

The combination of the method proposed in [96] and algorithm1 is presented in [97]. While it has not the above described restriction of [96], the proposed structure is able to utilize the side information not only for avoiding drift but also for central decoding higher quality. The performance curves are about 0.25 dB higher than that achieved by algorithm1 at low loss rates. This PSNR gain is however decreased for higher loss rates.

5 Summary

Based on the above descriptions and criteria, the MDCs and their functionalities are summarized in Table 1. In the table, the standard compatibility property of methods is indicated by “Yes” or “No”. For some MDCs, side decoding can be done with a standard decoder while it is not the case for central decoding; these states are also indicated in Table 1. For example SD:Yes, means that only side decoder is standard compatible, and for central decoding the decoder must be modified.

In Table 1, there are three levels of “Low”, “Moderate”, and “High”, indicating the capability of the algorithm for adapting its redundancy. In cases where redundancy can be continuously changed from minimum to maximum, they are indicated by “High”; “Low” indicates that the method has no provision for redundancy adaptation. The MDCs in which the redundancy are controllable but not from minimum to maximum or not continuously, are indicated by “Moderate”.

The next column of the table is complexity. The less complex algorithms are indicated by “Low”, and algorithms for which the complexity is scaled by the number of descriptions are indicated by “High”. The rest of the algorithms which fall somewhere in between are designated as “Moderate”.

The second-last column specifies the capability of each algorithm for higher number of descriptions. In the table, "High" indicates that creating a high number of descriptions, say more than four, using this algorithm is easily executable, whereas "Low" indicates that having more than two descriptions is not efficient. The other MDCs by which up to four descriptions is practical, are indicated by "Moderate".

Table 1 MDC methods comparison

MDC domain	MDC approach	References	Standard compatibility	Redundancy tunability	Complexity	Capability to increase the number of descriptions	Summary
Spatial-domain MDC	Zero padding	[26-28]	Yes	High	High	Moderate	Zero padding provides an efficient way for redundancy adding.
	Duplicating least predictable data	[28]	Yes	High	High	Low	Having available the least predictable data in each description leads to higher side quality, at the cost of additional rate.
	Filtering at the encoder and/or decoder	[29]	Yes	Low	High	Moderate	The unreceived description is estimated from the received description using the filters designed at the encoder.
	Sending the difference of descriptions	[31-32]	No	High	Moderate	Low	The signal of description 2(1) is subtracted from description 1 (2) and the residual signal is quantized and send together with description 1 (2), leading to higher side quality at the cost of additional rate.
	Nonlinear polyphase subsampling	[33-34]	Yes	Low	High	Moderate	The redundancy is assigned to the more important regions leading to more efficient compression at the cost of incurred computational complexity.
	Multi-rate MDC	[35]	No	High	Moderate	Low	A coarse quantized version of description 2 (1) is included in description 1 (2).
Tempo ral domai n MDC	Duplicating the motion information in both descriptions	[37-38]	No	Moderate	[37]: Low [38]: High	Low	MVs help to conceal the lost description more efficiently, with the cost of additional rate for sending MVs.

	Duplicating the less recoverable data	[40]	Yes	Moderate	High	Low	The frames that can hardly be estimated are simply copied into all descriptions.
	Multi-rate MDC	[41-42]	Yes	High	Moderate	Moderate	A coarser quantized version of the frames of description 1 (2) are copied into description 2 (1).
	Changing the local frame rate	[43-44]	No	Moderate	High	Low	Making the inter-frame motion smooth helps estimation of the lost frames more efficiently in the side decoder, at cost of additional complexity and rate.
Frequency Domain	MDSQ	[45] [46] [47] [48] [49]	No	High	Low	Low	The output of the scalar quantizer is not the same at the descriptions; the levels are designed to be complementary at the central decoder, in order to reconstruct the original DCT coefficients will less quantization error.
	Coefficients partitioning	[53-55] [56-58]	SD: Yes	High	Low	Low	The most important coefficients are duplicated and the rest are split among descriptions.
	MDTC	[59-61] [62-64] [65-67]	No	Moderate (Except for [65] which is High)	Moderate	Moderate	DCT coefficients, instead of partitioning directly, are again transformed by a correlating transform and the new coefficients which now are correlated are split between descriptions. This correlation helps estimating the unavailable coefficients of the lost description.
	Multi-rate MDC	[68-72]	No	High	Moderate	Moderate	The coarser quantized of DCT coefficients are included into the descriptions leading to better side fidelity compared to coefficients partitioning, at the cost of additional rate.
	MLMDC	[74]	No	Moderate	High	Moderate	Each description composed of combination of base and enhancement layers; they are decoded like as SVC two-layer decoding at central decoder. The layers cannot be separated at the side decoder and hence estimation is used.
Compressed domain MDC	MDC-FEC	[76-78]	Yes	Moderate	High	High	The compressed layered data after addition of FEC codes are divided into multiple descriptions, the higher the number of received descriptions, the more layers of video are decodable.
Hybrid domain MDC	Temporal+Spatial	[79]	Yes	Low	Low	High	Hybrid domain used in order to generate more than two descriptions more efficiently than the single domain approaches. After for

