Redundant slice optimal allocation for H.264 multiple description coding

Tammam Tillo, *Member, IEEE,* Marco Grangetto, *Member, IEEE,*
Gabriella Olmo, *Senior Member, IEEE*

Abstract—In this paper a novel H.264 multiple description technique is proposed. The coding approach is based on the redundant slice representation option, defined in the H.264 standard. In presence of losses, the redundant representation can be used to replace missing portions of the compressed bitstream, thus yielding a certain degree of error resilience. This paper addresses the creation of two balanced descriptions based on the concept of redundant slices, while keeping full compatibility with the H.264 standard syntax and decoding behavior in case of single description reception. When two descriptions are available still a standard H.264 decoder can be used, given a simple pre-processing of the received compressed bitstreams. An analytical setup is employed in order to optimally select the amount of redundancy to be inserted in each frame, taking into account both the transmission condition and the video decoder error propagation. Experimental results demonstrate that the proposed technique favorably compares with other H.264 multiple description approaches.

Index Terms—Multiple description coding, H.264, rate allocation

I. INTRODUCTION

VIDEO compression is mainly based on block-based motion prediction and compensation, which exploits the temporal correlation between subsequent frames. However, the presence of a prediction loop in the video codec makes the compressed sequence highly vulnerable to errors, due to the dependency among successive coded frames. As a consequence of the possible error propagation, video transmission over networks subject to packet losses is very challenging. In order to overcome this vulnerability, one can consider the use of multiple description coding (MDC), where different non hierarchical representations (or *descriptions*) of the same data, yielding mutually refinable information, are generated in order to be transmitted over independent paths. In this paper we will limit ourself to the practical situation of two descriptions.

The most popular methods to generate MD are based on the pioneering MD scalar quantizer (MDSQ) proposed in [1]; this principle has been applied to video coding in [2]. Another class of methods employs correlating transforms

[3]; this approach has been applied to motion compensated MD video coding in [4]. These methods, although providing good performance, are conceived as stand alone codecs, and generate descriptions that are incompatible with video standards such as H.264 [5]. This consideration has led to the proposal of MDC schemes that can be configured as pre- and post-processing stages to be associated with *any* video co-decoder. Among them we can mention [6], [7], [8], where the correlation among descriptions is obtained by means of 1 or 2-dimensional oversampling of the original image by zero padding in the DCT transform domain; the inverse transform is then operated, and the resulting image is split into sub-images representing the descriptions. In [9], the 1-dimensional zero-padding scheme proposed in [10] for images is generalized to video, with a frame by frame approach; the scheme is shown to outperform 2-D oversampling for video, as explained in [11]. In [12], the authors suggest to use the *slice group* coding tool available in H.264 in order to create two balanced description. Although the generated descriptions are indeed H.264 compliant, the use of the slice group modality impairs the compression efficiency. In order to mitigate this effect, in [13] the authors suggest using three-loop slice group MDC; nevertheless this latter solution still exhibits a performance impairment in terms of coding efficiency at the two side encoders.

In this paper we propose a novel MDC coding approach that generates descriptions that are fully compliant with the H.264 video coding standard. This goal is achieved by exploiting the redundant slice coding option available in the standard. Nevertheless, the proposed approach is general and can be used with any other hybrid video codec even if, in such a case, compatibility with the original standard definition may not be guaranteed. The most important contributions of the present work are the derivation of an optimal redundancy allocation strategy and its implementation within an open loop rate control technique, based on the selection of the quantization parameter. It is worth noticing that the proposed technique is able to optimally allocate the MDC redundancy according to the network status and the error propagation characteristic of the employed codec.

The rest of the paper is organized as follows. In Section II the basic MDC principles are recalled. The proposed method is described in Sect. III. Sect. IV presents the analytical optimization approach and its actual implementation using the H.264 syntax. In Sect. V results are presented. Finally, in Section VI some conclusions are drawn.
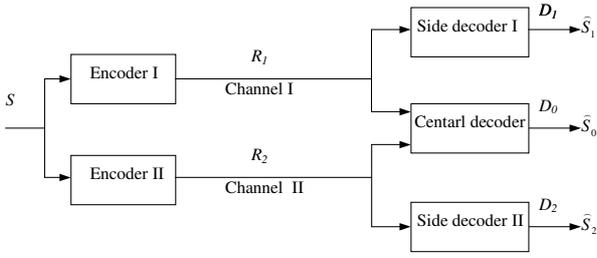
Figure 1.   MD co-decoder (2 descriptions).

II. PRINCIPLES OF MULTIPLE DESCRIPTION

In the MDC approach, two or more independently decodable representations (*descriptions*) of the same data are generated. If these descriptions are transmitted over independent channels and if any subset of them is received, the *side decoder* can reconstruct the data with a given *side distortion*. The more descriptions are received, the higher is the quality of the decoded signal. The minimum distortion *(central distortion)* is obtained when all descriptions are received [14]. The general scheme of a MD system with two descriptions, which is the practical situation addressed in this paper, is reported in Fig. 1.

In case of 2 descriptions, the performance of the technique can be evaluated in terms of the central distortion D_0, and the two side distortions D_1, D_2 when either description is received, as functions of the total bit rate $R_1 + R_2$, where R_1, R_2 are the bit rates devoted to the encoding of either description. Alternatively, for image and video applications, the central and side peak signal-to-noise ratios, which are related to the image visual quality, can be evaluated. In case of *balanced descriptions*, as it is always assumed in the following, we can assume $D_1 = D_2$, or, equivalently, $PSNR_1 = PSNR_2$ and $R_1 = R_2$. The same quality yielded by the central decoder can be obtained by the reference single description coding (SDC) scheme with a rate $R_1 + R_2 - \Delta R$, where ΔR is the extra rate introduced by the MD scheme as the price to be paid in order to allow for multiple quality levels. For a practical MD scheme it is very important to offer an easy way to tune this redundancy as a function of the actual network conditions.

III. PROPOSED ALGORITHM

The proposed algorithm aims at creating two balanced descriptions of a video sequence, each of which is an H.264 compliant bitstream. This objective is fulfilled using the concepts of *primary* and *redundant* slices, defined in the H.264 standard [15]. Primary slices are used to code the primary picture, and are associated to a normative decoding procedure. On the other hand, redundant slices represent an alternative representation of a picture; the H.264 recommendation does not specify a normative decoder behavior in presence of redundant slices. Clearly, when some of the samples in the decoded primary picture cannot be correctly decoded due to errors or transmission losses, whereas the redundant slice can be correctly decoded, the decoder shall replace the samples of the decoded primary picture with the corresponding ones of the decoded redundant slice.

The proposed coding approach is summarized in Fig. 2. First of all, a single H.264 bitstream is generated, where a redundant representation of each coded picture is encoded. A redundant picture, which is generally composed of several redundant slices, is obtained employing a quantization parameter QP_r, different from that used in the corresponding primary picture. The value of the redundant QP_r parameter allows one to shape the amount of coding redundancy inserted in the compressed bitstream. It is important to notice that each redundant P slice is predicted using only the previously encoded primary slices.

In other words, at the encoder side, the redundant slices are not used to predict the subsequent pictures. At the decoder side, when a primary slice is not available, its redundant counterpart is used as a backup; as a consequence, because of the mismatch between the primary and the redundant representations, an error is introduced in the decoder prediction loop. The error due to the decoding of a redundant slice propagates for a certain number of pictures; nevertheless, in the H.264 case the generated decoder drift does not produce strong artifacts, as reported in [16] for the case of bitstream switching. The objective of the proposed algorithm is the development of an effective allocation strategy able to tune the coding redundancy so as to limit the decoder drift when redundant slices need to be used because of transmission losses. Clearly, the optimal redundancy allocation is a function of the transmission conditions and the amount of drift distortion.

The obtained H.264 bitstream with redundant slices can be used to form two balanced descriptions of the original video sequence, by simply rearranging the compressed data.

