DRIFT-FREE MULTIPLE DESCRIPTION INTRA VIDEO CODING

Pedro D. F. Correia[1,2], *Pedro A. A. Assunção*[1,3] *and Vitor M. M. da Silva*[1,4]

[1] Instituto de Telecomunicações, Portugal; [2]Instituto Politécnico de Tomar/EST, Portugal
[3]Instituto Politécnico de Leiria/ESTG, Portugal; [4]Universidade de Coimbra/DEEC, Portugal
{*pcorreia, amado, vitor*}*@co.it.pt*

ABSTRACT

This paper proposes a multiple description coding scheme based on a multi-loop structure, which prevents drift distortion accumulation in intra predicted slices. The drift is compensated by generating a controlled amount of side information used by the decoder whenever any description is lost in the corresponding network path. The experimental results show that intra predicted slices do not suffer from drift and their quality is significantly improved (*e.g.*, 6-8 dB) at reduced redundancy cost, (*e.g.*, 0.2-0.25), in comparison with the open loop implementation. Error propagation in subsequent frames is also evaluated when one description is lost for a whole frame and for two rows of macroblocks. In both cases the overall video quality is significantly improved and considering normal GOP sizes the redundancy introduced by the proposed scheme is negligible.

Index Terms— Multiple Description Coding, MDSQ, drift distortion, H.264/AVC

1. INTRODUCTION

Current communication systems such as internet and wireless networks, are characterised by using unreliable channels exposed to high packet loss rates due to either routing delay and congestion or poor wireless channel conditions. In order to cope with such error prone channels the multimedia services based on the most recent video coding standard H.264/AVC [1] include different error resilience tools and coding techniques to ease error recovery at source level, *e.g.*, data partition, redundant slices and flexible macroblock ordering.

A different approach to cope with lossy channels, particularly those that can provide multiple paths between the sender and the receiver, is Multiple Description Coding (MDC) where the encoder produces several independent bitstreams (descriptions) of the same source video [2]. An interesting feature in MDC is the possibility of any single description to be independently decoded. If all descriptions are available, then they are jointly decoded yielding high quality reconstructed video, but if only one description is received, then this is still decodable though achieving lower quality.

In MDC the main problem is to control the amount of redundancy introduced among descriptions to achieve an overall good quality, in comparison with the rate-distortion performance of the traditional single description coding scheme (SDC). The Redundancy-Rate Distortion (RRD) [3] is an MDC performance parameter used to measure the amount of redundancy introduced in MDC compared with SDC at the same quality.

This work has been supported by Fundação Para a Ciência e Tecnologia (FCT), under Grant SFRH/BD/30087/2006

MDC algorithms can be classified into different categories according to the method used for generating the descriptions: MD scalar quantisation (MDSQ) [4], MD by subsampling the source signal in different domains (e.g., spatial, temporal, frequency) [5] [6], MD transform coding using correlating transforms [3], MD of motion information [7] and partitioning of transform coefficients [8].

In general, MDSQ encoders decode all descriptions to reconstruct a high quality frame to form predictions. Therefore, when a description is lost in the network, the predictions reconstructed by the decoder do not match those originally used by the encoder for the same coded blocks. This generates a distortion which is accumulated in the decoder reconstruction loop and propagated throughout all subsequent predicted blocks. This distortion increases with the number of decoded blocks predicted from previous ones. This problem is the origin of drift and it happens in both spatial and temporal predictive coding algorithms. In the case of temporal domain, several multi-loop architectures were proposed to deal with prediction mismatch by sending side information to the decoder, for each description[9].

In recent work, the advanced video coding tools and features provided in H.264/AVC to form multiple MD schemes was used. In [10] a slice group scheme is presented with three motion compensation loops. The video signal is normally encoded in the central encoder and then divided into two descriptions, each one corresponding to one slice group. Each slice group includes redundant information from the other one. A similar approach is proposed in [11] where the temporal and spatial correlations between macroblocks are exploited to achieve efficient redundancy coding. In [12] the redundant slice feature of H.264/AVC is exploited in order to form two different descriptions with controlled redundancy and in [13] an H.264/AVC MDSQ open loop scheme is proposed without drift control assuming that I frames are received without errors.

Most of the MD coding schemes developed so far, deal with the temporal prediction problems assuming that intra frames are always correctly decoded. Since H.264/AVC intra frame coding is highly predictive, the error accumulation by drift cannot be ignored because it is also present when decoding such frames. In case of transmission error or packet loss, the poor quality of an intra frame will be propagated through all dependent pictures. Therefore it is necessary to guarantee drift free intra frames in order to avoid distortion propagation throughout the video sequence. This paper proposes a three loop drift-free MDSQ scheme based on H.264/AVC, which can be used to increase error robustness of intra frames in multipath communication environments.