	Spatial+Frequency	[80-81]	No	Moderate	Low	High	
Non-partitioning methods	Generation two different presentation of video	[82-86]	Yes (except for [85])	Low	High	Low	Two descriptions are generated so that their reconstruction error is uncorrelated; these two noisy representations of the video are then merged to better reconstruct the original video at the central decoder.

6 Discussion

Our goal in this paper was to review existing MDC schemes and provide a taxonomy and analysis to aid practitioners and researchers for better understanding and selection of the most suitable MDC scheme for error-resilient video delivery over lossy networks. We saw that MDC is one of a number solutions for error-resiliency, and that for cases with moderate to high packet loss rates, such as wireless video streaming, or for live applications where packet retransmission or buffering is unacceptable, such as live video conferencing, MDC is a better solution to counter video packet loss compared to existing methods. We also saw that MDC can be coupled with P2P transport to support video transmission to heterogeneous receivers. Finally, a comprehensive taxonomy and discussion of MDC techniques was presented in section 3 and summarized in table 1.

Based on what was presented in the paper, we can suggest a 5-question process to choose an MDC scheme for a specific video delivery application, as shown in Fig. 24.

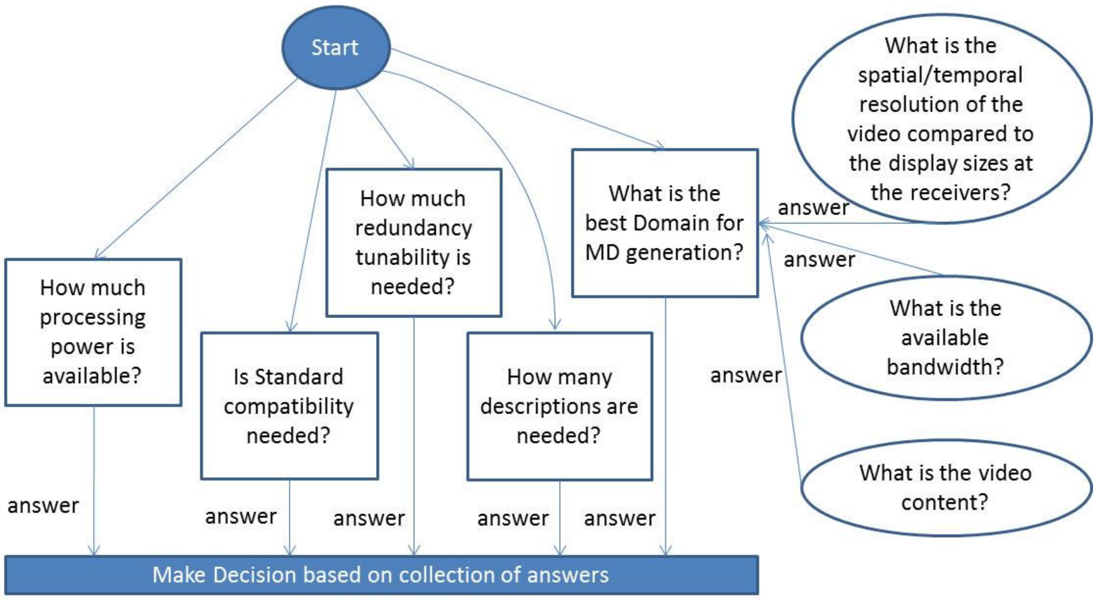

Fig. 24 Proposed decision process for choosing an appropriate MDC method for a specific video application.

The first question to be answered is whether standard compatibility is needed or not? For standard compatible MDCs, there is no need to change the decoder, and the default SDC decoder can be used. To merge the descriptions when more than one description is received, some post processing tasks are needed which are done outside the decoder engine.

The second question is that, based on the specific application, how many descriptions do we need? To combat packet loss in the real world, most of the time two-description coding is sufficient. However, if MDC is also used for scalability to heterogeneous receivers, four, eight, sixteen or even more descriptions might be needed. As seen in Table 1, compressed and hybrid domain MDCs easily provide higher number of descriptions. For more than eight descriptions, compressed domain MDCs are preferred.

