Proceedings of the 5th WSEAS Int. Conf. on MULTIMEDIA, INTERNET AND VIDEO TECHNOLOGIES, Corfu, Greece, August 17-19, 2005 (pp154-159)

Multiple Description Coding for Error-Resilient Video Transmission

Mengyao Ma†, Oscar C. Au‡, S.-H. Gary Chan†
†Dept. of Computer Science, ‡Dept. of Electrical and Electronic Engineering
Hong Kong University of Science and Technology
{myma,eeau,gchan}@ust.hk

Abstract: - Since the quality of compressed video is vulnerable to errors, video transmission over unreliable Internet is very challenging today. Multiple Description Coding (MDC) is an error resilience technique to encode the source into multiple streams, which are sent to the destination through different channels. The strategy of MDC is that an arbitrary set of descriptions can be received to construct the video with acceptable quality. In this paper, we review some MDC approaches, categorized by the dimension of description sub-sampling; namely, spatial or temporal. In spatial MDC, the blocks of each frame, or the motion vector and the residues of each block are coded into multiple descriptions; while in temporal MDC, the video sequence is coded into multiple decodable streams, each with its own prediction process.

Key-Words: - Error Resilience, Multiple Description Coding, MDSQ, PCT, Three-Loop Structure, VRC, Error Concealment.

I. INTRODUCTION

Delivering video of good quality over the Internet or wireless networks is very challenging today, due to the use of predictive coding and *Variable Length Coding* (VLC) in video compression [1][2]. In block-based video coding method, each frame is divided into some rectangular blocks (macroblocks). If we use INTER prediction mode, each block is predicted from a previously coded frame by *Motion Compensation*. Then the *Residue* (prediction error) is transformed (e.g. using DCT), variable length coded, and sent to the decoder along with the *Motion Vector*. If data loss occurs during the transmission, the corresponding frame will be corrupted, and this error will be propagated to the following frames because of inter-prediction, until the next INTRA-coded frame is correctly received. For example, a simple bit error in VLC can cause desynchronization; as a result, all the following bits cannot be used until a synchronization code arrives [1]. Due to these facts, it is useful to develop some schemes to improve the *Error Resilience* (ER) property of the compressed video.

The coding efficiency of an ER codec is lower than a normal codec, because the encoder needs to introduce some redundancy to the stream. In case of error, the decoder would use this additional information to reconstruct the video. The key of designing an ER codec is to make a best compromise between the gain in error resilience and the amount of redundancy sent [1]. In evaluating the effectiveness of an ER technique, we should also consider about the reconstructed video quality, the delay between the two endpoints, and the processing complexity [2].

Several error resilience methods have been developed for video communication. One such method is *Forward Error Correction* (FEC), typically applied at channel coding stage. In this method, FEC codes are added to the video stream by the encoder, and the decoder uses these codes to correct some bit errors. FEC techniques can be jointly used with other error resilience methods such as MDC and subband-based coders [3][4][5]. Another method for error resilience is *Layered* (*Scalable*) *Coding* (LC). LC refers to partitioning the video stream into more than one layer. The base layer is protected and transmitted with higher priority; it contains the most important information for the video and can be used to provide acceptable video quality. Each enhancement layer incrementally improves the video quality [5][6]. LC provides different video qualities according to channel bandwidth, but the layers have to be obtained incrementally, leading to inflexibility and low video quality when a lower layer is lost. In addition to this approach, *Multiple Description Coding* (MDC) is often used. Contrary to layered coding, MDC divides the video stream into equally important *streams* (*descriptions*). These descriptions are sent to the destination through different channels. Error may occur in the channels. Suppose the failure probability of each channel is independently and identically distributed with probability p. If we use the conventional *Single Description Coding* (SDC) method, the entire description will be lost with probability p; if we use M descriptions and send them on M different channels, the probability of losing the entire description is p^M, which is much less than p. For two descriptions, if both descriptions are received at the destination, a high quality video is reconstructed; if any one description is received, the decoder can still reconstruct a video with a lower but acceptable quality. In this paper, we will address this category of error resilience method and illustrate using two descriptions as example.

Multiple description coding has been studied for about twenty years. It was original invented at Bell Laboratories, along with speech communication over the telephone network [7]. The first technique used a simple odd/even separation [8]. The idea of channel splitting was also studied by information theorists in the early days [9][10]. Today, MDC can be used in not only audio coding [11], but also image [12][13] and video coding [14][15]. More information on the history and applications of MDC can be found in [7]. In this paper, we

This work has been supported in part by the Research Grants Council of the Hong Kong Special Administrative Region, China (project no. HKUST6203/02E and HKUST6156/03E), the Innovative Technology Commission (project no. ITS/122/03) and HKUST (project no. DAG04/05.EG34).

mainly focus on the development and technique of MDC as an ER method for video coding.

