

Multiple Description Wavelet Coding of Layered Video Using Optimal Redundancy Allocation

Nikolaos V. Boulgouris,¹ Konstantinos E. Zachariadis,² Angelos Kanlis,³ and Michael G. Strintzis^{4,5}

¹Department of Electronic Engineering, Division of Engineering, King's College London, WC2R 2LS London, United Kingdom

²The Kellogg School of Management, Northwestern University, IL 60208, USA

³The European Patent Office, Munich 80298, Germany

⁴The Informatics and Telematics Institute, Thessaloniki GR-57001, Greece

⁵The Electrical and Computer Engineering Department of the University of Thessaloniki, Thessaloniki GR-54124, Greece

Received 2 March 2005; Revised 30 August 2005; Accepted 1 September 2005

We present a wavelet-based framework for the encoding of video in multiple descriptions. Using the proposed methodology, the generation of multiple descriptions is performed so that drift is eliminated at the decoder regardless of the number of received descriptions. Moreover, the proposed framework is flexible in the sense that it allows the encoding of video into an arbitrary number of descriptions. We also present a thorough analysis of rate allocation issues and propose three algorithms for the optimal allocation of redundancy. Experimental results for the transmission of video using two descriptions demonstrate the efficiency of the proposed method.

Copyright © 2006 Hindawi Publishing Corporation. All rights reserved.

1. INTRODUCTION

Multiple description (MD) coding [1, 2] offers an attractive framework for the transmission of multimedia over heterogeneous networks. In MD coding, a source is encoded into multiple independently decodable bitstreams which are mutually refining and equally important. At the decoder side, the reconstruction quality is dependent on the number of descriptions that was errorlessly received. Due to its flexibility, multiple description coding is considered a very robust and reliable tool for information transmission.

Multiple description coding has been investigated for image [3–5] and video transmission [6–11]. In the particular case of video transmission, the study of MD systems becomes more complicated due to the uncertainty about the information that will be available at the decoder of an MD system.

In [12], a methodology was presented for the design of two-channel orthonormal filter banks based on the Lagrangian optimization of the redundancy rate-distortion performance of MD subband coding. In [7], an MD predictive quantization system was introduced, appropriate for the encoding of correlated information sources such as video and speech. The proposed system was used to construct a balanced twin-description interframe MD video coder, and performance results are presented using two packetization

strategies. A review on MD coding was recently presented in [13].

In [6], MD video coders were proposed which use motion-compensated prediction. These systems utilize MD transform coding, three separate prediction paths, and side information in order to accommodate all possible scenarios at the decoder. For this reason, three different algorithms for redundancy allocation were implemented, and experimental results were presented. An improved algorithm based on the same principles was presented in [10] where the encoding of the side information was modified in order to be useful even if no drift occurs. In [14], a novel scheme for double-description coding was proposed, which is built in the H.263 coder and replicates some selected DCT coefficients in both descriptions. The selection is based on a threshold determined using rate-distortion techniques. In [8], a novel way to deal with redundancy was devised. Temporal redundancy was used to control the tradeoff between drift and redundancy. However, this method does not inherently eliminate drift, that is, the cumulative distortion which occurs whenever the reference frames used at the decoder are not identical to the ones used by the encoder.

In [9], a drift-free wavelet-based MDC video coding scheme was proposed. However, the redundancy allocation algorithm did not take into consideration the impact of the temporal redundancy into the design of the system, thus

resulting in suboptimal coding. The above problem was dealt with in [15], where an improved version of the method in [9] was presented.

In [16], a multiple description coding method for video streaming was presented. The method in [16] was based on a 3D discrete wavelet transform. Redundancy was allocated by applying Lagrangian optimization techniques for the appropriate selection of subband quantizers. In [17], an MDC scheme for video coding was presented based on a spatiotemporal multiresolution analysis. Correlation between the two descriptions was introduced in the temporal domain by using an oversampled motion-compensated filter bank.

In the present paper, the intraframe and the motion compensated prediction residual frames are wavelet-coded and divided into a redundant and an enhancement part with the redundant part encoded in all descriptions and the enhancement part distributed in several descriptions. The “repeat or split” strategy was chosen over other proposed techniques, such as that presented in [2] since, in our case, drift-free reconstruction is straightforward. Using the above framework, we present and evaluate two techniques for the multiple description coding of video sequences.

(i) In the first technique, *only the redundant part* is used for the construction of reference frames and thus the resulting video coding scheme is able to perform drift-free reconstruction. Since the quality of the reference frame affects the coding efficiency of the system, an algorithm incorporating the impact of temporal correlation is also presented for the allocation of redundancy among multiple descriptions.

(ii) In the second technique, *both the redundant and the nonredundant parts* of the stream are used for the creation of the reference frame. This technique uses high-quality reference frames but the reconstructed video suffers from drift in case of transmission over channels with severe loss.

Additionally, in the present paper the problem of optimal redundancy allocation, that is, the appropriate selection of the redundant and the enhancement parts for each frame, is investigated. Specifically, this problem is formulated as the maximization of the average video quality under the constraint of a target total rate. Three variations of an optimization algorithm are proposed and evaluated in terms of their complexity. It should be noted here that, in our system, the compression and the optimization steps are distinct. In this manner, our redundancy allocation algorithm is applied directly to compressed source layers, that is, the algorithm actually parses the compressed stream to multiple descriptions. This clearly differentiates our algorithm from the method in [16] in which the generation of descriptions is performed by application of appropriate quantizers to the transform coefficients.

The structure of the paper is as follows. In Section 2, the proposed framework for multiple description coding of video is presented. Section 3 describes the wavelet coding of intraframes and motion compensation residuals. In Section 4, the exploitation of temporal correlation during the optimization process is discussed. In Section 5, the redundancy allocation problem is formulated. The complexity of the redundancy allocation algorithm is studied in Section 6,

and a faster algorithm is presented in Section 7 based on the Equivalent Continuous Problem. In Section 8, experimental results are presented and finally conclusions are drawn in Section 9.

2. PROPOSED FRAMEWORK FOR MULTIPLE DESCRIPTION GENERATION

The proposed system for the generation of multiple descriptions is depicted in Figures 1 and 2. Initially, the available bit budget is evenly allocated to the frames in a group of pictures (GOP). The first frame in each GOP is intra-coded using block-based wavelet coding. The resulting coded stream is distributed over a number of descriptions. A portion of the bitstream is redundant in all descriptions. The correlation between consecutive frames is subsequently removed using overlapped block motion compensation (OBMC) [18]. The reference frames used to calculate motion vectors are the original frames in order to ensure good precision in the estimation of the motion vectors. Motion vectors are losslessly coded using the techniques in [19] and are included in *all* descriptions.

Using the previously estimated half-pixel accurate motion vectors, the procedure for the generation of multiple descriptions for the interframes continues as follows: initially, the first interframe is compensated. No intra-coding is used in interframes. We employ two different mechanisms for the derivation of reference frames that are used during motion compensation. In the first, a version of the I-frame, reconstructed using *only the redundant part* of the bitstream so far coded, is used as reference for the compensation process. In the second, *both redundant and nonredundant parts* are used for the derivation of reference frames in motion compensation. The prediction error is derived by subtracting the compensated prediction from the original interframe. The prediction error is wavelet transformed and coded into multiple descriptions. A version of the error frame is reconstructed using either the redundant part or both redundant and nonredundant information of the coded bitstream depending on which of the two mechanisms described above is used. The reconstructed error frame is added to the compensated frame. The resulting interframe (instead of the original) will serve as the reference frame for the compensation of the next interframe. The same procedure is iterated until all frames in a GOP are treated.

Using the above methodology, the proposed multiple description video coding scheme is able to produce an arbitrary number of descriptions at the cost of reduced compression efficiency whenever the number of descriptions is large. In each description, there is a *redundant part*, which is always used for the derivation of the reference frame in the motion compensation process, and a complementary *refinement part*, which is used to improve the quality of each description and may or may not be used for the derivation of the reference frame. When both redundant and nonredundant information is used, reference frames of high quality are available. When only the redundant part is used, the motion compensation process performed at the encoder can be identically

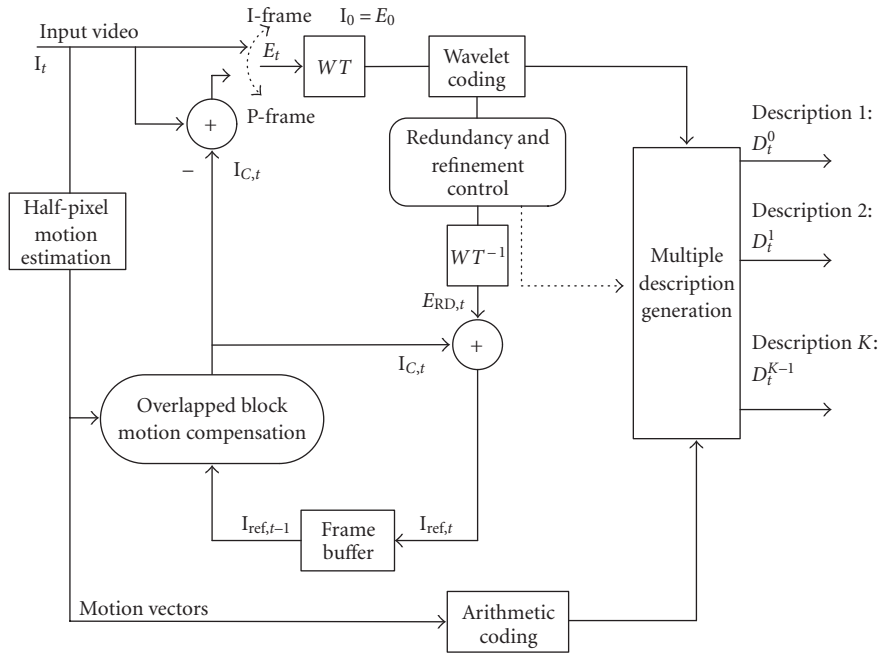


FIGURE 1: Block diagram of the coder.

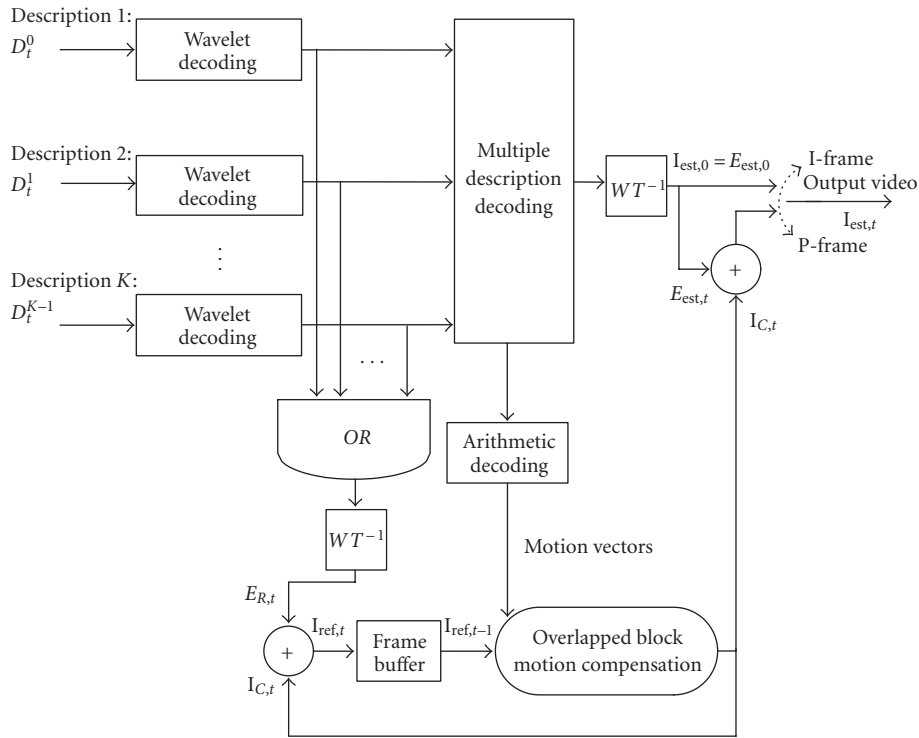


FIGURE 2: Block diagram of the decoder.

replicated at the decoder even if only one description is received. This is a very important feature of our coder since, if the decoder is unable to use the same reference frames, errors will accumulate in the decoded video sequence causing the aforementioned drift distortion [20]. With the proposed methodology, which relies only on the redundant part for

motion compensation, the possibility of facing drift at the decoder is eliminated and thus a reconstructed sequence of high quality is obtained even if only some (or even a single) descriptions are received.