In particular, the descriptions are formed by interlacing primary and redundant slices, so as to create two H.264 bitstreams which contain alternatively the primary and the redundant representation of each slice, as depicted in Fig. 2. At the decoder side, if both descriptions are received, the decoding of the primary representation of every slice is guaranteed. On the other hand, if a description is lost, the received one is a compliant H.264 bitstream, that can be decoded yielding inferior quality because of the drift generated by the redundant slices. The distortion experienced by the side decoder can be controlled by tuning the quality of the redundant slices, i.e. the value of QP_r.

It is worth noticing that, as in every MDC scheme, the introduced redundancy is beneficial in the case of single description reception, whereas it impairs the overall performance in case of two descriptions reception. Therefore, the introduced redundancy should match the network condition in order to achieve the best average performance. Moreover, the redundancy allocated to each frame should take into account the picture position within the *Group of Pictures* (GOP). In fact, a decoding error propagates till the next I frame; as a consequence, errors and losses that occur on the first frames of the GOP have a higher impact on the distortion. This observation suggests that the coding redundancy should be carefully tuned over the whole GOP.
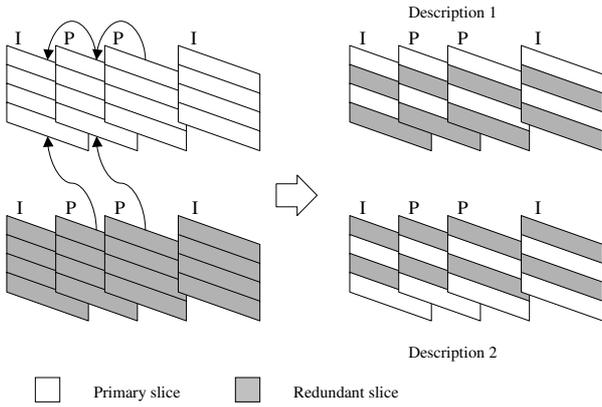
Figure 2. Multiple description scheme for H.264 using redundant slice.

IV. OPTIMAL ALLOCATION PROBLEM

In this section we develop a mathematical framework for the redundancy allocation problem.

In order to optimize the amount of redundancy of the MDC scheme, one needs to estimate the expected decoder distortion at coding time, depending on the video characteristics, encoding modes and network state. The well known results in [17], [18] show that it is possible to accurately model the expected distortion at pixel level; nevertheless, these techniques require intense computation. In the following of the present paper, the redundancy will be tuned at frame level. Given this objective, an approach aiming at modeling the expected distortion at frame level is more adequate and preferable in terms of computational cost. In [19] the decoder expected distortion is shown to exhibit a frame level statistical behavior; based on this observation, the expected distortion is estimated on a frame basis and applied to the problems of adaptive intra mode selection and source/channel rate control. A similar approach is adopted in [20]. In this section, the estimation of the expected distortion in presence of redundant slices is tackled from the same standpoint.

Let us assume that the k-th primary slice in frame $i = M(k)$ has been lost, where $M(\cdot)$ is a many-to-one function that maps the slice index over the corresponding frame index. We want to evaluate the impact of this loss on the total distortion of the current GOP with Intra period N. This distortion $d_{t,k}$, caused by the loss of the k-th primary slice, can be evaluated as:

$$d_{t,k} = d_{r,k} + \sum_{j=M(k)+1}^{N} d_{m,k}[j],$$

where $d_{r,k}$ is the distortion due to the k-th redundant slice, which is used to replace the lost primary slice, and $d_{m,k}[j]$ is the propagated distortion at time j due to the mismatch generated in frame $M(k)$. The propagated distortion decays over time due to leakage in the prediction loop and, according to the model proposed in [21], we can write

$$\sum_{j=M(k)+1}^{N} d_{m,k}[j] = d_{m,k} \sum_{n=1}^{N-M(k)} f[n]$$

where $d_{m,k} = d_{m,k}[M(k) + 1]$ and $f[n]$ is the so called *power transfer function*. This latter function is used to model the distortion attenuation experienced at distance n from the mismatch. The distortion decay is determined by the fact that some macroblocks (MB) are encoded in Intra mode, by the presence of the de-blocking loop filter and because of the spatial filtering used for sub-pixel motion estimation. In fact, both the de-blocking filter and the spatial interpolation attenuate the high spatial frequency components of the superimposed transmission error [21], [22].

The term $d_{m,k}$ can be approximated as shown in Appendix A:

$$d_{m,k} \approx \begin{cases} d_{r,k}\left[1 - \left(\frac{\Delta_{p,k}}{\Delta_{r,k}}\right)^2\right] & \text{for aligned quantizers} \\ d_{r,k} & \text{for not aligned quantizers} \end{cases} \tag{1}$$

where $\Delta_{r,k}$ and $\Delta_{p,k}$ are the redundant and primary quantization steps. As far as $\Delta_{r,k} \gg \Delta_{p,k}$, the total GOP distortion can be approximated as:

$$\begin{aligned} d_{t,k} &= d_{r,k} + d_{m,k} \sum_{n=1}^{N-M(k)} f[n] \\ &\approx d_{r,k} \sum_{n=0}^{N-M(k)} f[n] = d_{r,k}\phi_{M(k)} \end{aligned}$$

where $\phi_{M(k)} = \sum_{n=0}^{N-M(k)} f[n]$, with $f[0] = 1$. In the following we assume that every slice fits into the *maximum transfer unit* (MTU), which is transmitted over a network characterized by independent packet losses with probability p. Moreover, p is considered constant for the transmission of a whole GOP. This assumption is not critical, since estimates of p are usually obtained through control packets, e.g. *Real Time Control Protocol* (RTCP) packets, which are periodically sent across a feedback channel to probe the network status. Clearly, it is unfeasible to refresh network statistics for the transmission of a single frame, i.e. 15/30 reports per second. As an example, the interval between RTCP packets is required to be larger than 5 seconds [23]. At the same time, in the case of video streaming applications through unreliable links, it is reasonable to use GOP length in order of seconds. Under the hypothesis that p is known, each primary slice is received with probability $p(1 - p) + (1 - p)^2 = (1 - p)$. Analogously, the redundant slice is received alone with probability $p(1 - p)$. Otherwise, with probability p^2, no slice representation is received.

As a consequence, the expected distortion of the GOP, caused by the loss of a slice k in the frame $M(k)$, can be evaluated as:

$$\overline{d}_k = (1 - p)d_{p,k} + p(1 - p)d_{t,k} + p^2 d_{0,k} \tag{2}$$

where $d_{p,k}$ is the distortion due to the primary slice and $d_{0,k} = d_{l,k} \sum_{n=0}^{N-M(k)} f[n]$ is the distortion when both representations of the k-th slice are lost; $d_{l,k}$ is the slice mismatch distortion and depends on the adopted concealment strategy. Nevertheless, if the packet loss probability is low, the term $p^2 d_{0,k}$ in (2) can be neglected. As a consequence, the following optimization algorithm will not take into account

error concealment. Thus, (2) turns out to be

$$\overline{d}_k \approx (1-p)d_{p,k} + p(1-p)d_{r,k}\phi_{M(k)} \tag{3}$$

The final goal is the estimation of the expected distortion of the GOP, when a random number of slices are lost. Assuming that the error, caused by the loss of a given k-th slice, are uncorrelated, one can evaluate the total expected distortion of the GOP by summing all the slice contributions:

$$
\begin{aligned}
\overline{D} &= \sum_k \overline{d}_k \\
&= \sum_{i=1}^{N} \sum_{k:M(k)=i} \left[(1-p)d_{p,k} + p(1-p)d_{t,k} + p^2 d_{0,k} \right] \\
&\approx (1-p)\sum_{i=1}^{N} \sum_{k:M(k)=i} (d_{p,k} + p d_{r,k}\phi_i)
\end{aligned}
\tag{4}
$$