MDC can be divided into two categories according to the dimension a video stream is sampled: spatial sub-sampling and temporal sub-sampling. In spatial sub-sampling, each frame is divided into two descriptions using, for example, the motion vectors of the blocks [16], the residues [17][18], or even the blocks themselves [19]. In temporal sub-sampling, the video is coded into multiple decodable streams, each with its own prediction process [20][21]. For example, an even frame is predicted from the previous even frame, and an odd frame is predicted from the previous odd frame. There are other more efficient temporal methods, which we will discuss in detail later. We will also see that different MDC methods can be combined together, or with other ER techniques. One simple way to do spatial sub-sampling is to partition the stream into several sets, then compress and transmit them independently. When error happens, interpolation is used to reconstruct the sample from the available descriptions [19]. This approach introduces much redundancy because it does not utilize the correlation between adjacent samples [2]. There are some other efficient methods in the literature: *Multiple Description Scalar Quantizer* (MDSQ) [22] and *Pair-Wise Correlating Transform* (PCT) [23][24]. MDSQ can be extended to *Multiple Description Lattice Vector Quantizer* (MDLVQ) [25][26], quantizers whose codebook is a lattice. Because MDLVQ is not used widely in ER for image and video coding, we will mainly focus on MDSQ and PCT in this paper.

The remainder of the file is organized as follows. We discuss in section II spatial MDC. We first present the two basic MDC methods: MDSQ and PCT, followed by a *Three-Loop* MDC model [17]. Section III is about temporal MDC. We show how this technique can be further improved by, for example, using the feedback from the channels, by *Temporal Interpolation*, or by the combination with other ER technique. We conclude in section IV.

II. SPATIAL MDC

Figure 1 is a typical two-description MDC coder [2]. The encoder generates two descriptions D_1 and D_2, which are sent to decoder through different channels. The decoder has three parts, and only one works at a time. If both D_1 and D_2 are received, decoder d_0 will reconstruct the same signal as the one at the encoder side. If only D_1 (D_2) is received, decoder $d_1(d_2)$ will be used to reconstruct a signal with a lower but acceptable quality. Because the information carried in D_2 (D_1) is lost, the reconstructed signal would have some error. We want to decrease this error. Decoder d_0 is referred to as *Central Decoder*, and $d_1(d_2)$ is *Side Decoder*. Suppose D_1 and D_2 are the same, which are the duplicate of the signal at the encoder. Then the receival of either D_1 or D_2 will make the decoder to reconstruct the signal correctly, at the expense of much redundancy. If bit-rate R is used to send the same signal in SDC, this simple duplication costs 2R bit-rate. As the correlation between the two descriptions decreases, the redundancy decreases, but the reconstructed video quality at the side decoder decreases. From this point of view, the design goal in MDC is to decrease the distortion at side decoder $d_1(d_2)$ and to reduce the bit-rate for sending D_1 and D_2.

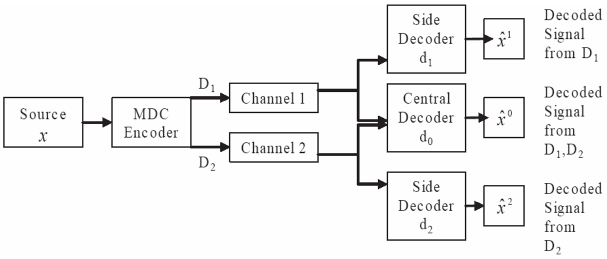


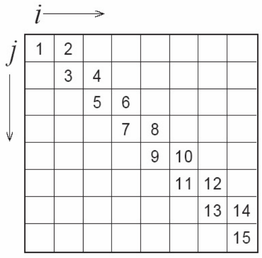
Fig. 1. Diagram of multiple description encoding and decoding.

Two MDC methods are widely used as the basic approaches to generate two descriptions: *Multiple Description Scalar Quantizer* (MDSQ) [22][27][18] and *Pair-Wise Correlating Transform* (PCT) [23][24][28]. In the following, we will first introduce these two approaches, and then present some MDC models developed for them. A common feature of these models is using a *Three-Loop* structure proposed in [17].

