Balanced Multiple Description Video Coding Using Optimal Partitioning of the DCT Coefficients

Kiran R. Matty and Lisimachos P. Kondi, *Member, IEEEf*

***Abstract*—In this paper, we propose a balanced multiple description video coding scheme which is based on the partitioning of the discrete cosine transform (DCT) coefficients. Our scheme splits the single-layered stream generated by a standard coder into two correlated and balanced (virtually identical rates and distortions) substreams. The optimization is in the redundancy-rate-distortion sense using Lagrangian relaxation for optimum allocation of redundancy among the blocks in the frame. A greedy algorithm to meet the equal distortions criterion and an algorithm to optimally achieve equal bit rates for the two descriptions are proposed. Our simulation results substantiate our claim of achieving balanced descriptions at high peak signal-to-noise ratios values for any specified redundancy and any target bit rate. Our design complies with the existing motion-compensated DCT-based standards, i.e., each description can be independently decoded by any standard decoder.**

***Index Terms*—Greedy algorithm, H.263, multiple description coding, partitioning algorithm, rate distortion-theory, video coding.**

I. Introduction

IN RECENT times, transmission of multimedia over "best-effort" networks such as the Internet has become ubiquitous. But, due to heterogeneity of the networks, congestion and a plethora of other reasons, packet losses and delays have become pervasive. For example, a typical heterogeneous network scenario might demand that the packets move from the higher capacity fiber link to the lower capacity wireless link. This kind of hopping might ensue dropping of packets. If the network supports preferential treatment of some packets, then the use of multiresolution or layered source coding is the obvious choice, but this might not hold true if the networks are oblivious to the contents of the packets and fail to discriminate. Also, packet retransmission (supported by transport protocols like TCP) can be unwarranted due to real-time constraints and absence of feedback channels. Another common problem is network congestion, for example, when packets have to be dropped upon arrival at the switches owing to the fullness of the local buffers. This fosters the need for better source coding algorithms to make the received data meaningful.

Multiple description coding (MDC) [1] is an error resilient source coding technique which can be extended to any source (audio, video and data). Some of the early theoretical work on MDC appears in [2]–[4]. A number of schemes have been proposed leading to the development of balanced or unbalanced systems [5]–[9]. In MDC, a single source of data is represented with several subsets of data, popularly called descriptions, such that the source can be approximated from any of these subsets. If these subsets carry identical data over each channel and both channels are operational, then half of the information received is redundant. On the other hand, if very little correlation exists between the data carried by the subsets, loss of any of the subsets would severely impair the decoding at the receiver. The primary objective of a MDC coder is to achieve a minimum fidelity reconstruction in the event that only a single channel is operational and, should more channels be working, higher fidelity reconstruction should be achievable depending on the number of channels received correctly. Sending information over multiple channels with independent failure events increases the odds of receiving some information from at least one of the channels. Some of the desirable qualities of the multiple descriptions are that they should be mutually refining and that each description should be independently decodable thereby yielding video of acceptable quality.

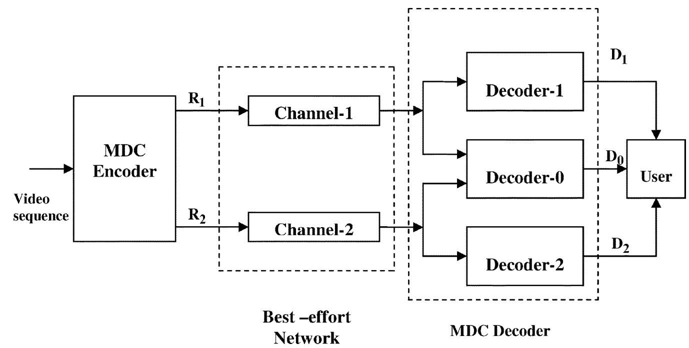

Fig. 1. Generic MDC coder system.

Fig. 1 depicts a generic MDC encoder/decoder system wherein R^* refers to bit rate that would be required to encode the video sequence using a single description coding (SDC) scheme (for example, baseline H.263) and D_0 represents the two-channel or SDC distortion, i.e., when both descriptions arrive at the receiver. (R_1, D_1) and (R_2, D_2) are the rate-distortion pairs corresponding to the two descriptions. Reibman *et al.* in [10] proposed a MDC video coder based on rate-distortion splitting wherein the output of the standard video coder is split into two correlated streams. Header information, motion vectors and discrete cosine transform (DCT) coefficients with magnitudes above a certain threshold are duplicated into the two descriptions whereas the remaining coefficients are alternated between the two streams. Thresholds for the blocks are computed optimally in the rate-distortion sense. At the decoder, if both descriptions are received, redundant information is discarded, else the received description is decoded.

Manuscript received February 11, 2004; revised August 26, 2004. This paper was recommended by Associate Editor R. Lancini.

The authors are with the Department of Electrical Engineering, University at Buffalo, The State University of New York, Buffalo, NY 14260 USA (e-mail: krmatty@eng.buffalo.edu; lkondi@eng.buffalo.edu).

