# Multiple Description Video Coding Based on Frame-Rotation

Yubai Zhang[1] ,Yao Zhao[2] ,Huihui Bai[3] ,Chunyu Lin[4]
Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China
yzhao@bjtu.edu.cn

Koh Hang[5],Anhong Wang[6]
A*STAR Institute for Infocomm Research(12R) [5]
Taiyuan University of Science and Technology,Taiyuan 030024,China[6]

***Abstract*—Multiple Description Coding is an effective way to improve the robustness of real-time data transmission over unreliable channels. In this paper, we propose a novel MDC scheme based on rotation which is compatible with the current video standard. In the MDC approach, two or more independently decodable descriptions of the same data are generated. In the case of two descriptions, one of the descriptions is formed by the original video sequence and the other is formed by the rotated video sequence. The decoder will reconstruct the video sequence according to the different content received. Experimental results show that the performance of the proposed scheme is superior to the classical RS-MDC scheme about 1 dB.**

***Keywords-MDC; Video-Rotation; Error-Concealment; RS-MDC***

## I. INTRODUCTION

Modern video transmission and storage systems are typically characterized by a wide range of access network connection qualities and end terminal capabilities. The current Internet is the "best-effort" network which has many problems such as channel interference, network congestion and routing delay. These problems will lead to data errors or packet loss. Otherwise, random bit errors, consecutive sudden errors and other problems of the wireless communication channel further deteriorate the channel conditions, which resulting in the loss of bit-stream. The packet loss of one frame will impact the following frames due to the motion estimation and motion compensation in H.264/AVC and MPEG video coding standard. These issues limit the real-time Multimedia-communications in the network [1].

## II. MULTIPLE DESCRIPTION CODING

Multiple Description Coding (MDC) is an effective method to solve the problems above. In the MDC approach, two or more independently decodable descriptions of the same data are generated. Suppose the error probability of each channel is $p$, the error probability of all the channels at the same time will be $p^n$. Under the state of that all descriptions are transmitted over independent channels, if just part of them is received, the side decoder can reconstruct the complete data with a given side distortion. The more descriptions are received, the higher is the quality of the decoded signal [2]; when all the descriptions are received, and the central decoder will be called. In reality, there have been many important applications about MDC, which involves speech coding, image coding, video coding, distributed storage systems, and low-latency system. MDC is also suitable for the application in non-priority protected network environment due to that each description is equally important. Moreover it has better real-time advantages compared with the mechanism for fault-tolerant encoding scheme [3, 4] which is based on the FEC-error protection and automatic repeat request (ARQ).

As we know, in the traditional video coding the temporal correlation between adjacent frames is used to improve the data compression capability. Most of the video coding standards which are based on motion estimation and motion compensation [5] may lead to the motion mismatch. Mismatch is an error caused by the inconsistency between the encoder and decoder references. The traditional mismatch-control is that the separate prediction loop for each description is used to encode and reconstruct the video sequence. The conventional MDC is to down-sample the video sequence into two sub-sequences as odd frames and even frames. Each sub-sequence has its own prediction, coding, transmission and reconstruction. We can reconstruct the missing frames through motion compensation and temporal interpolation of adjacent frames when one description has been received, and the video sequence can be reconstructed through recombination of the received parity frames when the two descriptions are all received. Although this algorithm can be directly applied to the traditional the single description program, it is not flexible in redundancy adjustment. The sampling in time, space and frequency domain reduces the correlation between adjacent frames and the corresponding blocks, which increases the prediction residual and decreases the compression efficiency. Vaishampayan [7] avoid the mismatch by adopting joint quantization of the prediction error in the two prediction rings. However, the coarse quantization of the two side encoders will lead to low coding efficiency. Besides, the algorithm which uses the mismatch between the forecast of Center Road and sidewalks as the side information and encodes the side information into each description to control the redundancy and mismatch is proposed in [8]. Although the redundancy is controllable in such algorithms, the structure is complex and the side information has no effect when the two descriptions are all received. The H.264 redundant slices (slice) [1] provides an encoding scheme compatible with the H.264 standard by optimizing the quantization step of the original piece and

978-0-7695-4738-1/12 $26.00 © 2012 IEEE
DOI 10.1109/CMCSN.2012.7



IEEE computer society

the redundant slice. In [9], a method is proposed to optimize the allocation of redundancy in the macroblock layer, which achieves improvement compared with [1].

All of the MDC approaches mentioned above try to use the inherent redundancy as the spatial sampling program, and inject more redundancy as the MDC scheme based on redundant slice. The quality of the decoded video is worse than the corresponding single description video coding scheme in case of one description reception. In addition, these methods often require changes of the traditional video coding schemes to adapt to the multiple descriptions coding, and the system complexity will be dramatically increased. In this paper, a low complexity MDC scheme based on rotation is proposed to solve the above problems. This scheme can obtain the same performance as the traditional single description coding scheme when only one description reception. The rotation scheme is similar with the multiple-description image coding algorithm in [10], but the way of coding and the structure of the packet are different.