The third question is how much redundancy tunability is needed, which depends on channel behavior. As discussed, redundancy is not needed to be high for low loss channels but needs to be high for high loss channels. Since in the real world, channels dynamically vary, MDCs with high redundancy tunability are of more interest. Note that MDCs are resilient against variable loss rates, and this can be used to get the best performance; so as not to waste the available bandwidth.

The fourth question is what the best suited domain for MD generation is. Generally, there is no straightforward solution for this, and the answer needs a qualitative approach. The situation is similar to choosing the best SVC modality among spatial, temporal and SNR scalable modalities. As an example for SVC, if I have 3 receivers with different bandwidth and I need to create 3 layers, should my base layer consists of the video at a lower screen size (spatial), lower frames per second (temporal), lower SNR quality, or a combination of the three? This is a very difficult problem to solve methodically and there is no general rule and mostly it is case dependent. A similar scenario applies when choosing and MDC domain. For example, when considering spatial domain MDC, we encounter the following questions:

a) What is the original video size compared to the display sizes at the receivers? If the image size is itself small, spatial subsampling in order to generate the descriptions makes the situation worse and is not recommended.
b) What is the available bandwidth? The available bandwidth determines the compression degree. If we have a strong compression, the resolution of the image at the receivers is less important and so spatial MDC could be a candidate even for low resolution original sources.
c) What is the video content? If the video content is such that for example temporal resolution is more stringent than spatial resolution, spatial MDC is preferred, and vice versa. Whether spatial MDC provides subjectively satisfactory quality depends on the video content and what the viewers are looking for.

Finally, as the fifth question, we must check whether the encoder has any processing power constraints. If it does, for example as in mobile encoders, the complexity issue must also be taken into account and a less complex scheme must be selected.

Based on the above 5 questions, in practice we can significantly narrow down the choice of the specific MDC technique for a given video delivery application.

7 Conclusions

In this paper we presented an overview and survey on video MDC. We showed that for video transmission over channels with moderate to high and fluctuating loss rate, MDC is a good solution due to the redundancy inherent in MDC streams. We also saw that because of this redundancy, MDC in not a good choice in very low or low loss rate environments.

Based on the domain at which the MDC separation is done, we categorized existing MDC methods into six groups and the corresponding papers were also cited in Table 1. In the table, the performance of MDCs of each group with respect to four criteria, namely standard compatibility, redundancy tunability, complexity, and extendibility to n-description coding were compared. We also presented a 5-question process that can be used with the table to help scientists and practitioners choose the best suited MDC for their specific application.

For future work, there are a number of directions. For example, to solve the drift problem, there can be some solutions which mitigate drift without using two or three prediction loops as discussed in section 4. The allocated redundancy determines the mismatch of the side and central pictures. The higher the allocated redundancy, the less mismatch the side and central pictures have and hence the smaller error propagation. Therefore, at higher loss rates, the need for higher redundancy is increased. Moreover, the influence of early frames of the GOP in error propagation is more than that of end frames. So it is not reasonable to allocate the same amount of redundancy to all frames of the GOP for all channel conditions. Although there are some works in this regard [87] [88], they are for a specific MDC and not applicable to other MDC methods. The issue of optimal redundancy allocation is therefore understudied and requires more research.

There are also some works discussing ME, mode decision, and rate control in SDC. However, their results are not applicable to MDC. For example, the best MVs for SDC are not necessarily the best MVs for MDC. Taking into account the effect of side and central mismatch, packet loss rate, and error propagation of MDC, ME and mode decision in the codec's rate control become topics of interest which are less studied and need further research as well.