The previous assumption on slice error uncorrelation is quite reasonable. In fact, the quantization error, e.g. the one yielded by the primary slice, can be considered as uncorrelated with the pixel values, especially at high coding rates (see [24] in the case of lossy coding of correlated samples). The mismatch errors generated by the decoding of a redundant slice, being the difference between two quantization errors, can also be regarded as uncorrelated with the pixels. In conclusion, it is reasonable to assume that both the quantization error and the mismatch error of different slices are uncorrelated with the pixels and therefore they are uncorrelated with each other. Nevertheless, the video encoding process encompasses a multitude of coding options, whose effect is hardly modeled by high rate quantization [24]. In order to validate our assumption also in this situation, the additive distortion model has been verified experimentally, as well. A GOP of 100 frames of the Foreman QCIF sequence has been encoded with $QP = 26$ and three slices per frame are created. The distortion of the primary slice representations $d_{p,k}$ is measured. Then, the first slice of the second frame ($k = 4$) is replaced by a redundant representation with $QP = 34$ and the corresponding $d_{t,k}$ is measured. The same experiment is repeated with $k = 7$, i.e. the co-located slice in the following frame. Finally, the GOP distortion obtained when decoding redundant representations for $k = 4$ and $k = 7$ jointly, is measured; this latter, being 1539.9, turns out to be accurately approximated by the additive model $\sum_{i=1}^{N} [\sum_{k:M(k)=i, k \neq 4,7} d_{p,k} + d_{t,4} + d_{t,7}] = 1594.2$, used in (4). Therefore, the additive distortion model turns out to be accurate for the estimation of the whole GOP distortion. It is worth noticing that similar additive models are used in [19] and [20] in order to evaluate the expected distortion over a GOP when lost frame are recovered by concealment.

Defining $D_{r,i} = \sum_{k:M(k)=i} d_{r,k}$ and $D_{p,i} = \sum_{k:M(k)=i} d_{p,k}$ as the distortion of the redundant and the primary representations of picture i, and $R_{r,i}$, $R_{p,i}$ the corresponding rates measured in bits, we can simplify the previous equation as:

$$\overline{D} = (1-p)\left\{ \sum_{i=1}^{N} D_{p,i} + p \sum_{i=1}^{N} D_{r,i}\phi_i \right\} \tag{5}$$

At this point, the optimization problem can be formulated as the following constrained minimization:

$$\begin{cases} \min \overline{D} \\ \text{subject to} \sum_{i=1}^{N} (R_{p,i} + R_{r,i}) = R_{GOP} \end{cases} \tag{6}$$

where R_{GOP} is the overall rate per GOP, which is limited by the available bandwidth. The problem can be solved by means of the standard Lagrangian approach by minimizing the cost function

$$J = \overline{D} + \lambda \sum_{i=1}^{N} (R_{p,i} + R_{r,i}) \tag{7}$$

where λ is the Lagrangian multiplier. Imposing $\nabla J = 0$ we get

$$\frac{\partial J}{\partial R_{p,i}} = (1-p)\frac{\partial D_{p,i}}{\partial R_{p,i}} + \lambda = 0 \tag{8}$$

$$\frac{\partial J}{\partial R_{r,i}} = p(1-p)\phi_i \frac{\partial D_{r,i}}{\partial R_{r,i}} + \lambda = 0 \tag{9}$$

From (8) and (9) we can conclude that in order to minimize $\overline{D}$ the following condition must be satisfied:

$$\frac{\partial D_{p,i}}{\partial R_{p,i}} = p\phi_i \frac{\partial D_{r,i}}{\partial R_{r,i}} \ , \ \forall i \tag{10}$$

Moreover, observing (8) one can notice that $\frac{\partial D_{p,i}}{\partial R_{p,i}}$, i.e. the optimal R-D slope of the primary representation, does not depend on the frame index i; consequently, evaluating (10) for $i = 1$ we get $\frac{\partial D_{p,i}}{\partial R_{p,i}} = \frac{\partial D_p}{\partial R_p} = p\phi_1 \frac{\partial D_{r,1}}{\partial R_{r,1}}$. This latter expression can be plugged into (10) in order to obtain the optimal R-D slopes for the redundant slices as a function of the frame index i. To summarize, the expected GOP distortion is minimized under the following conditions:

$$\frac{\partial D_p}{\partial R_p} = p\phi_1 \frac{\partial D_{r,1}}{\partial R_{r,1}} \tag{11}$$

$$\frac{\partial D_{r,i}}{\partial R_{r,i}} = \left(\frac{\phi_1}{\phi_i}\right) \frac{\partial D_{r,1}}{\partial R_{r,1}} \tag{12}$$

Equation (11) imposes a constant ratio between the rate distortion (R-D) slopes of the primary frames and the redundant representation of the first frame of the GOP. This result is graphically represented in Fig. 3. Given a certain R-D point of the primary frame representation (black dot in Fig. 3-(a)), (11) determines the optimal R-D slope of the first redundant frame. In particular, the R-D slope of the redundant representation is determined by the scaling factor $p\phi_1$. Note that, the larger the value of $p\phi_1$, the higher rate $R_{r,1}$ must be devoted to the redundant representation. This result is confirmed by intuition; in fact, large values of either the packet loss probability p or the propagated distortion ϕ_1 require a high amount of redundancy. Since the redundant representation must be characterized by a coding rate inferior to that of the primary picture, it turns out that it must be verified that $p\phi_1 \leq 1$. This condition is a consequence of the approximations employed to work out the analytical optimization, namely the hypothesis $\Delta_{r,k} \gg \Delta_{p,k}$ assumed to simplify (1), the assumption that ϕ_i does not depend on the rate and the condition $p << 1$ in (2). If the condition $p\phi_1 \leq 1$ is not satisfied, one should modify the model employed or

adopt numerical solutions of the optimization problem. In practice, this may happen when p is very large or there is very little leakage in the prediction loop; both these situations lead to MDC where the redundant information is almost a replica of the primary one. It is worth pointing out that this is a limit case not treated in the present work.

On the other hand, (12) determines the relationship between the optimal R-D point of the first and the i-th redundant frame. In this case, the scaling parameter between such two slopes is given by the ratio ϕ_1/ϕ_i, which is only related to the error propagation phenomenon. Since by construction $\phi_i \geq \phi_{i+1}, \forall i$, it turns out that $\left|\frac{\partial D_{r,i+1}}{\partial R_{r,i+1}}\right| \geq \left|\frac{\partial D_{r,i}}{\partial R_{r,i}}\right|, \forall i$. This condition means that, if all the frames in the GOP have similar R-D curves, the redundant rates $R_{r,i}$ decrease as a function of the frame index, as shown in Fig.3-b. This result is again confirmed by intuition, since the impact of a redundant slice on the total GOP distortion decreases with the frame index.

Finally, we can describe the effects of the proposed optimization on the MDC coding redundancy, defined as the ratio between the extra rate and the overall rate:

$$\rho = \frac{\sum_i R_{r,i}}{\sum_i \left(R_{p,i} + R_{r,i}\right)} \tag{13}$$

In particular, it is very important to note that at high coding rates, and given a certain value of p, the optimal value of ρ decreases as the rate increases. As an example, let us consider the optimal R-D points selected as in Fig.3-b for a given rate budget R_{GOP}. Under the hypothesis of high rate, the derivative $\frac{\partial D_p}{\partial R_p}$ does not change significantly when augmenting the rate. As a consequence, for larger R_{GOP} the primary rate can be incremented whereas the amount of redundant rate remain almost constant. In fact, the optimization results impose a linear dependence between the redundant R-D slopes and the primary one; thus, at high rate, the redundant R-D points remain almost unaltered. In other words, as the coding rate increases, the primary rate can be incremented more than the corresponding optimal redundant rate, causing a reduction of ρ.