A. Multiple Description Scalar Quantizer (MDSQ)

The design of multiple description scalar quantizer (MDSQ) is first proposed by *Vaishampayan* in [22]. The objective is to minimize the average distortion when both descriptions are received, subject to the constraints on the average distortion when only one description is received. In this method, two descriptions are obtained by index assignments of two quantizers. Suppose source $x \in \Re$. The first step of the encoder is a regular quantizer. It partitions $\Re$ into N cells, and each cell is an interval with an *interval index*. If x belongs to the l^{th} interval, the quantizer will produce l. Then l is passed to the *index assignment* block, which maps l into a *codeword pair* (i, j). Then i and j are sent to channel 1 and channel 2, respectively. At the decoder side, if both i and j are received, central decoder d_0 is used to map (i, j) to $\hat{x}^0$. If only i (j) is received, side decoder d_1 (d_2) is used to reconstruct $\hat{x}^1$ ($\hat{x}^2$) from i (j). Suppose the input to MDSQ encoder is a random variable X, and the output of the *index assignment* are random indexes I and J. If square-error distortion measure is used, the optimal reconstruction is: $\hat{x}^0 = d_0(i, j) = E[X|I = i, J = j]$, $\hat{x}^1 = d_1(i) = E[X|I = i]$, and $\hat{x}^2 = d_2(j) = E[X|J = j]$.

It is shown in [22] that a good design of *index assignment* is crucial for the performance of MDSQ, because it determines the mean square error (MSE) for both central and side decoder. However, since the optimal index assignment is difficult to get, it develops two index assignment algorithms at high rates. One example for low bit rate applications is shown in Figure 2 [27]. The row index i and the column index j compose a codeword pair (i, j), which corresponds to the interval index of the quantizer. Note that the interval indexes lie on the main diagonal of the codebook, increasing from left to right. Index assignment in Figure 2 is referred as A2 index assignments, which contains two diagonals. For a given rate R, as the number of diagonals increases, the percentage of empty cells decreases. As a result, the redundancy caused by this coding



1	2						
	3	4					
		5	6				
			7	8			
				9	10		
					11	12	
						13	14
							15

Fig. 2. Index assignments of MDSQ: A2 ([18]).

scheme is reduced, leading to decreasing MSE at the central decoder and increasing MSE at the side decoder.

The MDSQ analyzed in [22] assumes that the encoded information (i, j) is sent to the channels with fix length coding. It is extended to variable length coding with significant improvements in performance in [27]. For example, at rate R = 4.0 bpss, a gain of 4.8 dB is achieved. In [18], MDSQ is applied to image and video coding for lossy networks. The central idea is to encode the transform coefficients with MDSQ. Reconstruction technique at the side decoder is also developed. The MDSQ for M-channel transmission is generalized in [29] and [30], which are optimized for the channel states[1] and embedded for progressive image transmission, respectively.

B. Pair-Wise Correlating Transform (PCT)

Another way to get two descriptions is to use a pair-wise correlating transform (PCT), which can be implemented by a rotation matrix [23][24]. Applying PCT to a pair of uncorrelated variables will produce two correlated coefficients, which are sent to the destination through different channels. If one of these coefficients is lost, a receiver can estimate it from the other by using the correlation information. A PCT codec tries to minimize distortion both at the central decoder d_0 and at the side decoder d_1 (or d_2), given equal loss probability of channel. When introducing correlation between coefficients in different streams, a PCT codec should also make the coefficients in the same stream uncorrelated.

Suppose (S_1, S_2) is a pair of uncorrelated coefficients obtained from the *Karhunen-Loeve Transform* (KLT), and its PCT is represented by a 2×2 matrix T. Applying T to (S_1, S_2) produces a pair of correlated coefficients (D_1, D_2). In [24], T is an orthogonal transform parameterized by rotation angle θ:

$$\begin{pmatrix} D_1 \\ D_2 \end{pmatrix} = T \begin{pmatrix} S_1 \\ S_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} S_1 \\ S_2 \end{pmatrix} \quad (1)$$

In this method, the correlation between D_1 and D_2 is controlled by angle θ, and maximized at $\theta = \pi/4$. The advantage of using an orthogonal transform is that we can quantize the coefficients in (D_1, D_2) domain instead of (S_1, S_2) domain, so that the quantization error added to the transformed coefficients is the same as that added to the original ones [23]. Suppose the quantized value of (D_1, D_2) is $(\tilde{D}_1, \tilde{D}_2)$.

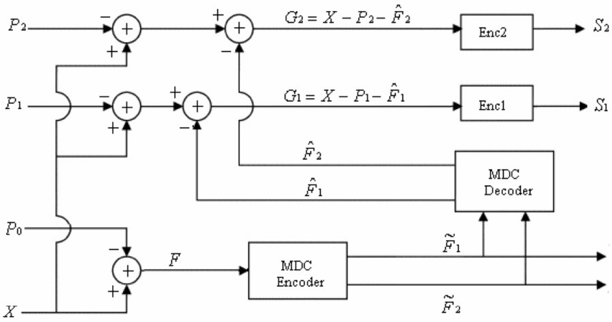


Fig. 3. The Three-Loop Structure for INTER mode [17].

