OPEN ACCESS

University of BRISTOL

Wang, D., Canagarajah, C. N., & Bull, D. R. (2005). Slice group based multiple description video coding with three motion compensation loops. In IEEE International Symposium on Circuits and Systems, 2005 (ISCAS 2005) Kobe, Japan. (Vol. 2, pp. 960 - 963). Institute of Electrical and Electronics Engineers (IEEE). 10.1109/ISCAS.2005.1464749

Link to published version (if available):
10.1109/ISCAS.2005.1464749

Link to publication record in Explore Bristol Research
PDF-document

University of Bristol - Explore Bristol Research

General rights

SLICE GROUP BASED MULTIPLE DESCRIPTION VIDEO CODING WITH THREE MOTION COMPENSATION LOOPS

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. Three motion compensation loops, one central and two side loops, are maintained for two descriptions in the encoder. The central encoder is exactly the same as basic single description encoder, while side encoder uses information from main encoder to do the motion compensation with easily controlled amount of redundancy. If there is no loss, this scheme is completely standard compliant which doesn't need any further processing at decoder side. The rate-distortion of side encoder can be controlled in a very large dynamic range, so that redundancy can be easily managed with specific requirement of bitrate or side quality, unlike some other MDC methods with fixed redundancy. Results show that our scheme works with better performance and more flexibility than [8], which is 2 loop slice group based MDC, and keeps subjective quality very good.**

I. 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] using pairwise correlating transforms to transform a vector of DCT coefficients into another vector of correlated components, which introduces additional correlations between components. This was used in motion compensated video coding [4]. A spatial way of generating MDC is proposed in [5] through pre- and post- processing. Redundancy is introduced by padding zeros in frequency domain which results in more correlations. After this processing the video is sub-sampled into two descriptions. The two descriptions are independently coded at the encoder. In [6], which is a temporal scheme, video sequence is divided into two by means of odd and even frames and different concealment methods are used to estimate lost frames. In [7] odd and even frames compose two descriptions, which is similar to [6], but three motion compensation (MC) loops are maintained. It performs well on ideal MDC environments and packet lossy network. But it can only use previous two frames as reference with constant weights of two motion vectors. In [8] another spatial approach is proposed based on slice group coding tools in H.264. Two slice groups compose main information in two descriptions respectively. The other slice group is encoded very coarsely to maintain basic information as redundancy. It performs very well with good error resilient property, but due to the restriction of fixed quality setting for main slice group, which is to keep required central quality, it's not so flexible to control redundancy and results for low bitrate is not good.

We propose a new scheme based on slice group, but with three MC loops. All video data is encoded normally in the central encoder. Using motion information and residual data of central encoder, video is further encoded with controlled redundancy. The proposed scheme is able to encode with a very large dynamic range of controlled redundancy and proves to have good rate-distortion results.

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

II. DESCRIPTION OF THE PROPOSED SG-MDC

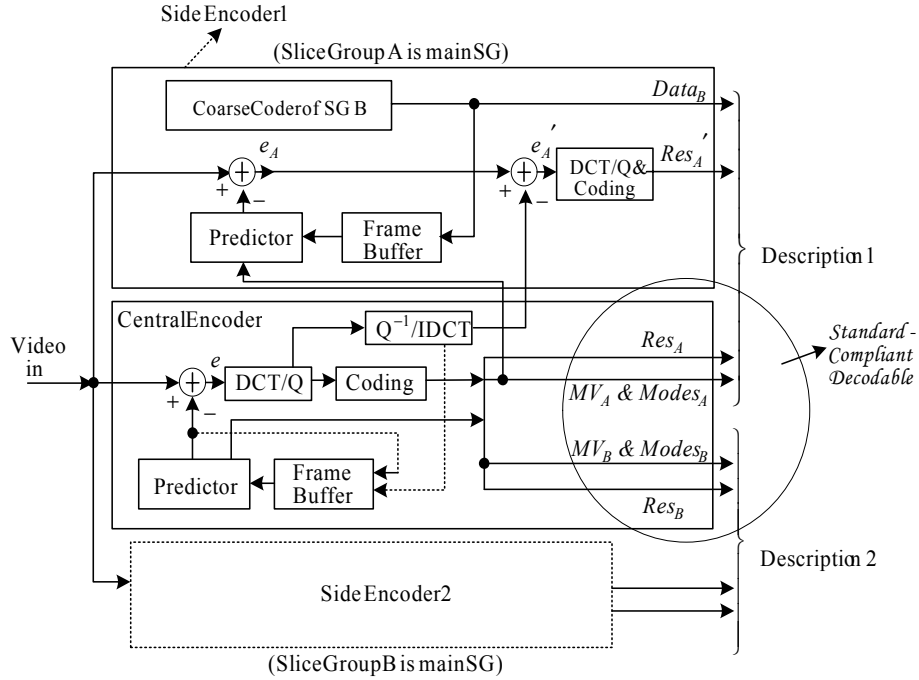Slice Group is a new coding tool in H.264. Picture is divided

