Wang, D., Canagarajah, CN., & Bull, DR. (2004). Slice group based multiple description video coding using motion vector estimation. In *2004 International Conference on Image Processing (ICIP 2004) Singapore* (Vol. 5, pp. 3237 - 3240). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/ICIP.2004.1421803

2004 International Conference on Image Processing (ICIP)

SLICE GROUP BASED MULTIPLE DESCRIPTION VIDEO CODING USING MOTION VECTOR ESTIMATION

D. Wang[†]*, N. Canagarajah and D. Bull*

Centre for Communications Research,
University of Bristol, Bristol, BS8 1UB, UK
Tel: +44 (0)117 9545202, Fax: +44 (0)117 9545206
[†]Email: Dong.Wang@bristol.ac.uk

ABSTRACT

This paper proposes a novel scheme for multiple description video coding approach using slice group coding tool proposed in H.264. Independent motion compensation loops are maintained for two descriptions in the encoder. In each description, one slice group is encoded as main information, while the other slice group is encoded very coarsely, as redundancy, to keep basic information. This coarse slice group can be encoded using normal encoding with coarse quantizer, or using predicted motion vectors by the main slice group. Mode decision is made to select best encoding method. Results show that our scheme works very well and keeps subjective quality very good for middle and higher bitrate. Redundancy and side quality can be controlled by changing parameters of coarsely coded slice group, unlike some other MDC methods with fixed redundancy.

1. INTRODUCTION

Video transmission over lossy network is a challenging problem. In video compression, due to predictive coding, any bit loss may cause great quality degradation. Multiple description coding is one approach to address this problem, where several sub bit streams called descriptions are generated from source video. Each description can reconstruct video of acceptable quality and all the descriptions together can reconstruct higher quality video.

Unlike layered video coding techniques, each description generated by MDC can independently be decoded and reconstructed to acceptable quality. This can give a graceful degradation of received video with loss, while avoiding catastrophic failure of layered coding due to loss of base layer.

An MDC system consists of two kinds of decoders. One is the central decoder which is used when all the descriptions are received, and the other is side decoder which just uses one or a subset of descriptions to reconstruct video of acceptable quality.

More correlations in descriptions will result in higher quality of side decoded video. At the same time central decoder must perform with lower efficiency because more redundancy is introduced. Extensive research on MDC to increase the efficiency has been conducted.

MDC based on Scalar Quantization is developed in [1] to divide a signal by two coarser quantizers, and it's applied to predictive video coding in [2]. Output of each quantizer is the approximation of single description. Any one description can use its coarse data to generate a basic video and both of them can be combined to reconstruct higher quality video. Another approach on image coding is addressed in [3] and [4] using pairwise correlating transforms to transform a vector of DCT coefficients into another vector of correlated components, which introduces additional redundancy between components. This was used in motion compensated video coding [5]. Another simple way of generating MDC is that through pre- and post- processing, as in [6]. Redundancy is introduced by padding zeros in frequency domain. The source video frames are transformed using DCT. Certain number of zeros are padded in frequency domain, and after inverse transform, the video is sub-sampled into two descriptions. The two descriptions are independently coded at the encoder. In [7], video sequence is divided into two by means of odd and even frames and different concealment methods are used to estimate lost frames. In [8] odd and even frames compose two descriptions, which is similar to [7], but three MC loops are maintained. It performs well on ideal MDC environments and packet lossy network. A restriction is that it can only use previous two frames with constant weights of two motion vectors. Overlapping technique is used on motion vectors in [9] to achieve more accurate prediction of lost data.

We propose a new method, with two descriptions, based on slice group (SG-MDC), the new coding tools developed in H.264. In each description, one slice group is encoded normally as main information, while the other slice group is encoded very coarsely, as redundancy, to keep basic information. Two descriptions are symmetric. To make this redundancy coding less costly, motion vector estimation technique is used and additional coding optimization is made between this technique and the normal encoding. The proposed scheme is able to encode with little redundancy, which can varies by changing the quality of the redundancy slice group.

The rest of this paper is organized as follows. In Section 2 our slice group based MDC is described. Section 3 gives the results and analysis of experiments. Conclusions are presented in section 4.

2. DESCRIPTION OF THE PROPOSED SG-MDC

Slice Group is a new coding tool in H.264. Picture is divided into slice groups and it can be further divided into slices in scan order. Not like slice, slice group allows encoding picture not by scan order which make it much more flexible. There are totally 7 types of macroblock allocation for slice group, i.e. slice group

0-7803-8554-3/04/$20.00 ©2004 IEEE.