The determination of the portion of the bitstream that is redundant in all descriptions is performed after the wavelet

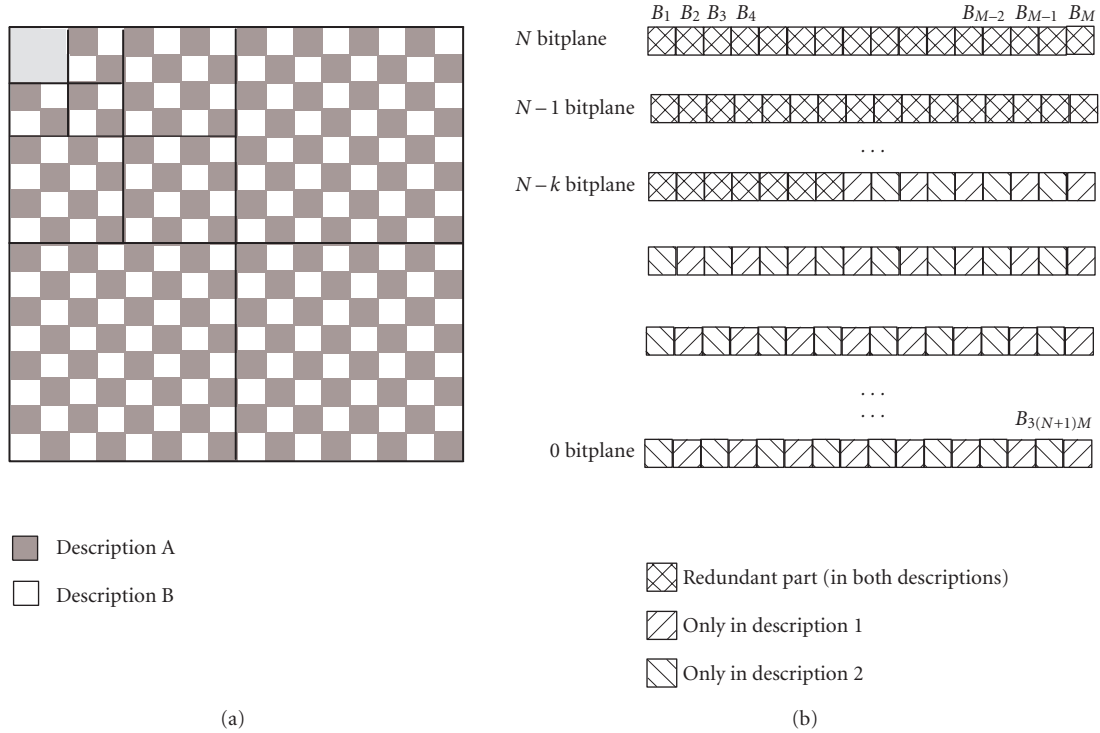


FIGURE 3: (a) Assignment of the blocks of a wavelet representation for the case of two descriptions. The bitstreams corresponding to the blocks may be included in one or more descriptions. (b) Representation of redundant and nonredundant part of the stream for the case of two descriptions.

coding of the intra and the residual error frames. The wavelet coefficients are coded using a simple bitplane encoder, based on the context models in [21]. Specifically, the decomposed frame is divided into blocks of equal dimensions. Each block may be included in some or all descriptions. Thus, some blocks may appear in all descriptions whereas some other blocks appear in only one of the descriptions. The inclusion of blocks in one or more descriptions is done so as to maximize the average quality at the decoder, subject to a total rate constraint, and attain fairly equal bitrate and fairly equal quality descriptions. Such an assignment is depicted in Figure 3(a). A representation of the redundant and nonredundant part of the coded bitstream for a two-description system is shown in Figure 3(b).

The generation of descriptions can be achieved by including appropriate blocks of wavelet coefficients in one or both of the descriptions. In the case of two descriptions, this is achieved by using the checkerboard pattern which we originally proposed in [9]. This approach bears some resemblance with the flexible macroblock ordering (FMO) approach in H.264 (see, e.g., [22]). However, there are fundamental differences between FMO and our approach which arise from the fact that our method operates in the wavelet domain whereas FMO is applied in the spatial domain. Since the FMO approach uses spatial blocks, the loss of a block would mean complete loss of information for that spatial region. This is why in FMO at least a coarsely quantized version of a chess-block need be included in each description. Clearly,

this means that using FMO there is much less control over redundancy since information about all blocks need be encoded in both descriptions. Moreover, since redundancy is introduced by the use of different quantizers, and not by explicitly including the same portion of the bitstream in all descriptions, the elimination of drift is not a trivial task. Finally, in FMO there is a need for error concealment in case the reconstructed quality in a spatial region is not good. Unlike the FMO approach, in our system, a loss of a wavelet block (due to the loss of the description in which the block is encoded) causes only the loss of some detail in the reconstructed frame. Moreover, in our method, most wavelet blocks are included in only one of the descriptions and only a few important blocks are included in both descriptions. This is possible since the wavelet transform compacts the important information in a few blocks (subbands) of transform coefficients. This strategy seems to be naturally more suitable for MD coding since it allows better manipulation of redundancy and generally achieves lower redundancy levels.

Throughout our manuscript we assume that no B-frames are encoded (see Figure 4). However, this assumption does not affect the significance of our work, which can also be applied when using B-frames. Suppose that we have an intra-coded frame, several (unidirectionally predicted) interframes, and some other frames that are to be bidirectionally predicted using the intra- and interframes. Apparently, our MD generation methodology is directly applicable to the sequence of intra- and interframes. In each description,

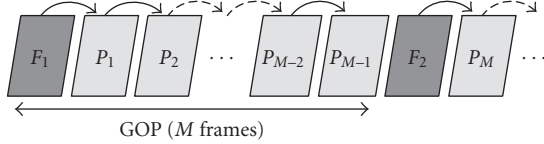


FIGURE 4: Structure of a group of pictures (GOP) in the proposed coding scheme where F_1, F_2 are intra-coded frames and P_1, P_2, \dots, P_M are interframes.

bidirectionally predicted frames could be encoded based on the reconstructions of intra- and interframes which are achieved using the bitstream in the *same* description. Note that, since B-frames do not propagate errors and do not cause drift, the reconstructed versions of intra- and interframes can be obtained using not only the redundant part of the description but also using the nonredundant part as well. An interesting and desirable result of this strategy is that, as these reconstructions will be different in the two descriptions, the associated residuals of the bidirectionally predicted frames will be inherently different in the two descriptions. This is perfectly consistent with the MD coding principle of encoding different versions of the information in each description.

In the ensuing section, the complete wavelet coding method, used for both intra- and interframes, is described.

3. BLOCK-BASED WAVELET CODING OF MOTION COMPENSATION RESIDUALS

The intra-frame and the motion-compensated residuals are decomposed using a wavelet transform based on the 9–7 biorthogonal filter bank [23]. The maximum absolute coefficient in each subband is placed in the image header. All subband maxima are arithmetically encoded. The transmission of information takes place in a bitplane-wise manner starting from the most significant bit (MSB) to the least significant bit (LSB). Within each bitplane, subbands are encoded in a predefined scanning order from the lowest to the highest resolution.

Each subband is divided into a set of blocks. The default block size is $(W/2^{L+1}) \times (H/2^{L+1})$, where W, H are the width and height of the frame, respectively, and L is the maximum level of the wavelet decomposition. For each block, first the coefficients whose most significant bit is on the bitplane currently coded are identified by comparison to a threshold $T = 2^n$, where n is the index of the bitplane that is being coded. If a coefficient becomes significant, that is, it is found to be greater than or equal to T for the first time, then its sign is coded. This process is often called *significance identification* [24] and the compressed significance map for a block is termed *significance layer*. Similarly, the *refinement layer* is defined as the one containing the n th bitplane of coefficients (in a block) found significant in previous passes. In our coder, refinement layers for the n th bitplane are transmitted immediately after the transmission of significance layers for the same bitplane. Note that each layer contains significant or refinement information for a single block and that the even-

tual allocation of layers in descriptions is performed by taking into consideration the fact that the decoding of a layer is possible only when all its predecessor layers in the same block are also included in the description.

The n th bit in the binary representation of a coefficient f in subband \mathcal{B} is coded if the maximum coefficient in the subband \mathcal{B} is greater than or equal to the current threshold

$$\max_{f \in \mathcal{B}}(f) \geq 2^n. \quad (1)$$

The deployment of the above rule reduces drastically the number of coefficients whose significance is tested during the coding of a significance identification layer. For this reason, subband maxima are included in all descriptions. However, in order to further reduce the number of symbols that have to be coded during the layer coding stage, a single bit is initially coded to indicate whether all coefficients in a block are insignificant. A value of “1” of this bit indicates that the block contains no significant coefficients and no further information is coded for this block.

The symbol streams described above are coded using adaptive arithmetic codes [25]. The context modelling strategy in [21] is followed for the coding of significance identification layers. Refinement bits are entropy coded using a single adaptive arithmetic model. The max frequency count of the arithmetic coder was set equal to 512 in order to allow fast adaptation of the coder to the statistics of the incoming symbol stream.

In order to apply an efficient redundancy allocation algorithm that takes into account the actual rate-distortion characteristics of the compressed stream, the distortion decrease achieved by the transmission of each bitplane should be calculated [21, 26] for each layer. The distortion decrease caused by the transmission of the i th layer is given by

$$D_i = \sum_t \left((\hat{f}_t^{n+1} - f_t)^2 - (\hat{f}_t^n - f_t)^2 \right), \quad (2)$$

where n is the index of the bitplane included in the layer, t is the coefficient index, and c, \hat{c} denote the original and the reconstructed wavelet coefficients, respectively. Each layer corresponding to a specific block of wavelet coefficients cause different reduction in the distortion. Analytical expressions for the distortion reduction caused by the transmission of layers can be found in [26]. Let R_i be the number of bits required for the coding of the i th layer. When all pairs (D_i, R_i) are determined, the redundancy allocation algorithm can be applied. This is examined in the following sections.

4. TEMPORAL CORRELATION COMPUTATION

An optimization algorithm should take into consideration the temporal correlation linking adjacent video frames. Modelling the dependency of adjacent frames in a video sequence is a nontrivial problem. In this paper, in order to deal with this issue, we introduce a temporal correlation coefficient a_i , $0 \leq a_i < 1$, meant to incorporate the effect of temporal correlation of layer i into the optimization algorithm. Specifically, we assume (a similar conclusion was

drawn in [27]) that the distortion reduction in frame $m + 1$ is $a_i D_i$, where m is the frame index. In the same manner, the additional distortion reduction $a_i D_i$ in frame $m + 1$ stimulates additional distortion reduction $a_j(a_i D_i)$ in frame $m + 2$, $a_k(a_j(a_i D_i))$ in frame $m + 3$ and so on, where a_j, a_k, \dots are the temporal correlation coefficients for frames $m + 1, m + 2, \dots$ correspondingly. We further assume that a_i, a_j, a_k are approximately equal for all frames in a GOP since the dependency between consecutive frames in the same GOP is not expected to exhibit significant variations. In general, the distortion reduction in frame n caused by the transmission of the i th layer in frame m , $m < n$, is $a_i^{n-m} D_i$. Thus, as the temporal distance $n - m$ between m and n increases the additional distortion reduction decreases exponentially. Assuming that the total number of frames in a GOP is M , the total distortion decrease is given by

$$D_i + a_i D_i + a_i^2 D_i + \dots + a_i^{M-m} D_i, \quad (3)$$

where $a_i D_i$ is the distortion reduction caused in the $m + 1$ frame, $a_i^2 D_i$ is the distortion reduction in the $m + 2$ frame, and so forth. The above quantity is equivalently written as the sum

$$D_i + (a_i + a_i^2 + \dots + a_i^{M-m}) D_i, \quad (4)$$

where the first term is the distortion reduction in the current frame and the second term denotes the distortion reduction in all subsequent frames. If

$$C_i = a_i + a_i^2 + \dots + a_i^{M-m} = \sum_{n=1}^{M-m} a_i^n = \frac{a_i - a_i^{M-m}}{1 - a_i}, \quad (5)$$

the total distortion reduction caused by the transmission of the i th layer in the m th frame can now be expressed as

$$D_i + D_i C_i, \quad (6)$$

where $D_i C_i$ is the cumulative distortion reduction¹ that is caused in the subsequent frames due to the higher quality of the current (reference) frame m . Clearly, with this formulation, layers in frames lying in the beginning of a GOP are more important than layers of frames at the end of the GOP since the quality of the former affects the quality of the latter. The coefficients a_i , and hence C_i , which quantify the impact of the current frame on the quality of subsequent frames were calculated using the methods in [27].