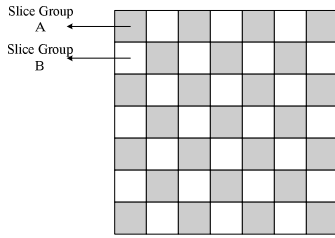Figure. 2 Structure of the proposed 3 Loop MDC scheme

Figure. 1 Macroblock Map
For Dispersed Slice Group

into slice groups and it can be further divided into slices in scan order. There are totally 7 types of macroblock allocation for slice group, i.e. slice group map types, in which type 1 is called 'dispersed' slice group map. For two slice groups A and B, macroblock allocation map is as fig. 1, which is like a checkerboard. It is very effective for error concealment, and this map type with two slice groups is chosen as the basis of our MDC scheme. The basic idea is that, in each description, only one slice group is finely encoded, and the other slice group is encoded coarsely to keep basic quality of reconstruction and the MC loop.

3L-SGMDC encoder consists of 3 encoders of which one is central encoder and the others are side encoders. Central encoder is the same as single-description standard encoder. One description consists of streams from one side encoder and corresponding half of central encoder. If two descriptions are correctly received at decoder side, data generated from side encoder will be just redundancy and data from central encoder is fed into standard decoder. If one description is lost, the corresponding side decoder is used to decode the received

description. It is obvious that these two descriptions are symmetric and two side coders are independent from each other. Any of them can reconstruct video without drift problem by itself. In the following we will only discuss description 1 consisting of side encoder 1 and central encoder, since the two descriptions are symmetric.

As in fig.2, information from central encoder is splitted into two which are for two descriptions. Based on this existed data, we are expecting to design two side encoders with low redundancy and high flexibility. It is known that the residual data will take most of bitrate for the standard encoder. So result will be better if we utilize the data of central encoder to reduce the bits consumed for residual data. Hence we transfer motion and mode information of central encoder to side encoder and use them to make predictions. This prediction is coarser than central prediction, but residual data will have most correlations with that from central encoder. We do the subtraction between them to get a further residual data. The equations are:
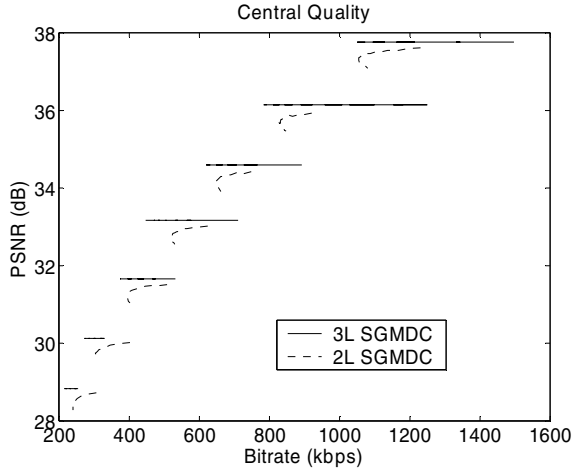
$$e_A{}' = \{\Psi - P(1)\} - Q^{-1}\{IDCT\{Res_A\}\}$$