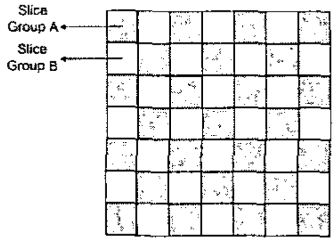


Fig. 1 Macroblock Map For Dispersed Slice Group

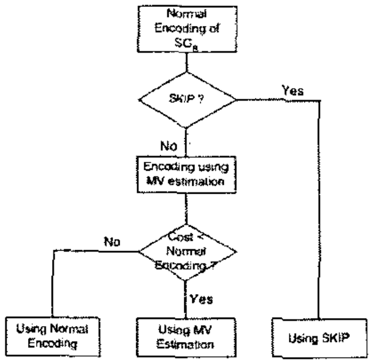


Fig. 3 Mode decision for SGB encoding

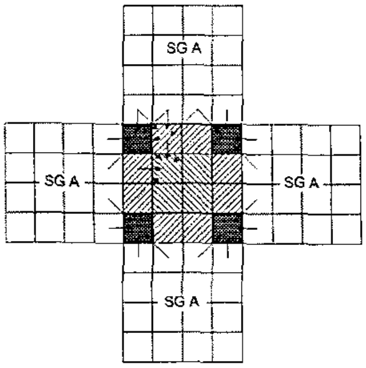


Fig. 4 Motion Vector Estimation

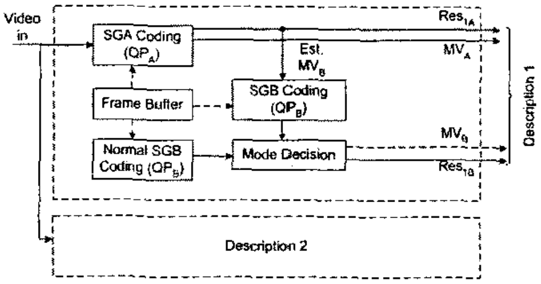


Fig. 2 Structure of proposed SG-MDC

map types, in the current H.264 drafts in which type 1 is called 'dispersed' slice group map. For two slice groups A and B, macroblock map is as fig. 1. They are distributed as checkerboard, which is very effective for error concealment, and this map type with two slice groups is chosen as the basis of our MDC scheme.

In our scheme, each slice group composes main information in one description, as depicted in fig. 2. We call this slice group as main slice group. For each picture, the main slice group is encoded normally and the other slice group is also encoded. This slice group is completely redundancy, for it also exists in the other description as main slice group. The basic idea is to encode it using a big quantizer step size to keep basic information and try to make the redundancy as little as possible. If two descriptions are received correctly, the redundancy is discarded and using main information we can reconstruct video of good quality by standard decoder. If only one description is received, we can reconstruct main slice group of good quality, and use the redundancy to reconstruct the other slice group of acceptable quality by the side decoder. It is obvious that these two descriptions are symmetric and independent from each other and any of them can reconstruct video without drift problem by itself.

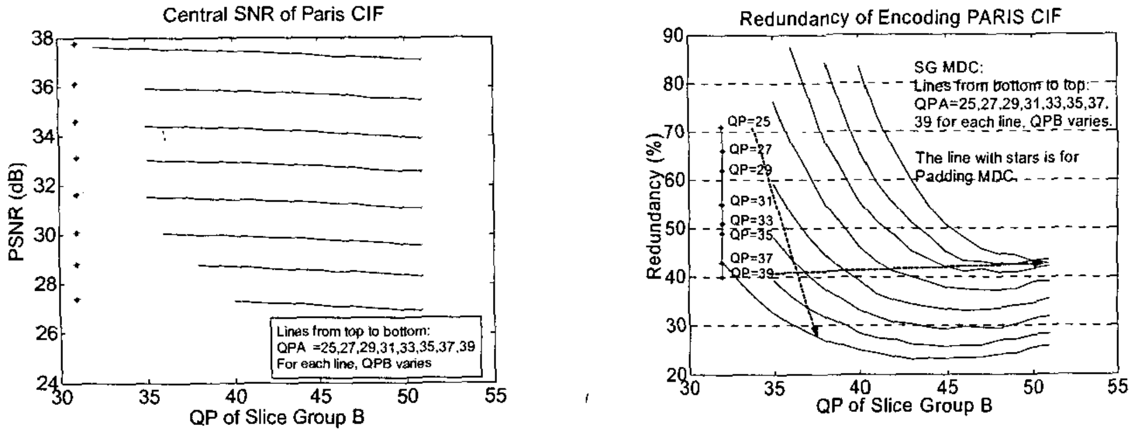
In the following we will only focus on the first description which uses slice group A (SGA) as main information, for the two descriptions are symmetric and we got the same performance out of each description.