The optimal conditions expressed by (11) and (12) allow one to design an optimal redundancy allocation strategy. This can be accomplished using a closed loop rate-control technique, able to optimally partition the rate budget between the primary and redundant slices. This goal can be accomplished as far as good estimates of the R-D slopes are available. However, this objective is beyond the scope of the present work. In this paper, we propose a simpler yet effective open loop approach based on the optimal selection of the quantization parameter. This is achieved using the following standard H.264 R-D approximation [25]:

$$\frac{\partial D}{\partial R} = -0.85 \times 2^{\left(\frac{QP-12}{3}\right)} \tag{14}$$

Plugging (14) into (11) we obtain the QP values for the primary and redundant frames:

$$QP_p = QP_{r,1} + 3\log\left(p\phi_1\right) \tag{15}$$

$$QP_{r,i} = QP_{r,1} + 3\log\left(\frac{\phi_1}{\phi_i}\right)$$
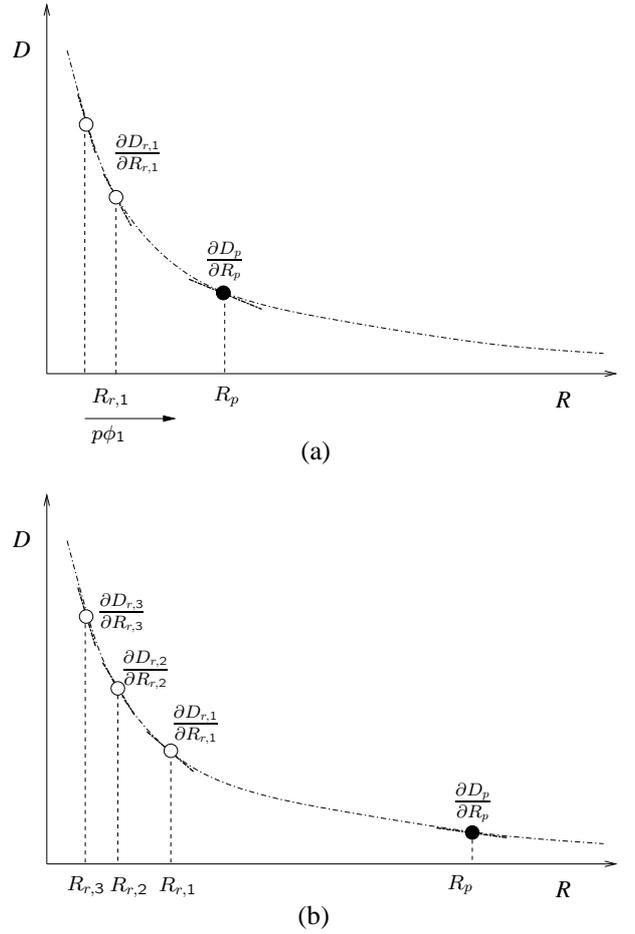
Figure 3. Graphical interpretation of optimal conditions (11), (12).

Equation (15) states that the optimal quantization parameter $QP_{r,i}$, used to code the i-th redundant picture, varies as a function of i and the packet loss probability p; in particular, the quantization step is larger for higher values of i, i.e. while approaching the GOP end. On the other hand, p determines larger values for $QP_{r,i}$, i.e. smaller redundancy; this is a very sensible condition.

A. Algorithm implementation

The above described optimization procedure has been implemented in the JM9.4 H.264 reference software. The most important design choices are detailed in the following.

Since the compressed video is going to be transmitted across a packet network, it is important to adopt a proper data partitioning strategy. To this end, each picture in the video sequence is partitioned into a certain number of slices so as to guarantee that each compressed H.264 *Network Access Level* (NAL) unit is smaller than the network MTU. Moreover, the picture slicing yields a certain degree of error resilience; in fact, packet losses will appear as partial picture losses at the decoder side, thus making the subsequent concealment procedure more effective. As a consequence, a limit on the maximum number of MBs per slice is enforced, as well.

Each picture in the video sequence is encoded in two passes, i.e. the primary and the redundant pass. During the

(a)

(b)

Figure 4. $\Delta(QP)_i$ as a function of frame index i and packet loss probability p (a) and as a function of i and α (b).

two descriptions alternatively. The output of this process are two balanced descriptions, which are fully H.264 compliant. As a consequence, any standard decoder can be used to retrieve the video sequence at the receiver side. In particular, if a single description is received, half of the primary slices will be missing, and will be replaced by their redundant counterparts; this process represents the side decoder and permits to reconstruct the video signal with a reduced level of fidelity, assuming that a sufficient amount of redundancy has been allocated. If both descriptions are received, a standard H.264 decoder will pick-up the primary representation of each slice, retrieving the full quality video. In fact, any standard H.264 decoder is going to skip the redundant slices in presence of their primary representation. This task represents the central decoder, and requires a low-complexity pre-processing operation, where the two descriptions are merged into a single H.264 bitstream. The pre-processing task requires simple parsing of the slice headers in order to get the picture index so as to merge the two descriptions synchronously. In general, the two H.264 descriptions will be transmitted across two independent physical or virtual channels. At the receiver side, the best representation of each slice, available within its play-out deadline, can be decoded. Only in the case that both descriptions are lost, the decoder must invoke a concealment algorithm. The experimental results presented in the following section are obtained using the temporal concealment available in JM9.4 reference software for the missing predicted MB. When losses occur in an I slice, the simple MB replacement with the co-located MB in the previous frame is used.

It is worth noticing that the final splitting into two descriptions is not essential if one has a single channel or connection. In such a case, the allocated redundancy is able to guarantee error resilience under the hypothesis that packet losses are still independent. In case of burst errors, such as those due to node congestion or wireless links, it is mandatory to adopt a packet scheduling procedure able to guarantee the maximum possible amount of delay between the transmission of the primary and its corresponding redundant slice data, so as to reduce the probability that both representations fall in the same burst. Nevertheless, packet scheduling is beyond the scope of the present study.

V. EXPERIMENTAL RESULTS

In this section, the proposed technique is validated by means of a number of experimental trials.

First of all, it is important to analyze the behavior of the proposed redundancy allocation strategy. The redundant slice quantization parameters are selected according to (15), which depends on the value of the parameters N, p and α. It is well known that setting the GOP size N, one can obtain a trade-off between coding efficiency and resynchronization capability in presence of transmission losses. The optimization of N is not included in the previous analytical model; in fact, N should be selected taking into account other application constraints, such as random access to the video sequence in presence of multiple users and/or packet losses. However, in order to evaluate the effects of this parameter

first pass, the primary picture is encoded with a constant quantization parameter QP_p, which is specified by the user. During this pass, a certain number of slices are created so as to respect the two constraints on the MTU limit and the maximum number of MBs per slice. During the redundant pass, a redundant representation of each primary slice is created. The quantization parameter $QP_{r,i}$ for a given redundant slice of the i-th frame is selected according to (15). As for the power transfer function $f[n]$, for the sake of simplicity the function $f[n] = e^{-\alpha n}$ is employed; this expression approximates, at low levels of attenuation, the function $f[n] = (1 + \gamma_1 n)^{-1}$ reported in [22] and $f[n] = (1 + \gamma_2 n)^{-\frac{1}{2}}$ reported in [21], α, γ_1 and γ_2 being parameters to be selected. Using this representation of the power transfer function, the ϕ_i parameter introduced in (15) becomes $\phi_i = \frac{1-e^{-\alpha(N-i+1)}}{1-e^{-\alpha}}; i \in [1, N]$.