$$Res_A = Q\{DCT\{\Psi - P(0)\}\}$$
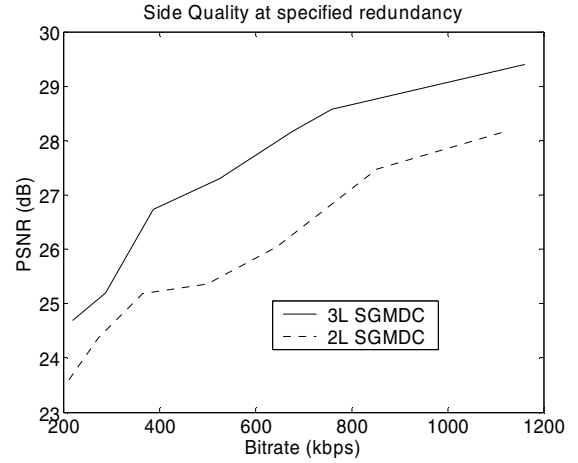
$$Res_A{}' = Q\{DCT\{e_A{}'\}\}$$

if we ignore the effect of quantization, it will be clearer that:

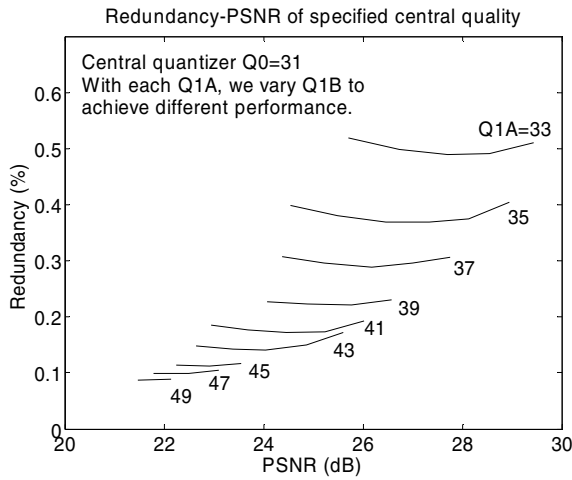$$e_A{}' = \{\Psi - P(1)\} - \{\Psi - P(0)\} = P(0) - P(1)$$

where Ψ is the video data for slice group A fed in, and $P(i)$ is the prediction for slice group A from central encoder (i=0) and side

Central Quality

(a)

Side Quality at specified redundancy

(b)

Figure.3 Performance of 3L-SGMDC. (a) Central quality vs. total bitrate compared to 2L-SGMDC; (b) Side quality vs. total bitrate at redundancy of minimum point of 2L-SGMDC; (c) Redundancy-Quality curve at central bitrate 412.87kbps.

Redundancy-PSNR of specified central quality

Central quantizer Q0=31
With each Q1A, we vary Q1B to achieve different performance.

(c)

encoder (i=1), with the same prediction mode as in central encoder. The residual in side encoder is actually the difference between two predictions in central and side encoder.

At the side decoder side, the reconstruction is got by:

$$Recon = Q^{-1}\left\{IDCT\left\{Res_A^{'}\right\}\right\} + Q^{-1}\left\{IDCT\left\{Res_A\right\}\right\} + P(1),$$

which is the same as reconstruction in the encoder.

Since slice group B takes very few bits, the motion estimation and coding is done separately for it in side encoder.

For each side encoder there are three quantisation parameters affecting the results which are Q0 for central encoder, Q1A for slice group A in side encoder and Q1B for slice group B in side encoder. Q0 will be fixed for specified central quality. Q1A and Q1B can be varied to achieve different side qualities. This brings much flexibility and it is separate to control side and central quality.