If both $\tilde{D}_1$ and $\tilde{D}_2$ are received, we can use the inverse transform T^{-1} to reconstruct the original variables. If only $\tilde{D}_1$ is received, $\tilde{D}_2$ is reconstructed by an optimal linear estimator, which minimizes the mean square estimation error. $\tilde{D}_1$ is similarly reconstructed when only $\tilde{D}_2$ is received. A special case of orthogonal transform, leading to two balanced streams for $\theta = \pi/4$ is implemented in [23]. By studying how the correlation (introduced by transform T) affects the bit rates and the distortion at the side decoder, *Yao et al.* propose an optimal correlating transform in [28]. Since this optimal transform is non-orthogonal, quantization must be applied to (S_1, S_2) first before they are correlating transformed.

Another design issue for PCT is how to select the coefficients to pair, and how much redundancy should be added to each pair. Because when two coefficients are correlating transformed as in (1), the coding efficiency is reduced. The pairing strategy should minimize the increase in bitrate, while making reconstruction at the side decoder of satisfactory quality [23]. Analysis and simulation show that paring a large coefficient with a small one is optimal. On the other hand, when a pair of variables is too small relative to the quantization error, it needs not be paired and can be simply split among the two streams [23][24][31].

Simulation results in [31] shows that PCT can reach a lower redundancy range than MDSQ.[2] But when the percentage of redundancy is high, MDSQ will achieve a higher PSNR at the side decoder than PCT, for the same central distortion and the same rate constraint.

C. Three-Loop Structure

Although MDC can be applied to image coding to achieve acceptable quality when loss occurs, there is still some problem when it is applied to video coding. For two descriptions, the decoder may receive both of them, or only one; so there are three possible scenarios at the decoder: reconstruction at the central decoder d_0, or at the side decoder $d_1(d_2)$. When the video is reconstructed at a side decoder, there is a *prediction mismatch (drift)*, since the reference in the following motion compensation is different from the one at the encoder. A model with two independent prediction loops is proposed in [32]

---

[1] Each channel can either deliver its description or lost its description, so a group of M channels would always be in one of 2^M possible states.

[2] When using the bit rate required by the SDC codec as the reference, redundancy is defined as the additional rate (in bits/pixel or bpp) required by the coders in comparison [31].

to avoid the drift between the encoder and decoder. Each prediction loop performs its own prediction based on a single-description reconstruction. An additional coarse quantizer is added to each loop to avoid the mismatch at the side decoder. As a result, the coding efficiency degrades a lot, especially when the step size of the quantizer is large.

A more efficient MDC model for video coding is proposed by *Reibman et al.* in [17], named *Multiple Description Transform Coding* (MDTC). The main idea is to use three separate prediction loops at the video encoder to mimic the three possible scenarios at the decoder: both descriptions are received (*central prediction loop*) or only one is received (*side prediction loop*). The mismatch between the encoder and decoder is sent as the *side information*. In this approach, only the DCT coefficients are coded by multiple descriptions; the motion vectors and other header information are duplicated on both channels. Figure 3 is the three-loop structure for INTER mode. For each block X, the encoder generates three predicted blocks P_0, P_1 and P_2, whose references are reconstructed from both descriptions, Description 1 (D_1) and Description 2 (D_2), respectively. The prediction error (DCT coefficients) when both descriptions are available, $F = X - P_0$ is paired and coded into two descriptions, $\tilde{F}_1$ and $\tilde{F}_2$, using PCT (labelled as MDC Encoder). Let the reconstructed signal F from only $\tilde{F}_i$ be $\hat{F}_i$, for $i = 1, 2$. Then the reconstruction of X with only one description is $P_i + \hat{F}_i$, for $i = 1, 2$. The mismatch (side information) G_i is coded by a single description coder (Enc1 and Enc2) and sent to the decoder. Define:

$$\begin{aligned} G_i &= X - P_i - \hat{F}_i, \\ S_i &= Q_i\{T\{G_i\}\},\ i = 1, 2. \end{aligned} \tag{2}$$

where $T\{\cdot\}$ and $Q_i\{\cdot\}$ are DCT and quantization operators. At the decoder side, if both descriptions are available, the block is reconstructed using

$$\hat{X} = P_0 + \hat{F}, \tag{3}$$

where $\hat{F}$ is the reconstructed signal from $\hat{F}_1$ and $\hat{F}_2$ using inverse PCT. If only one description is received, the reconstruction is

$$\hat{X} = P_i + \hat{F}_i + \hat{G}_i,\ i = 1, 2, \tag{4}$$

where $\hat{G}_i$ is the decoded side information from S_i.