Finally, the obtained H.264 bitstream is post-processed so as to split it into two balanced H.264 descriptions, as already described in Sect. III. In order to guarantee that the two bitstreams can be independently decoded, the crucial information contained in the *Sequence Parameter Set* (SPS) and *Picture Parameter Set* (PPS) NAL units is duplicated in both descriptions. All the other NAL units, which represent either a primary or a redundant slice, are queued in the

Figure 5. Average PSNR with $p = 0.01, 0.05, 0.10$, and $N = 11, 21, 45$ on Foreman CIF sequence at 30 fps versus coding rate R.
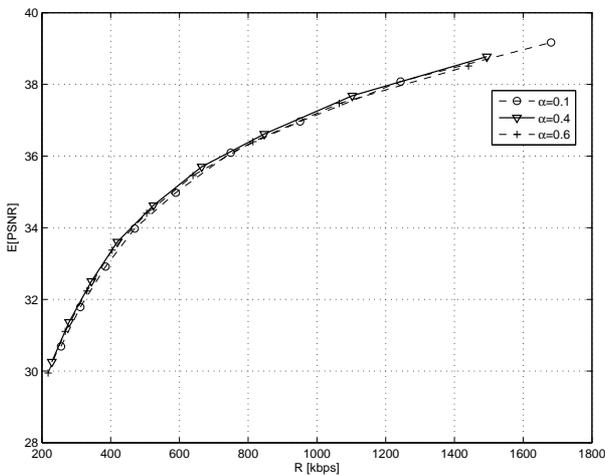
Figure 6. Average PSNR with $p = 0.05$ and variable α on Foreman CIF sequence at 30 fps versus coding rate R.

on the algorithm performance, we carried out simulations using different values of N. The value of the packet loss probability p is assumed to be known at encoding time; this information can be determined from a feedback channel, such as the client feedback feature included in the Release 6 of the recently standardized 3GPP streaming standard PSS [16], [26], or by estimating the transmission characteristics. In Fig. 4-(a) the redundant quantization parameter increment $\Delta(QP)_i = QP_{r,i} - QP_p$ is plotted as a function of p and the frame index i in the case $N = 50$ and $\alpha = 0.2$; it is evident that $\Delta(QP)_i$ increases as long as p approaches 0, i.e. the coding redundancy tends to 0 for low packet loss probability. The parameter α, which determines the propagation error attenuation, merits further attention; in fact, α depends on the video sequence content, the used coding options and the decoder concealment capabilities [21], [22]. In the context of the proposed algorithm, the value of α affects the allocation of the coding redundancy; in fact, a small α models a long error propagation phenomenon, which can be compensated for with a significant amount of redundancy. This behavior is depicted

in Fig. 4-(b), where $\Delta(QP)_i$ is reported as a function of i and α in the case $p = 0.05$ and $N = 50$. Further discussion on the settling of α will be presented later in this section.

The experimental results reported in this section are obtained with the following settings. The primary slice quantization parameter QP_p is selected in the interval $(22, 38)$, in order to span a considerable range of coding rates. A GOP structure containing only P slices is adopted and a frame buffer containing the last 5 coded pictures is employed. Slices are formed according to the policy described in Sect. IV-A, i.e. each compressed H.264 slice is delivered by one MTU; moreover, each slice is constrained to contain a maximum number of 80 MBs for CIF sequences (i.e. 5 slices per frame), and 33 MBs for QCIF sequences (i.e. 3 slices per frame). Finally, each packet, which contains a single slice, is lost using the Bernoulli model with probability p. It is worth pointing out that the H.264 standards includes other resilience options that may further improve the results reported in following. Nevertheless, the optimization and the joint benefits of other features, such as Intra MB refresh, flexible MB ordering, B slices, data partitioning, etc. are beyond the scope of the present work. The performance is measured in terms of average luminance PSNR, obtained with 50 independent transmission trials. This amounts to the transmission of more that $2 \cdot 10^4$ packets, which yields significant results from the statistical point of view.

A first set of simulations is carried out using the first 90 frames of the standard CIF *Foreman* sequence at 30 fps. It aims at assessing the consequences of using different values of N on the algorithm performance. In Fig. 5 the average PSNR versus the rate R for all the nine combinations given by $p = 0.01, 0.05, 0.10$, and $N = 11, 21, 45$ with fixed $\alpha = 0.4$, is shown. We can notice that, in the case $p = 0.10$, using a long GOP $N = 45$ causes a significant gap with respect to the performance obtained for $N = 11$; in fact at high coding rate the performance impairment is about 2 dB. On the contrary, using $N = 11$ when $p = 0.01$ causes a 1dB impairment with respect to the case with $N = 45$. It is evident that it is better to shorten the GOP length when the probability of packet loss increases. As a consequence, in all the following experiments N will be empirically selected as the best tradeoff between compression efficiency and error robustness; in particular, given the adopted simulation conditions, it turns out that the best settings are $(p, N) \in \{(0.01, 45), (0.05, 21), (0.10, 11)\}$, which correspond to the bold lines in Fig. 5. However, from Fig. 5 we can also notice that using $N = 21$ for all the tested values of p results in a maximum impairment, over the previously defined pairs, limited to 0.5 dB. This means that the proposed allocation algorithm is able to shape the values of $QP_{r,i}$ so as to optimally adapt to the used N.

The second set of simulations aims at assessing the importance of the α parameter selection on the algorithm performance. In Fig. 6 the average PSNR versus the rate R in the case $p = 0.05$, $N = 21$ and $\alpha = 0.1, 0.4, 0.6$ respectively is shown. It is worth pointing out that the allocation algorithm is not very sensitive to the actual value of α; in fact, the best average performance is obtained in the case $\alpha = 0.4$, even if the other simulations provide only slightly inferior results.
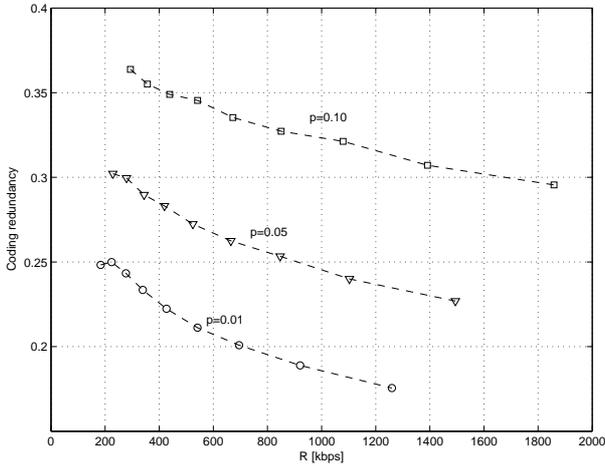
Figure 7. Foreman CIF sequence: coding redundancy ρ versus rate R in the cases $p = 0.01, p = 0.05, p = 0.10$.
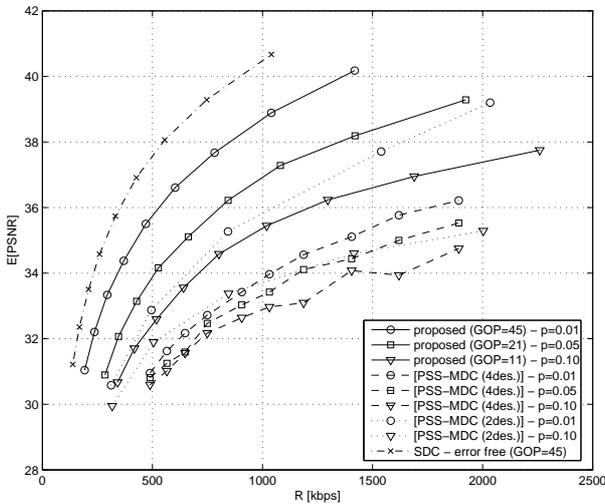
Figure 8. Foreman CIF sequence: average PSNR with $p = 0.01, p = 0.05, p = 0.10$, compared to PSS-MDC [27], [28], and SDC error free performance.

In Fig. 7 the redundancy ρ, as defined in (13), yielded by the optimal allocations for $p = 0.01, 0.05, 0.10$, $\alpha = 0.4$ and $N = 45, 21, 11$ respectively, is shown versus R. Obviously, ρ decreases for low values of packet loss probability. It is important to notice that the allocated redundancy decreases at higher coding rates; this effect is due to the fact that, in this case, the R-D curve becomes almost flat, as already mentioned when commenting (11). In other words, it turns out that, given a certain value of p, the optimal value of redundancy also depends on the actual coding rate.