Digital Object Identifier 10.1109/TCSVT.2005.848343

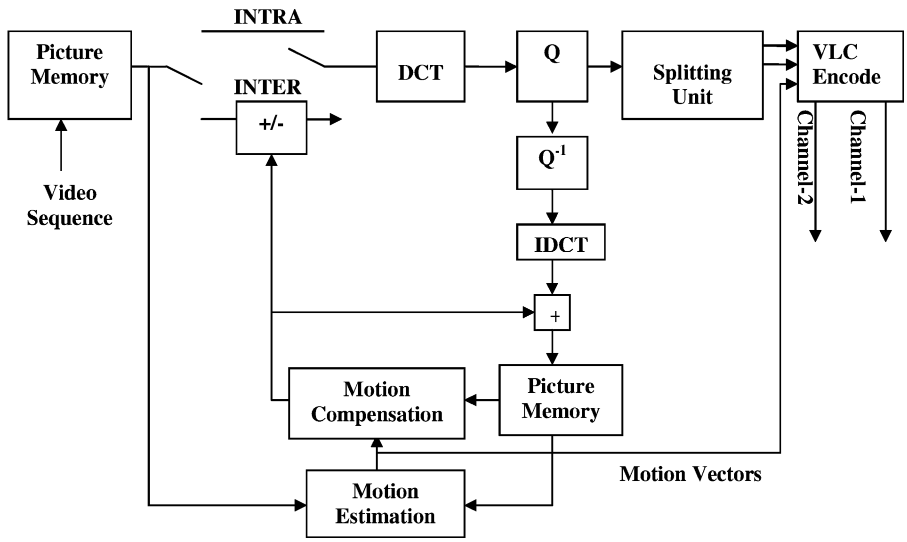

Fig. 2. MD-BRDS encoder. Q: Quantization, DCT: Discrete cosine transform; IDCT: inverse discrete cosine transform; VLC: Variable length coding.

However, the alternation of nonzero DCT coefficients spawns unbalanced descriptions. The balanced approach, i.e., encoding the descriptions at equal rates and distortions, is desirable from the standpoint of transmissions over the Internet, wherein the packets aren't awarded preferential treatment. Subsequent sections would offer a thorough treatment of our proposed MDC scheme which we term as multiple description using balanced rate distortion splitting (MD-BRDS). Our foregoing research in this direction appears in [11].

The rest of the paper is organized as follows. In Section II, the structure of the proposed multiple description codec is presented. In Section III, the algorithm for the partitioning of the DCT coefficients is discussed. In Section IV, experimental results are presented. Last, in Section V, conclusions are drawn.

II. Proposed MD-BRDS System

Our primary objective is to encode a video sequence into two video substreams with equal bit rates that are capable of achieving identical peak signal-to-noise ratios (PSNRs) at the receiver, after being transmitted over two independent channels. Our proposed MD-BRDS uses the redundancy-rate-distortion (RRD) [12] criterion at the encoder end to split the output stream of a standard video coder like H.263 into two correlated substreams as shown in Fig. 2.

The encoding algorithm, which sits in the splitting unit that is embedded in the existing H.263 encoder, uses Lagrangian relaxation [13], [14] to obtain the optimal threshold values for each block. As shown in Fig. 2, the splitting unit operates on the quantized DCT coefficients. Two descriptions are created by duplicating: 1) header information; 2) motion vectors; and 3) quantized dc coefficients of intrablocks, after which nonzero ac coefficients whose quantization levels equal or exceed the threshold value are duplicated into the two descriptions and then the remaining coefficients are split in accordance with our partitioning algorithm to meet the balanced condition. The splitting is done in such a way that when a quantized coefficient value is sent to one of the descriptions, a corresponding zero is sent to the other description. In this work, the VLC tables and the syntax of the standard H.263 codec [15] are used.

The operation of the MD-BRDS decoder is influenced by two conditions: 1) if both descriptions arrive at the decoder, then the coefficients from the two streams are simply merged into a single stream that can be decoded by a standard decoder 2) If one of the descriptions arrives at the receiver, then the received stream is simply decoded by a standard decoder like the H.263 decoder. It is noteworthy to observe that the MD-BRDS decoder is oblivious to threshold gymnastics. As a result, no overheads of transmitting threshold values are incurred thereby making our design simple. Also, our design complies to the H.263 syntax since each description can be decoded as is by a standard H.263 decoder. Although baseline H.263 is used as the basis of our coder, any motion-compensated DCT-based coder can be used.

The concept of redundancy serves as the cornerstone of the MDC paradigm. The redundancy ρ is defined as

$$\rho = \frac{\sum_{i=1}^{n} (R_{1i} + R_{2i} - R_i^*)}{\sum_{i=1}^{n} R_i^*} \tag{1}$$

where ρ is defined over a group of blocks (GOB) which comprises of 66 blocks for QCIF-size video, i.e., $n = 66$. R_i^* is the number of bits that would be required to encode the ith block using an SDC scheme which in our case is realized using the baseline H.263. R_{1i} and R_{2i} are the total bits expended in encoding the description-1 and description-2 apiece, corresponding to the ith block, such that $R_{1i} \leq R_i^*$ and $R_{2i} \leq R_i^*$. In the redundancy calculations, headers, motion vectors and other overheads are taken into reckoning.