Encoding SGB can be made by encoding normally as SGA, but with a bigger quantizer step size (QP). We name it as normal encoding here. But results show that redundancy is too much, since motion and header information take up many bits in SGB, which is normally around 50%. For SGA is always encoded firstly which means we have SGA data before encoding SGB, we employ motion vector (MV) estimation for SGB using the data of SGA, hereby reducing overhead for header and motion information. However, with the inaccurately estimated motion vectors which results in more bits for SGA, encoding will not be always optimal in spite of reducing bits for SGB header. If the costs are higher than normal encoding, there is no reason to use this estimation method. Based on this consideration, additional mode decision is made as fig. 3.

We set the highest priority to skip mode. If the result of mode decision in normal encoding is skip mode, we select skip mode. If not, encoding using MV estimation is compared to normal encoding. If the cost is lower than normal encoding, MV estimation based encoding is selected. Otherwise normal encoding is used instead. One bit in syntax should be used to flag this mode decision, if it is not skip mode.

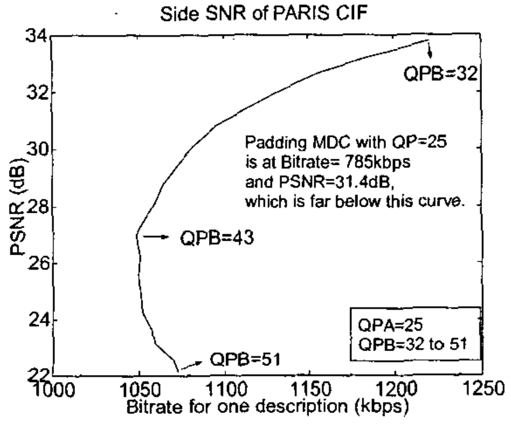
As to motion vector estimation, it is very flexible and lots of estimation methods can be used. The method indicated in fig. 4 is selected for our scheme. For dispersed slice group type, there are five macroblocks of which the middle one is for SGB, if the macroblock to be encoded in SGB is not the border or corner of picture. Each macroblock consists of 16 4x4 blocks. We estimate motion vectors for each block. At first, motion vectors of four blocks at the corner are estimated using four neighbour blocks in SGA. Then motion vectors of blocks at border of macroblock are estimated using neighbouring 3 blocks in SGA and one adjacent block at the corner. At last, the motion vectors of the four middle blocks are estimated using neighbouring 5 blocks. We use median method to calculate the motion vector value. The reason we select 4x4 motion vectors is that generally 4x4 motion estimation is more accurate, and since we don't need to encode these motion vectors, it should be of the best performance. For border MB, similar median methods are used.

Encoding of main slice group is completely normal encoding and can use multiple reference frames. This will make the encoding very efficient. Currently we use the most recent reference frame for the redundancy slice group. The best motion vectors for each reference frame in SGA will be generated during mode decision and we can use it directly without adding more



(a)

(c)



(b)

Fig. 5 Performance of SG-MDC. (a) Central Quality for different QPA with varying QPB; (b) Side Quality for QPA=25 with varying QPB; (c) Redundancy of different QPA with varying QPB.

computation.

3. RESULTS AND ANALYSIS

We examine the performance of our proposed SG-MDC. All the data are got by assuming that one entire description is lost, which is the ideal MDC environment. For all the cases, we use the video coding standard H.264 [10][11] as the basic coder. H.264 is the latest video coding standard. It has many new features, which make it much better than previous standards. Slice group coding tools make our scheme very convenient. Fixed frame rate (30frames/second) and constant quantizer step size are used for each slice in all frames of sequence. No B frame is used. Entropy coding is CAVLC. Sequence 'paris' is used. Currently very few papers on MDC use H.264 as platform. We implement [6] on H.264 for comparison, which is called Padding MDC here.

Central quality of SG-MDC is presented in fig. 5 (a). Each line is one series of encoding with constant QP of SGA (QPA) and different QP of SGB (QPB). From top to bottom QPA is 25 to 39. The several stars are the quality of normal single-description encoding using the same QP as corresponding QPA. We can see that with the same QP, the central quality is almost the same as that of single-description encoding, and with increasing of QPB, the central quality changes very slowly, which means changing side quality won't affect the quality of SGA much (But with redundancy, the R(D) curve must be worse than the single-description encoding). Therefore, we can evaluate redundancy through increasing of bitrate with the same QP.

Fig. 5 (b) gives an example of the property of side quality with changing QPB for constant QPA=25. We consider the bitrate of one description vs. Side PSNR. It's clear that with smaller QPB, the side quality is better. With QPB increasing, the side quality and bitrate are lower. But when the QPB is very big, which means the quality of SGB is too low, the side quality keeps decreasing but the bitrate increases. This is because the prediction of SGA is too bad due to the very bad quality of SGB in previous frames and encoding is much less efficient. This property results in a minimum point for the bitrate and a turning point for the R(D) curve. This can also be seen in Fig. 5 (c), in which the redundancy curve is similar as the R(D) curve in Fig. 5 (b). In order to see the details of the curve in this figure we haven't draw the Padding MDC, for which the point is at 1570kbps and 31.4dB, which is far below our curve.