5. FORMULATION OF THE REDUNDANCY ALLOCATION PROBLEM

In order to address the problem of optimal allocation in MD video coding, it is important to derive expressions for the average video quality at the decoder and the total rate used in terms of the assignment strategy. Although in the experimental results section we consider the average PSNR over the entire sequence, in this section we will attempt to maximize

the distortion reduction incurred by each frame of the GOP separately. This simplification will not significantly affect the optimality of the strategy derived here, while it will serve in addressing the problem of optimal assignment in a more rigorous way and in providing useful insight into the optimization procedure.

Let us assume that each frame is coded into L layers, each using R_i bits and contributing a reduction of distortion equal to D_i relative to the quality of the current frame and $C_i D_i$, $i = 1, \dots, L$, to the quality of the next frames in the GOP,² when used for motion compensation for the next frames. We further assume that the curve appearing in Figure 5(a) is concave, namely,

$$\frac{D_1}{R_1} \geq \frac{D_2}{R_2} \geq \dots \geq \frac{D_L}{R_L}. \quad (7)$$

This assumption is generally valid for the case of our coder (a curve based on real data is shown in Figure 5(b)). We further note that lower-indexed layers correspond to coarse image information whereas high-indexed layers correspond to detail information. Between adjacent frames, coarse information is much more correlated than detail information. Thus, a_i is fully expected to decrease with i . Since C_i is obviously a monotone function of a_i , this implies that:

$$C_1 \geq C_2 \geq \dots \geq C_L, \quad (8)$$

an observation which is also verified experimentally. This ensures that (7) will still hold, if we replace the D_i 's with $D_i(1 + C_i)$, that is,

$$\frac{D_1(1 + C_1)}{R_1} \geq \frac{D_2(1 + C_2)}{R_2} \geq \dots \geq \frac{D_L(1 + C_L)}{R_L}. \quad (9)$$

We wish to encode the initial video sequence into K descriptions, each of which will either provide a coarse reconstruction of the initial sequence by itself or improve a reconstruction based on one of the other descriptions. To this end, for every frame in the GOP we will assign a number of layers to each description in a way so as to maximize the distortion reduction incurred under a limited-rate constraint. We will consider the case of double-description coding ($K = 2$). The general case is studied in Appendix B.

Let $I = \{1, \dots, L\}$ denote the set of the possible values that the layer indices may assume. The problem of providing two descriptions for each frame in the GOP is equivalent to assigning a set of layer indices $I_1 \subset I$ to the first and a set $I_2 \subset I$ to the second description. Subsequently, the two descriptions will be transmitted over two communication links to the decoder. If A_k represents the event that description k reaches the decoder and p denotes the probability that each stream is successfully delivered to the decoder (i.e.,

¹ Even though all coefficients D_i , a_i , and C_i depend on the frame index m , this dependence will in the sequel be omitted for convenience.

² For the last frame in the GOP $C_i = 0$, $i = 1, \dots, L$.

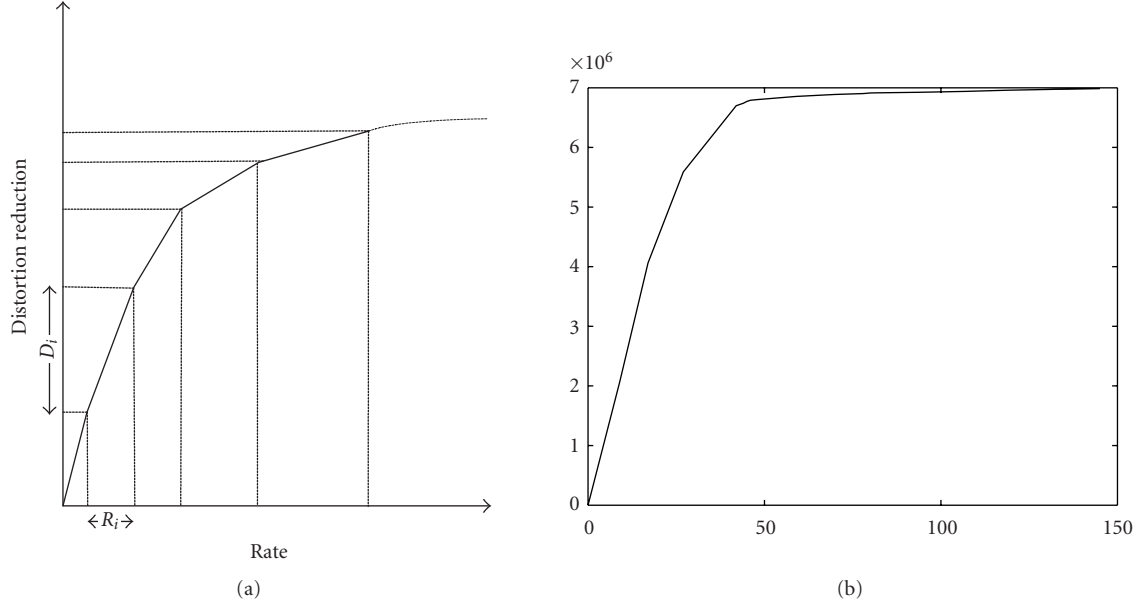


FIGURE 5: (a) Comprising layers and induced distortion reduction, (b) distortion reduction as a function of rate for a frame of “Akiyo” using the source coder of Section 3.

$p = \Pr\{A_k\}$, $k = 1, 2$), four events exist for each frame:

- $\mathbf{B}_1 \triangleq \mathbf{A}_1 \setminus \mathbf{A}_2$: only the first description is delivered
- $\mathbf{B}_2 \triangleq \mathbf{A}_2 \setminus \mathbf{A}_1$: only the second description is delivered
- $\mathbf{B}_{12} \triangleq \mathbf{A}_1 \cap \mathbf{A}_2$: both descriptions are delivered
- $\mathbf{B}_0 \triangleq A_1^c \cap A_2^c$: no descriptions are delivered.

The probability of each of these events may be easily derived if we make the reasonable assumption that the events A_1 and A_2 are independent:

$$\begin{aligned} \Pr\{B_1\} &= p(1-p), & \Pr\{B_2\} &= p(1-p), \\ \Pr\{B_{12}\} &= p^2, & \Pr\{B_0\} &= (1-p)^2. \end{aligned} \quad (10)$$

Let $d(B_1)$, $d(B_2)$, $d(B_{12})$, $d(B_0)$ denote, respectively, the distortion reduction at the decoder for the current frame when each of the events B_1 , B_2 , B_{12} , and B_0 occurs. Their values may be calculated as

$$\begin{aligned} d(B_1) &= \sum_{i \in I_1} D_i, & d(B_2) &= \sum_{i \in I_2} D_i, \\ d(B_{12}) &= \sum_{i \in I_1 \cup I_2} D_i, & d(B_0) &= 0. \end{aligned} \quad (11)$$

Moreover, when at least one of the descriptions arrives at the decoder, the layers common to all descriptions will be used for the motion compensation of the next frame in the GOP, incurring an additional distortion reduction of $C_i D_i$ for each layer. Let $B_{12} \triangleq B_0^c$ denote the event that at least one description reaches the decoder and $I_\cap \triangleq I_1 \cap I_2$ denote the set of indices common to both descriptions. Then, $\Pr\{B_{12}\} = p(2-p)$ and the corresponding distortion reduc-

tion will be

$$d(B_{12}) \triangleq \sum_{i \in I_\cap} C_i D_i. \quad (12)$$

Consequently, the expected distortion reduction, $\bar{D}_e(I_1, I_2)$, incurred at the decoder, when the index-assignment policy (I_1, I_2) is used, will be

$$\begin{aligned} \bar{D}_e(I_1, I_2) &= \Pr\{B_1\}d(B_1) + \Pr\{B_2\}d(B_2) \\ &\quad + \Pr\{B_{12}\}d(B_{12}) + \Pr\{B_0\}d(B_0) \\ &= p(1-p) \sum_{i \in I_1} D_i + p(1-p) \sum_{i \in I_2} D_i \\ &\quad + p^2 \sum_{i \in I_1 \cup I_2} D_i + p(2-p) \sum_{i \in I_\cap} C_i D_i, \end{aligned} \quad (13)$$

and after some simple manipulations we arrive at

$$\bar{D}_e(I_1, I_2) = p(2-p) \sum_{i \in I_\cap} D_i(1+C_i) + p \sum_{i \in I_\cup} D_i, \quad (14)$$

where $I_\cup \triangleq (I_1 \cup I_2) \setminus I_\cap$ is the set of indices contained in exactly one of the descriptions.

The total rate, $\bar{R}(I_1, I_2)$, used by the two streams is

$$\bar{R}(I_1, I_2) = \sum_{i \in I_1} R_i + \sum_{i \in I_2} R_i, \quad (15)$$

and may also be expressed as

$$\bar{R}(I_1, I_2) = 2 \sum_{i \in I_\cap} R_i + \sum_{i \in I_\cup} R_i. \quad (16)$$

Assuming that the total rate used may not exceed a pre-defined rate budget R_B , our purpose is to identify the index-assignment sets I_1 and I_2 , which do not violate the rate constraint and maximize the expected distortion reduction at the decoder

$$\max_{I_1, I_2: \bar{R}(I_1, I_2) \leq R_B} \bar{D}_e(I_1, I_2). \quad (17)$$

It is clear from (14) and (16) that the expected distortion reduction and total rate depend upon the sets I_\cap and I_\cup . Furthermore, the factor p in the expected distortion reduction (14) may be ignored for the optimization procedure for the sake of simplicity. Therefore, the maximization problem may be rephrased as

Maximization problem

Find disjoint sets $I_\cap, I_\cup \subset I$ maximizing

$$\bar{D}(I_\cap, I_\cup) = (2 - p) \sum_{i \in I_\cap} D_i(1 + C_i) + \sum_{i \in I_\cup} D_i \quad (18)$$

subject to the constraint

$$\bar{R}(I_\cap, I_\cup) = 2 \sum_{i \in I_\cap} R_i + \sum_{i \in I_\cup} R_i \leq R_B. \quad (19)$$

The solution of the above problem will yield the optimal sets I_\cap and I_\cup , where I_\cap will contain the indices of the layers assigned to both streams and I_\cup will contain the indices assigned only to one of the streams. In order to obtain the optimal I_1, I_2 , we need to further partition I_\cup into two disjoint index-assignment sets, one for each stream. It is clear from (14), however, that any such partition will yield sets I_1, I_2 , inducing the same expected distortion reduction at the decoder; hence, the partition of I_\cup may be arbitrary (we may even assign the whole set I_\cup to only one of the streams). However, since balanced MD coding is sought, an acceptable partitioning should result in fairly equal total rates of I_1 and I_2 . In order to achieve this, the indices in I_\cup may be ordered in terms of decreasing corresponding rates R_i and be assigned alternately to each stream.

6. COMPLEXITY ANALYSIS

If we were to solve the maximization problem (17) by exhaustively examining all possible realizations of I_1 and I_2 , this would involve 2^{2L} possibilities, since there are 2^L subsets of the index set I . Clearly, the optimal solution will be achieved by choosing any pair of sets I_1 and I_2 resulting in the same sets I_\cap^* and I_\cup^* , which solve the maximization problem described by (18) and (19). Hence, we only need to examine all possible realizations of disjoint sets $I_\cap, I_\cup \subset I$.

Note that since there are 2^L possible subsets of the index set I , any subset $A \subset I$ may be expressed as the binary

```

max D = 0 (maximum distortion originally 0)
I_\cap^* = I_\cup^* = \emptyset (optimal sets originally empty)
for I_\cap = 0, \dots, 2^L - 1 (all possible realizations of I_\cap)
  for I_\cup = 0, \dots, 2^L - 1 (all possible realizations of I_\cup)
    if (I_\cap AND I_\cup) \neq \emptyset (check if sets are disjoint)
      if (19) is satisfied (check rate constraint)
        Calculate expected distortion reduction
        \bar{D}(I_\cap, I_\cup) from (18)
        if \bar{D}(I_\cap, I_\cup) > max D, update max D, I_\cap^* and I_\cup^*
          (update optimal sets)
        endif
      endif
    endif
  endfor
endfor
Partition I_\cup^* into two fairly equal-rate subsets I_\cup^{*(1)} and I_\cup^{*(2)}.
The optimal index assignment is given by I_1^* = I_\cap^* \cup I_\cup^{*(1)},
I_2^* = I_\cap^* \cup I_\cup^{*(2)}.

```

ALGORITHM 1: Exhaustive search algorithm.

representation of a number between 0 and $2^L - 1$, with the i th bit being 1, if $i \in A$ and 0 otherwise. An exhaustive search algorithm which will determine the optimal solution I_\cap^*, I_\cup^* to the maximization problem is shown in Algorithm 1.