III. Coefficient Partitioning Algorithm

In this section, we address the following two problems underlying our MD-BRDS scheme.

1) Determination of the threshold values for deciding which coefficients should be duplicated in both descriptions.

2) Splitting of the remaining DCT coefficients between the two descriptions.

The rates and distortions of each description that result from splitting the DCT coefficients using all possible thresholds need to be known in order to determine the optimal thresholds. Thus, the steps involved to effect the splitting of the DCT coefficients are also used toward the determination of the threshold values, as explained in Section III-A.

A. Determination of the Threshold

The redundancy induced and the distortions associated with the descriptions are functions of the threshold. The problem underlying the design of the MDC coder is the selection of the threshold values for optimal allocation of redundancy among the blocks to meet the redundancy budget. Our optimization is at the GOB level. In H.263 with QCIF-sized frames, each GOB comprises of one line of 16 $\times$ 16 macroblocks (11 macroblocks). Each macroblock consists of six blocks (four luminance and two chrominance 8 $\times$ 8 blocks). Thus, each GOB comprises of 66 blocks. DCT coefficients are quantized in accordance with the H.263 standard specifications.

For a given threshold x_i, $R_{1i}(x_i)$, $R_{2i}(x_i)$ and $D_i(x_i)$ are found by partitioning the DCT coefficients as in Section III-B and computing the resulting rates and distortions. The partitioning of the DCT coefficients given a threshold is done in two steps. In the first step, an initial partitioning is determined that minimizes the difference in the energies of the two descriptions, while in the second step, the difference in the rates of the two descriptions is minimized while maintaining the energies determined in the first step.

The threshold values for the optimal allocation of redundancy among the blocks within a GOB are selected by solving the constrained optimization problem formally stated as

$$\{x_1^*, x_2^*, \ldots, x_n^*\} = \arg\min \sum_{i=1}^{n} D_i(x_i) \tag{2}$$

subject to

$$\sum_{i=1}^{n} \xi_i\left(x_i^*\right) \leq \xi_{\text{budget}} = \left\lceil \rho * \sum_{i=1}^{n} R_i^* \right\rceil; \quad 0 \leq \rho \leq 1 \tag{3}$$

where $D_i(x_i)$ and $\xi_i(x_i)$ are the distortion and redundancy functions, respectively, for an arbitrary block indexed as i. $D_i(.)$ is defined as the mean square error (MSE) between the dequantized transform coefficients of description-1 and the original unquantized transform coefficients of block i. We could have very well used the description-2, since the distortions associated with the two descriptions would be virtually identical. x_i is a threshold value selected from among all possible threshold choices for the ith block. x_i^* is the optimal value of the threshold corresponding to the ith block. ξ_{budget} is the redundancy budget constraint on a GOB. $\xi_i(x_i)$ for the ith block is defined as: $\xi_i(x_i) = R_{1i}(x_i) + R_{2i}(x_i) - R_i^*$.

To solve the problem in (2) and (3), we consider a RRD framework with MSE as our distortion criterion. Redundancy and distortion measures over the blocks within a GOB are additive. As a result, the constrained problem defined by (2) and (3) can be converted to a simpler equivalent unconstrained problem as in (4) by combining redundancy and distortion through the Lagrange multiplier λ. We solve this unconstrained problem for different positive values of the λ which result in tracing out of the convex hull points of the RRD curve. We pursue the point on the convex hull which yields minimum distortion while meeting the redundancy budget as described in [14]

$$J_{\text{GOB}}(\lambda) = \sum_{i=1}^{n} D_i(x_i) + \lambda \sum_{i=1}^{n} \xi_i(x_i) \tag{4}$$

$$= \sum_{i=1}^{n} \left\{D_i(x_i) + \lambda \xi_i(x_i)\right\}. \tag{5}$$

In (4), we have relaxed the budget constraints. The unconstrained problem now becomes the minimization of the Lagrangian cost function defined by (4). As seen in (5), it is possible to express $J_{\text{GOB}}(\lambda)$ as a sum of individual Lagrangian subcosts per block and then perform the minimization independently for each block by holding λ constant for each block. Then, if the redundancy budget is met for a specific λ, it would ensure that the optimization of the individual Lagrangian costs would yield an optimal redundancy allocation across the whole GOB. The appropriate value of the Lagrangian multiplier λ^* is not known *a priori* but can be determined using the bisection algorithm [14], [16]. At λ^*

$$\sum_{i=1}^{n} \xi_i\left(x_i^*\right) \simeq \xi_{\text{budget}}$$

where x_i^* is the optimal operating point on the convex hull of the RRD function of the i^{th} block for $i = 1, \ldots, n$.