Note that in this model, the side information S_1 and S_2 is totally redundant when no error occurs [17]. This portion of redundancy can be controlled by varying the quantization step size of S_i. Another source of redundancy comes from the correlating transform of F, which can be controlled by changing the rotation angle of the transform matrix [24]. In order to use the side information to enhance the reconstruction quality when the channel is error-free, a MDC scheme using matching pursuits is proposed in [33]. In [14], a two-stage MDC is proposed, with an additional MDSQ introduced into the central prediction.

III. TEMPORAL MDC

The conventional INTER mode approach for video coding (SDC) is illustrated in Figure 4(a), where each P-frame is predicted from its immediate previous frame. Although the compression efficiency of this approach is high, it is vulnerable to errors in the transmission channel. If one frame is lost or corrupted (for example: P_4) during the transmission, the error in the reconstructed frame at the decoder will propagate to the remaining frames until the next I-frame (I_1) is received. An implementation of multiple description for this kind of situation is to separate the original video sequence into several streams, as illustrated in Figure 4(b) with odd and even sequence.[3] Each stream has a different prediction loop and is independently decodable from the other. Since the reference frames are farther in time, the prediction is not as good as SDC and the bit rate needed is higher. If error occurs in one stream, i.e. P_4 is lost as in Figure 4(c), the other stream is still correct. So we can only display $P_5 P_7 P_9 \ldots$ until the next I-frame is received. The video is reconstructed at half of its original frame rate. The drawback of this simple approach is that when error occurs in both streams, all the following frames will be corrupted until the next I fame is received, similar as the single stream approach. The *Video Redundancy Coding* (VRC) approach is similar to this method, and has been adopted into H.263+ (Annex N: Reference Picture Selection Mode) [34]. The video sequence is divided into two or more *threads*, and the frames are assigned to one thread in a round-robin fashion. A *Sync Frame* is inserted at regular intervals to restart all the threads. An example with two threads is illustrated in Figure 4(d).

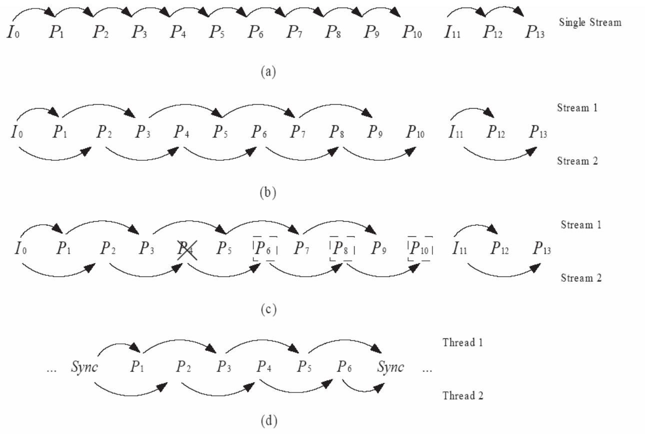


Fig. 4. Illustration of different approach for video coding (the arrow means that the previous frame is used as the reference of the latter). (a) One stream approach. (b) Multiple stream approach with odd and even sequence. (c) Error occurs in (b). (d) Multiple Threads and Sync Frame in H.263+.

To improve this basic implementation, several approaches can be used. One way is to select reliable transmission paths over the Internet for video packets [20]. In this approach, a *path diversity transmission system* is used where the server explicitly sends different packets over different paths, by specifying different intermediate routers. As a result, the performance provided by multiple paths is better than using a random path. The second way is to dynamically select the reference picture of each frame according to the reconstructed video quality [21][35][36]. This approach requires feedback

---

[3] Temporal MDC approach is sometimes referred to as *Multiple Threads* [34], *Multiple State* [20], or *Multiple Stream* [35].

information from the channel. The third way is to use *Error Concealment* (EC) to reconstruct the corrupted frames, so that the propagated error to the following frames is reduced [20]. In Figure 4(c), since the frames before and after P_4 (P_3 and P_5) are correct, we may use them to recover P_4 by *temporal interpolation* [37][38]. Another way is to combine the basic odd/even sub-sampling with other error resilience methods [39][40]. In this section, we will introduce the latter three approaches.

A. Reference Picture Selection Using Feedback

When the feedback information is available, the encoder can encode the current frame with reference to the correctly received frame by the decoder, instead of always using the immediate frame as the reference. The advantage of this approach is that it can stop error propagation without introducing much redundancy, while the drawback is requiring a feedback channel [21].

