Multiple Description Coding Scheme for H.264/AVC Intra Slices

Pedro D. F. Correia[†,‡], *Pedro A. A. Assunção*[†,ℏ] *and Vitor M. M. da Silva*[†,§]

[†] Instituto de Telecomunicações, [§]Universidade de Coimbra/DEEC
Pólo II, FCTUC, Pinhal de Marrocos, 3030-290 Coimbra,Portugal
[‡]Instituto Politécnico de Tomar/EST, Portugal; [ℏ]Instituto Politécnico de Leiria/ESTG, Portugal
e-mail:{*pcorreia, amado, vitor*}*@co.it.pt*

***Abstract* — This paper proposes a Multiple Description Scalar Quantisation (MDSQ) coding scheme for H.264/AVC intra slices which is based on a multi-loop structure to prevent distortion accumulation at the decoder due to drift in intra predicted blocks. The drift distortion is reduced by sending a controlled amount of redundant information to be used whenever any of the descriptions fails to reach the decoder. The experimental results show that the quality of intra coded slices is significantly improved (*e.g.*, 6-8 dB) at reduced redundancy cost, (*e.g.*, 0.2-0.25), in comparison with the open loop MDSQ implementation.**

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 these error prone channels, most multimedia services based on the newest advanced video coding standards, such as H.264/AVC [1], include different error resilience tools and coding techniques to 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]. At the receiver side any of these descriptions can be independently decoded. If all descriptions are available, then they are combined to achieve high quality reconstructed video, but a lower quality, though still acceptable, is obtained if only one description is decoded.

When transmission errors occur in any description, it is necessary to deal with the mismatch problem between the encoder and decoder predictions. This applies to both spatial and temporal predictive coding algorithms. In order to deal with prediction mismatch in the temporal domain, several multi-loop architectures were proposed to send side information for each description to the target decoder [3].

Recent work use the advanced video coding tools and features provided in H.264/AVC to form multiple description encoding schemes. In [4] a slice group scheme is presented with three motion compensation loops. All video data is normally encoded in the central encoder and 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 [5] where the temporal and spatial correlations between macroblocks are exploited to achieve efficient redundancy coding. In [6] the redundant slice feature of H.264/AVC is exploited in order to form two different descriptions with controlled redundancy and in [7] 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 address the temporal prediction problem and assume that intra frames are always correctly decoded. Since H.264/AVC intra frame coding is highly predictive, the error accumulation by drift 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 intra frames with good quality in order to avoid distortion propagation along the video sequence. Considering that H.264/AVC Baseline profile was designed for low delay applications in packet loss environments, as well as for platforms with low processing power, it is crucial to protect intra frames in order to achieve an acceptable quality at end user terminals. This paper proposes a three loop MDSQ scheme based on H.264/AVC, which can be used to increase error robustness of intra frames. The proposed scheme can be also applied to either P or B frames.

The paper is organized as follows. In section 2 the MDSQ coding mechanism is introduced. Section 3 describes the H.264/AVC based MDSQ architecture. Section 4 describes the drift effect in intra prediction and section 5 presents the experimental results. Section 6 concludes the paper.

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

2. MULTIPLE DESCRIPTION SCALAR QUANTISATION

In MDSQ the main principle is to construct two uniform quantisers with coarse reconstruction values if separate inverse quantisation is done. If joint inverse quantisation is performed then finer reconstruction values are obtained. Considering only two different descriptions, the goal of MDSQ is to find coarse scalar quantisers for each side decoder such that after being combined together a finer central quantiser is obtained producing lower distortion than each of the side decoders.

Figure 1 shows the MDSQ process which is based on two different functions: central quantisation and index assignment. In the first function, the sample x is quantised using the central quantiser which generates a quantiser index l.

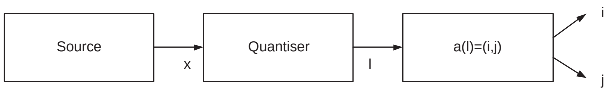

Fig. 1. Generic MDSQ Scheme.

The index assignment function consists in mapping l into a two dimensional space (i, j). The index assignment is often represented by an index assignment matrix, whose elements are the central quantiser index, while the output is an index pair comprised of the respective column and row indices. The index assignment matrix is characterised by the index spread which is the number of diagonals and by the scanning path of the central indices in the matrix. The combination between both of these functions determines the amount of redundancy and the rate of each description.