Although this algorithm will always produce an optimal solution, the number of possible realizations of I_\cap and I_\cup , over which the search will be performed, is 3^L , still prohibitive even for moderate values of L . The NP-completeness of the maximization problem described by (18) and (19) can also be shown by formulating it as an integer (0–1) programming problem as shown in Appendix A.

In view of these remarks, it would be desirable to establish some optimality results that will narrow the number of possible candidate solutions or devise techniques that would search through a smaller set of possible near-optimal solutions. To this end, the following will prove helpful.

Lemma 1. *If I_\cap and I_\cup are fixed and $j \in I_\cap$ or $j \in I_\cup$, replacing layer j with layers of higher indices, such that their total rate does not exceed R_j , would result in smaller expected distortion reduction.*

Proof. Assume that $j \in I_\cap$ (the proof for $j \in I_\cup$ is similar) and $j_1, \dots, j_k \notin I_\cap, I_\cup$ with $j \leq j_1 \leq \dots \leq j_k$ and

$$\sum_{i=1}^k R_{j_i} \leq R_j. \quad (20)$$

If I_\cap is replaced by the set $\hat{I}_\cap \triangleq (I_\cap \setminus \{j\}) \cup \{j_1, \dots, j_k\}$, then the rate constraint (19) would still be satisfied and the expected distortion reduction (18) would decrease by

$$\begin{aligned} & \bar{D}(I_\cap, I_\cup) - \bar{D}(\hat{I}_\cap, I_\cup) \\ &= (2 - p) \left[D_j(1 + C_j) - \sum_{i=1}^k D_{j_i}(1 + C_{j_i}) \right]. \end{aligned} \quad (21)$$

Using (9) and (20) it is straightforward to show that the outcome of (21) is nonnegative; hence, this replacement would prove inefficient. \square

The same also holds if we were to replace more than one lower-indexed layers with higher-indexed ones of smaller total rate. In other words, Lemma 1 suggests that, if possible (i.e., if the rate constraint is not violated), we should replace higher-indexed layers with lower-indexed ones with appropriate total rate. However, Lemma 1 might mislead us to assume that the optimal solution would consist of sets I_\cap^* and I_\cup^* comprising the lower-indexed layers, that is,

$$I_\cap^* = \{1, \dots, L_\cap^*\}, \quad I_\cup^* = \{L_\cap^* + 1, \dots, L_\cup^*\}, \quad L_\cap^* \leq L_\cup^*. \quad (22)$$

This would not be true in case the rate margin $R_M \triangleq R_B - 2 \sum_{i \in I_\cap} R_i - \sum_{i \in I_\cup} R_i$ can be filled by replacing one (or more) of the lower-indexed layers j with one or more higher-indexed layers $j \leq j_1 \leq \dots \leq j_k$, such that $2 \sum_{i=1}^k R_{j_i} \leq 2R_j + R_M$. It is possible that in this case the resulting expected distortion reduction actually be larger, as shown in the example below.

Counterexample 1. Let $R_B = 21.5$, $p = 0.8$, $C_i = 0$, $i = 1, \dots, L$, and R_i, D_i given by the following table:

i	1	2	3	4	5
R_i	5	4	3	2	1.5
D_i	0.9	0.7	0.4	0.25	0.18

It turns out that the optimal sets I_\cap, I_\cup of the form (22) are $I_\cap = \{1\}$ and $I_\cup = \{2, 3, 4, 5\}$ ($L_\cap^* = 1, L_\cup^* = 5$) resulting in total rate 20.5 and expected distortion reduction 2.61. There is, however, a rate margin $R_M = R_B - 20.5 = 1$ that may be taken advantage of, if I_\cap or I_\cup is properly chosen. In fact, if the sets $I_\cap = \{2, 4\}$ and $I_\cup = \{1, 4, 5\}$ are used, the total rate matches the rate budget R_B and the expected distortion reduction increases slightly to 2.62.

This counterexample verifies that the optimal solution will not always be of the form (22); however, extensive experimentation showed that in most cases the sets I_\cap and I_\cup given by (22) provide a near-optimal solution, as was indeed the case in the previous example.

An improved exhaustive search algorithm, which stems from this remark, would consider only sets I_\cap, I_\cup of the form (22). The number of possible candidates may be further reduced based on the following lemmas.

Lemma 2. L_\cup^* cannot exceed any certain value beyond which the sum $\sum_{i=1}^{L_\cup^*} R_i$ exceeds the rate budget R_B .

Proof. This lemma is a direct consequence of the total rate constraint (19) for $L_\cap^* = 0$. \square

Lemma 3. L_\cup^* cannot be smaller than any value for which the sum $\sum_{i=1}^{L_\cup^*} R_i$ does not exceed $R_B/2$.

```

max D = 0 (maximum distortion originally 0)
L_\cap^* = L_\cup^* = 0 (optimal sets originally empty)
L_1 = max{l \in I : \sum_{i=1}^l R_i \leq R_B/2} (smallest value for L_\cup^*)
L_2 = max{l \in I : \sum_{i=1}^l R_i \leq R_B} (largest value for L_\cup^*)
L_\cap = L_1 (initial value for L_\cap)
for L_\cup = L_1, \dots, L_2 (all possible values of L_\cup)
  while \sum_{i=1}^{L_\cap} R_i > R_B - \sum_{i=1}^{L_\cup} R_i
    decrease L_\cap
  endwhile
  I_\cap = \{1, \dots, L_\cap\}, \quad I_\cup = \{L_\cap + 1, \dots, L_\cup\}
  (corresponding index-assignment sets)
  Calculate expected distortion reduction
  \bar{D}(I_\cap, I_\cup) from (18)
  if \bar{D}(I_\cap, I_\cup) > max D update max D, L_\cap^*, and L_\cup^*
  (update optimal values)
endfor
I_\cap^* = \{1, \dots, L_\cap^*\}, I_\cup^* = \{L_\cap^* + 1, \dots, L_\cup^*\}
(optimal index-assignment sets)

```

ALGORITHM 2: Improved exhaustive search algorithm.

Proof. If $\sum_{i=1}^{L_\cup^*} R_i \leq R_B/2$, the best choice for L_\cap^* is $L_\cap^* = L_\cup^*$, since the rate constraint will still be met. If there exists a $l > L_\cup^*$ with $\sum_{i=1}^l R_i \leq R_B/2$, then setting $L_\cap^* = L_\cup^* = l$ improves $\bar{D}(I_\cap, I_\cup)$. \square

Lemma 4. For a given L_\cup^* , the optimal value of L_\cap^* is the largest integer $l \leq L_\cup^*$, for which the total rate for I_\cap does not exceed the remaining available rate, $2 \sum_{i=1}^l R_i \leq R_B - \sum_{i=l+1}^{L_\cup^*} R_i \Leftrightarrow \sum_{i=1}^l R_i \leq R_B - \sum_{i=1}^{L_\cup^*} R_i$.

Proof. It is straightforward to prove that the more layers I_\cap comprises, the better the distortion reduction will be. Therefore, we should try to “fit” as many layers as possible in the remaining available rate. \square

Lemmas 2–4 may be used to narrow down the exhaustive search space. In particular, Lemmas 2 and 3 suggest that we should examine values of L_\cup^* , in a set $\{L_1, \dots, L_2\}$, while Lemma 4 suggests that for each of these values of L_\cup^* there is a unique optimal value of L_\cap^* ; hence, it suffices to examine only $L_2 - L_1 + 1 < L$ cases. In view of these results, we can describe the improved exhaustive search procedure in Algorithm 2.

The while loop in this algorithm searches for the maximum value of L_\cap fitting in the rate margin, since, as can be easily verified, the corresponding value of L_\cap for $L_\cup + 1$ will be smaller than that for L_\cup (the previous value of L_\cap). Hence, the search is performed over $L_2 - L_1 + 1$ possible values of L_\cup^* and L_1 possible values of L_\cap^* and the complexity of the algorithm will be linear in L .

In general, the improved exhaustive search algorithm will result in sets I_\cap^* and I_\cup^* , which do not exactly meet the rate constraint. In this case, there will be a rate margin $R_M \triangleq R_B - 2 \sum_{i \in I_\cap^*} R_i - \sum_{i \in I_\cup^*} R_i$, which can be “filled” with smaller segments outside I_\cap^* or I_\cup^* . A further improvement would search for possible augmentations of I_\cap^* or I_\cup^* , so that the total rate be closer to the rate budget R_B .

As already stated, this algorithm will, in general, yield suboptimal yet near-optimal solutions to the maximization problem. A further (and more important) disadvantage of this algorithm is that, when applied in the general case of $K > 2$ descriptions, its complexity will be even higher. If we are to construct a low-complexity algorithm for the general case, we may resort to heuristics emanating from a continuous-case consideration of the problem. This is explored in the next section.

7. EQUIVALENT CONTINUOUS PROBLEM

By examining closely the discrete maximization problem described by (18) and (19), we first note that the sums $\sum_{i \in I_\cap} D_i(1 + C_i)$, $\sum_{i \in I_\cap} R_i$ and $\sum_{i \in I_\cup} D_i$, $\sum_{i \in I_\cup} R_i$ are the distortion reduction and rate “measures” of I_\cap and I_\cup respectively. A further restriction arises from the requirement that I_\cap and I_\cup have to comprise intervals dictated by the available blocks and that partial blocks may not be used. If we relax this restriction, we may formulate a corresponding Continuous Maximization Problem, which is easier to solve.

Assume that the curve appearing in Figure 5 represents a continuous, differentiable, nondecreasing, and concave function $D(R)$ of the rate R . Then the derivative $D'(R)$ will be a well-defined, continuous, positive, and decreasing function of R , for every $R \in \mathbb{R}_+$. In a similar fashion, assume that the fraction of distortion reduction due to motion compensation is provided by a continuous decreasing function $c(R)$ and that the curve corresponding to the products $D_i C_i$ defines a function $C(R)$ with derivative $C'(R) = D'(R)c(R)$, which will have properties similar to those of $D'(R)$.³ For any rate interval $[r_1, r_2]$, let μ_R, μ_D, μ_C denote the following quantities:

$$\begin{aligned}\mu_R([r_1, r_2]) &= \int_{r_1}^{r_2} dr = r_2 - r_1, \\ \mu_D([r_1, r_2]) &= \int_{r_1}^{r_2} D'(r) dr = D(r_2) - D(r_1), \\ \mu_C([r_1, r_2]) &= \int_{r_1}^{r_2} c(r) D'(r) dr = C(r_2) - C(r_1).\end{aligned}\quad (23)$$

In practice, the number of intervals of the form $[r_1, r_2]$ is always finite (with an upper bound equal to the number of bits in the compressed bitstream). Obviously, the measure of a union of a finite number of disjoint intervals of the form $[r_1, r_2]$ would equal the sum of the measures of these intervals. Thus, a continuous version of the discrete maximization problem described by (18) and (19) may now correspondingly be formulated as follows.

Continuous maximization problem

Find disjoint sets $S_\cap, S_\cup \subset \mathbb{R}_+$ maximizing

$$\bar{D}(S_\cap, S_\cup) = (2 - p)[\mu_C(S_\cap) + \mu_D(S_\cap)] + \mu_D(S_\cup) \quad (24)$$

subject to the constraint

$$\bar{R}(S_\cap, S_\cup) = 2\mu_R(S_\cap) + \mu_R(S_\cup) \leq R_B. \quad (25)$$

With the further reasonable assumption that S_\cap and S_\cup are unions of closed intervals, properties stronger than Lemma 1 may be established for the continuous problem, leading to optimal solutions.

Lemma 5. *If S_\cup is fixed, the optimal S_\cap comprises the “smallest-rate region” of the remaining space $\mathbb{R}_+ \setminus S_\cup$, that is,*

$$S_\cap^* = [0, R_\cap] \cap (\mathbb{R}_+ \setminus S_\cup), \quad (26)$$

for some positive rate R_\cap .