References

1. Moo Young K, Kleijn WB (2006) Comparative rate-distortion performance of multiple description coding for real-time audiovisual communication over the Internet. Communications, IEEE Transactions on 54 (4):625-636. doi:10.1109/tcomm.2006.873071
2. Bernardini R, Durigon M, Rinaldo R, Vitali A Comparison Between Multiple Description And Single Description Video Coding With Forward Error Correction. In: Multimedia Signal Processing, 2005 IEEE 7th Workshop on, Oct. 30 2005-Nov. 2 2005 2005. pp 1-4. doi:10.1109/MMSP.2005.248546
3. Baccaglini E, Tillo T, Olmo G (2010) A Comparison Between ULP and MDC With Many Descriptions for Image Transmission. Signal Processing Letters, IEEE 17 (1):75-78. doi:10.1109/lsp.2009.2032976
4. Dán G, Fodor V, Karlsson G (2005) Are Multiple Descriptions Better Than One? In: Boutaba R, Almeroth K, Puigjaner R, Shen S, Black J (eds) Networking 2005. Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems, vol 3462. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp 684-696. doi:10.1007/11422778_55
5. Lee Y-C, Kim J, Altunbasak Y, Mersereau RM (2003) Layered coded vs. multiple description coded video over error-prone networks. Signal Processing: Image Communication 18 (5):337-356
6. Chakareski J, Han S, Girod B (2005) Layered coding vs. multiple descriptions for video streaming over multiple paths. Multimedia Systems 10 (4):275-285. doi:10.1007/s00530-004-0162-3
7. Reibman AR, Yao W, Xiaoxin Q, Zhimei J, Chawla K Transmission of multiple description and layered video over an EGPRS wireless network. In: Image Processing, 2000. Proceedings. 2000 International Conference on, 10-13 Sept. 2000 2000. pp 136-139 vol.132. doi:10.1109/icip.2000.899246
8. Singh R, Ortega A, Perret L, W. J (2000) Comparison of multiple description coding and layered coding based on network simulations. Paper presented at the SPIE Conf. Visual Comm. Image Processing.,
9. Javadtalab A, Omidyeganeh M, Shirmohammandi S, Hosseini M A rate control algorithm for x264 high definition video conferencing. In: Multimedia and Expo (ICME), 2011 IEEE International Conference on, 11-15 July 2011 2011. pp 1-6. doi:10.1109/icme.2011.6012225

10. Greco C, Cagnazzo M, Pesquet-Popescu B (2012) Low-Latency Video Streaming With Congestion Control in Mobile Ad-Hoc Networks. Multimedia, IEEE Transactions on 14 (4):1337-1350. doi:10.1109/tmm.2012.2195480
11. Yiting L, Gibson JD (2011) Routing-Aware Multiple Description Video Coding Over Mobile Ad-Hoc Networks. Multimedia, IEEE Transactions on 13 (1):132-142. doi:10.1109/tmm.2010.2089504
12. Min Q, Zimmermann R (2010) An Adaptive Strategy for Mobile Ad Hoc Media Streaming. Multimedia, IEEE Transactions on 12 (4):317-329. doi:10.1109/tmm.2010.2046275
13. Jiann-Jone C, Shih-Chieh L, Ching-Hua C, Chen-Hsiang S, Jyun-Jie J, Chi-Chun L (2012) A Multiple Description Video Codec With Adaptive Residual Distributed Coding. Circuits and Systems for Video Technology, IEEE Transactions on 22 (5):754-768. doi:10.1109/tcsvt.2011.2179459
14. Padmanabhan VN, Wang HJ, Chou PA Resilient peer-to-peer streaming. In: Network Protocols, 2003. Proceedings. 11th IEEE International Conference on, 4-7 Nov. 2003 2003. pp 16-27. doi:10.1109/icnp.2003.1249753
15. Wu JC, Peng KJ, Lu MT, Lin CK, Cheng YH, Huang P, Yao J, Chen HH (2006) HotStreaming: Enabling scalable and quality IPTV services. Paper presented at the Proc. IPTV Workshop Conjunction 15th Int. World Wide Web Conf.,
16. Castro M, Druschel P, Kermarrec AM, Nandi A, Rowstron A, Singh A (2003) SplitStream: High-bandwidth multicast in cooperative environments. Paper presented at the ACM SOSP'03, New York, USA.,
17. Padmanabhan VN, Wang HJ, Chou PA, Sripanidkulchai K (2002) Distributing streaming media content using cooperative networking. Paper presented at the ACM/IEEE NOSSDAV, Miami, FL, USA.,
18. Magnetto A, Gaeta R, Grangetto M, Sereno M (2010) TURINstream: A Totally pUsh, Robust, and effIcieNt P2P Video Streaming Architecture. Multimedia, IEEE Transactions on 12 (8):901-914. doi:10.1109/tmm.2010.2077623
19. Kobayashi M, Nakayama H, Ansari N, Kato N (2009) Robust and Efficient Stream Delivery for Application Layer Multicasting in Heterogeneous Networks. Multimedia, IEEE Transactions on 11 (1):166-176. doi:10.1109/tmm.2008.2008933
20. Goyal VK (2001) Multiple description coding: compression meets the network. Signal Processing Magazine, IEEE 18 (5):74-93. doi:10.1109/79.952806
21. Wang Y, Reibman AR, Lin S (2005) Multiple Description Coding for Video Delivery. Proceedings of the IEEE 93 (1):57-70. doi:10.1109/jproc.2004.839618
22. Wolf J, Wyner A, Ziv J (1980) Source coding for multiple descriptions. Bell Syst Tech J 59 (8):1417-1426
23. Witsenhausen HS, Wyner AD (1980.) Source coding for multiple descriptions II: A binary source. Bell Labs.,Tech. Rep. TM-80-1217,
24. Ozarow L (1980) On a source coding problem with two channels and three receivers. Bell Syst Tech J 59:1909-1921
25. Apostolopoulos JG, Trott MD (2004) Path diversity for enhanced media streaming. Communications Magazine, IEEE 42 (8):80-87. doi:10.1109/mcom.2004.1321395