The paper is organized as follows. Section 2 describes the H.264/AVC-based MDSQ architecture. Section 3 describes the drift effect in intra prediction, section 4 presents the experimental results of the proposed scheme. Finally, section 5 concludes the paper.

2. DRIFT-FREE INTRA MDSQ ARCHITECTURE

The proposed H.264/AVC intra drift-free MD video encoding scheme is based on a three prediction loop architecture as shown in figure 1. In order to keep the block diagram simple only intra coding related functions are shown. This MD architecture produces two distinct independent decodable H.264/AVC bitstreams that can be combined in order to achieve different output quality levels. For each of this streams, the encoder also produces side information for drift compensation. Only the index assignment module of the MDSQ process is non-normative. The side information can be transmitted as redundant slices and then it can be used by the decoder to prevent drift.

The drift-free MD encoder of figure 1 is composed by one central encoder and two side encoders. The central encoder includes the MDSQ module that produces two descriptions and the corresponding streams. These two distinct bitstreams are run-length and entropy encoded using CAVLC. All the syntax elements, namely headers and prediction modes are duplicated in both descriptions.

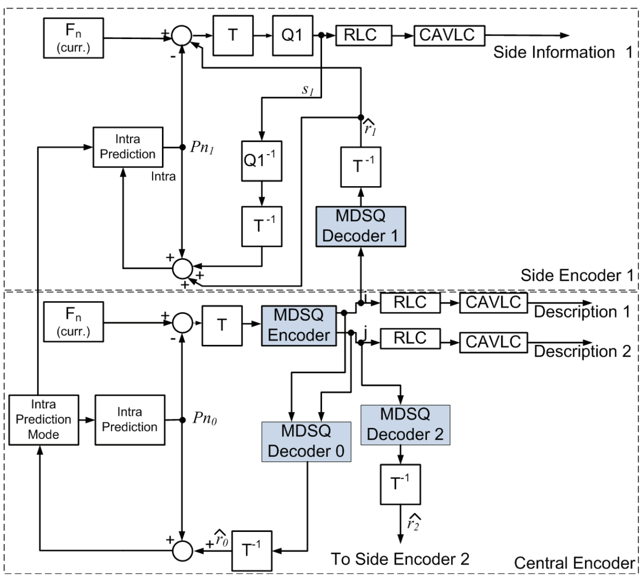

Fig. 1. MDSQ intra encoder with drift compensation.

The two side encoders control the redundant information generated at the output in order to guarantee a drift free acceptable decoded frame. The side information s_i can be defined as,

$$s_i = Q_i\{T\{F_n - P_{n,i} - \widehat{r}_i\}\}, i = 1, 2 \tag{1}$$

Q_i is the quantisation operation with side quantiser QP_i that determines the amount of redundant information, T is the transform operation. F_n is the current frame and $P_{n,i}$ are the prediction value from each respective side encoders. $\widehat{r}_i$ is defined as

$$\widehat{r}_i = T^{-1}\{Q_0^{-1}\{A_i^{-1}(r_i)\}\}, i = 1, 2 \tag{2}$$

which represents the residue available at decoder if only one description is correctly received. A_i^{-1} represents the inverse index assignment operation if only one description exists. In this case it is necessary to define a reconstruction rule for each side index because each one corresponds to several possible central indices. In this paper, the main diagonal central value is used as reconstruction rule. Q_0^{-1} is the inverse quantisation with the central encoder quantisation parameter QP_0.

At the decoder, if both descriptions are available then an Inverse Index Assignment is made to restore the central description index. Then inverse quantisation and inverse transform are applied. If some description is lost then the side decoder will decode a drift free frame, though with poorer quality than if both descriptions were received. The reconstructed frame $\widehat{F}_{n,i}$ can be defined as

$$\widehat{F}_{n,i} = P_{n,i} + \widehat{r}_i + T^{-1}\{Q_i^{-1}\{s_i\}\}, i = 1, 2 \tag{3}$$

.

The drift free decoded frames are used to decode the subsequent central frames whenever both descriptions are correctly decoded. The decoder architecture is represented in figure 2.