Proof. We will outline the general concept behind (26). Assume that (26) does not hold. Then there exist $\delta > 0$ and $r_2 > r_1 \geq 0$ such that the interval $[r_1, r_1 + \delta]$ is lying outside S_\cap (i.e., $[r_1, r_1 + \delta] \cap S_\cap = \emptyset$) and the interval $[r_2, r_2 + \delta]$ is contained in S_\cap (i.e., $[r_2, r_2 + \delta] \subset S_\cap$). If we replace S_\cap with $\tilde{S}_\cap \triangleq (S_\cap \setminus [r_2, r_2 + \delta]) \cup [r_1, r_1 + \delta]$ (remove the second interval and add the first), then the rate constraint will still be met and the increase in expected distortion reduction (24) will be

$$\begin{aligned}\bar{D}(\tilde{S}_\cap, S_\cup) - \bar{D}(S_\cap, S_\cup) &= (2 - p)(\mu_C([r_1, r_1 + \delta]) + \mu_D([r_1, r_1 + \delta]) \\ &\quad - \mu_C([r_2, r_2 + \delta]) - \mu_D([r_2, r_2 + \delta])) \\ &= (2 - p) \left(\int_{r_1}^{r_1 + \delta} D'(r)[1 + c(r)] dr - \int_{r_2}^{r_2 + \delta} D'(r)[1 + c(r)] dr \right) \\ &\stackrel{(\alpha)}{\geq} (2 - p) \left(\int_{r_1}^{r_1 + \delta} D'(r + r_2 - r_1)[1 + c(r + r_2 - r_1)] dr \right. \\ &\quad \left. - \int_{r_2}^{r_2 + \delta} D'(r)[1 + c(r)] dr \right) \\ &\stackrel{(\beta)}{=} (2 - p) \left(\int_{r_2}^{r_2 + \delta} C'(\rho) d\rho - \int_{r_2}^{r_2 + \delta} C'(r) dr \right) = 0,\end{aligned}\quad (27)$$

where (α) results from $r_2 - r_1 > 0$ and the fact that $D'(\cdot)$ and $c(\cdot)$ are decreasing and (β) involves a simple change of integration variable. It follows, therefore, that S_\cap will not be optimal (since it is outperformed by \tilde{S}_\cap) unless it is given by (26) for some R_\cap . \square

In a similar manner, it is possible to establish an equivalent property for S_\cup .

Lemma 6. *If S_\cap is fixed, the optimal S_\cup comprises the “smallest-rate region” of the remaining space $\mathbb{R}_+ \setminus S_\cap$, that is, $S_\cup^* = S_\cap \cup [0, R_\cup]$ for some positive rate R_\cup .*

³ In other words, $D'(R)$ corresponds to the ratios D_i/R_i and $c(R)$ to the coefficients C_i .

Furthermore, the concavity of $D(\cdot)$ implies the following.

Lemma 7. *If $r_1 < r_2$, $\delta > 0$ and $[r_1, r_1 + \delta] \in S_{\sqcup}$, $[r_2, r_2 + \delta] \in S_{\cap}$, then the sets $\tilde{S}_{\cap} \triangleq S_{\cap} \setminus [r_2, r_2 + \delta] \cup [r_1, r_1 + \delta]$ and $\tilde{S}_{\sqcup} \triangleq S_{\cap} \setminus [r_1, r_1 + \delta] \cup [r_2, r_2 + \delta]$ yield smaller expected distortion reduction.*

Proof. This is true because the contribution of S_{\cap} in the expected distortion reduction (24) involves the factor $2 - p > 1$ and the function $C(R) \geq D(R)$, $R \in \mathbb{R}_+$. Hence, incorporating the smaller-rate interval $[r_1, r_1 + \delta]$ in S_{\cap} and the higher-rate interval $[r_2, r_2 + \delta]$ in S_{\sqcup} will yield smaller expected distortion, as is easily be verified. \square

Lemmas 5, 6, and 7 suggest that the jointly optimal sets S_{\cap}^* , S_{\sqcup}^* will be intervals of the form

$$S_{\cap}^* = [0, R_{\cap}], \quad S_{\sqcup}^* = [R_{\cap}, R_{\sqcup}], \quad (28)$$

for some $R_{\sqcup} \geq R_{\cap} \geq 0$.

In terms of the original maximization problem, (28) would provide the optimal solution if the (0–1) constraint for \mathbf{x} is relaxed, namely, if assignment of partial blocks is allowed.

In view of (28), the equivalent continuous problem may be restated as follows.

Continuous maximization problem

Find positive rates $R_{\sqcup} \geq R_{\cap} \geq 0$ maximizing

$$\bar{D}(R_{\cap}, R_{\sqcup}) = (2 - p)[C(R_{\cap}) + D(R_{\cap})] + D(R_{\sqcup}) - D(R_{\cap}) \quad (29)$$

subject to the constraint

$$\bar{R}(R_{\cap}, R_{\sqcup}) = R_{\cap} + R_{\sqcup} \leq R_B. \quad (30)$$

This is a simple Lagrangian maximization problem with optimal solution R_{\cap}^* , R_{\sqcup}^* satisfying the constraint (30) at the boundary. The optimal R_{\cap}^* should satisfy

$$(2 - p)[C'(R_{\cap}^*) + D'(R_{\cap}^*)] - D'(R_B - R_{\cap}^*) - D'(R_{\cap}^*) = 0, \quad (31)$$

which after some simple manipulations translates to the condition

$$\phi(R_{\cap}^*) \triangleq \frac{D'(R_B - R_{\cap}^*)}{D'(R_{\cap}^*)} - (2 - p)c(R_{\cap}^*) = 1 - p. \quad (32)$$

Observe that, since $D'(\cdot)$ and $c(\cdot)$ are decreasing, $\phi(\cdot)$ will be continuous and increasing in the interval $[0, R_B/2]$ and the continuous maximization problem will not involve local maxima. Also, the smallest value of $\phi(\cdot)$ will be $\phi(0) = D'(R_B)/D'(0) - (2 - p)c(0)$ and the largest value will be

```

 $L_1 = \max\{l \in I : \sum_{i=1}^l R_i \leq R_B/2\}$  (index corresponding to  $R_B/2$ )
 $L_2 = \max\{l \in I : \sum_{i=1}^l R_i \leq R_B\}$  (index corresponding to  $R_B$ )
if  $\phi(1, L_2) > 1 - p$ , set  $L_{\cap}^* = 0$ ,  $L_{\sqcup}^* = L_2$  and exit (case  $1 - p < \phi(0)$  in (33))
if  $\phi(L_1, L_1) < 1 - p$ , set  $L_{\cap}^* = L_1$ ,  $L_{\sqcup}^* = L_1$  and exit (case  $1 - p > \phi(R_B/2)$  in (33))
 $L_{\cap} = 1$ ,  $L_{\sqcup} = L_2$  (initial values)
while  $L_{\cap} \leq L_{\sqcup}$ 
  while  $\sum_{i=1}^{L_{\sqcup}} R_i > R_B - \sum_{i=1}^{L_{\cap}} R_i$  decrease  $L_{\sqcup}$ 
    (find largest  $L_{\sqcup}$  satisfying rate constraint)
  if  $\phi(L_{\cap}, L_{\sqcup}) > 1 - p$  set  $L_{\cap}^* = L_{\cap}$ ,  $L_{\sqcup}^* = L_{\sqcup}$  and exit
    (crossed  $1 - p$  line)
  increase  $L_{\cap}$  (next value of  $L_{\cap}$ )
endwhile
 $L_{\cap}^* = L_1$ ,  $L_{\sqcup}^* = L_1$  (if this point is reached,  $\phi(L_{\cap}, L_{\sqcup})$  never
crossed the  $1 - p$  line)

```

ALGORITHM 3: Fast search algorithm.

$\phi(R_B/2) = 1 - (2 - p)c(R_B/2)$. Therefore, if $(1 - p) \in [\phi(0), \phi(R_B/2)]$, the optimal value for R_{\cap}^* will be $\phi^{-1}(1 - p)$. Otherwise (32) does not have a solution and optimality is achieved either at 0 or $R_B/2$. In general, we can write

$$R_{\cap}^* = \begin{cases} 0, & \text{if } 1 - p < \phi(0), \\ \phi^{-1}(1 - p), & \text{if } 1 - p \in [\phi(0), \phi(R_B/2)], \\ R_B/2, & \text{if } 1 - p > \phi(R_B/2), \end{cases} \quad (33)$$

while $R_{\sqcup}^* = R_B - R_{\cap}^*$.

Returning to the discrete maximization problem, it is reasonable to assume that a near-optimal solution will resemble that of the equivalent continuous maximization problem, especially for large values of L . This means that a near-optimal choice for the index assignment sets would be $I_{\cap} = \{1, \dots, L_{\cap}^*\}$, $I_{\sqcup} = \{L_{\cap}^* + 1, \dots, L_{\sqcup}^*\}$, where L_{\cap}^* and L_{\sqcup}^* would be such that

$$\sum_{i=1}^{L_{\cap}^*} R_i \simeq R_B - \sum_{i=1}^{L_{\sqcup}^*} R_i, \quad (34)$$

$$\phi(L_{\cap}^*, L_{\sqcup}^*) \triangleq \frac{D_{L_{\sqcup}^*}/R_{L_{\sqcup}^*}}{D_{L_{\cap}^*}/R_{L_{\cap}^*}} - (2 - p)C_{L_{\cap}^*} \simeq 1 - p, \quad (35)$$

This consideration suggests Algorithm 3 above.

The advantage of this algorithm lies in that it involves fewer calculations and terminates sooner than the improved exhaustive search algorithm. It is clear, however, that the price paid for its reduced complexity, which is important in cases of real-time applications, is its inferior performance compared to the exhaustive search algorithms.

Let us also note that the implementation of the fast search algorithm involves serial search through all values from 0 to the terminating, estimated optimal, value of L_{\cap}^* . A further improvement would involve a binary search modification of this algorithm, according to the actual values of $\phi(L_{\cap}, L_{\sqcup})$ at the boundaries of the binary-search interval.

TABLE 1: Description size (bytes) ratio and PSNR ratio of the two descriptions, for several frames of the sequence “Foreman” ($p = 0.9$, $R_{\text{total}} = 128$ Kbps).

Frame	Foreman	
	Bytes ratio	PSNR ratio
1	1.015	1.005
2	1.020	1.006
3	0.950	0.985
4	0.972	0.993
5	0.989	1.003

8. EXPERIMENTAL RESULTS

The proposed multiple description video coding scheme was experimentally evaluated for the transmission of the Y component (15 frames/second) of the standard test sequence “Foreman” over two channels. Each frame was coded in two descriptions. Motion vector information was duplicated in both descriptions. The proposed redundancy allocation Algorithm 3 of the preceding section was applied for video transmission over two channels of *total* capacity 128 Kbps and for three different probabilities of description arrival: $p = 0.8, 0.9, 0.95$, or equivalently three probabilities of description loss equal to 20%, 10%, 5%. The number of frames in each GOP was chosen with respect to p as suggested in [28]. The target rate R_B for each frame was determined by allocating to intra-frames a rate equal to four times the rate allocated to interframes. The resulting descriptions, as shown in Table 1 for the first five frames of the sequence, are remarkably “balanced,” that is, they have approximately equal size and yield almost equal reconstruction qualities.

In the present work, we assume that descriptions that arrive at the decoder do not contain bit errors. We examine two types of transmission scenarios: in the first scenario, we assume that the channels retain their status during the entire transmission. In this case, the parameter p serves as a means to control the redundancy and is not directly associated with the condition of the channel. In the second scenario, we assume that the channels go on and off during transmission. In the latter scenario, it is possible that both descriptions of a frame are lost. In such a case, the decoder uses the most recent reference frame that is available. For each frame, the peak-signal-to-noise-ratio is used as a measure of the reconstruction quality (in dB)

$$\text{PSNR} = 10 \log_{10} \frac{255^2}{\text{MSE}}. \quad (36)$$

Following the approach adopted in [29, 30], the reported mean PSNR values are computed by averaging decoded MSE values and then converting the mean MSE to the corresponding PSNR value rather than averaging the PSNR values directly.

In the first transmission scenario, the coding of the “Foreman” sequence into two descriptions is simulated under the respective assumption that the channels are available or unavailable during the entire transmission. As expected, the central distortion in the proposed scheme that allows

drift accumulation, which we will term multiple description wavelet video coder (MDWVC), is superior in comparison to the proposed drift-free system, termed DF-MDWVC. This was expected since when both descriptions are available, drift is eliminated anyway. On the other hand, the side distortion appears to be lower in the drift-free system. The performance of MDWVC is shown in Figure 6. The redundancy rate-distortion performance of our coders is shown in Figure 7. As seen, DF-MDWVC and MDWVC reach similar performances for redundancy greater than 15%. For lower redundancies, the drift-free system performs worse due to the very low quality of the reference frames.

In the second simulation, in which the channels may go on and off from frame to frame, we tested our systems under identical description loss patterns. For each frame, one, two, or none of the descriptions was lost. As seen from Figure 8 and Tables 2 and 3, the drift-free system is much more reliable and demonstrates no abrupt changes in its performance, contrary to MDWVC which demonstrates significant variations in the video quality it delivers. In addition, both schemes demonstrate significant gains over the single description scheme which appears to collapse very frequently due to description losses. In Figure 8(d), we report the performance of a scheme that is based on H.264 and uses the FMO for transmission of video over two channels. This scheme uses P-frames and two FMO slices. As seen, despite the fact that the H.264-based scheme uses advanced error concealment techniques at the decoder, the reconstruction quality it delivers exhibits significant variations in comparison to the quality achieved by our drift-free scheme.