26. Shirani S, Gallant M, Kossentini F Multiple description image coding using pre- and post-processing. In: Information Technology: Coding and Computing, 2001. Proceedings. International Conference on, Apr 2001 2001. pp 35-39. doi:10.1109/itcc.2001.918761
27. Gallant M, Shirani S, Kossentini F Standard-compliant multiple description video coding. In: Image Processing, 2001. Proceedings. 2001 International Conference on, 2001 2001. pp 946-949 vol.941. doi:10.1109/icip.2001.959203
28. Tillo T, Olmo G (2007) Data-Dependent Pre- and Postprocessing Multiple Description Coding of Images. Image Processing, IEEE Transactions on 16 (5):1269-1280. doi:10.1109/tip.2007.891799
29. Yapıcı Y, Demir B, Ertürk S, Urhan O (2008) Down-sampling based multiple description image coding using optimal filtering. Journal of Electronic Imaging, SPIE 17 (3)
30. Ates C, Urgun Y, Demir B, Urhan O, Erturk S Polyphase downsampling based multiple description image coding using optimal filtering with flexible redundancy insertion. In: Signals and Electronic Systems, 2008. ICSES '08. International Conference on, 14-17 Sept. 2008 2008. pp 193-196. doi:10.1109/icses.2008.4673390
31. Jing W, Jie L H.264 Intra Frame Coding and JPEG 2000-based Predictive Multiple Description Image Coding. In: Communications, Computers and Signal Processing, 2007. PacRim 2007. IEEE Pacific Rim Conference on, 22-24 Aug. 2007 2007. pp 569-572. doi:10.1109/pacrim.2007.4313300
32. Zhe W, Kai-Kuang M, Canhui C (2012) Prediction-Compensated Polyphase Multiple Description Image Coding With Adaptive Redundancy Control. Circuits and Systems for Video Technology, IEEE Transactions on 22 (3):465-478. doi:10.1109/tcsvt.2011.2168131
33. Fumagalli M, Lancini R, Stanzione A Video transmission over IP by using polyphase downsampling multiple description coding. In: Multimedia and Expo, 2001. ICME 2001. IEEE International Conference on, 22-25 Aug. 2001 2001. pp 1095-1098. doi:10.1109/icme.2001.1237917
34. Shirani S (2006) Content-based multiple description image coding. Multimedia, IEEE Transactions on 8 (2):411-419. doi:10.1109/tmm.2005.864349
35. Jiang W, Ortega A Multiple description coding via polyphase transform and selective quantization. In: SPIE Conf. Visual Commun. And Image Processing, San Jose, CA., 1999. pp 998-1008
36. Apostolopoulos JG Error-resilient video compression through the use of multiple states. In: Image Processing, 2000. Proceedings. 2000 International Conference on, 2000 2000. pp 352-355 vol.353. doi:10.1109/icip.2000.899393
37. Kibria R, Kim J (2008) H.264/AVC-based multiple description coding for wireless video transmission. Paper presented at the International Conference on Communications,
38. Apostolopoulos JG (2001) Reliable video communication over lossy packet networks using multiple state encoding and path diversity. Paper presented at the Visual Communications and Image Processing (VCIP),
39. Thomas GA (1987) Television motion measurement for DATV and other applications," 1987/11. BBC Research Report BBC RD,
40. Tillo T, Olmo G Low complexity pre postprocessing multiple description coding for video streaming. In: Information and Communication Technologies: From Theory to Applications,