B. Splitting the DCT Coefficients

Toward this end, we formulate as follows. Given a threshold value corresponding to a block, we wish to partition the quantized coefficients of that block into two subsets subject to the constraint that their bit rates and distortions are as close as possible. We describe a simple partitioning algorithm which operates in two phases. Let $\hat{\mathbf{C}} = \{\hat{c}_0, \hat{c}_1, \ldots, \hat{c}_{63}\}$ be a set of quantized coefficients (quantization levels) of a block to be split. The quantized dc coefficient of the intrablocks and the quantized ac coefficients whose quantization levels equal or exceed the threshold value are duplicated and the remainder of the quantized ac coefficients are split in accordance with our partitioning algorithm. $\hat{c}_i$ is the ith quantized coefficient of a block, while $\hat{c}_0$ is the quantized dc coefficient of the block. Let $\hat{\mathbf{C}}_1$ and $\hat{\mathbf{C}}_2$ be the two subsets generated after splitting the contents of $\hat{\mathbf{C}}$ and $\mathbf{S}(.)$ be a function which defines our splitting strategy such that

$$[\hat{\mathbf{C}}_1, \hat{\mathbf{C}}_2] = \mathbf{S}(\hat{\mathbf{C}}) \tag{6}$$

$$\hat{\mathbf{C}} = \mathbf{S}^{-1}(\hat{\mathbf{C}}_1, \hat{\mathbf{C}}_2). \tag{7}$$

Recall that the splitting is done in such a way that when a quantized coefficient value is sent to one of the descriptions, a corresponding zero is sent to the other description such that the cardinalities $|\hat{\mathbf{C}}|$, $|\hat{\mathbf{C}}_1|$, and $|\hat{\mathbf{C}}_2|$ are the same, including the zero coefficients. At the decoder, the inverse operation is merely a simple merger of the two streams into a single stream as shown in (7).

The proposed coefficient splitting algorithm operates in two phases. In the first phase, an initial partitioning of the coefficients is found in order to minimize the difference in energies between the two partitions. The algorithm in the first phase is greedy in the sense that it strives for low complexity and does not guarantee that the energy difference is the minimum possible, although in practice it performs very well. In the second phase, the difference in the bit rates between the two partitions is minimized while the energies of the partitions are kept the same as those obtained in the first phase. The algorithm in the second phase is optimal in the sense that it guarantees the minimum difference in the bit rates of the two partitions given the energy partitioning obtained in the first phase. We next explain the splitting algorithm in detail.

Using Parseval's theorem, it is possible to calculate the energy of a signal in the transform domain, as long as the transform is unitary. In the first phase, we search for the initial partitions such that the difference in energies of the partitions in the transform domain is minimized to achieve virtually identical PSNRs at the receiver

$$\Delta_1 = \left| \sum_{i=1}^{|\mathbf{C}|} a_i^2 - \sum_{i=1}^{|\mathbf{C}|} b_i^2 \right| \quad (8)$$

where, $a_i \in \mathbf{C}_1, b_i \in \mathbf{C}_2, \forall i \in \{1, \ldots, |\mathbf{C}|\}$ and $\mathbf{C}$, $\mathbf{C_1}$ and $\mathbf{C_2}$ are the "dequantized" versions of $\hat{\mathbf{C}}$, $\hat{\mathbf{C}}_1$ and $\hat{\mathbf{C}}_2$, respectively.

The total energy content of a block is computed, then the coefficients which remain after duplication are assigned to one of the subsets until half the total energy content of the block is reached, after which the remainder of the coefficients are pushed to the other subset thereby equalizing the energy contents of the two subsets. If required, the set $\hat{\mathbf{C}}$ is permuted to generate a new combination of subsets. We iterate until the energy contents of the two subsets are almost equal. Thus, initial partitions of coefficients are obtained. The proposed algorithm is greedy in the sense that it proceeds in the direction of achieving the smallest difference in energies at each iteration and iterates until the least difference in energies is achieved.

In the second phase, the difference in the bit rates between the two partitions Δ_2 is minimized while the energies of each partition are kept the same as those obtained in the first phase. Δ_2 is defined as

$$\Delta_2 = |R_{1i} - R_{2i}| \quad (9)$$

where R_{1i} and R_{2i} are the total bit rates associated with the subsets $\hat{\mathbf{C}}_1$ and $\hat{\mathbf{C}}_2$, respectively, corresponding to the ith block. Examination of the standard VLC tables for DCT coefficients reveals that replacing a nonzero coefficient by a zero disturbs the run-length which subsequently changes the bit rate. We exploit this observation to achieve virtually identical bit rates for the two descriptions. We start our search by selecting the first nonzero coefficient from either of the initial partitions $\hat{\mathbf{C}}_1$ or $\hat{\mathbf{C}}_2$ and parse the other subset to check for its presence by comparing the quantization level of the coefficients. The number of matches is counted. Depending on the number of matches, coefficients are swapped between the two subsets to generate new combinations of subsets corresponding to each match. If no matches were to be found for any selected nonzero coefficient, then the next nonzero coefficient in the subset is considered. All the subset combinations assume the same energies as in the initial partitions obtained in the first phase. Of all the combinations, the combination of subsets with the least difference in bit rates is retained. The search is stopped, when the difference in bit rates equates to zero, otherwise the combination which achieves lower difference is retained and the algorithm proceeds with the next nonzero coefficient of the selected initial partition until the least difference in bit rates is achieved. The algorithm in the second phase is optimal given the initial partitioning in the first phase, in the sense that it optimally minimizes Δ_2 over all possible subsets that have the same energy per subset as in the initial partitions obtained in the first phase.