Reconstructed frames obtained by simulating the transmission of 180 frames of the “Foreman” sequence at 15 frames/second over two channels of total capacity 128 Kbps and probability of description arrival equal to 0.9 using the above systems are displayed in Figure 9. The reconstruction displayed in Figure 9(c), achieved using the drift-free system, is qualitatively more pleasant than the reconstruction using MDWVC. This proves that, in practical cases, the drift-free system can be a better choice even though MDWVC operates better at low error rates. The image reconstructed using the single description scheme exhibits the worst performance.

In Figure 10, we present the reconstruction quality obtained using the drift-free system for the case of transmission over four channels of total capacity 128 Kbps and probabilities of description loss equal to 20%.

9. CONCLUSIONS

We presented a wavelet-based framework for the encoding of video in multiple descriptions. The generation of multiple descriptions was performed so that drift is eliminated at the decoder side. The proposed framework is flexible and allows the encoding of video into an arbitrary number of descriptions. The resulting framework is endowed with the capability for drift-free reconstruction regardless of the number of descriptions that arrived at the decoder. Three algorithms were also presented for the optimal allocation of

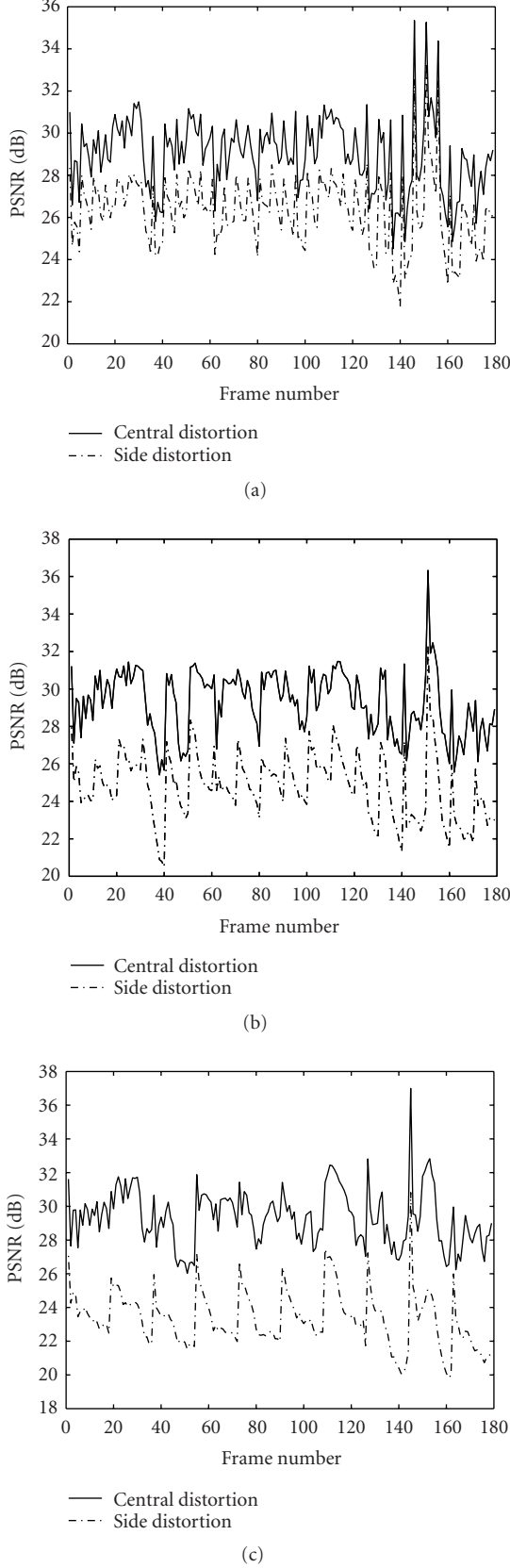


FIGURE 6: Reconstruction quality for the “Foreman” sequence using the MDWVC coder: (a) 20% probability of loss, (b) 10% probability of channel loss, and (c) 5% probability of loss.

TABLE 2: Performance comparison. PSNR is reported.

Packet loss	Single	Drift	Drift-free
20%	17.95	26.60	26.90
10%	21.70	27.32	27.49
5%	22.96	27.60	27.84

TABLE 3: Performance comparison. Standard deviation of reconstruction quality is reported.

Packet loss	Single	Drift	Drift-free
20%	7.07	2.58	1.94
10%	4.67	2.32	1.46
5%	5.05	2.00	1.34

redundancy. Experimental results for the transmission of video using two descriptions demonstrated the efficiency of the proposed method.

APPENDICES

A. INTEGER (0–1) PROGRAMMING FORMULATION

Let $x_i^\cap = \mathbf{1}_{I_\cap}(i)$, $x_i^\sqcup = \mathbf{1}_{I_\sqcup}(i)$, $i = 1, \dots, L$ denote binary-valued variables, where $\mathbf{1}_A$ is the indicator function of set A , that is, $\mathbf{1}_A(x) = 1$ if $x \in A$ and $\mathbf{1}_A(x) = 0$ if $x \notin A$. Then, the sets I_\cap and I_\sqcup are determined by the vectors $\mathbf{x}^\cap \triangleq [x_1^\cap, \dots, x_L^\cap]^T$ and $\mathbf{x}^\sqcup \triangleq [x_1^\sqcup, \dots, x_L^\sqcup]^T$, respectively, where \mathbf{A}^T denotes the transpose of matrix \mathbf{A} . If we adopt this notation, (18) may be written as

$$\bar{D}(I_\cap, I_\sqcup) = (2 - p)\mathbf{c}^T \cdot \mathbf{x}^\cap + \mathbf{d}^T \cdot \mathbf{x}^\sqcup, \quad (\text{A.1})$$

with $\mathbf{c} \triangleq [D_1(1 + C_1), \dots, D_L(1 + C_L)]^T$, $\mathbf{d} \triangleq [D_1, \dots, D_L]^T$, and constraint (19) as

$$\bar{R}(I_\cap, I_\sqcup) = 2\mathbf{r}^T \cdot \mathbf{x}^\cap + \mathbf{r}^T \cdot \mathbf{x}^\sqcup \leq R_B, \quad (\text{A.2})$$

with $\mathbf{r} \triangleq [R_1, \dots, R_L]^T$. Property $I_\cap \subset I_\sqcup$ may be written as

$$\mathbf{x}^\cap + \mathbf{x}^\sqcup \leq \mathbf{1}_L, \quad (\text{A.3})$$

where $\mathbf{1}_L$ is the $L \times 1$ unity vector and inequalities involving vectors are meant in the percomponent sense.

In order to find the optimal solution, it suffices to find binary-valued vectors \mathbf{x}^\cap and \mathbf{x}^\sqcup minimizing (A.1) subject to the constraints (A.2) and (A.3). This is an integer (0–1) programming problem and can be formulated by defining

$$\begin{aligned} \mathbf{x} &\triangleq \begin{bmatrix} \mathbf{x}^\cap \\ \mathbf{x}^\sqcup \end{bmatrix}, & \bar{\mathbf{d}} &\triangleq \begin{bmatrix} (2 - p)\mathbf{c} \\ \mathbf{d} \end{bmatrix}, \\ \mathbf{C} &\triangleq \begin{bmatrix} \mathbf{I}_L & \mathbf{I}_L \\ \mathbf{r}^T & \mathbf{r}^T \end{bmatrix}, & \mathbf{b} &\triangleq \begin{bmatrix} \mathbf{1}_L \\ R_B \end{bmatrix}, \end{aligned} \quad (\text{A.4})$$

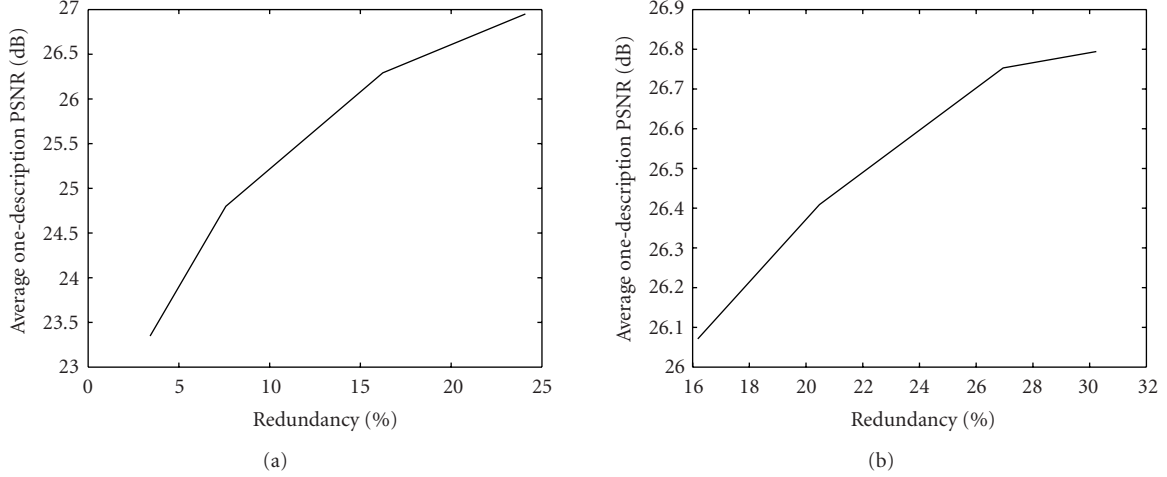


FIGURE 7: Redundancy rate-distortion performance of the proposed schemes: (a) drift, (b) drift-free.

where \mathbf{I}_L is the $L \times L$ identity matrix, \mathbf{x} and $\bar{\mathbf{d}}$ are $2L \times 1$ vectors, \mathbf{C} is a $(L+1) \times L$ matrix, and \mathbf{b} is a $(L+1) \times 1$ vector. In view of these definitions, the maximization problem may be expressed as an integer-programming problem.

Integer (0–1) programming problem

Find (0–1)-valued vector \mathbf{x} such that

$$\begin{aligned} \max \bar{\mathbf{d}}^T \cdot \mathbf{x}, \\ \mathbf{C}\mathbf{x} \leq \mathbf{b}. \end{aligned} \quad (\text{A.5})$$

Although several techniques exist for the solution of integer-programming problems, it is well known that integer-programming problems are, in general, NP-complete and, most of the times, exhaustive search over all possible realizations of binary-valued vector \mathbf{x} is the only procedure that guarantees optimal solution. Even if a cutting-plane or branch-and-bound technique is used, it does not guarantee that the number of operations will be less than exponential in L .

B. THE GENERAL MULTIPLE DESCRIPTION PROBLEM

In the general case, the original frame comprises L layers and we need to form $K \geq 2$ descriptions so that a rate constraint is met and the expected distortion reduction at the decoder is maximized. Conforming to the notation used for the double-description case, we define the index sets I_k , $k = 1, \dots, K$, where each I_k describes the assignment of layers to description k , and the events $A_k = \{\text{Description } k \text{ reaches the decoder}\}$, $k = 1, \dots, K$.

The index-assignment sets I_k , $k = 1, \dots, K$ define 2^K disjoint subsets of the index set $I = \{1, \dots, L\}$, which can be written as

$$\tilde{J}_{\mathbf{x}} = \bigcap_{k=1}^K \tilde{I}_k^{x_k}, \quad \mathbf{x} \in \{0, 1\}^K, \quad (\text{B.1})$$

where the subscript $\mathbf{x} = [x_1, \dots, x_K]^T$ is a $(K \times 1)$ binary-valued vector and

$$\tilde{I}_k^b = \begin{cases} I_k, & b = 1, \\ I_k^c, & b = 0, \end{cases} \quad k = 1, \dots, K. \quad (\text{B.2})$$

For every $\mathbf{x} \in \{0, 1\}^K$, the set $\tilde{J}_{\mathbf{x}}$ comprises the indices belonging to the sets I_j with $x_j = 1$. The original index-assignment sets I_k , $k = 1, \dots, K$ can then be expressed in terms of the collection $\{\tilde{J}_{\mathbf{x}}\}_{\mathbf{x} \in \{0, 1\}^K}$ as

$$I_k = \bigcup_{\mathbf{x} \in \{0, 1\}^K: x_k = 1} \tilde{J}_{\mathbf{x}}, \quad k = 1, \dots, K. \quad (\text{B.3})$$

Let $w(\mathbf{x}) \triangleq \sum_{i=1}^K x_i$ denote the weight of the binary-valued vector \mathbf{x} and for every index set $A \subset I$ define

$$R(A) \triangleq \sum_{i \in A} R_i, \quad D(A) \triangleq \sum_{i \in A} D_i, \quad (\text{B.4})$$

representing the total rate and distortion reduction of the layers with indices in A . The total rate sent to the decoder can be expressed as

$$\begin{aligned} \bar{R}(\{\tilde{J}_{\mathbf{x}}\}_{\mathbf{x} \in \{0, 1\}^K}) &= \sum_{k=1}^K \sum_{i \in I_k} R_i \stackrel{(\alpha)}{=} \sum_{k=1}^K \sum_{\mathbf{x} \in \{0, 1\}^K: x_k = 1} \sum_{i \in \tilde{J}_{\mathbf{x}}} R_i \\ &\stackrel{(\beta)}{=} \sum_{\mathbf{x} \in \{0, 1\}^K} w(\mathbf{x}) R(\tilde{J}_{\mathbf{x}}), \end{aligned} \quad (\text{B.5})$$

where (α) comes from (B.3) and the fact that the sets $\tilde{J}_{\mathbf{x}}$ are mutually disjoint and we can derive (β) by observing that each sum $\sum_{i \in \tilde{J}_{\mathbf{x}}} R_i$ appears exactly $w(\mathbf{x})$ times in the previous expression of the total rate.