In Fig. 8 the average PSNR obtained with the proposed allocation and $p = 0.01, 0.05, 0.10$, $\alpha = 0.4$, $N = 45, 21, 11$ respectively are shown versus R. In order to appreciate the robustness of the proposed scheme in Fig. 8 the error free SDC curve is reported as a performance upper-bound. Obviously, in this case no redundant slices are employed; moreover, only the performance obtained with $N = 45$, which offers the best compression efficiency with respect to the other tested GOP sizes, is shown. The slicing policy and the MTU limit are kept

unchanged for all simulations so as to enable fair comparisons of the results under the same transmission constraints. In the same figure the results available in [27] and [28] are shown, as well. These latter are MDC algorithms based on polyphase spatial sub-sampling (PSS-MDC). In [27], four descriptions are created by H.264 encoding of four QCIF sequences, obtained by sub-sampling an original CIF video, whereas in [28] two sub-sequences are generated by separating odd and even rows of the frames. The two generated sub-sequences are then encoded using H.264. To the best of our knowledge, the PSS-MDC is one of the first examples of H.264 compatible MDC technique, and therefore it represents a good benchmark for our algorithm performance. The results in [27] and [28] are obtained with a GOP $N = 21$, using both P and B pictures and slices of 400 bytes; the transmission conditions are the same as those used in this paper. Another H.264 based MDC technique has been recently presented [13]; however, in this latter paper some coding parameters, such as the redundancy and the coding rate, are not specified, thus preventing from fair comparisons. From Fig. 8 it turns out that the proposed MDC approach exhibits a noticeable performance improvement with respect to [27] and [28]. The average PSNR gain ranges from about 3 dB up to 5 dB over four descriptions PSS-MDC scheme and 2 dB up to 3 dB for the two descriptions case. The most significant improvement is obtained in the case $p = 0.01$. This gain is mainly due to the fact that the proposed approach permits to easily tune the redundancy according to the network conditions, so allowing one to introduce an amount of redundancy as low as needed to match the case of small p. On the other hand, [27] and [28] is constrained to a rigid spatial sub-sampling, which does not allow one to control the amount of redundancy. This turns out to depend on the spatial correlation among the sub-sampled sequences.

Finally, the proposed scheme favorably compares with [27] and [28] in terms of implementation complexity at the decoder side. In fact, the sub-sampling scheme requires the use of an additional interpolation stage to recover from losses. On the other hand, the proposed scheme is based on a single H.264 standard decoder.

The performance of the proposed scheme can also be evaluated in terms of central and side distortion (measured as the average of the two side distortions). In Fig. 9 the central and side performance in the three cases previously analyzed ($p = 0.01, 0.05, 0.10$) are shown versus R. It is clear that p affects the inserted redundancy, which means that higher p correspond to higher redundancy. As a consequence, the quality obtained when both descriptions are received decreases at higher p, whereas the quality delivered by only one description increases. In other words, the gap between the central and the side PSNR decreases as long as the redundancy increases. Fig. 10 shows in more detail the performance of the proposed allocation for $p = 0.05$; in particular, both the side distortions are reported; this permits to verify that the two descriptions are indeed balanced in terms of both rate and quality. Moreover, in this figure the central distortion is compared with the distortion obtained in case of transmission across links with $p = 0.05$ (corresponding to the design parameter). For completeness, in Fig. 11 the luminance PSNR

versus the frame number, yielded by the proposed technique at 1 Mbps with $p = 0.05$, is shown. In particular, side and central PSNR, along with the performance when packet losses are present, are reported. This figure permits to appreciate the distortion behavior within each GOP. It turns out that the punctual distortion introduced when $p = 0.05$, that is the loss probability for which the redundancy has been allocated, is limited; the side decoder is reported for comparison, and represents the limit case when an entire link is missing. On the same graph we report the allocated redundancy for each frame along with the average redundancy for all the sequence (the values of ρ are reported on the right hand vertical axis). As expected, the frame redundancy decreases as a function of the frame index even if it is not a monotonic decreasing function due to the non stationary R-D characteristics of the frames.

In Fig. 12-13 more experimental results carried out on QCIF *Foreman* and *Coastguard* sequences at 15 fps are reported. The average PSNR obtained in the case $p = 0.01, 0.05, 0.10$ is compared with the error free SDC performance. These results confirm that the proposed technique exhibits a good behavior independently of the video content, size and frame rate.

Finally, it is very important to test the sensitiveness of the allocation procedure with respect to channel mismatch. In Fig. 14 all the curves refer to the case of transmission over a link with $p = 0.05$; on the other hand, the redundancy is allocated assuming $p = 0.01, 0.05$ and 0.10, respectively. As expected, the allocation with the correct value of p (square marker) offers the best average performance in the considered rate range. When the packet loss probability is overestimated ($p = 0.10$), the allocated redundancy is larger than the optimal one; as a consequence, the average PSNR is impaired, even if the performance loss is limited to less than 1 dB. On the other hand, for a design parameter $p = 0.01$, the redundancy is too low and the average PSNR significantly drops especially at high rates. The fact that the curve $p = 0.05$ and $p = 0.01$ almost overlap at low rate is due to the adopted high rate approximation, that allowed us to design an open loop allocation, which in turn is not guaranteed to be optimal at very low bit rates.

VI. CONCLUSIONS

In this paper the redundant slice concept has been employed in order to design a novel MDC technique, producing two balanced descriptions of a video source. The proposed approach presents a number of advantages with respect to competing algorithms, namely the capability to flexibly tune the inserted redundancy, the use of standard H.264 encoder and decoder and the limited added computational cost. Moreover, the adopted analytical optimization approach results in a simple, yet powerful, open loop redundancy allocation technique. The proposed close form solution is worked out in the limit of high rates; the solution for low bit-rate video coders is left for future research. Other improvements are expected in a better modeling of the power transfer function, so as to be able to automatically match the video sequences and codec characteristics.
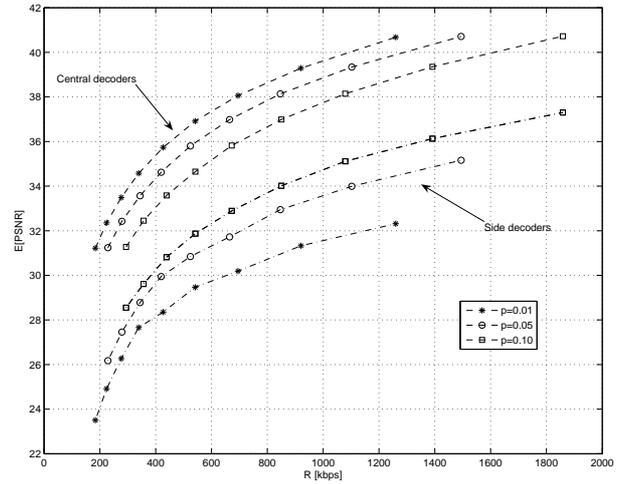
Figure 9. Foreman CIF sequence: central and side decoder performance for different optimizations ($p = 0.01, p = 0.05, p = 0.10$).
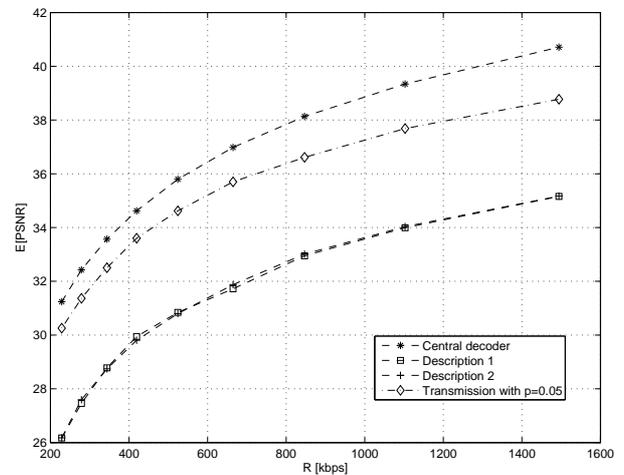
Figure 10. Foreman CIF sequence with optimal redundancy allocation for $p = 0.05$: average PSNR of central decoder, side decoder and transmission with $p = 0.05$.
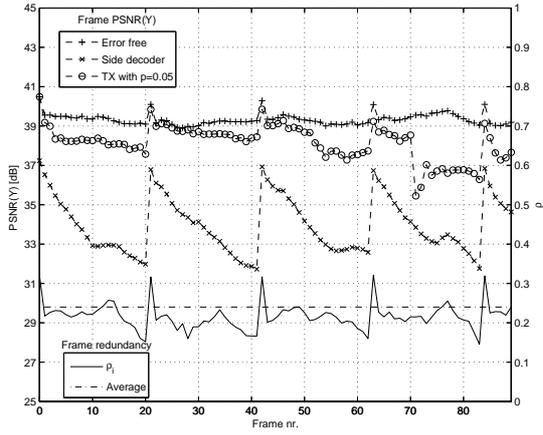
Figure 11. Foreman CIF sequence with optimal redundancy allocation for $p = 0.05$ at $R = 1$ Mbps: luminance PSNR of central decoder, side decoder and transmission with $p = 0.05$; and the allocated redundancy per-frame with the average redundancy.
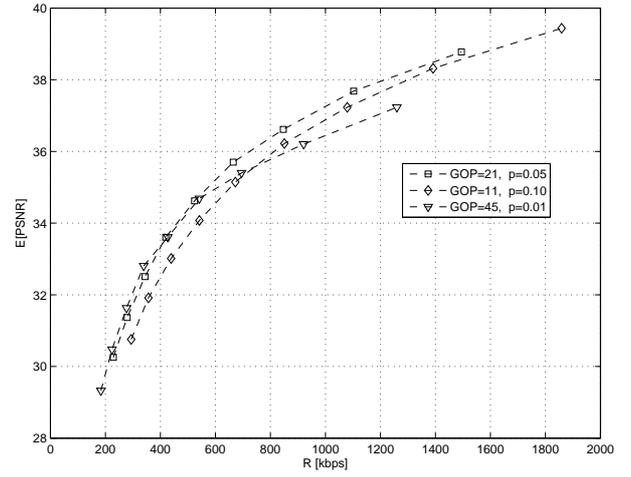
Figure 14. Foreman CIF: performance of the proposed technique with design parameter $p = 0.01, p = 0.05$ and $p = 0.10$ when transmitted with $p = 0.05$.
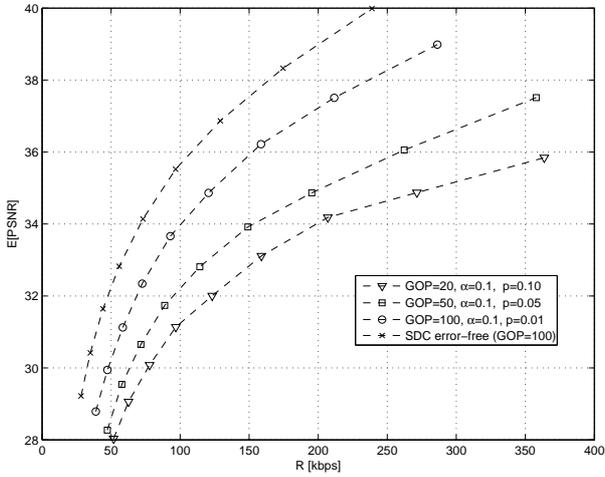
Figure 12. Foreman QCIF sequence at 15 fps: average PSNR of proposed algorithm with $p = 0.01, p = 0.05, p = 0.10$, compared to SDC (error free performance).