2004. Proceedings. 2004 International Conference on, 19-23 April 2004 2004. pp 519-520. doi:10.1109/ictta.2004.1307860
41. Tillo T, Baccaglini E, Olmo G (2010) Multiple Descriptions Based on Multirate Coding for JPEG 2000 and H.264/AVC. Image Processing, IEEE Transactions on 19 (7):1756-1767. doi:10.1109/tip.2010.2045683
42. Radulovic I, Frossard P, Ye-Kui W, Hannuksela MM, Hallapuro A (2010) Multiple Description Video Coding With H.264/AVC Redundant Pictures. Circuits and Systems for Video Technology, IEEE Transactions on 20 (1):144-148. doi:10.1109/tcsvt.2009.2026815
43. Huihui B, Yao Z, Ce Z Multiple Description Video Coding using Adaptive Temporal Sub-Sampling. In: Multimedia and Expo, 2007 IEEE International Conference on, 2-5 July 2007 2007. pp 1331-1334. doi:10.1109/icme.2007.4284904
44. Zhang M, Liu W, Wang R, Bai H (2008.) A Novel Multiple description video coding Algorithm. Paper presented at the International Conference on Computational Intelligence and Security,
45. Parameswaran V, Kannur A, Li B (2009) Adapting quantization offset in multiple description coding for error resilient video transmission. Journal of Visual Communication and Image Representation 20 (7):491-503
46. Vaishampayan VA (1993) Design of multiple description scalar quantizers. Information Theory, IEEE Transactions on 39 (3):821-834. doi:10.1109/18.256491
47. Campana O, Contiero R, Mian GA (2008) An H.264/AVC Video Coder Based on a Multiple Description Scalar Quantizer. Circuits and Systems for Video Technology, IEEE Transactions on 18 (2):268-272. doi:10.1109/tcsvt.2008.918113
48. Chao T, Hemami SS (2005) A new class of multiple description scalar quantizer and its application to image coding. Signal Processing Letters, IEEE 12 (4):329-332. doi:10.1109/lsp.2005.843764
49. Samarawickrama U, Jie L A two-stage algorithm for multiple description predictive coding. In: Electrical and Computer Engineering, 2008. CCECE 2008. Canadian Conference on, 4-7 May 2008 2008. pp 000685-000688. doi:10.1109/ccece.2008.4564622
50. Vaishampayan VA, Batllo JC (1998) Asymptotic analysis of multiple description quantizers. Information Theory, IEEE Transactions on 44 (1):278-284. doi:10.1109/18.651044
51. Chao T, Hemami SS (2004) Universal multiple description scalar quantization: analysis and design. Information Theory, IEEE Transactions on 50 (9):2089-2102. doi:10.1109/tit.2004.833344
52. Vaishampayan VA, Domaszewicz J (1994) Design of entropy-constrained multiple-description scalar quantizers. Information Theory, IEEE Transactions on 40 (1):245-250. doi:10.1109/18.272491
53. Reibman A, Jafarkhani H, Yao W, Orchard M Multiple description video using rate-distortion splitting. In: Image Processing, 2001. Proceedings. 2001 International Conference on, 2001 2001. pp 978-981 vol.971. doi:10.1109/icip.2001.959211
54. Matty KR, Kondi LP (2005) Balanced multiple description video coding using optimal partitioning of the DCT coefficients. Circuits and Systems for Video Technology, IEEE Transactions on 15 (7):928-934. doi:10.1109/tcsvt.2005.848343

55. Conci N, De Natale F (2007) Real-time multiple description intra-coding by sorting and interpolation of coefficients. Signal, Image and Video Processing 1 (1):1-10. doi:10.1007/s11760-007-0009-4
56. Comas D, Singh R, Ortega A, Marques F (2003) Unbalanced multiple description video coding with rate-distortion optimization. EURASIP J Appl Signal Process:81-90. doi:10.1155/S1110865703211215
57. Doe-Man C, Yao W (1999) Multiple description image coding using signal decomposition and reconstruction based on lapped orthogonal transforms. Circuits and Systems for Video Technology, IEEE Transactions on 9 (6):895-908. doi:10.1109/76.785727
58. Guoqian S, Samarawickrama U, Jie L, Chao T, Chengjie T, Tran TD (2009) Multiple Description Coding With Prediction Compensation. Image Processing, IEEE Transactions on 18 (5):1037-1047. doi:10.1109/tip.2009.2013068
59. Yao W, Orchard MT, Vaishampayan V, Reibman AR (2001) Multiple description coding using pairwise correlating transforms. Image Processing, IEEE Transactions on 10 (3):351-366. doi:10.1109/83.908500
60. Yao W, Orchard MT, Reibman AR Multiple description image coding for noisy channels by pairing transform coefficients. In: Multimedia Signal Processing, 1997., IEEE First Workshop on, 23-25 Jun 1997 1997. pp 419-424. doi:10.1109/mmsp.1997.602671
61. Orchard MT, Wang Y, Vaishampayan V, Reibman AR Redundancy rate-distortion analysis of multiple description coding using pairwise correlating transforms. In: Image Processing, 1997. Proceedings., International Conference on, 26-29 Oct 1997 1997. pp 608-611 vol.601. doi:10.1109/icip.1997.647986
62. Goyal VK, Kovacevic J Optimal multiple description transform coding of Gaussian vectors. In: Data Compression Conference, 1998. DCC '98. Proceedings, 30 Mar-1 Apr 1998 1998. pp 388-397. doi:10.1109/dcc.1998.672173
63. Goyal VK, Kovacevic J, Arean R, Vetterli M Multiple description transform coding of images. In: Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on, 4-7 Oct 1998 1998. pp 674-678 vol.671. doi:10.1109/icip.1998.723588
64. Yao W, Orchard MT, Reibman AR Optimal pairwise correlating transforms for multiple description coding. In: Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on, 4-7 Oct 1998 1998. pp 679-683 vol.671. doi:10.1109/icip.1998.723589
65. Wang Y, Reibman AR, Orchard MT, Jafarkhani H (2002) An improvement to multiple description transform coding. Signal Processing, IEEE Transactions on 50 (11):2843-2854. doi:10.1109/tsp.2002.804062
66. Reibman AR, Jafarkhani H, Yao W, Orchard MT, Puri R (2002) Multiple-description video coding using motion-compensated temporal prediction. Circuits and Systems for Video Technology, IEEE Transactions on 12 (3):193-204. doi:10.1109/76.993440
67. Goyal VK, Kovacevic J (2001) Generalized multiple description coding with correlating transforms. Information Theory, IEEE Transactions on 47 (6):2199-2224. doi:10.1109/18.945243
68. Samarawickrama U, Jie L, Chao T (2010) M-Channel Multiple Description Coding With Two-Rate Coding and Staggered Quantization. Circuits and Systems for Video Technology, IEEE Transactions on 20 (7):933-944. doi:10.1109/tcsvt.2010.2045820