In [21], *Shunan et al.* propose always selecting the last correctly received frame as the reference. The coded frames are transmitted along different paths depending on the available path bandwidth. The feedback message is sent back to the encoder for each frame: NACK for loss and ACK for correct reception. When a NACK message is received for a frame delivered on one path, this path is assumed to be "bad" until an ACK is received. Similarly, a path is assumed to be "good" until a NACK is received. In [35], a multiple stream approach is proposed using active probing. It also divides the channels into two kinds of states: good (G) and bad (B). The difference between this approach and the one in [21] is that it tries to make the descriptions on different paths independent whenever it is possible, while the previous one prefers to increase the correlation between the two streams to reduce the redundancy.

When feedback information is used, both the encoder and the decoder need a long-term memory buffer to store some past pictures that may be used as references for the following frames. The efficiency of such approaches depends on feedback delay [21]. The shorter the delay is, the less is the distance of the reference frame, and the lower the redundancy is. An optimal picture selection scheme is proposed in [36], which minimizes the expected end-to-end distortion given a certain rate constraint.

B. Temporal Interpolation for EC

Temporal interpolation was originally used to generate one or more frames between two received frames so as to improve the effective frame rate, and make the object motions in the video smoother. A *Motion Compensated Temporal Interpolation* (MCTI) method is proposed in [37], which uses block-based motion estimation to track motions of the objects between adjacent received frames. Both forward and backward motion estimation are performed to find the motion vector. However, the interpolated video frames tend to be blocky. Improvement methods are proposed in [41] and [42]. In [41], the motion vectors of surrounding blocks are further utilized to reconstruct the motion vectors of interpolated block. The correlation between spatially adjacent pixels is used for the refinement of reconstruction in [42].

The objective to introduce temporal interpolation is that it can be well combined with temporal MDC methods. Recall that in Figure 4(c), when frame P_4 is corrupted during transmission, its surrounding frames (P_3 and P_5) would be correct if stream 1 is error-free. So we can utilize P_3 and P_5 to interpolate P_4 with good quality. Since MCTI based interpolation methods need high computational requirement, the algorithm called *Unidirectional Motion Compensated Temporal Interpolation* (UMCTI) is more favorable, which performs the forward motion estimation instead of bidirectional motion estimation [38]. Furthermore, the advantage of using UMCTI in temporal MDC is that the exhaustive motion estimation needs not be performed at all, since the motion vectors from blocks of the n^{th} frame to the corresponding blocks in the $(n-2)^{th}$ frame is known. In other words, the motion vector from P_5 to P_3 is conserved in stream 1. Actually, experimental results show that the interpolation using the motion vectors in the original stream has better quality than the one using exhaustive search. The reason is that the motion vectors in the transmitted stream are estimated from the original video sequence, while the motion vectors got by exhaustive search is estimated from the compressed video with lower quality. As a result, the former one is better.

C. Combing Temporal MDC with Other ER Methods

Temporal MDC may be combined with other error resilience methods. For example, in [39], the authors propose to combine temporal sub-sampling with special sub-sampling. The video sequence is split into odd and even frames, or the video picture is split into odd and even lines. Neural network is used to determine which scheme is used. Temporal MDC can also be combined with multi-hypothesis motion compensated prediction, which can suppress the propagated error to some extent. In [40], a *Multiple Description Motion Compensation* (MDMC) method is proposed, which allows the encoder to perform prediction from both past even and odd frames such that less redundancy is required than VRC. Similar to the three-loop model in Figure 3, this approach uses one central prediction loop and two side prediction loops. The main difference is that the sub-sampling of this approach is temporal instead of spatial. The central predictor uses both even and odd past frames for the prediction, while the side predictors use only even or odd past frames. Comparison between MDMC and VRC is given in the paper. For same redundancy (compared to the SDC of original H.263+ codec), MDMC works better for high motion sequence, while VRC works better for low motion sequence.[4] The reason is that for low motion sequence, the gain from the slightly better error concealment used in MDMC cannot compensate for its extra redundancy (two motion vectors for each block and additional side prediction error). Comparison to MDTC (*Reibman*'s three-loop structure [17]) is also given. For "Foreman" QCIF sequence with bit rate 50.16 kbps, MDMC can achieve the

---

[4] Note that the redundancy of VRC, compared to original SD H.263+ codec, cannot be controlled.

same side distortion with 30% less redundancy [40]. Please refer to the paper for more details.