Vaishampayan [8] proposed some heuristic methods in order to design good index assignments. The scanning methods are the linear and the nested scan. These methods use balanced descriptions, which mean that the rates and distortions resulting in both descriptions are identical. The amount of redundancy is controlled by a spread parameter k where $2k + 1$ is the number of diagonals in the index matrix. In order to increase the encoder efficiency by mapping small values onto zeros, a modified index assignment nested scan is employed, as shown in the functional diagram of figure 2.

Description 1 \ Description 2	-4	-3	-2	-1	0	1	2	3	4	5
-4	-20	-16	-14	0	0	0	0	0	0	0
-3	-17	-15	-12	-8	0	0	0	0	0	0
-2	-13	-11	-10	-6	-4	0	0	0	0	0
-1	0	-9	-7	-5	-1	0	0	0	0	0
0	0	0	-3	-2	0	1	4	0	0	0
1	0	0	0	0	2	5	6	8	0	0
2	0	0	0	0	3	7	10	12	14	0
3	0	0	0	0	0	9	11	15	16	18
4	0	0	0	0	0	0	13	17	20	...
5	0	0	0	0	0	0	0	19	...	...

Fig. 2. Modified Nested Index assignment with k=2.

3. H.264/AVC-BASED MDSQ ENCODER/DECODER ARCHITECTURE

The proposed H.264/AVC drift-free MD video encoding scheme is based on a three prediction loop architecture as shown in figure 3. 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 reduce the drift distortion.

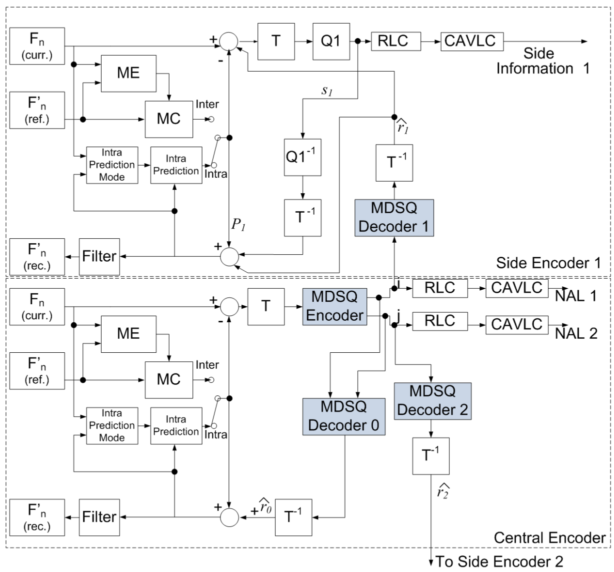

Fig. 3. Multiple Description Video Encoder Architecture

The drift free MD encoder of figure 3 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, prediction modes and motion vectors are duplicated in both descriptions.

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

$$s_i = Q_i\{T\{F_n - P_i - \widehat{r}_i\}\}, i = 1, 2 \quad (1)$$

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_i is the prediction value from each respective side encoders. $\widehat{r}_i$ is defined as

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

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.

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 can be defined as

$$\widehat{F}_i = P_i + \widehat{r}_i + T^{-1}\{Q_i^{-1}\{s_i\}\}, i = 1, 2 \quad (3)$$

.

The side prediction frame is used to decode the subsequent central frames whenever both descriptions are correctly decoded. The decoder architecture is represented in figure 4.

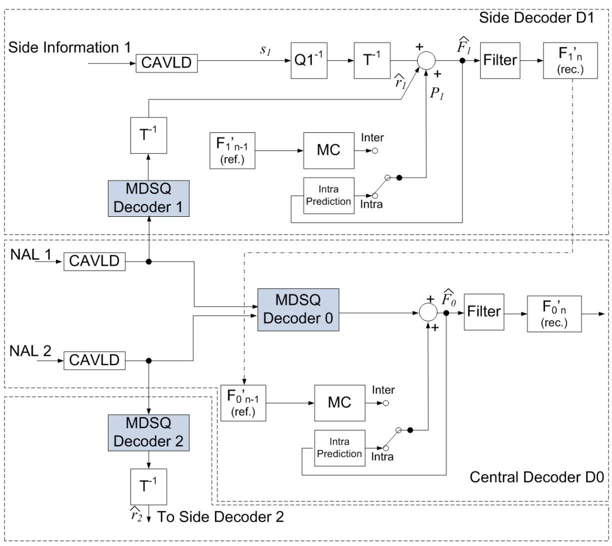

Fig. 4. Multiple Description Video Decoder Architecture

4. 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 one description gives a coefficient value that is different from the one computed in the central encoder. This will generate different prediction blocks and consequently drift distortion when decoding intra slices.