69. Wang D, Canagarajah N, Bull D Slice group based multiple description video coding using motion vector estimation. In: Image Processing, 2004. ICIP '04. 2004 International Conference on, 24-27 Oct. 2004 2004. pp 3237-3240 Vol. 3235. doi:10.1109/icip.2004.1421803
70. Che-Chun S, Yao JJ, Chen HH H.264/AVC-Based Multiple Description Coding Scheme. In: Image Processing, 2007. ICIP 2007. IEEE International Conference on, Sept. 16 2007-Oct. 19 2007 2007. pp IV - 265-IV - 268. doi:10.1109/icip.2007.4380005
71. Tillo T, Grangetto M, Olmo G (2008) Redundant Slice Optimal Allocation for H.264 Multiple Description Coding. Circuits and Systems for Video Technology, IEEE Transactions on 18 (1):59-70. doi:10.1109/tcsvt.2007.913751
72. Peraldo L, Baccaglini E, Magli E, Olmo G, Ansari R, Yao Y Slice-level rate-distortion optimized multiple description coding for H.264/AVC. In: Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on, 14-19 March 2010 2010. pp 2330-2333. doi:10.1109/icassp.2010.5496045
73. Kazemi M, Sadeghi K, Shirmohammadi S A high video quality Multiple Description Coding scheme for lossy channels. In: Multimedia and Expo (ICME), 2011 IEEE International Conference on, 11-15 July 2011 2011. pp 1-6. doi:10.1109/icme.2011.6012009
74. Kazemi M, Sadeghi KH, Shirmohammadi S (2012) A Mixed Layer Multiple Description Video Coding Scheme. Circuits and Systems for Video Technology, IEEE Transactions on 22 (2):202-215. doi:10.1109/tcsvt.2011.2159431
75. Albanese A, Blomer J, Edmonds J, Luby M, Sudan M (1996) Priority encoding transmission. Information Theory, IEEE Transactions on 42 (6):1737-1744. doi:10.1109/18.556670
76. Mohr AE, Riskin EA, Ladner RE (2000) Unequal loss protection: graceful degradation of image quality over packet erasure channels through forward error correction. Selected Areas in Communications, IEEE Journal on 18 (6):819-828. doi:10.1109/49.848236
77. Puri R, Ramchandran K (1999) Multiple description source coding through forward error correction codes. Paper presented at the 33rd Asilomar Conf.Signals, Systems and Computers, CA.,
78. El Essaili A, Khan S, Kellerer W, Steinbach E Multiple Description Video Transcoding. In: Image Processing, 2007. ICIP 2007. IEEE International Conference on, Sept. 16 2007-Oct. 19 2007 2007. pp VI - 77-VI - 80. doi:10.1109/icip.2007.4379525
79. Meng-Ting L, Jui-Chieh W, Kuan-Jen P, Huang P, Yao JJ, Chen HH (2007) Design and Evaluation of a P2P IPTV System for Heterogeneous Networks. Multimedia, IEEE Transactions on 9 (8):1568-1579. doi:10.1109/tmm.2007.907456
80. Chia-Wei H, Wen-Jiin T (2010) Hybrid Multiple Description Coding Based on H.264. Circuits and Systems for Video Technology, IEEE Transactions on 20 (1):76-87. doi:10.1109/tcsvt.2009.2026973
81. Zhiming X, Zhiping L, Anamitra M Multiple Description Image Coding With Hybrid Redundancy. In: Circuits and Systems, 2006. APCCAS 2006. IEEE Asia Pacific Conference on, 4-7 Dec. 2006 2006. pp 382-385. doi:10.1109/apccas.2006.342450
82. Xing W, Au OC, Jiang X, Zhiqin L, Yi Y, Weiran T A novel multiple description video coding based on H.264/AVC video coding standard. In: Circuits and Systems, 2009. ISCAS 2009. IEEE International Symposium on, 24-27 May 2009 2009. pp 1237-1240. doi:10.1109/iscas.2009.5117986