IV. CONCLUSION

In this paper, we review the technique of Multiple Description Coding (MDC), an error resilient method for video transmission over unreliable networks. Based on the dimension of how the multiple descriptions are obtained, two categories of MDC methods are introduced: spatial MDC and temporal MDC. For spatial MDC, *Multiple Description Scalar Quantizer* (MDSQ) and *Pair-Wise Correlating Transform* (PCT) are introduced, followed by some MDC models based on the *Three-Loop* structure. For temporal MDC, techniques that can improve the basic *Video Redundancy Coding* (VRC) approach are introduced: using the feedback information from the channels, enhancing the reconstruction by *Temporal Interpolation*, and combing MDC with other error resilience techniques. Comparison is also given. Since MDC introduces relatively high redundancy, it is suitable for high loss transmission media, where the reconstruction quality at the side decoder is critical [2]. If the channel loss is small, the reconstruction at the central decoder dominates, so *Single Description Coding* (SDC) can be used instead of MDC.

REFERENCES

[1] Y. Wang, S. Wenger, J. Wen, and A. K. Katsaggelos. (2000, July) Review of error resilient coding techniques for real-time video communications. SPM_W3K2.pdf. [Online]. Available: http://vision.poly.edu:8080/paper/

[2] Y. Wang and Q. F. Zhu, "Error control and concealment for video communication: a review," in *Proceedings of the IEEE*, May 1998, pp. 974 – 997.

[3] R. Puri and K. Ramchandran, "Multiple description source coding using forward error correction codes," in *the Thirty-Third Asilomar Conference on Signals, Systems, and Computers*, Oct. 1999, pp. 342 – 346.

[4] H. Cai, B. Zeng, G. Shen, and S. Li, "Error-resilient unequal protection of fine granularity scalable video bitstreams," in *Proc. IEEE International Conference on Communications*, June 2004, pp. 1303 – 1307.

[5] C.-M. Fu, W.-L. Hwang, and C.-L. Huang, "Efficient post-compression error-resilient 3D-scalable video transmission for packet erasure channels," in *Proc. IEEE ICASSP*, Mar. 2005, pp. 305 – 308.

[6] R. Aravind, M. Civanlar, and A. Reibman, "Packet loss resilience of MPEG-2 scalable video coding algorithms," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 426 – 435, Oct. 1996.

[7] V. Goyal, "Multiple description coding: compression meets the network," in *IEEE Signal Processing Mag.*, Sept. 2001, pp. 74 – 93.

[8] N. Jayant, "Subsampling of a dpcm speech channel to provide two self-contained half-rate channels," in *Bell Syst. Tech. J.*, vol. 60, Apr. 1981, pp. 501–509.

[9] J. K. Wolf, A. Wyner, and J. Ziv, "Source coding for multiple descriptions," in *Bell Syst. Tech. J.*, vol. 59, Oct. 1980, pp. 1417–1426.

[10] A. A. E. Gamal and T. M. Cover, "Achievable rates for multiple descriptions," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 851–857, Nov. 1982.

[11] S. Voran, "A multiple-description PCM speech coder using structured dual vector quantizers," in *Proc. IEEE ICASSP*, Mar. 2005, pp. 129 – 132.

[12] A. Gavrilescu, A. Munteanu, P. Schelkens, and J. Cornelis, "Embedded multiple description scalar quantizers for progressive image transmission," in *Proc. IEEE ICME*, July 2003, pp. I – 525–8.

[13] C. Tian and S. Hemami, "A new class of multiple description scalar quantizer and its application to image coding," *IEEE Signal Processing Lett.*, vol. 12, pp. 329 – 332, Apr. 2005.

[14] Y. C. Lee, Y. Altunbasak, and R. Mersereau, "An enhanced two-stage multiple description video coder with drift reduction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, pp. 122 – 127, Jan. 2004.

[15] Y.-C. Lee, Y. Altunbasak, and R. Mersereau, "Coordinated application of multiple description scalar quantization and error concealment for error-resilient mpeg video streaming," vol. 15, pp. 457 – 468, Apr. 2005.

[16] C. S. Kim and S. U. Lee, "Multiple description coding of motion fields for robust video transmission," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, pp. 999 – 1010, Sept. 2001.

[17] A. Reibman, H. Jafarkhani, Y. Wang, M. Orchard, and R. Puri, "Multiple description coding for video using motion compensated prediction," in *Proc. IEEE ICIP*, Oct. 1999, pp. 837 – 841.

[18] V. Vaishampayan, "Application of multiple description codes to image and video transmission over lossy networks," in *The 7th International Workshop on Packet Video*, Apr. 1991, pp. 2857 – 2860.