In fig. 5 (c) the redundancy curves are presented for different QPA. Each line is for a constant QPA and varying QPB. For high quality video such as QPA being around 25, the redundancy is very small and minimum point is 23.1%. And for low bit rate, the redundancy is around 40%. Also due to the property of central quality described above, we can change the side quality by changing redundancy through varying QPB without affecting central quality. We can see that redundancy is higher for bigger QPA, which means that SG-MDC is more suitable for higher quality video. With higher QPA, the difference between QPA and QPB is smaller and slice group B takes more percentage of bits in the bitstream which results in more redundancy. And for big QP such 39, we can see that there is no minimum point for redundancy due to this small difference. For the Padding MDC, the property of redundancy is opposite to SG MDC. It's high for high quality video and lower for low bit rate. From the comparison, it's obvious that SG MDC is better than Padding MDC. At middle and higher bitrate, its redundancy is much higher than SG MDC, and with low bitrate it performs just similar as SG MDC. Moreover, Padding MDC is fixed redundancy, i.e. side quality, while SG MDC can vary side quality.

Fig. 6 shows a comparison of pictures of frame number 82 in *paris* sequence. The first one is the decoded frame using normal single-description encoding. The second one is the side decoded frame with only half rate received. The side quality is 27dB, but the subjective quality is still acceptable.

This MDC method doesn't work so good for the sequences with much motion such as *foreman* sequence. This is due to the inaccurate motion vector estimations. Better estimation methods may improve the accuracy further. Nonlinear interpolation techniques can also be used, for there are many motion vectors available in the neighbouring macroblocks. Moreover, currently we only use the most recent reference frame for redundant slice group. Results are expected to be better if we select reference according to reference frame numbers of neighbouring macroblocks.

4. CONCLUSIONS

In this paper we introduced a novel MDC approach based on slice group coding tools in H.264. Two slice groups are used of which each composes main information for one description. Two independent MC loops are maintained which is well suited for ideal MDC environments. We add redundancy by encoding the other slice group coarsely and with motion vector estimation. It is shown through simulations that our SG-MDC performs very well especially for high quality video.

Normal Encoding with QP=25
PSNR=38dB

MDC with QPA=25 and QPB=40
Redundancy=25% Side PSNR=27dB
Half rate received

Fig. 6 Decoded Pictures of Frame no.82

REFERENCES

[1] V. Vaishampayan, "Design of multiple description scalar quantizers," IEEE Trans. On information Theory, vol. 39, no.3, pp.821-834, May 1993.

[2] V. Vaishampayan and S. John, "Interframe balanced-multiple-description video compression," Packet Video '99, New York, NY, USA, Apr. 1999.

[3] M. Orchard, Y. Wang, V. Vaishampayan, and A. Reibman, "Redundancy rate distortion analysis of multiple description coding using pairwise correlating transforms," Proc. IEEE Int. Conf. Image Proc, Santa Barbara, CA, USA, Oct. 1997.

[4] Y. Wang, M. Orchard, and A. Reibman, "Optimal pairwise correlating transforms for multiple description coding,", Proc. IEEE Int. Conf. Image Proc, Chicago, Illinois, USA, Oct. 1998.

[5] A. Reibman, H. Jafakhani, Y. Wang, and M. Orchard, "Multiple description coding for video using motion compensated prediction," Proc. IEEE Int. Conf. Image Proc, Kobe, Japan, Oct. 1999.

[6] D. Wang, N. Canagarajah, D. Redmill, D. Bull, "Multiple Descriptoin Video Coding Based on Zero Padding," Proc. IEEE Int. Symposium on Circuits and Systems, Vancouver, Canada, May 2004.

[7] J. G. Apostolopoulos, "Error-resilient video compression through the use of multiple states," Proc. IEEE Int. Conf. Image Proc, Vancouver, CA, USA, Sept. 2000.

[8] Y. Wang, S. Lin, "Error-Resilient Video Coding Using Multiple Description Motion Compensation,", IEEE Trans. On Circuits and Systems for Video Technology, Vol.12, No.6, June 2002.

[9] C. -S. Kim and S. -U. Lee, "Multiple description motion coding algorithm for robust video transmission," IEEE Int. Symp. on Circuits and Syst., Geneva, Switzerland, May 2000.

[10] H.264 standard, JVT-G050, 7th meeting, Pattaya, Thailand, 7-14 March, 2003

[11] Thomas Wiegand, Gary J. Sullivan, Gisle Bjontegaard, and Ajay Luthra, "Overview of the H.264/AVC Video Coding Standard," IEEE Transactions on Circuits and Systems for Video Technology, pp. 560- 576, July 2003.