Figure 15. The *pdf* function of the mismatch error of: (a) aligned quantizers ($\frac{\Delta_r}{\Delta_p} = 7$); (b) not aligned quantizers.

Figure 13. Coastguard QCIF at 15 fps: average PSNR of proposed algorithm with $p = 0.01, p = 0.05, p = 0.10$, compared versus SDC (error free performance).
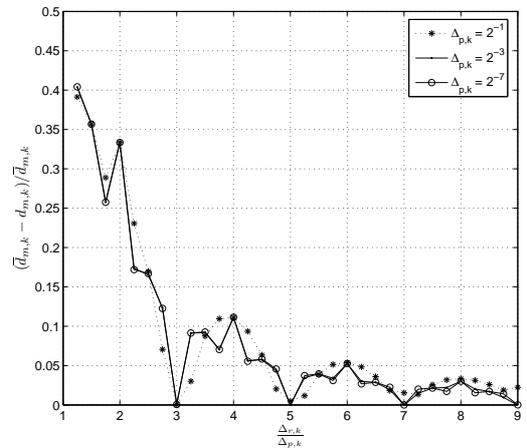
Figure 16. Relative estimation error of the mismatch distortion versus $\frac{\Delta_{r,k}}{\Delta_{p,k}}$, for a quantized iid Gaussian process with two mid-tread uniform quantizers.

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, "Balanced interframe multiple description video compression," in *Proc. of IEEE International Conference on Image Processing (ICIP)*, Oct. 1999, vol. 3, pp. 812–816.

[3] Y. Wang, M.T. Orchard, V. Vaishampayan, and Reibman A.R., "Multiple description coding using pairwise correlating transforms," *IEEE Trans. on Image Processing*, vol. 10, no. 3, pp. 351–367, Mar. 2001.

[4] Reibman A.R., H. Jafarkhani, Y. Wang, M.T. Orchard, and M. Puri, "Multiple description coding for video using motion compensated prediction," in *Proc. of IEEE International Conference on Image Processing (ICIP)*, Oct. 1999, vol. 3, pp. 837–841.

[5] T. Wiegand, G.J. Sullivan, G. Bjntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, July 2003.

[6] S. Shirani, M. Gallant, and F. Kossentini, "Multiple description image coding using pre- and post-processing," in *Proc. of International Conference on Information Technology: Coding and Computing)*, Apr. 2000.

[7] N. Franchi, M. Fumagalli, and R. Lancini, "Flexible redundancy insertion in a polyphase down sampling multiple description image coding," in *Proc. of IEEE International Conference on Multimedia and Expo (ICME)*, Aug. 2002.

[8] T. Tillo and G. Olmo, "A low complexity pre-post processing multiple description coding for video streaming," in *Proc. of the IEEE International Conference on information and communication technologies: From Theory to Applications (ICTTA 04), Damascus, Syria*, Apr. 2004, pp. 519–520.

[9] D. Wang, N. Canagarajah, D. Redmill, and D. Bull, "Multiple description video coding based on zero padding," in *Proc. of IEEE 2004 International Symposium on Circuits and Systems, (ISCAS)*, 2004.

[10] G. Olmo and T. Tillo, "Directional multiple description scheme for still images," in *Proc. of 10th International Conference on Electronics, Circuits and Systems (ICECS)*, Dec. 2003.

[11] M. Gallant, S. Shirani, and F. Kossentini, "Standard-compliant multiple description video coding," in *Proc. of IEEE International Conference on Image Processing*, Oct. 2001.

[12] D. Wang, N. Canagarajah, and D. Bull, "Slice group based multiple description video coding using motion vector estimation," in *Proc. of IEEE International Conference on Image Processing (ICIP)*, Sept. 2004.

[13] D. Wang, N. Canagarajah, and D. Bull, "Error concealment for slice group based multiple description video coding," in *Proc. of IEEE International Conference on Image Processing (ICIP)*, Sept. 2005.

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

[15] Joint Video Team JVT of ISO/IEC MPEG and ITU-T VCEG, *International Standard of Joint Video Specification (ITU-T Rec. H.264 , ISO/IEC 14496-10 AVC)*, Mar 2003.

[16] T. Schierl, T. Wiegand, and M. Kampmann, "3GPP compliant adaptive wireless video streaming using H.264/AVC," in *Proc. of IEEE International Conference on Image Processing (ICIP)*, Sept. 2005.

[17] R. Zhang, S. L. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience," *IEEE J. Sel. Areas Commun.*, vol. 18, pp. 966–976, June 2000.

[18] R. Zhang, S. L. Regunathan, and K. Rose, "End-to-end distortion estimation for RD-based robust delivery of pre-compressed video," in *Proc. of Thirty-Fifth Asilomar Conference on Signals, Systems and Computers*, Nov. 2001.

[19] Z. He, H. Cai, and C. W. Chang, "Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 12, no. 6, pp. 511–523, June 2002.

[20] Y. Wang, Z. Wu, and J.M. Boyce, "Modeling of transmission-loss-induced distortion in decoded video," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 16, no. 6, pp. 716–732, June 2006.

[21] B. Girod and N. Farber, "Feedback-based error control for mobile video transmission," in *Proc. of the IEEE*, Oct. 1999, pp. 1707–1723.