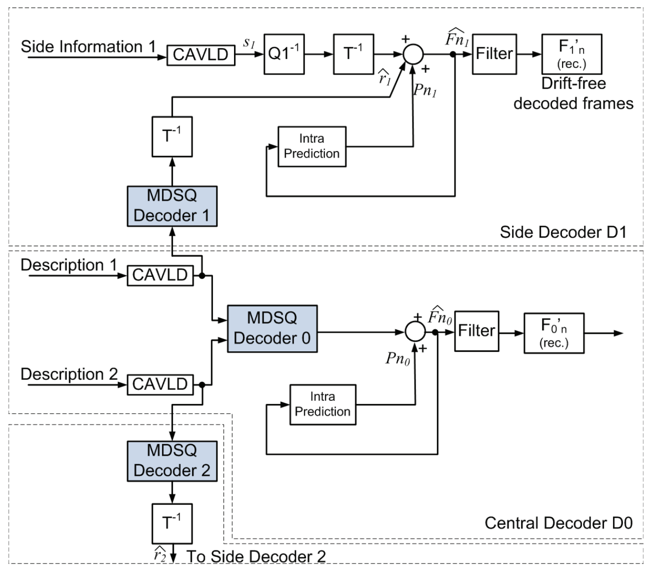

Fig. 2. MD intra decoder.

3. DRIFT ANALYSIS IN INTRA SLICES

As mentioned before, if side information does not exist and any description is lost, then a decoding mismatch occurs in intra predicted blocks because the original predictions generated in the central encoder loop can no longer be replicated in the decoder reconstruction loop. This is because the reconstruction rule for a single description gives coefficient values that are different from the ones computed in the central encoder. Therefore, different prediction blocks are generated and consequently drift distortion is accumulated when decoding intra slices.

Considering coefficient block k, defining $\widehat{r}_0$ as the decoded residue of the central decoder and $\widehat{r}_i$, $i = 1, 2$, the residue resulting from MDSQ decoding with only one description without side information, we can define

$$e_k = (\widehat{r}_0 - \widehat{r}_i) + (P_{n,0} - \widehat{P}_{n,i}), \tag{4}$$

as the error between the central and side decoders without side information. Since drift distortion is due to error propagation resulting from intra prediction mismatch and not because of the error generated immediately after inverse quantisation, $\widehat{P}_i$ carries the error accumulated from previous blocks. One can define a drift measure for one macroblock, assuming 4×4 block prediction as the following mean square error expression:

$$drift = \frac{1}{16} \sum_{j} (P_{n,0j} - \hat{P_{n,ij}})^2 \quad (5)$$

where $P_{n,0}$ and $\widehat{P_{n,i}}$ are prediction values of the central and side decoders respectively.

Figure 3 shows the drift distortion of one frame for central quantiser $QP_0 = 16$. As shown in the figure, the effect of drift is very obvious since the distortion increases as more macroblocks are decoded and then resets after each row of macroblocks. Such reset is because the first macroblock of each row does not use predictions from the last macroblocks of previous rows.

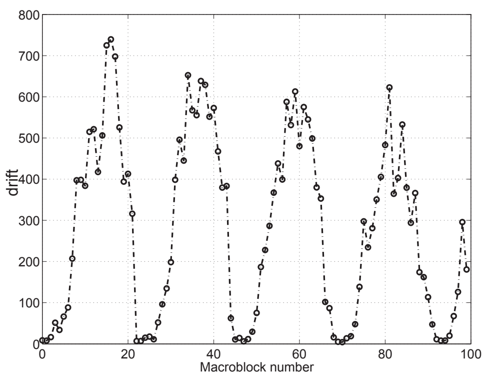

Fig. 3. Macroblock drift distortion for one intra frame-*coastguard* sequence.

4. EXPERIMENTAL RESULTS

The proposed MDC scheme is implemented on JM11.0, the reference software of H.264/AVC. All intra prediction modes are enabled and the GOP structure is IPPP... without B-frames. The test sequences are of CIF resolution and frame rate 30Hz.

Figure 4 shows the PSNR *vs* redundancy for the *coastguard* sequence using MDSQ with 3 diagonals in the worst case scenario where one description is lost for the entire frame. Two cases are compared: *a)* with and *b)* without drift compensation. Both cases use the same central quantiser QP_0. Drift compensation is done by adding the side information generated with $QP_i, i = 1, 2$ which also determines the amount of side redundancy. The curves were obtained for different combinations of QP_0 and $QP_i, i = 1, 2$ and it was assumed that $QP_1 = QP_2$. The PSNR-redundancy pairs were computed for QP_0 =10,16 and 20 and for each of these, the range of quantisation parameters QP_1, QP_2 shown in the figure was used. The redundancy generated depends on the quantisation parameters and the number of diagonals of the index assignment matrix. The isolated points shown in figure 4 correspond to the values of redundancy and PSNR obtained without side information for QP_0 =10,16 and 20. These results show that open loop decoding quality is greatly affected by drift distortion.