We can see that, since the motion compensation in side encoder for main slice group is the further one based on that in central encoder, the residual data to be encoded must be fewer and smaller, and the side result is related to both Q0 and Q1. Another important property of our 3L-SGMDC is that the central decoding is completely standard compliant. Bitstream from central encoder is directly fed into standard decoder and no further processing is

needed.

III. RESULTS AND ANALYSIS

We examine the performance of our proposed 3L-SGMDC. 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 [9] as the basic coder. 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' of CIF format is used. We implement [8] on H.264 for comparison, which is called 2L-SGMDC here.

Central quality of our 3L-SGMDC is presented in fig. 3 (a). It's obvious that the central quality is not affected by the side encoder, as discussed above. It is only associated with the Q0, not like 2L-SGMDC in which the other quantizer for side information, QPB, also affect central quality. Each line is one series of encoding with constant quantizer parameter (QP) of central information, Q0 for 3L-SGMDC and QPA for 2L-SGMDC. From top to bottom this QP is 25 to 37. We can see that with higher QPB, 2L-SGMDC has a little lower quality than 3L-SGMDC. With lowest redundancy 2L-SGMDC is about 0.5dB worse than our 3L-SGMDC. This is mainly because the worse slice group affects accuracy of motion estimation of normal slice group which is part of central information, while our 3L-SGMDC has independent central encoder and will maintain quality as good as normal single-description coding.

Since there are three parameters to control bitrate and quality, it is not easy to compare the performance with 2L-SGMDC. Here we choose the R-D point with the minimum redundancy for each specified central quality of 2L-SGMDC, and find the corresponding values with the same redundancy and central quality in 3L-SGMDC to compare. It is clear in fig.3 (b) that our 3L-SGMDC outperforms 2L-SGMDC much, which is around 1-2 dB.

One problem of 2L-SGMDC is that it cannot achieve very low redundancy, and with lower bitrate, the minimum redundancy will be higher. But it is not the case in our 3L-SGMDC. In fig.3 (c) we specify the central bitrate to be 412.87kbps with central quality 33.16dB. We vary the two quantizers in side encoder to achieve different redundancy. It can be seen that dynamic range of redundancy can be very broad, from very high (higher if we decrease Q1A) to less than 10%. When we increase Q1A, the redundancy is smaller. With same Q1A, the redundancy normally fluctuates in a range of about 10%. Even if the redundancy is very small, the subjective quality of side decoder is still acceptable. Performance of this flexible redundancy control makes it very useful to various applications. But 2L-SGMDC must have fixed quality setting of main slice group with specified central quality, for they are not only part of side information but also part of central information. This decides the performance of it, which has very small range of controlling redundancy. And with lower bitrate, the range is even smaller and redundancy becomes higher.

Fig.4 shows a comparison of pictures of frame number 95 in *paris* sequence. The first one is the decoded frame from central decoder, which has the same quality of the normal single-description coding. The second one is the side decoded frame with only half rate received. The side quality is 27.3 dB, but the subjective quality is still acceptable. For 2L-SGMDC, with the same redundancy, the quality is only 25.4dB.

It should be noted that for other sequences there have similar results.

IV. CONCLUSIONS

In this paper we introduced a novel MDC approach based on slice group coding tools in H.264 with three motion compensation loops. Two slice groups are used of which each composes main information for one side encoder, with the support from central information. Three independent MC loops are maintained which is well suited for ideal MDC environments. It is shown through simulations that our 3L-SGMDC performs very well and is very flexible of redundancy controlling. With half rate received, the side quality is still acceptable and outperforms 2L-SGMDC proposed in [8] for 1-2 dB.

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] 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.

[5] 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.

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

[7] 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.

[8] D. Wang, N. Canagarajah, D. Bull, "Slice Group Based Multiple Description Video Coding Using Motion Vector Estimation," Proc. IEEE Int. Conf. Image Proc, Singapore, Oct, 2004.

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

Central decoding with Q0=31
PSNR=33.2 dB

Side decoding with Q1A=37 and QPB=41
Redundancy=28%   Side PSNR=28.0 dB

Figure. 4   Decoded Pictures of Frame no.95