Rate-Distortion Optimized Bit Rate Control Scheme for A Wavelet Video Coder

Po-Yuen Cheng

Physical Optics Corporation, Applied Technology Division
Torrance, California 90505-5228

Jin Li and C.-C. Jay Kuo

Integrated Media Systems Center and Department of Electrical Engineering-Systems
University of Southern California, Los Angeles, California 90089-2564

ABSTRACT

The rate control problem can be greatly simplified by using a wavelet coder, since the desired bit rate for a particular frame can be easily reached by using its embedding property. We convert the rate control problem to a bit allocation problem for each frame, and solve it with two models. Frames are assumed to be independent in the basic model while their dependency is taken into account in the advanced model. Two computationally efficient rate control algorithms are then derived. Extensive experiments are performed to demonstrate the superior performance of the wavelet coder with the proposed rate control schemes over the MPEG standard with test model 5 (TM5).

Keywords: rate control, video coding, wavelet coding.

1 INTRODUCTION

Rate control is essential to video transmission through constant bit rate (CBR) channels such as ISDN, T1 and broadcasting channels. It is also important in other applications such as video editing and video storage with CD-ROM or digital video disk (DVD). The rate control problem can be roughly stated as: the determination of proper coding parameters so that decoded video quality is optimized with respect to a certain fixed channel rate. Since the relationship between coding parameters, coding rate and decoded video quality is not obvious, it can be complicated to meet the coding rate requirement by adjusting coding parameters. It is even more difficult to optimize the quality besides satisfying the coding rate constraint.

There are several existing rate control schemes for videos coded by the MPEG standard. A well known example is MPEG test model 5 (TM5) [9], which adopts a simple rate control scheme by adjusting the quantization step size based on buffer occupancy. Even though this scheme is easy to implement, the quality of decoded videos is poor since it does not pay any attention to distortion. To improve the video quality, the rate control can be formulated as a constrained optimization problem and solved by the Lagrangian or minimax technique [7]. Ortega *et al.* [10] proposed a rate control scheme which measured the coding distortion and rate through simulation, and used dynamic programming to search for the true global optimal solution. This approach is very complex since the rate-distortion (R-D) performance of the coder has to be measured by repeatedly encoding the source. Lin *et al.* [8] speeded up the scheme by using a spline interpolation (for I frames) and a piecewise linear interpolation (for P frames) to avoid the extensive measurement of R-D characteristics for all quantization settings. However, the complexity of this rate control algorithm is still high since the source video has to be encoded several times. Frimout [5] and Chen [1] proposed the use of exponential functions to model the relationship between rate, distortion, and quantization step size in a MPEG macroblock. Their schemes are computationally efficient, but the resulting performance is poor since the proposed model does not characterize the MPEG coder well. Although motion compensated predictive coding plays an important role in MPEG, the effect of frame prediction has been seldom taken into consideration in rate control. Lin *et al.* [8] observed from empirical data that the variance of the motion compensated residue grows linearly with the coding error of the reference frame. However, no analysis of this phenomenon was performed. Other research on rate control can be found in [2], [3], [15].

Wavelet coders have demonstrated an excellent performance in still image compression as evidenced by a sequence of papers by Shapiro [13], Taubman and Zakhor [14], Said and Pearlman [12], Ramchandran *et al.* [11] and Li *et al.* [6]. In addition to providing a better R-D tradeoff and a more pleasant subjective appearance, the wavelet coder has an embedding property in the sense that the bit stream can be truncated at any point without significant perceptible distortion. Compared to rate control for MPEG, research work on rate control for wavelet video coders is relatively few. Rate control for an embedded wavelet video coder is studied in this work. It is shown that the embedding property of the wavelet coder can greatly simplify the rate control problem and, as a result, it can be reformulated as a bit allocation problem for each frame. In this framework, we would like not only to meet the constant bit rate constraint but also to minimize the distortion with a basic and an advanced model. Frames are assumed to be independent in the basic model, and their dependency is considered in the advanced model. Two computationally efficient rate control algorithms are then derived by using the Lagrangian method. Extensive experiments are performed to demonstrate the superior performance of the wavelet coder with the proposed rate control schemes over the MPEG standard with test model 5 (TM5).

The paper is organized as follows. An embedded wavelet video coder is first described in Section 2. The coder uses motion compensation to reduce the temporal redundancy in image sequences and embedded wavelet coding to encode the residue. The R-D characteristic of the coder is also investigated. Then, the rate control problem is formulated and two rate control algorithms are developed in Section 3. The relationship between the reference and predictive frames is examined carefully. It is shown that the energy of the motion compensated residue depends linearly on the coding quality of the reference frame, and an advanced rate control scheme is derived based on this observation. Finally, experimental results and concluding remarks are given in Sections 4 and 5, respectively.