It should be noted that the header information, motion vectors and other overheads for a block are not to be taken into reckoning in the rate calculations in the partitioning algorithm as they would be duplicated anyway. Nevertheless, they need to be accounted for in the threshold computations. Variable length coding (VLC) tables of H.263 standard are used for the rate calculations.

TABLE I
"FOREMAN," TARGET BIT RATE FOR THE H.263 ENCODER (SDC) $= 128$ kb/s, actual SDC rate $= 128.36$ kb/s, SDC distortion $= 35.28$ dB

ρ	MD-RDS		MD-BRDS	
	$R_1(kbps)$	$R_2(kbps)$	$R_1(kbps)$	$R_2(kbps)$
0.45	96.06	95.53	94.58	94.47
0.60	103.99	101.74	102.82	102.69
0.70	110.47	107.75	109.14	109.07
0.80	116.76	114.28	115.54	115.47
0.90	122.84	121.01	121.93	121.93
1.00	128.36	128.36	128.36	128.36

IV. EXPERIMENTAL RESULTS

A number of experiments were conducted, some of which are presented here. We use the "Foreman" (10 s) and the "Akiyo" (10 s) sequences to compare the performance of our MD-BRDS algorithm with the MDC using rate-distortion splitting (MD-RDS) scheme proposed in [10]. The frame rate was 10 frames/s and one intraframe was forced for every ten encoded frames. Video quality is measured as the average PSNR (decibels) over the encoded frames. The TMN8 rate control algorithm [17] is used to maintain a target bit rate R^* for encoding the video sequence using the SDC scheme. Then, $R_1 \simeq R_2 \simeq ((\rho + 1) \cdot R^*/2)$, where ρ is the desired redundancy. Thus, our encoder receives two rate parameters, namely R^* and ρ. The descriptions generated by the MD-BRDS encoder are H.263 syntax compliant and can be decoded by the H.263 decoder. We have not used any of the H.263 annexes for the sake of simplicity, but they can very well be supported. Table I presents the bit rates achieved for the SDC, MD-RDS and the MD-BRDS encoders, respectively, for various values

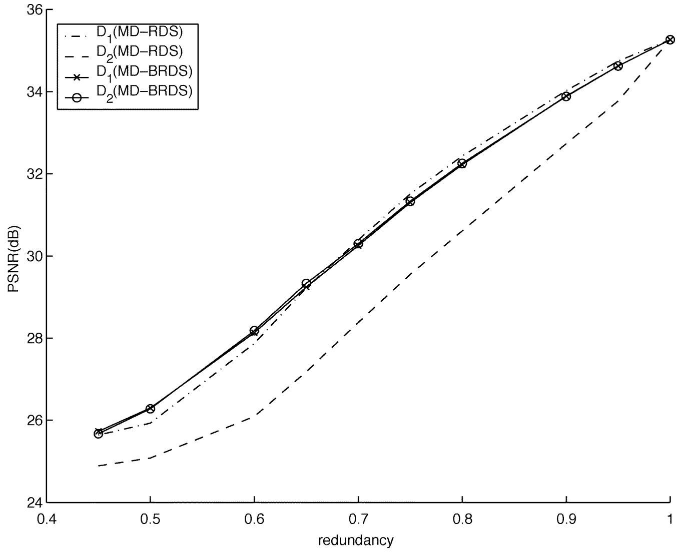

Fig. 3. Comparison of description distortions for different ρ values ("Foreman"), $R^* = 128.36$ kb/s, SDC distortion $= 35.28$ dB.

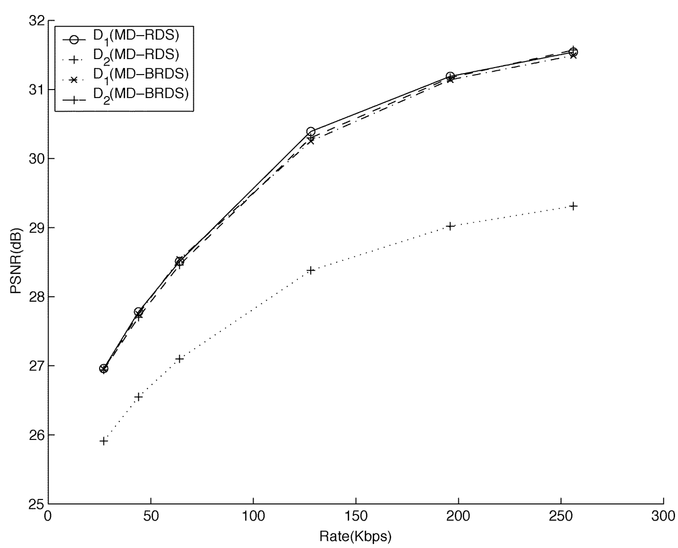

Fig. 4. Comparison of description distortions at different R^* using the "Foreman" sequence, $\rho = 0.70$.

of ρ using "Foreman." It can be seen that the rates of the two descriptions generated by the proposed MD-BRDS codec are virtually identical while there is some discrepancy in the rates generated by MD-RDS. Fig. 3 shows the corresponding description distortions for MD-BRDS and MD-RDS. It can be seen that the distortions of the two descriptions generated by MD-RDS are quite different while the descriptions of the proposed MD-BRDS algorithm have virtually identical distortions. Furthermore, these distortions are close to the distortion of the "best" description of MD-RDS. Therefore, when used in a packet network with independent erasures per packet, it is expected that the proposed algorithm would yield video a lower mean distortion than MD-RDS. Furthermore, the variance of the distortion would also be lower.