## MDC BASED ON ROTATION SCHEME

*The Encoder*

The MDC based on rotation scheme is easy to be implemented and can be adopted into different video codecs.

Figure 1 shows the diagram of the encoder for the two description case. The first description can be formed by encoding the input sequence with a normal H.264 encoder. For the second description, each frame of the input video sequence is firstly rotated 180 degrees. After that, the normal H.264 encoder is used to encode the rotated sequence to form descriptions2.

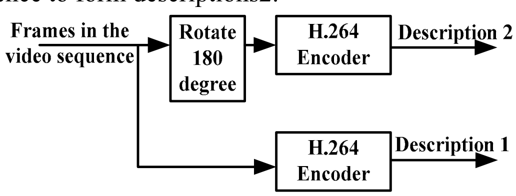


Figure 1. The encoder of multiple description video coding scheme based on rotation.

Figure 2 and Figure 3 show the package diagram of the normal frame and the rotated frame, and it ensures the correspondence of the normal frame and the rotated frame.

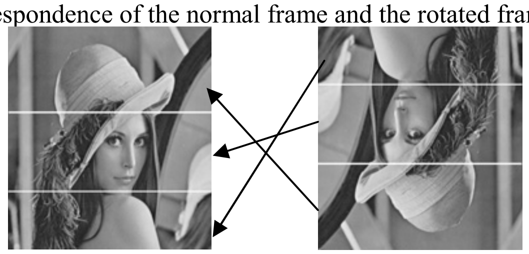

Figure 2. The rotation program schematic.

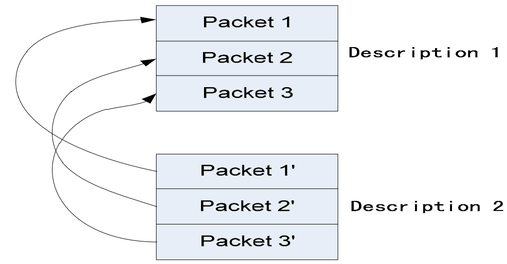


Figure 3. The corresponding packets.

The basic idea of this rotational multiple description coding is that each macroblock of the frames in the two descriptions uses the different reference block as prediction to get the different residual error. Other methods such as mirror and upside down placed frames are similar to the rotation and can also be used to form a multiple description system.

*The Decoder*

There are two different cases in the decoder. The first case is that the two descriptions are all received. We can reconstruct the sequence by averaging the corresponding pixel values of the two received descriptions. Then central performance obtains larger gain when the residuals of the two descriptions are more irrelevant or closer to negative correlation.

The theoretical model is as follows, setting $f$ as the original video frame. $\hat{f}_1$ and $\hat{f}_2$ are the reconstruction value of the two side decoders, $P$ indicates the prediction part, $Q$ indicates the quantization part, $e$ indicates the quantization error. The two side decoders have the following form that the two descriptions use different reference frame and different block for the encoded frame. This approach leads to different residual error and the different quantization error $e_1(n),e_2(n)$.

$$f(n) = P(\hat{f}_1(m)) + Q\{f(n) - P(\hat{f}_1(m))\} + e_1(n)$$
$$= \hat{f}_1(n) + e_1(n), m \le n$$
$$f(n) = P(\hat{f}_2(k)) + Q\{f(k) - P(\hat{f}_2(k))\} + e_2(n)$$
$$= \hat{f}_2(n) + e_2(n), k \le n$$

The corresponding reconstruction value of the central decoder is as follows:

$$\hat{f}(n) = 0.5(\hat{f}_1(n) + \hat{f}_2(n)) = f(n) - 0.5(e_1(n) + e_2(n))$$

The reconstruction error will be smaller when $e_1(n)$ and $e_2(n)$ are less relevant or closer to negative correlation.

Another case is that just one description reception. The system will use the corresponding reconstructed value of the other description as alternative when partial packets of one description are lost. The mismatch will be greatly reduced in the process of replacement because the two descriptions have the same bit rate, the mean value and the variance of the error are almost equal. Figure 2 above shows that one frame can be divided into several packets. We can use the corresponding reconstructed video packet in the other description as a replacement when the packet in one description is lost.