2 EMBEDDED WAVELET VIDEO CODER

An embedded wavelet video coder is described in this section. Motion compensated predictive coding is adopted in this coder. We use the block-based motion compensation adopted by the MPEG standard and then encode the residual image with an embedded wavelet coder as described below.

The coder first decomposes the motion compensated residue with an L-scale wavelet transform ($L = 5$ in our experiments). The maximum absolute value of the wavelet coefficient over the entire image is searched and denoted by

$$T_0 = \max_{b,i,j} \mid w^b_{i,j} \mid, \tag{1}$$

where b is the scale parameter and i, j are spatial domain indices. We first identify and record the most significant bit (MSB) of the wavelet coefficient. The collection of all these bits is called the first bit layer. The threshold of this layer is set as

$$T_1 = \frac{T_0}{2}. \tag{2}$$

For all wavelet coefficients whose absolute values are greater than T_1, they are significant in the first layer and assigned bit '1'. Otherwise, they are insignificant and assigned bit '0'. Significant bits are encoded with the context adaptive arithmetic coder [4], where the context of a coding position is defined by the significant bits of its spatial neighbors and that of its parent band coefficient. For significant coefficients , their signs are recoded right after the coding of the significant bit. Then, we proceed to bit layer 2, where the threshold is reduced by half, i.e.

$$T_2 = \frac{T_1}{2} = 2^{-2}T_0. \tag{3}$$

	Flower	Mobile	Tennis	Cheer
β_I	2.07	1.65	1.08	1.72
β_P	1.50	1.50	0.86	1.43
β_I/β_P	1.38	1.10	1.27	1.20

Table 1: Coding efficiency parameters of the proposed embedded wavelet video coder with respect to the I and P frames.

For the second layer, the coding is split into two stages: significance identification and refinement coding. Significance identification deals with coefficients that are not significant in the previous layer. Their absolute values are compared with threshold T_2 and encoded with '1' (significant) and '0' (insignificant). If significant, the sign bits are encoded as well. Refinement coding deals with coefficients that have been identified to be significant in the previous layer. They are refined with the precision of one more bit. We place the significance identification stage before refinement coding stage in each bit layer since it offers a better R-D tradeoff if the coding bit stream is cut during the coding of this layer. The same process repeats for bit layers 3, 4 and so on until the allocated rate is reached. The embedded wavelet coder used in the work is actually an improvement of the layered zero coder (LZC) proposed by Taubman and Zakhor [14]. For more details, we refer to [6].

The embedding property allows a smooth degradation of the image quality when the coding bits are reduced. Suppose that the coded bit stream is truncated at the middle of layer i. Then, a certain number of wavelet coefficients are encoded with i bits while others encoded with $i-1$ bits. The embedding property greatly simplifies the rate control task.

We performed extensive experiments for the evaluation of the R-D characteristic of the embedded wavelet coder with respect to I and P frames. We plot the coding distortion versus the coding rate in Fig. 1 for test videos Flower, Mobile, Tennis and Cheer of the CIF format. Empirically, the R-D performance of the wavelet video coder can be closely approximated by an exponentially decaying function as

$$D = D_{\max} 2^{-\beta R} = \sigma^2 2^{-\beta R}, \tag{4}$$

where $D_{\max}$ is the coding distortion at coding rate $R = 0$, which is also equal to the variance σ^2 of the wavelet coefficients before coding, and the coding efficiency parameter β characterizes the decaying rate of the distortion as the bit rate increases. It is clear that a coder is more efficient, if the corresponding β is larger. Parameters β for I and P frames are denoted by β_I and β_P and listed in Table 1. The ratio between β_I and β_P is between 1.1 and 1.4.

3 TWO RATE CONTROL SCHEMES

3.1 Problem formulation

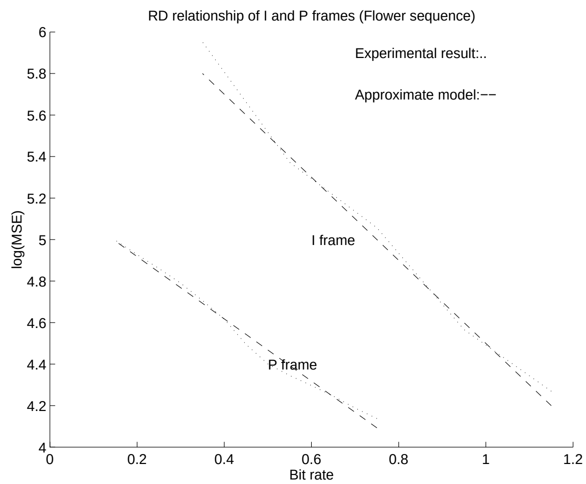
The rate control problem for the embedded wavelet coder can be formulated as follows. Suppose that the transmitted video can be partitioned into groups of pictures (GOPs), with each GOP starts with an intra-coded frame (I) and followed by $N-1$ predictively-coded frame (P), the channel capacity is C bits per second, and the duration of one GOP is T seconds. The allocated rate for one GOP is