In Figs. 4 and 5, a comparison of the distortions of the two descriptions generated by the two algorithms is shown as a function of R^*. In Fig. 4, the "Foreman" sequence is used with

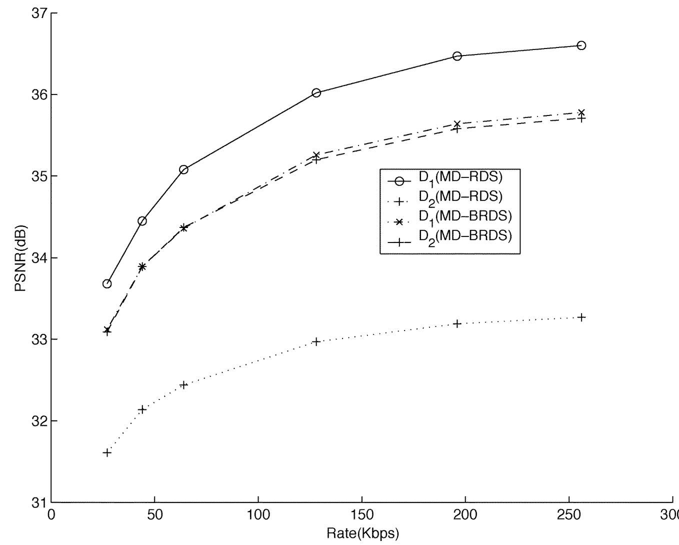

Fig. 5. Comparison of description distortions at different R^* using the "Akiyo" sequence, $\rho = 0.85$.

TABLE II
COMPARISON OF MD-BRDS AND SDC AT THE SAME BIT RATES. THE PSNR OF MD-BRDS FOR RATE $R_1 + R_2$ IS $D_0 = 35.28$ dB IN ALL CASES

ρ	R_1 (kbps)	D_1 (dB)	H.263 at rate R_1	H.263 at rate $R_1 + R_2$
0.45	94.58	25.73	34.10	37.09
0.60	102.82	28.13	34.48	37.55
0.70	109.14	30.25	34.67	37.85
0.80	115.54	32.22	34.84	38.12
0.90	121.93	33.89	35.07	38.38
1.00	128.36	35.28	35.28	38.63

$\rho = 0.70$ while, in Fig. 5, the "Akiyo" sequence is used with $\rho = 0.85$. Similar observations can be made.

To show the effects of the added redundancy, we compare the results of the MD-BRDS to results (using the same bit rates) of SDC (baseline H.263) using "Foreman." These results are presented in Table II. We compare D_1 with standard H.263 encoded at rate R_1, and D_0 (the two-channel distortion) with H.263 encoded at rate $R_1 + R_2$. As expected, D_1 gets closer to H.263 at rate R_1 as ρ increases.

Fig. 6 presents subjective results for the MD-BRDS algorithm observed for frame 18 of the "Foreman" sequence for a target bit rate of 128 kb/s for the encoder. Fig. 6(a) is obtained by decoding the single layered bit stream (SDC) generated by baseline H.263 encoder, R^* represents the corresponding bit rate in bits per pixel (b/p) and D^* is the corresponding distortion.

We next discuss the computational complexity of the proposed encoder. For a given threshold, the MD-RDS encoder generates one partitioning of the DCT coefficients per block. In contrast, our proposed MD-BRDS encoder generates a number of partitions, one of which is finally selected. It should be pointed out that, in terms of memory usage, our encoder retains only two sets of partitions per block at a time in its pursuit to find the optimal partition. Fig. 7 displays a histogram showing the number of partitions generated per block versus the number of blocks generating this number of partitions for