[19] A. Tom, C. Yeh, and F. Chu, "Packet video for cell loss protection using deinterleaving and scrambling," in *Proc. IEEE ICASSP*, Apr. 1991, pp. 2857 – 2860.

[20] J. Apostolopoulos, "Reliable video communication over lossy packet networks using multiple state encoding and path diversity," in *Visual Communications and Image Processing*, Jan. 2001, pp. 392 – 409.

[21] S. Lin, S. Mao, Y. Wang, and S. Panwar, "A reference picture selection scheme for video transmission over ad-hoc networks using multiple paths," in *Proc. IEEE ICME*, Aug. 2001, pp. 96 – 99.

[22] V. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. Inform. Theory*, vol. 39, pp. 821 – 834, May 1993.

[23] Y. Wang, M. Orchard, and A. Reibman, "Multiple description image coding for noisy channels by pairing transform coefficients," in *Proc. IEEE First Workshop on Multimedia Signal Processing*, June 1997, pp. 419 – 424.

[24] M. Orchard, Y. Wang, V. Vaishampayan, and A. Reibman, "Redundancy rate-distortion analysis of multiple description coding using pairwise correlating transforms," in *Proc. IEEE ICIP*, Oct. 1997, pp. 26–29.

[25] S. Servetto, V. Vaishampayan, and N. Sloane, "Multiple description lattice vector quantization," in *Data Compression Conference*, Mar. 1999, pp. 13 – 22.

[26] V. Vaishampayan, N. Sloane, and S. Servetto, "Multiple-description vector quantization with lattice codebooks: design and analysis," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1718 – 1734, July 2001.

[27] V. Vaishampayan and J. Domaszewicz, "Design of entropy-constrained multiple-description scalar quantizers," *IEEE Trans. Inform. Theory*, vol. 40, pp. 245 – 250, Jan. 1994.

[28] Y. Wang, M. Orchard, and A. Reibman, "Optimal pairwise correlating transforms for multiple description coding," in *Proc. IEEE ICIP*, Oct. 1998, pp. 679 – 683.

[29] A. Gavrilescu, A. Munteanu, P. Schelkens, and J. Cornelis, "Embedded multiple description scalar quantizers for progressive image transmission," in *Proc. IEEE ICME*, July 2003, pp. I – 525–8.

[30] S. Voran, "The channel-optimized multiple-description scalar quantizer," in *Proc. IEEE The 2nd Signal Processing Education Workshop*, Oct. 2002, pp. 400 – 405.

[31] Y. Wang, M. Orchard, V. Vaishampayan, and A. Reibman, "Multiple description coding using pairwise correlating transforms," *IEEE Trans. Image Processing*, vol. 10, pp. 351 – 366, Mar. 2001.

[32] V. Vaishampayan and S. John, "Balanced interframe multiple description video compression," in *Proc. IEEE ICIP*, Oct. 1999, pp. 812 – 816.

[33] X. Tang and A. Zakhor, "Matching pursuits multiple description coding for wireless video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, pp. 566 – 575, June 2002.

[34] S. Wenger, "Video redundancy coding in h.263+," in *Proc. Audio- Visual Services over Packet Networks*, Sept. 1997.

[35] E. Setton, Y. Liang, and B. Girod, "Adaptive multiple description video streaming over multiple channels with active probing," in *Proc. IEEE ICME*, July 2003, pp. I – 509–12.

[36] Y. Liang, M. Flierl, and B. Girod, "Low-latency video transmission over lossy packet networks using rate-distortion optimized reference picture selection," in *Proc. IEEE ICIP*, Sept. 2002, pp. II–181 – II–184.

[37] C.-K. Wong and O. Au, "Fast motion compensated temporal interpolation for video," in *Visual Communications and Image Processing*, May 1995, pp. 1108 – 1118.

[38] C.-W. Tang and O. Au, "Unidirectional motion compensated temporal interpolation," in *Proc. IEEE ISCAS*, June 1997, pp. 1444 – 1447.

[39] O. Lotfallah and S. Panchanathan, "Adaptive multiple description coding for internet video," in *Proc. IEEE ICASSP*, Apr. 2003, pp. V – 732–5.

[40] Y. Wang and S. Lin, "Error-resilient video coding using multiple description motion compensation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, pp. 438 – 452, June 2002.

[41] C.-K. Wong, O. Au, and C.-W. Tang, "Motion compensated temporal interpolation with overlapping," in *Proc. IEEE ISCAS*, May 1996, pp. 608 – 611.

[42] T. Chen, "Adaptive temporal interpolation using bidirectional motion estimation and compensation," in *Proc. IEEE ICIP*, Sept. 2002, pp. II–313 – II–316.