There are two cases in the decoder when partial packets of one description are lost, which are shown in Figure 4 and Figure 5 respectively. The red lines in the two figures represent the lost packet. Figure 4 shows the case that the same video packets in the two descriptions are not lost simultaneously, and the reconstructed packet in one description replaces the lost packet in the other description. The average of the two descriptions ensures the high quality reconstruction in the decoder. This case avoids the mismatch (i.e. inconsistency between the encoder reference block and the decoder reference block) because the two descriptions use the original description as the reference frame. Figure 5 shows the case that the same video packets in the two descriptions are all lost, and then we can use the error-concealment technique in the H.264 decoder to reconstruct the lost packets. The decoder averages the two descriptions received. The two descriptions make use of the average of the reconstructed frames as a reference which is closer to the original frame. Hence the reconstructed frame can further reduce the mismatch.

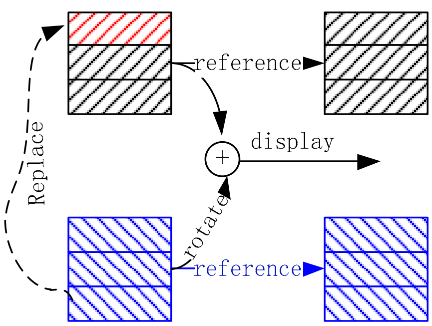


Figure 4. one of the same video content corresponding of the two descriptions is lost.

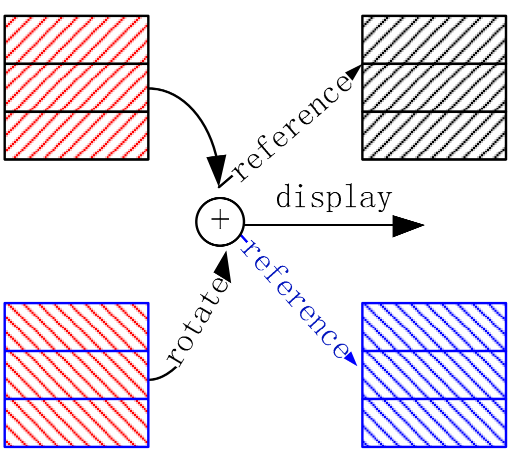


Figure 5. The same video content corresponding of the two descriptions are all lost.

In order to adjust the redundancy, the residuals can be received by subtracting the reconstructed average of the two descriptions from the original video sequence. We sample the encoded redundancy and generate parity packets to produce the second part of the two descriptions.

The new multiple-description coding system is shown in Figure 6. The first part of each description is decoded in the usual way. For the second part (reconstructed frame) of each description, we can sum the decoded result of the redundancy and the first part to get the reconstruction.

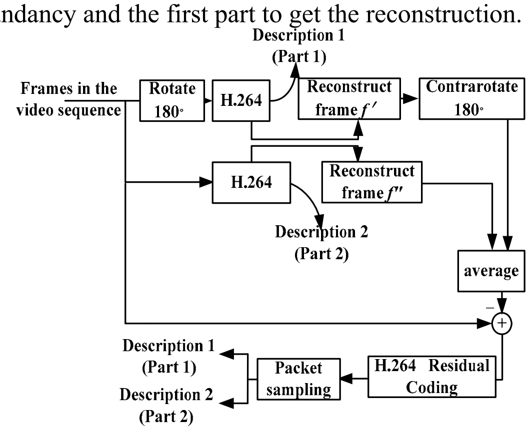


Figure 6. Multiple description coding system based on the residuals redundancy adjustment.

The entire system will ensure that the two descriptions can be received simultaneously with high probability when the bit rate is fixed and the channel condition is good. The residual signal takes up more bits in this condition, so that the entire coding schemes tend to provide a better two-way performance. However, the entire system can only receive one of the two descriptions most of times when the channel condition is poor. And the residual signal can be assigned with a relatively small bit-rate in this condition. Ideally, if the two channels are completely reliable, the system degrades into H.264 encoder and transfers data packets alternately in the two channels, i.e. the description contains only the second part of Figure 6. And the system in Figure 6 contains only the first described part which can provide a relatively bigger loss resistance when the channel transmission is relatively unreliable.

## III. EXPERIMENTAL RESULTS

We adopt the H.264 JM software to generate the video sequence with the fixed macro-block number. Taking a row of MBs as a video package can guarantee the normal encoded video packet and the rotated encoded video packets contain the same content. The structure of the H.264 group of picture (GOP) is IPPP and the test video sequence is the Foreman sequence in CIF format.

As we all know that RS-MDC is the classic MDC scheme at present which has a relatively good performance. We compare the performance of the Rotation-MDC with RS-MDC. Figure 7 shows the PSNR of the Center Road (the two descriptions are all received) and the side-channel (one of the two descriptions is received) and the experimental results of the multiple description coding scheme based on a redundant piece (RS-MDC) for comparison (restructure the sentence). Two of the GOP size is set to 45. The result of multiple description coding system is much better than the RS-MDC solution as shown in Figure 7.