Considering the 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

$$e_k = (\widehat{r}_0 - \widehat{r}_i) + (P_0 - \widehat{P}_i), \quad (4)$$

as the error between the central and side decoders without side information for macroblock k. P_0 and $\widehat{P}_i$ are prediction values of the central and side decoders respectively. 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 accumulation throughout all decoded blocks. One can define a drift measure for one macroblock, assuming 4×4 block prediction as the following mean square error expression:

$$drift = \frac{1}{16} \sum_j (P_{0j} - \hat{P}_{ij})^2, i = 1, 2 \quad (5)$$

Figure 5 represents the drift distortion of one frame for central QP0=16.

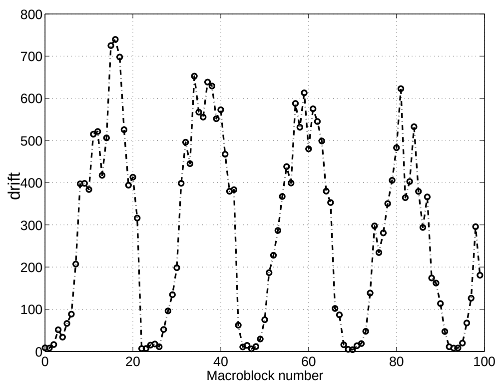

Fig. 5. Macroblock drift distortion for one intra frame-*coastguard* sequence.

As shown in the figure, the effect of drift is obvious since the distortion increases as more macroblocks are decoded and 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.

5. EXPERIMENTAL RESULTS

The proposed MDC scheme was implemented on JM11.0 reference software of H.264/AVC. Figure 6 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: with and without drift compensation. Both cases use the same central QP (QP0). The drift compensation is done by adding the side information generated with either QP1 or QP2, which also determines the amount of side redundancy. The results are obtained for different combinations of QP0 and side QP values. The use of some central QP will generate one natural MDSQ redundancy value that depends on the quantisation parameter range and the number of diagonals of the index assignment matrix. The isolated points shown in figure 6 correspond to the values of redundancy and PSNR obtained without side information for QP0=10,16 and 21. These results show that open loop decoding quality is greatly affected by drift distortion.

From the figure one can see that by adding side information at the expense of a small increase in redundancy, the decoded picture quality is significantly improved. For example, with QP0=10, the redundancy increase is about 0.2 and the decoded quality has almost 6dB gain. For higher QP0=16 and with an increase of 0.25 in redundancy more than 8 dB gain is obtained. These redundancy values are in line with other recent results for drift control in temporal dependency [4], [5]. Note that this amount 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. Therefore, the overall redundancy introduced by the proposed MDSQ scheme is negligible for normal GOP sizes.

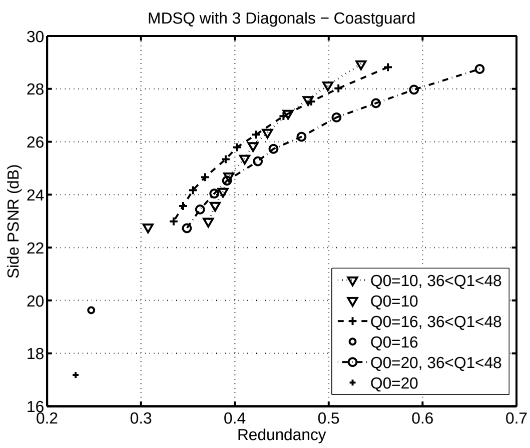

Fig. 6. Effect of side information: PSNR $vs.$ redundancy.

6. 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 8 dB at the expense of an acceptable redundancy increase, e.g., 0.25. Since intra slices greatly influence the overall quality of the video sequence the proposed MDSQ scheme might be useful to improve error resilience in multipath communications environments.

7. 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, 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

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