83. Ce Z, Minglei L (2009) Multiple Description Video Coding Based on Hierarchical B Pictures. Circuits and Systems for Video Technology, IEEE Transactions on 19 (4):511-521. doi:10.1109/tcsvt.2009.2013521
84. Zhao A, Wang W, Cui H, Tang K (2007) Efficient multiple description scalable video coding scheme based on weighted signal combinations. Tsinghua Science and Technology 12 (1):86-90. doi:10.1016/s1007-0214(07)70013-5
85. Tian S, Rajan PK Multiple description coding using transforms and data fusion. In: Information Technology: Coding and Computing, 2005. ITCC 2005. International Conference on, 4-6 April 2005 2005. pp 85-90 Vol. 81. doi:10.1109/itcc.2005.205
86. Zhao A, Yu Y, Sun G, Li G, Hui J (2009) Multiple description video coding based on residuum compensation. Science in China Series F: Information Sciences 52 (8):1439-1445. doi:10.1007/s11432-009-0133-6
87. Kamnoonwatana N, Agrafiotis D, Canagarajah CN (2012) Flexible Adaptive Multiple Description Coding for Video Transmission. Circuits and Systems for Video Technology, IEEE Transactions on 22 (1):1-11. doi:10.1109/tcsvt.2011.2129251
88. Chunyu L, Tillo T, Yao Z, Byeungwoo J (2011) Multiple Description Coding for H.264/AVC With Redundancy Allocation at Macro Block Level. Circuits and Systems for Video Technology, IEEE Transactions on 21 (5):589-600. doi:10.1109/tcsvt.2011.2129270
89. Sehgal A, Jagmohan A, Ahuja N Wireless video conferencing using multiple description coding. In: Circuits and Systems, 2001. ISCAS 2001. The 2001 IEEE International Symposium on, 2001 2001. pp 303-306 vol. 305. doi:10.1109/iscas.2001.922045
90. Vaishampayan VA, John S Balanced interframe multiple description video compression. In: Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on, 1999 1999. pp 812-816 vol.813. doi:10.1109/icip.1999.817235
91. Yilong L, Oraintara S Drift-free multiple description video coding with redundancy rate-distortion optimization. In: Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on, 23-26 May 2005 2005. pp 4034-4037 Vol. 4034. doi:10.1109/iscas.2005.1465516
92. Yao W, Shunan L (2002) Error-resilient video coding using multiple description motion compensation. Circuits and Systems for Video Technology, IEEE Transactions on 12 (6):438-452. doi:10.1109/tcsvt.2002.800320
93. Franchi N, Fumagalli M, Lancini R, Tubaro S (2005) Multiple description video coding for scalable and robust transmission over IP. Circuits and Systems for Video Technology, IEEE Transactions on 15 (3):321-334. doi:10.1109/tcsvt.2004.842606
94. Franchi N, Fumagalli M, Lancini R, Tubaro S Multiple description video coding for scalable and robust transmission over IP. In: Proc. PV 03, Nantes, France, 2003. Proc. PV 03, Nantes, France,
95. Correia P, Assuncao PA, Silva V (2012) Multiple Description of Coded Video for Path Diversity Streaming Adaptation. Multimedia, IEEE Transactions on 14 (3):923-935. doi:10.1109/tmm.2011.2182184
96. Yen-Chi L, Altunbasak Y, Mersereau RM A drift-free motion-compensated predictive encoding technique for multiple description coding. In: Multimedia and Expo, 2003. ICME '03.

Proceedings. 2003 International Conference on, 6-9 July 2003 2003. pp III-581-584 vol.583. doi:10.1109/icme.2003.1221378
97. Yen-Chi L, Altunbasak Y, Mersereau RM (2004) An enhanced two-stage multiple description video coder with drift reduction. Circuits and Systems for Video Technology, IEEE Transactions on 14 (1):122-127. doi:10.1109/tcsvt.2003.819182