[22] N. Farber, K. Stuhlmuller, and B. Girod, "Analysis of error propagation in hybrid video coding with application to error resilience," in *Proc. of the IEEE International Conference on Image Processing*, Oct. 1999, pp. 550–554.

[23] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A transport protocol for real-time applications," *Internet Engineering Task Force - RFC 1889*, Jan. 1996.

[24] N.S. Jayant and P. Noll, *Digital Coding of waveforms*, Prentice Hall, 1984.

[25] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. Sullivan, "Rate-constrained coder control and comparison of video coding standards," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 688–703, July 2003.

[26] 3GPP TSG-SA4 PSM SWG Internal working draft, *Transparent end-to-end Packet-switched streaming service (PSS); Protocols and codes (Release 6)*, Nov 2004.

[27] R. Bemardini, M. Durigon, R. Rinaldo, L. Celetto, and A. Vitali, "Polyphase spatial subsampling multiple description coding of video streams with H.264," in *Proc. of IEEE International Conference on Image Processing*, Oct. 2004, pp. 3213–3216.

[28] A. Vitali, F. Rovati, R. Rinaldo, R. Bemardini, and M. Durigon, "Low-complexity standard-compatible robust and scalable video streaming over lossy/variable bandwidth networks," in *Proc. of ICCE. 2005 Digest of Technical Papers. International Conference on Consumer Electronics*, Jan. 2005, pp. 7–8.

APPENDIX A

In the following we will evaluate the distortion $d_{m,k}$, generated by the mismatch between the k-th primary slice and its redundant representation. Let us assume to represent a given frame with a high resolution quantizer with quantization step $\Delta_{p,k}$ (primary representation) and a lower resolution quantizer with quantization step $\Delta_{r,k}$ (redundant representation). As far as the quantizers are sufficiently fine, we can assume a uniform distribution of the transformed coefficients within each quantization interval.

The distortion $d_{m,k}$, that arises when a primary representation is lost and replaced with the redundant one, can be evaluated as the expected value of the square of the difference e between the primary and redundant reconstruction levels:

$$d_{m,k} = E\{e^2\} = \int_{-\infty}^{\infty} e^2 f_\epsilon(e) de$$

where $f_\epsilon(e)$ is the probability density function (*pdf*) of the error e. For simplicity, the case of aligned quantizers with $\frac{\Delta_{r,k}}{\Delta_{p,k}} = 2n + 1$, where n is integer, is considered. Taking into account the previous assumptions we can model $f_\epsilon(e)$ for aligned quantizers as:

$$f_\epsilon(e) = \frac{\Delta_{p,k}}{\Delta_{r,k}} \delta(e) + \frac{\Delta_{p,k}}{\Delta_{r,k}} \sum_{j=1}^{n} [\delta(e - j\Delta_{p,k}) + \delta(e + j\Delta_{p,k})]$$

where $\delta(e)$ represents the Dirac delta function and $\Pi(\frac{e}{\Delta_{r,k}})$ represents the window function with full-width $\Delta_{r,k}$. In the case of aligned quantizers one obtains:

$$f_\epsilon(e) \approx \frac{1}{\Delta_{r,k}} \Pi(\frac{e}{\Delta_{r,k}})$$

A pictorial representation of the two quantizer *pdf* is reported in Fig. 15. At this point, with some algebraic manipulation, we can evaluate $d_{m,k}$ for the aligned quantizers case as:

$$d_{m,k} = \frac{\Delta_{p,k}}{\Delta_{r,k}} \sum_{j=1}^{n} \int_{-\infty}^{\infty} e^2 \left[\delta(e - j\Delta_{p,k}) + \delta(e + j\Delta_{p,k})\right] de$$

$$= 2\frac{\Delta_{p,k}^3}{\Delta_{r,k}} \sum_{j=1}^{n} j^2$$

$$= \frac{\Delta_{r,k}^2}{12} \left[1 - \left(\frac{\Delta_{p,k}}{\Delta_{r,k}}\right)^2\right]$$

Considering that $d_{r,k} = \frac{\Delta_{r,k}^2}{12}$, we obtain:

$$d_{m,k} = d_{r,k} \left[1 - \left(\frac{\Delta_{p,k}}{\Delta_{r,k}}\right)^2\right] \tag{16}$$

It is worth pointing out that when $\Delta_{p,k} \ll \Delta_{r,k}$ previous expression can be simplified as $d_{m,k} \approx d_{r,k}$. Following the same procedure in the case of not aligned quantizer $d_{m,k}$ can be simply evaluated as

$$d_{m,k} \approx d_{r,k} \tag{17}$$

In order to validate the obtained analytical results, simulation were carried out using a mid-tread uniform quantizer on a sequence of one million samples of a Gaussian random process with unit variance and zero mean. A high resolution representation of the sequence has been generated using different values for $\Delta_{p,k}$, namely $\Delta_{p,k} = 2^{-1}, 2^{-3}, 2^{-7}$. In correspondence with each value of $\Delta_{p,k}$, we have generated different low resolution representations of the sequence by using different $\Delta_{r,k}$. In Fig. 16, we report the relative estimation error defined as $(\overline{d}_{m,k} - d_{m,k})/\overline{d}_{m,k}$, where $\overline{d}_{m,k}$ represents the actual value of the mismatch distortion, against $\frac{\Delta_{r,k}}{\Delta_{p,k}}$. It is clear that in the case $\frac{\Delta_{r,k}}{\Delta_{p,k}} = 1$ we have $\overline{d}_{m,k} = d_{m,k} = 0$. From the reported results, we can notice that (16) represents with a good accuracy the mismatch error behavior for the aligned quantizers case, i.e. $\frac{\Delta_{r,k}}{\Delta_{p,k}}$ is odd. On the other hand, the accuracy of (17) increases as the $\frac{\Delta_{r,k}}{\Delta_{p,k}}$ augments.

Marco Grangetto (S'99 - M'03) received the "summa cum laude" degree in Electrical Engineering in 1999, and the Ph.D. degree in 2003, from Politecnico di Torino, Italy. He is currently a Post-Doctoral researcher at the Image Processing Lab, in the same University. He was awarded the premio Optime by Unione Industriale di Torino in Sept. 2000, and a Fulbright grant in 2001 for a research period with the Department of Electrical and Computer Engineering, University of California, San Diego. His research interests are in the fields of multimedia signal processing and communications. In particular, his expertise includes wavelets, image and video coding, data compression, video error concealment, error resilient video coding unequal error protection and joint source channel coding. He has participated to the ISO standardization activities on Part 11 of the JPEG 2000 standard. He has been member of the Technical Program Committee for several international conferences, including IEEE ICME, ICIP and ISCAS.

Gabriella Olmo (S'89 - M'91 - SM'06) has received the Laurea Degree "summa cum laude" and the PhD in Electronic Engineering at Politecnico di Torino, where she is currently Associate Professor. Her main recent interests are in the fields of wavelets, remote sensing, image and video coding, resilient multimedia transmission, joint source-channel coding, distributed source coding. She has coordinated several national and international research programs in the fields of wireless multimedia communications, under contracts by the European Community and the Italian Ministry of Education. Gabriella Olmo is member of IEEE Communications Society and IEEE Signal Processing Society, and is part of the Editorial Board of "Signal, Image and Video Processing", Springer journal. She has coauthored more than 110 papers in international technical journals and conference proceedings. She has been member of the Technical Program Committee and session chair for several international conferences.

Tammam Tillo (S'02 - M'06) was born in Damascus, Syria. He received the degree in Electrical engineering from the University of Damascus, Syria, in 1994, and the Ph.D. in Electronics and Communication Engineering from Politecnico di Torino, Italy, in 2005. From 1999 to 2002 he was with *Souccar For Electronic Industries*, Damascus, Syria. In 2004 he was visiting researcher at the EPFL (Lausanne, Switzerland). He is currently a Post-Doctoral researcher at Dipartimento di Elettronica, Politecnico di Torino. His research interests are in the areas of robust transmission, image and video compression, and hyperspectral image compression.