For a given $\mathbf{x} \in \{0, 1\}^K$, assume that $x_{j_1} = \dots = x_{j_{w(\mathbf{x})}} = 1$ and the rest are zero. In order to express the expected

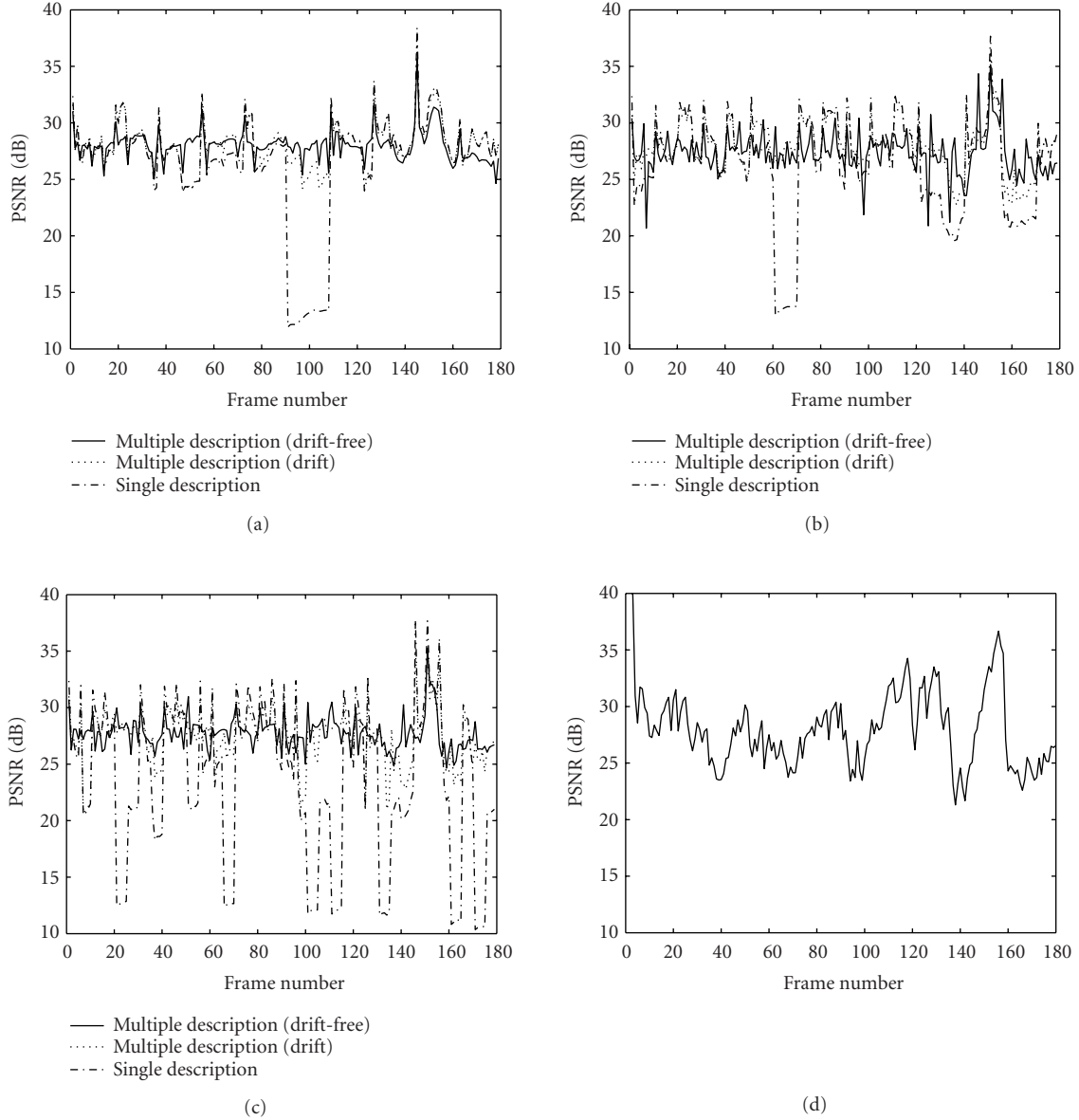


FIGURE 8: Reconstruction quality for the “Foreman” sequence when the channels go on and off during transmission and a probability of error equal to (a) 5%, (b) 10%, (c) 20%, and (d) transmission based on H.264 using flexible macroblock ordering.

distortion reduction at the decoder in terms of the collection $\{\tilde{J}_x\}_{x \in \{0,1\}^K}$, we observe that the distortion at the decoder will improve by $D(\tilde{J}_x)$ (layers with indices in \tilde{J}_x will be used) whenever the event $A_x \triangleq \{\text{description } j_1 \text{ description } j_2 \text{ or } \dots \text{ description } j_{w(x)} \text{ is delivered}\}$ occurs, that is,

$$A_x = \bigcup_{\ell=1}^{w(x)} A_{j_\ell}. \tag{B.6}$$

Assuming that the events $A_k, k = 1, \dots, K$ are independent and $\Pr\{A_k\} = 1 - \Pr\{A_k^c\} = p$, we can calculate its

probability

$$\begin{aligned} \Pr\{A_x\} &= 1 - \Pr\{A_x^c\} \\ &= 1 - \Pr\left\{\bigcap_{\ell=1}^{w(x)} A_{j_\ell}^c\right\} = 1 - (1 - p)^{w(x)}. \end{aligned} \tag{B.7}$$

If we also define $C(A) \triangleq \sum_{i \in A} D_i C_i$ for $A \subset I$, the distortion reduction due to motion compensation based on the layers common to all descriptions will be $C(\tilde{J}_K)$, $\mathbf{1}_K$ being the $(K \times 1)$ unity vector. The distortion reduction due to motion compensation is conditional on the event $A_{\mathbf{1}_K}$ (at least one

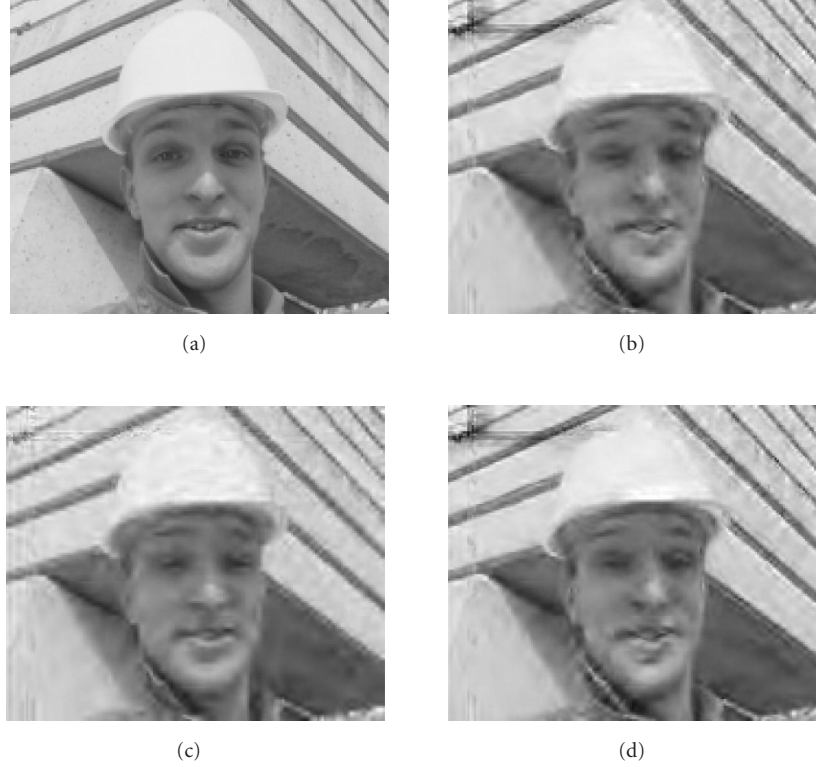


FIGURE 9: Reconstructed frame for the transmission of the “Foreman” sequence, $p = 0.9$, over two channels of total capacity 128 Kbps: (a) original “Foreman” frame, (b) reconstructed using the coder without drift control (25.84 dB), (c) reconstructed using the drift-free coder (28.81 dB), and (d) reconstructed using the single description coder (25.78 dB).

of the descriptions reaches the decoder) whose probability is $1 - (1 - p)^K$. Therefore, the overall expected distortion reduction at the decoder will be

$$\begin{aligned} \overline{D}(\{\tilde{J}_x\}_{x \in \{0,1\}^K}) &= \Pr\{A_{1_K}\}C(\tilde{J}_{1_K}) + \sum_{x \in \{0,1\}^K} \Pr\{A_x\}D(\tilde{J}_x) \\ &= [1 - (1 - p)^K]C(\tilde{J}_{1_K}) \\ &\quad + \sum_{x \in \{0,1\}^K} [1 - (1 - p)^{w(x)}]D(\tilde{J}_x). \end{aligned} \quad (\text{B.8})$$

At this juncture, observe that both the total rate (B.5) and the expected distortion reduction (B.8) can be expressed as linear functions of the $\{R(\tilde{J}_x)\}_{x \in \{0,1\}^K}$ and $\{D(\tilde{J}_x)\}_{x \in \{0,1\}^K}$, respectively, with coefficients depending only on the weight of the index vector \mathbf{x} . Therefore, we can group all sets \tilde{J}_x with the same weight and define the new (fewer) sets

$$J_k = \bigcup_{x \in \{0,1\}^K: w(x)=k} \tilde{J}_x, \quad k = 0, \dots, K, \quad (\text{B.9})$$

each set J_k containing the layer indices assigned to exactly k descriptions. Also, observe that the set $\tilde{J}_0 = J_0$ has a zero coefficient in both (B.5) and (B.8); hence, it does not contribute to the total rate or expected distortion reduction.

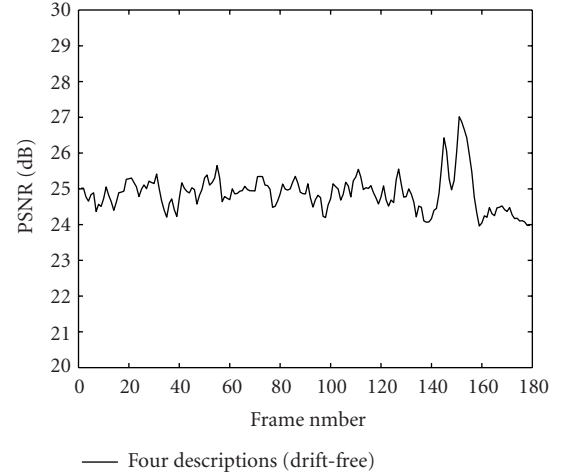


FIGURE 10: Reconstruction quality obtained using the drift-free system with four descriptions transmitted over channels with probability of loss equal to 20%.

By reformulating (B.5) and (B.8), the maximization problem for the general multiple description case may be stated as follows.