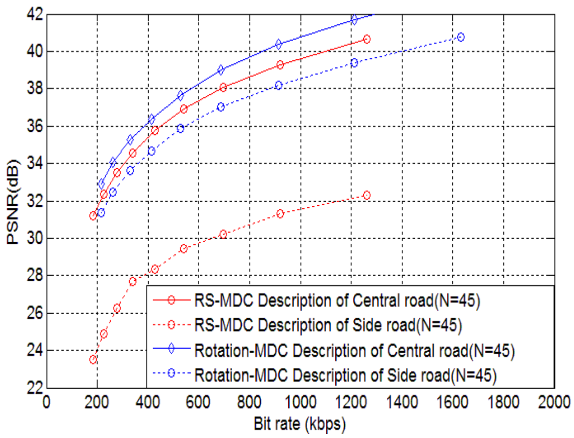


Figure 7. Comparison of the performance from side and central decoder.

Figure 8 shows the comparison of the experimental results based on the rotation scheme and the RS-MDC scheme under different packet loss rates and GOP structures. The GOP size is 11, 21 and 45 and the packet loss rate is 1%, 5% and 10% respectively. The performance of the rotation-MDC scheme is better than the RS-MDC scheme.

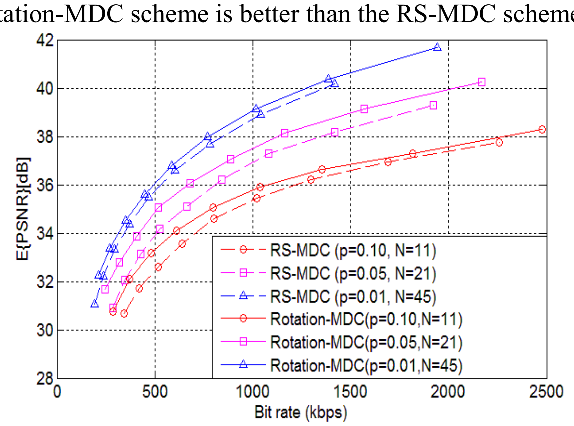


Figure 8. Comparison of the performance in the packet loss network.

## IV. CONCLUSION

As described above, the Rotation-MDC has two inputs, the original video sequence and the rotated sequence, which have been encoded with a standard H.264 encoder. The rotation method improves the robustness of the data transmission over unreliable channels. Experimental results show that the performance of the proposed scheme is superior to the classical RS-MDC scheme about 1 dB. Therefore the proposed Rotation-MDC is a promising scheme which has a superior performance compared with other MDC schemes.

## REFERENCES


[1] T. Tillo, M. Grangetto, G. Olmo. Redundant slice optimal allocation for H.264 multiple description coding [J]. IEEE Transactions on Circuits and Systems for Video Technology. 2008, 18 (1):59-70.

[2] V. K. Goyal, Multiple description coding: compression meets the network, IEEE Signal Processing Magazine, vol. 18, no. 5, pp.74–93, Sept. 2001.

[3] Yuan Zhang, Wen Gao, Yan Lu, Qingming Huang, and Debin Zhao, Joint source-channel rate-distortion optimization for h.264 video coding over error-prone networks. IEEE Transactions on Multimedia, 9(3):445–454, 2007.

[4] S. Soltani, K. Misra, and H. Radha, Delay constraint error control protocol for real-time video communication, IEEE Transactions on Multimedia, 2009,11(4):742 –751

[5] Y. Wang, A. R. Reibman, S. Lin, Multiple Description Coding for Video Delivery[J]. Proceedings of the IEEE. 2005, 93 (1):57-70.

[6] Z. Wei, C. Cai, K. Ma, A Novel H.264-based Multiple Description Video Coding Via Polyphase Transform and Partial Prediction[C]. In: International Symposium on Intelligent Signal Processing and Communications. 2006:151-154

[7] V. A. Vaishampayan, S. John, Balanced interframe multiple description video compression[C]. In: IEEE International Conference on Image Processing. 1999:812-816

[8] Y. Lee, Y. Altunbasak, R. M. Mersereau. An enhanced two-stage multiple description video coder with drift reduction [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2004, 14 (1):122-127

[9] Chunyu Lin, Tillo T., Yao Zhao and Byeungwoo Jeon, Multiple description coding for H.264/AVC with redundancy allocation at macro block level [J], IEEE Transactions on Circuits and Systems for Video Technology. 2011, 21 (5):589-600

[10] Chunyu Lin, Yao Zhao and Ce Zhu, Two-Stage diversity-based multiple description image coding [J], IEEE Signal Processing Letters, 2008, 15:837-840