From the figure one can realise that, by adding side information at the expense of a small increase in redundancy, the decoded picture quality is significantly improved. For example, for $QP_0 = 10$, the redundancy increase is about 0.2 and the decoded quality has almost

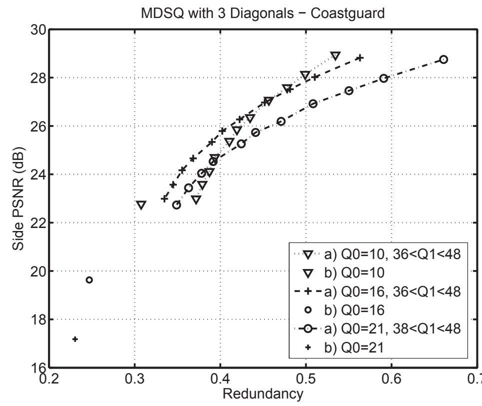

Fig. 4. Effect of side information: PSNR vs redundancy.

6dB gain. For higher $QP_0 = 16$ and redundancy of 0.25 more than 8dB gain is obtained. These redundancy values are in line with other recent results for drift control in temporal dependency [10], [11]. Note that this amount of redundancy only refers to one intra frame which, in general occurs every 15 or more frames, depending on the GOP size. Taking into account that all subsequent frames are dependent on the initial I frame of a GOP, the whole sequence benefits from drift compensation in the I frame.

Table 1 shows the redundancy introduced by an intra frame in the whole GOP for different GOP sizes. I frame and subsequent P frames are encoded with the same QP_0 and intra side information redundancy is set varying QP_1 and QP_2. The results show that little redundancy is added to the GOP. For instance, for 0.53 obtained with $QP_0 = 16$ and $QP_1 = 38$ of intra redundancy only 0.082 of overall redundancy is added in a GOP size of 10 frames and 0.029 in a GOP of 30 frames. Therefore, the overall redundancy introduced by the proposed MDSQ scheme is negligible for normal GOP sizes.

I frame redundancy	GOP redundancy		
	GOP=10	GOP=15	GOP=30
0.44	0.058	0.039	0.020
0.53	0.082	0.054	0.029
0.63	0.110	0.076	0.033

Table 1. Redundancy in the whole GOP for I frame redundancies of 0.44,0.53 and 0.63.

The effect of drift compensation in an intra frame is shown in figure 5 where the PSNR of each macroblock is compared for both cases, i.e., *a)* with and *b)* without drift compensation using $QP_0 = 16$ and $QP_1 = 38$ for case *a)* and $QP_0 = 16$ for case *b)*. The figure clearly shows that drift compensation produces much smoother PSNR along the frame than in the case where such compensation is not done. In the case of no drift compensation, the lowest PSNR is below 15dB, which is definitely not acceptable.

The other performance aspect evaluated for the proposed scheme was the influence of drift compensation in decoding a whole GOP as shown in figures 6 and 7, assuming that P frames were decoded without errors. As we can seen in the figure 6, the gain in quality is about 8dB. Figure 7 shows the PSNR considering that only the second and sixth macroblock rows are lost. Again, it can be seen that even with such a small loss the quality gain is still significant, about 7dB.

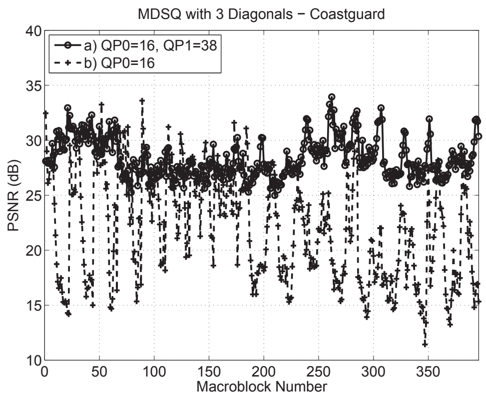

Fig. 5. Macroblock PSNR comparison: *a)* with drift compensation; *b)* without drift compensation.