General maximization problem

Find disjoint sets $J_1, \dots, J_K \subset I$ maximizing

$$\begin{aligned} \bar{D}(J_1, \dots, J_K) &= [1 - (1 - p)^K]C(J_K) \\ &+ \sum_{k=1}^K [1 - (1 - p)^k]D(J_k) \end{aligned} \quad (\text{B.10})$$

subject to the constraint

$$\bar{R}(J_1, \dots, J_K) = \sum_{k=1}^K kR(J_k) \leq R_B. \quad (\text{B.11})$$

The integer-programming formulation of the general maximization problem would involve K binary-valued $L \times 1$ vectors \mathbf{x}_k , $k = 1, \dots, K$, with

$$\mathbf{x}_{k,i} = \begin{cases} 1, & \text{if } i \in J_k, \\ 0, & \text{if } i \notin J_k, \end{cases} \quad k = 1, \dots, K, \quad i = 1, \dots, L, \quad (\text{B.12})$$

and the requirement that the J_k , $k = 1, \dots, K$ be disjoint can be written as

$$\mathbf{x}_1 + \dots + \mathbf{x}_K \leq \mathbf{1}_L. \quad (\text{B.13})$$

Let us define

$$\begin{aligned} \mathbf{x} &\triangleq [\mathbf{x}_1^T \ \dots \ \mathbf{x}_k^T \ \dots \ \mathbf{x}_K^T]^T, \\ \bar{\mathbf{d}}_K &\triangleq [p\mathbf{d}^T \ \dots \ (1 - (1 - p)^k)\mathbf{d}^T \ \dots \ (1 - (1 - p)^K)\mathbf{c}^T]^T, \\ \mathbf{C}_K &\triangleq \begin{bmatrix} \mathbf{I}_L & \dots & \mathbf{I}_L & \dots & \mathbf{I}_L \\ \mathbf{r}^T & \dots & k\mathbf{r}^T & \dots & K\mathbf{r}^T \end{bmatrix}, \\ \mathbf{b}_K &\triangleq [\mathbf{1}_L^T \ R_B]^T, \end{aligned} \quad (\text{B.14})$$

where \mathbf{x} and $\bar{\mathbf{d}}_K$ are $KL \times 1$ vectors, \mathbf{C}_K is a $(L + 1) \times KL$ matrix, \mathbf{b}_K is a $(L + 1) \times 1$ vector and the $L \times 1$ vectors \mathbf{r} , \mathbf{d} , \mathbf{c} are those defined in the double-description integer-programming formulation. Then, the integer-programming formulation of the general multiple description problem will be as follows.

General integer (0–1) programming problem

Find (0–1)-valued vector \mathbf{x} such that

$$\begin{aligned} \max \quad & \bar{\mathbf{d}}_K^T \cdot \mathbf{x}, \\ \text{C}_K \mathbf{x} & \leq \mathbf{b}_K. \end{aligned} \quad (\text{B.15})$$

As is clear from the integer-programming formulation, the complexity of the general maximization problem may be as high as 2^{KL} . Heuristics similar to those proposed for the double-description case may be used for an estimate of the optimal index-assignment scheme, based on the general equivalent continuous problem, which can be easily formulated from (B.10) and (B.11). It is reasonable to conjecture

that the heuristics stemming from the equivalent continuous general maximization problem will provide solutions deviating from the optimal one even more as K increases.

ACKNOWLEDGMENTS

The authors would like to thank Savvas Argyropoulos for his help with the H.264 results and the anonymous reviewers for their constructive comments. This work was partly presented in the IEEE International conference on image processing, 2004. This work was supported by the EU IST projects “3DTV,” “BOEMIE,” and “K-SPACE.”

REFERENCES

- [1] A. E. El Gamal and T. M. Cover, “Achievable rates for multiple descriptions,” *IEEE Transactions on Information Theory*, vol. 28, no. 6, pp. 851–857, 1982.
- [2] V. A. Vaishampayan, “Design of multiple description scalar quantizers,” *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 821–834, 1993.
- [3] A. C. Miguel, A. E. Mohr, and E. A. Riskin, “SPIHT for generalized multiple description coding,” in *Proceedings of IEEE International Conference on Image Processing (ICIP ’99)*, vol. 3, pp. 842–846, Kobe, Japan, October 1999.
- [4] W. Jiang and A. Ortega, “Multiple description coding via polyphase transform and selective quantization,” in *Visual Communications and Image Processing (VCIP ’99)*, vol. 3653, part 1-2 of *Proceedings of SPIE*, pp. 998–1008, San Jose, Calif, USA, January 1999.
- [5] A. E. Mohr, E. A. Riskin, and R. E. Ladner, “Generalized multiple description coding through unequal loss protection,” in *Proceedings of IEEE International Conference on Image Processing (ICIP ’99)*, vol. 1, pp. 411–415, Kobe, Japan, October 1999.
- [6] A. R. Reibman, H. Jafarkhani, Y. Wang, M. T. Orchard, and R. Puri, “Multiple description coding for video using motion compensated prediction,” in *Proceedings of IEEE International Conference on Image Processing (ICIP ’99)*, vol. 3, pp. 837–841, Kobe, Japan, October 1999.
- [7] V. A. Vaishampayan and S. John, “Interframe balanced multiple description video compression,” in *Proceedings of 9th Packet Video Workshop (PVW ’99)*, pp. 812–816, New York, NY, USA, April 1999.
- [8] A. Sehgal, A. Jagmohan, and N. Ahuja, “Wireless video conferencing using multiple description coding,” in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS ’01)*, vol. 5, pp. 303–306, Sydney, NSW, Australia, May 2001.
- [9] N. V. Boulgouris, K. E. Zachariadis, A. N. Leontaris, and M. G. Strintzis, “Drift-free multiple description coding of video,” in *Proceedings of 4th IEEE Workshop on Multimedia Signal Processing (MMSP ’01)*, pp. 105–110, Cannes, France, October 2001.
- [10] Y.-C. Lee, Y. Altunbasak, and R. M. Mersereau, “An enhanced two-stage multiple description video coder with drift reduction,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 122–127, 2004.
- [11] N. Franchi, M. Fumagalli, R. Lancini, and S. Tubaro, “Multiple description video coding for scalable and robust transmission over IP,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 3, pp. 321–334, 2005.

- [12] X. Yang and K. Ramchandran, "Optimal multiple description subband coding," in *Proceedings of IEEE International Conference on Image Processing (ICIP '98)*, vol. 1, pp. 654–658, Chicago, Ill, USA, October 1998.
- [13] Y. Wang, A. R. Reibman, and S. Lin, "Multiple description coding for video delivery," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 57–70, 2005.
- [14] A. R. Reibman, H. Jafarkhani, Y. Wang, and M. T. Orchard, "Multiple description video using rate-distortion splitting," in *Proceedings of IEEE International Conference on Image Processing (ICIP '01)*, vol. 1, pp. 978–981, Thessaloniki, Greece, October 2001.
- [15] N. V. Boulgouris, K. E. Zachariadis, A. Kanlis, and M. G. Strintzis, "Multiple description wavelet coding of layered video," in *Proceedings of IEEE International Conference on Image Processing (ICIP '04)*, vol. 4, pp. 2263–2266, Singapore, October 2004.
- [16] M. Pereira, M. Antonini, and M. Barlaud, "Multiple description coding for internet video streaming," in *Proceedings of IEEE International Conference on Image Processing (ICIP '03)*, vol. 3, pp. 281–284, Barcelona, Spain, September 2003.
- [17] T. Petrisor, C. Tillier, B. Pesquet-Popescu, and J.-C. Pesquet, "Redundant multiresolution analysis for multiple description video coding," in *Proceedings of 6th IEEE Workshop on Multimedia Signal Processing (MMSP '04)*, pp. 95–98, Siena, Italy, September–October 2004.
- [18] H. Watanabe and S. Singhal, "Windowed motion compensation," in *Visual Communications and Image Processing '91: Visual Communication*, vol. 1605, part 1 of *Proceedings of SPIE*, pp. 582–589, Boston, Mass, USA, November 1991.
- [19] N. V. Boulgouris, D. Tzovaras, and M. G. Strintzis, "Lossless image compression based on optimal prediction, adaptive lifting, and conditional arithmetic coding," *IEEE Transactions on Image Processing*, vol. 10, no. 1, pp. 1–14, 2001.
- [20] J. F. Arnold, M. R. Fracter, and Y. Wang, "Efficient drift-free signal-to-noise ratio scalability," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 1, pp. 70–82, 2000.
- [21] D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1158–1170, 2000.
- [22] D. Wang, N. Canagarajah, and D. Bull, "Slice group based multiple description video coding using motion vector estimation," in *Proceedings of IEEE International Conference on Image Processing (ICIP '04)*, vol. 5, pp. 3237–3240, Singapore, October 2004.
- [23] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Transactions on Image Processing*, vol. 1, no. 2, pp. 205–220, 1992.
- [24] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3445–3462, 1993.
- [25] I. H. Witten, R. M. Neal, and J. G. Cleary, "Arithmetic coding for data compression," *Communications of the ACM*, vol. 30, no. 6, pp. 520–540, 1987.
- [26] J. Li and S. Lei, "An embedded still image coder with rate-distortion optimization," *IEEE Transactions on Image Processing*, vol. 8, no. 7, pp. 913–924, 1999.
- [27] P.-Y. Cheng, J. Li, and C.-C. Jay Kuo, "Rate control for an embedded wavelet video coder," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 4, pp. 696–702, 1997.
- [28] A. R. Reibman, H. Jafarkhani, Y. Wang, M. T. Orchard, and R. Puri, "Multiple-description video coding using motion-compensated temporal prediction," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 3, pp. 193–204, 2002.
- [29] P. G. Sherwood and K. Zeger, "Error protection for progressive image transmission over memoryless and fading channels," *IEEE Transactions on Communications*, vol. 46, no. 12, pp. 1555–1559, 1998.
- [30] D. G. Sachs, R. Anand, and K. Ramchandran, "Wireless image transmission using multiple-description based concatenated codes," in *Proceedings of Data Compression Conference (DCC '00)*, p. 569, Snowbird, Utah, USA, March 2000.

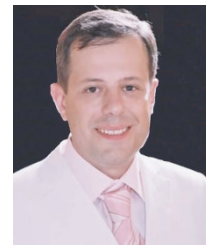
Nikolaos V. Boulgouris received the Diploma and the Ph.D. degrees from the Electrical and Computer Engineering Department of the University of Thessaloniki, Greece, in 1997 and 2002, respectively. Since December 2004, he has been a Lecturer with the Department of Electronic Engineering, Division of Engineering, at King's College London, United Kingdom. From September 2003 to November 2004 he was a Postdoctoral Fellow with the Department of Electrical and Computer Engineering, University of Toronto, Canada. Previously, he was affiliated with the Informatics and Telematics Institute in Greece. He has participated in several research projects in the areas of image/video communication, pattern recognition, multimedia security, and content-based indexing and retrieval. He is a Member of the IEEE and the British Machine Vision Association.



Konstantinos E. Zachariadis is currently a Ph.D. candidate of Managerial Economics and Strategy in Northwestern University's Kellogg School of Management. He received his M.S. in electrical and computer engineering, from Northwestern University, for work on source fidelity over fading channels using scalable codes and erasure codes; and his Diploma in electrical and computer engineering from the Aristotle University of Thessaloniki in Greece, for work on multiple description coding of images and video. For his graduate studies, he has received the IEEE Life Member Graduate Study Fellowship and a Fulbright Fellowship. His current research interests are on dynamic contract theory, stochastic control, services operations, and resource allocation in wireless communications.



Angelos Kanlis obtained the Diploma in electrical engineering from the Aristotle University of Thessaloniki, Thessaloniki, Greece in 1992, and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1994 and 1997, respectively. Since 2003, he has been a Patent Examiner at the European Patent Office, Munich, Germany. He was a Visiting Professor in the Department of Computer & Communication Engineering, Volos, Greece (2002–2003) and a Research Associate in the Informatics and Telematics Institute, Thessaloniki, Greece (2000–2002) and the Foundation of Research and Technology, Heraklion Crete, Greece (1997–1999).



Michael G. Strintzis received the Diploma in electrical engineering from the National Technical University of Athens, Athens, Greece, in 1967 and the M.A. and Ph.D. degrees in electrical engineering from Princeton University, Princeton, NJ, in 1969 and 1970, respectively. He joined the Electrical Engineering Department, University of Pittsburgh, Pittsburgh, PA, where he served as an Assistant Professor from 1970 to 1976



and an Associate Professor from 1976 to 1980. During that time, he worked in the area of stability of multidimensional systems. Since 1980, he has been a Professor of electrical and computer engineering at the Aristotle University of Thessaloniki, Thessaloniki, Greece. He has worked in the areas of multidimensional imaging and video coding. Over the past ten years, he has authored over 100 journal publications and over 200 conference presentations. In 1998, he founded the Informatics and Telematics Institute, currently part of the Centre for Research and Technology Hellas, Thessaloniki. He was awarded the Centennial Medal of the IEEE in 1984 and the Empirikeion Award for Research Excellence in Engineering in 1999.