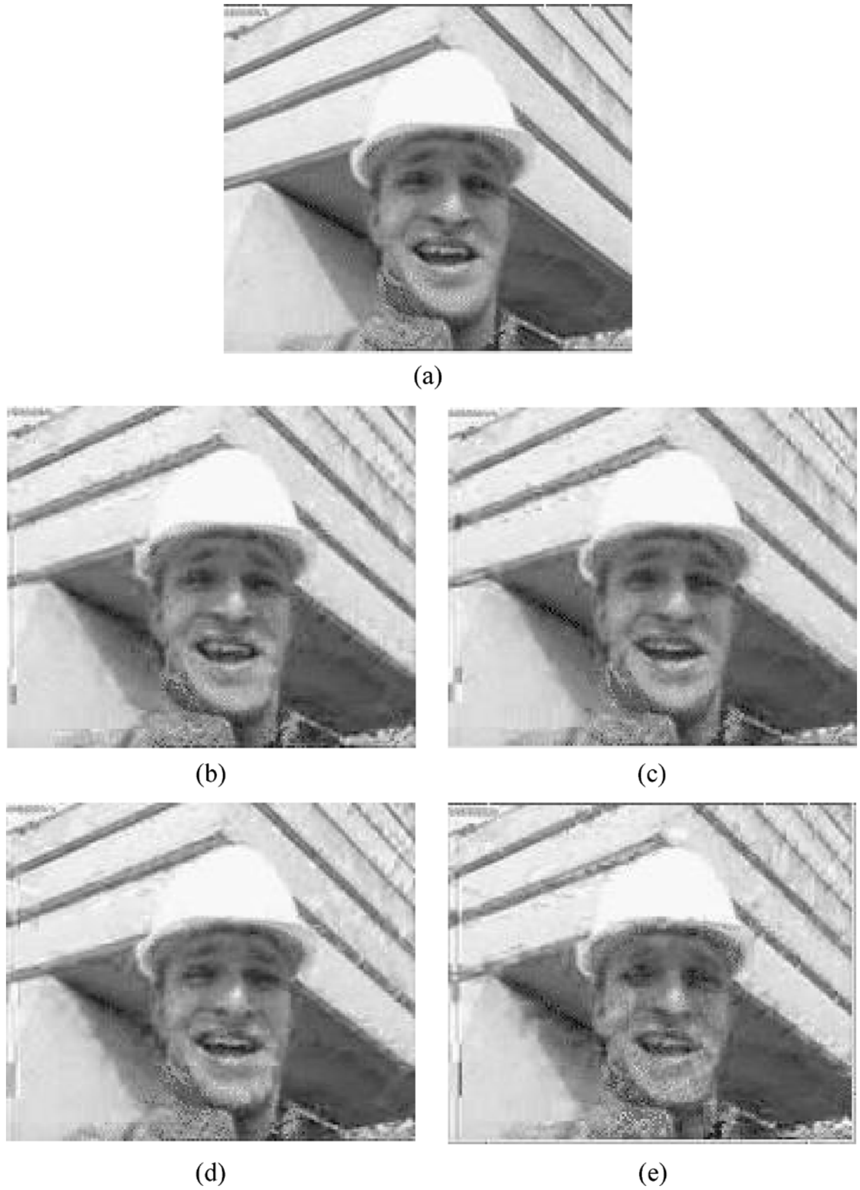

Fig. 6. Reconstruction results for frame 18 of the "Foreman" sequence, $\rho = 0.70$ (a) SDC, $R^* = 0.506$ b/p, $D^* = 35.28$ dB. (b) MD-BRDS, reconstructed from description-1, $R_1 = 0.431$ b/p, $D_1 = 30.30$ dB. (c) MD-BRDS, reconstructed from description-2, $R_2 = 0.430$ b/p, $D_2 = 30.25$ dB. (d) MD-RDS, reconstructed from description-1, $R_1 = 0.436$ b/p, $D_1 = 30.39$ dB. (e) MD-RDS, reconstructed from description-2, $R_2 = 0.425$ b/p, $D_2 = 28.28$ dB.

the "Foreman" sequence for different values of ρ. It can be seen that, in all cases, the vast majority of blocks in the video sequence generate no more than ten partitions. It can also be seen that, as ρ increases, the number of blocks with number of partitions between 1 and 10 also increases. As more coefficients are duplicated, the number of sets of partitions to be generated to achieve the optimal partitions is reduced. Thus, the computational complexity of the proposed encoder is reasonable and it decreases as ρ increases.

V. Conclusion

We have presented a balanced multiple description video codec. We have shown that the simple strategy of alternation and duplication of DCT coefficients spawns unbalanced descriptions which could be undesirable from the standpoint of transmissions over the "best-effort" networks like the today's Internet. We proposed a balanced MDC scheme to overcome this shortcoming. Our experimental results show that the proposed scheme outperforms the codec in [10] at the expense of reasonably larger complexity.

In this paper, H.263 was used as the basis for our algorithm, however any motion-compensated DCT-based codec could be used. The coefficient partitioning algorithm relies on Parseval's theorem, which is valid for all unitary transforms, including the DCT. The latest H.264 video compression standard [18] uses an integer approximation of the DCT that is not unitary (the rows of the transform matrix do not have the same norm). However, this is compensated for in the quantization process [19]. Thus, in the second phase of the coefficient splitting algorithm, swapping coefficients with the same quantization index between subsets will not significantly change the energy of each subset. Therefore, the proposed algorithm could be adapted to work with H.264.

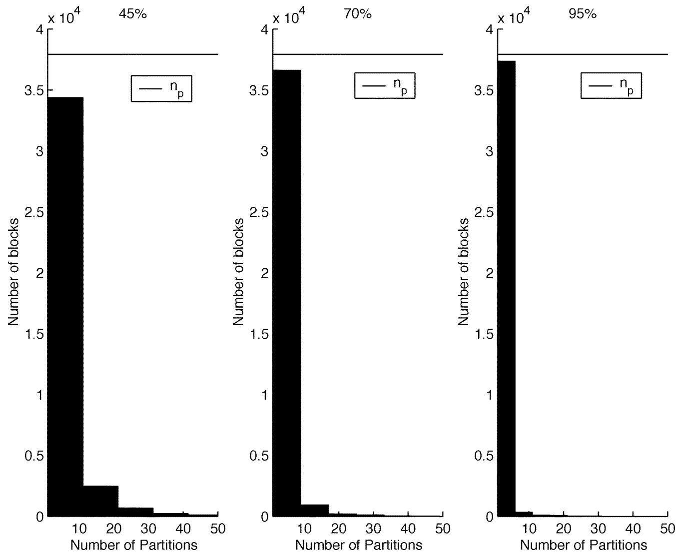

Fig. 7. Histogram showing the number partitions generated per block versus the number of blocks generating this number of partitions for "Foreman." n_p indicates the total number of blocks in the video sequence. Results are shown for $\rho = 0.45, \rho = 0.70$ and $\rho = 0.95$. The target bit rate was 128 kb/s.