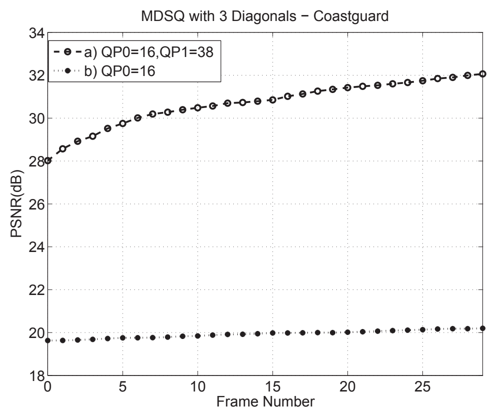

Fig. 6. One description is lost for one intra frame: *a)* with drift compensation; *b)* without drift compensation.

5. CONCLUSIONS

In this paper a drift-free H.264/AVC MD scheme for intra slices was presented. Two independent decodable descriptions are produced using MDSQ and side information to compensate the drift distortion. Experimental results show that the proposed MDC scheme can improve the quality of intra slices about 8dB at the expense of an acceptable redundancy increase. Drift compensation in intra slices is also shown to provide significant quality gains (*e.g.* 7-8dB) in the whole sequence. Therefore, the proposed scheme might be useful in MDSQ video encoding.

6. REFERENCES

[1] ITU-T-H.264, "Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T rec. H.264/ISO/IEC 14496-10 AVC)," 2003.

[2] Vivek Goyal, "Multiple description coding: Compression meets the network," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 74–93, September 2001.

[3] Yao Wang, Michael T. Orchard, Vinay Vaishampayan, and Amy R. Reibman, "Multiple description coding using pairwise

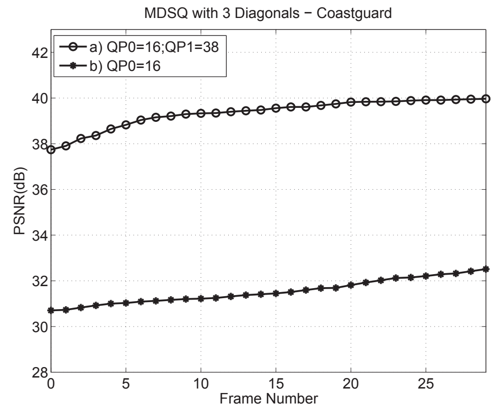

Fig. 7. One description is lost for the 2nd and 6th macroblock rows: *a)* with drift compensation; *b)* without drift compensation.

correlating transforms," *IEEE Transactions On Image Processing*, vol. 10, no. 3, pp. 351–366, March 2001.

[4] Vinay Anant Vaishampayan, "Design of Multiple Description Scalar Quantizers," *IEEE Transactions On Information Theory*, vol. 39, no. 3, pp. 821–834, May 1993.

[5] I. V. Bajic and J. W. Woods, "Domain based multiple description coding of images and video," *IEEE Transactions on Image Processing*, vol. 12, no. 10, pp. 1211–1225, October 2003.

[6] T. Petrisor, B. P.-Popescu, and J.-C. Pesquet, "Perfect reconstruction in reduced redundancywavelet-based multiple description coding of images," *Proc. of EUSIPCO, Antalia, Turkey*, September 2005.

[7] Chang-Su Kim and Sang-Uk Lee, "Multiple description coding of motion fields for robust video transmission," *IEEE Transactions On Circuits and Systems For Video Technology*, vol. 11, no. 9, pp. 999–1010, September 2001.

[8] Kiran R. Matty and Lisimachos P. Kondi, "Balanced multiple description video coding using optimal partitioning of the dct coefficients," IEEE *Transactions On Circuits and Systems For Video Technology*, vol. 15, no. 7, pp. 928–934, July 2005.

[9] Yao Wang, Amy R. Reibman, and Shuman Lin, "Multiple description coding for video delivery," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 57–70, January 2005.

[10] N. Canagarajah D. Wang and D. Bull, "Slice group based multiple description video coding wth three motion compensation loops," *Proceedings of the* IEEE *International Sysmposium on Circuits and Systems, Kobe, Japan*, pp. 960–963, May 2005.

[11] Che-Chun Su, Homer H. Chen, Jason J. Yao, and Polly Huang, "H.264/avc-based multiple description video coding using dynamic slice groups," *Image Commun.*, vol. 23, no. 9, pp. 677–691, 2008.

[12] Marco Grangetto Tamman Tillo and Gabriel Olmo, "Redundant slice optimal allocation for h.264 multiple description coding," IEEE *Transactions On Circuits and Systems For Video Technology*, vol. 18, no. 1, pp. 59–70, January 2008.

[13] Gian Antonio Mian Ottavio Campana, Roberto Contiero, "An h.264/avc video coder based on a multiple description scalar quantizer," IEEE *Transactions On Circuits and Systems For Video Technology*, vol. 18, no. 2, pp. 268–272, February 2008.