In this work, we concentrated on the RRD performance of the proposed multiple description codec. Future work involves the application of the MD-BMDC codec to video transmission over direct sequence code division multiple access (DS-CDMA) wireless systems. We are considering the problem of joint optimization of source coding and physical layer parameters along the lines of our previous work on scalable video transmission over DS-CDMA wireless systems [20]–[22].

REFERENCES

[1] V. K. Goyal, "Multiple description coding: compression meets the network," *IEEE Signal Process. Mag.*, vol. 18, pp. 74–93, Sep. 2001.

[2] J. Wolf, A. Wyner, and J. Ziv, "Source coding for multiple descriptions," *Bell Syst. Tech. J.*, vol. 59, pp. 1417–1426, Oct. 1980.

[3] L. Ozarow, "On a source coding problem with two channels and three receivers," *Bell Syst. Tech. J.*, vol. 59, pp. 1909–1921, Dec. 1980.

[4] A. E. Gamal and T. M. Cover, "Achievable rates for multiple descriptions," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 6, pp. 851–857, Nov. 1982.

[5] V. K. Goyal, J. Kovačević, R. Arean, and M. Vetterli, "Multiple description transform coding of images," in *IEEE Int. Conf. Image Processing*, Chicago, IL, Oct. 1998, pp. 674–678.

[6] A. Mohr, E. Riskin, and R. Ladner, "Generalized multiple description coding through unequal loss protection," in *IEEE Int. Conf. Image Processing*, Kobe, Japan, Oct. 1999, pp. 411–415.

[7] R. Puri and K. Ramchandran, "Multiple description source coding using forward error correction codes," in *Proc. 33rd Asilomar Conf. Signals and Systems*, 1999, pp. 71–79.

[8] D. Comas, R. Singh, and A. Ortega, "Rate-distortion optimization in a robust video transmission based on unbalanced multiple description coding," in *IEEE 4th Workshop on Multimedia Signal Processing*, Cannes, France, Oct. 2001, pp. 581–586.

[9] V. K. Goyal and J. Kovačević, "Generalized multiple description coding with correlating transforms," *IEEE Trans. Inf. Theory*, vol. 47, pp. 2199–2224, Sep. 2001.

[10] A. R. Reibman, H. Jafarkhani, Y. Wang, and M. T. Orchard, "Multiple description video using rate-distortion splitting," in *IEEE International Conference on Image Processing*, Thessaloniki, Greece, Oct. 2001, pp. 978–981.

[11] K. R. Matty and L. P. Kondi, "Balanced multiple description video coding using optimal partitioning of the DCT coefficients," in *IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Montreal, QC, Canada, May 2004, pp. 245–248.

[12] Y. Wang, M. T. Orchard, V. A. Vaishampayan, and A. R. Reibman, "Multiple description coding using pairwise correlating transforms," *IEEE Trans. Image Process.*, vol. 10, pp. 351–366, Mar. 2001.

[13] K. Ramchandran and A. Ortega, "Rate-distortion methods for image and video compression," *IEEE Signal Process. Mag.*, vol. 15, pp. 23–50, Nov. 1998.

[14] G. Schuster and A. Katsaggelos, *Rate-Distortion Based Video Compression*. Norwell, MA: Kluwer, 1997.

[15] *H.263 Video Coding for Low bit Rate Communications*, Sep. 1997.

[16] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in rate-distortion sense," *IEEE Trans. Image Process.*, vol. 2, pp. 160–175, Apr. 1993.

[17] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, pp. 172–185, Feb. 1999.

[18] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[19] H. S. Malvar, A. Hallapuro, M. Karczewicz, and L. Kerofsky, "Low-complexity transform and quantization in H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 598–603, Jul. 2003.

[20] L. P. Kondi, D. Srinivasan, D. A. Pados, and S. N. Batalama, "Layered video transmission over multirate DS-CDMA wireless systems," *IEEE Trans. Circuits Syst. Video Technol.*, to be published.

[21] D. Srinivasan, L. P. Kondi, and D. A. Pados, "Scalable video transmission over wireless DS-CDMA channels using minimum TSC spreading codes," *IEEE Signal Process. Lett.*, vol. 1, no. 10, pp. 836–840, Oct. 2004.

[22] L. P. Kondi, D. Srinivasan, D. A. Pados, and S. N. Batalama, "Layered video transmission over multirate DS-CDMA wireless systems," in *Proc. SPIE Conf. Image and Video Communications.*, vol. 5022, Santa Clara, CA, Jan. 2003, pp. 289–300.