$$R_{\text{GOP}} = CT. \tag{5}$$

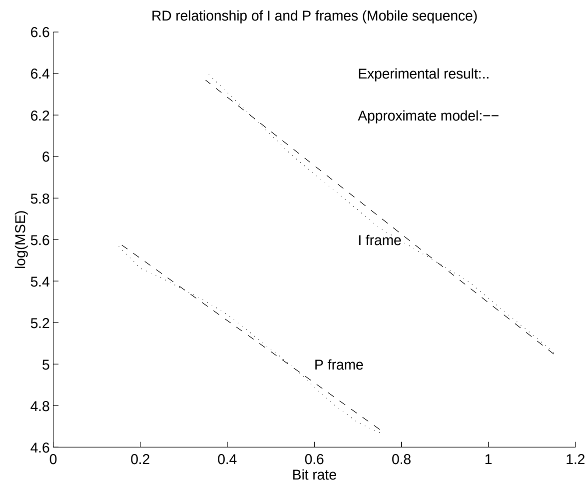
The rate-control problem for the embedded wavelet video coder can then be formulated as the bit allocation among N frames in a GOP.

$$\begin{cases} R_1 + R_2 + R_3 + \cdots + R_N = R_{\text{GOP}}, & \text{(bit rate constraint)} \\ \min(D_1 + D_2 + D_3 + \cdots + D_N), & \text{(distortion minimization)} \end{cases} \tag{6}$$

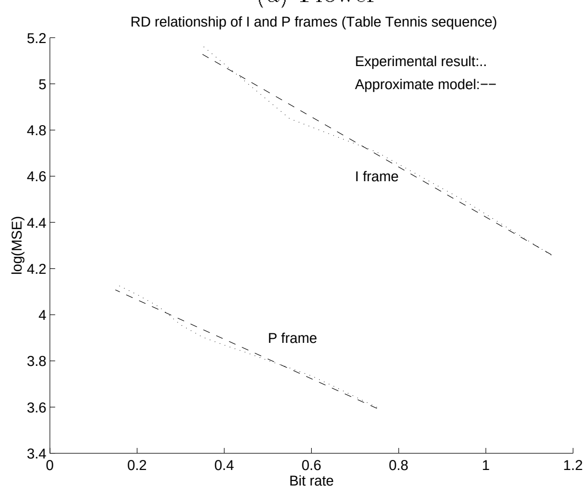
where R_i and D_i, $i = 1, \cdots, N$, are the coding rate and distortion for each frame, respectively. In contrast with previous work on rate control where the quantize step size is used as the coding control parameter, we now use the coding bit rate R_i directly as the coding control parameter by exploiting the embedded property of the wavelet coder.



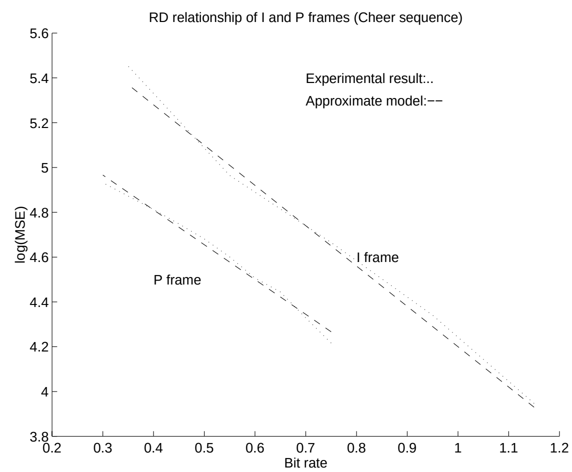
(a) Flower



(b) Mobile

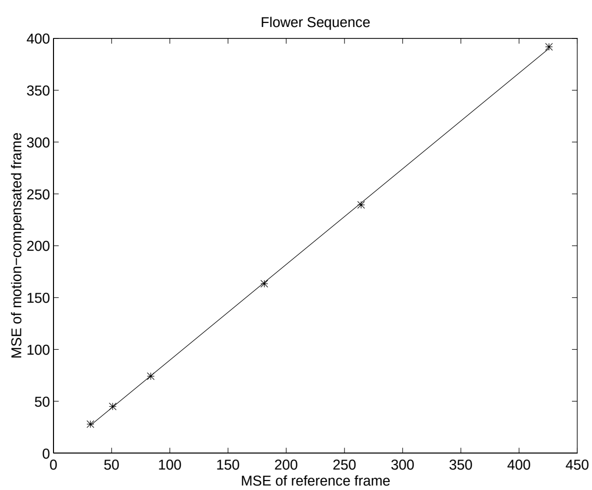


(c) Tennis

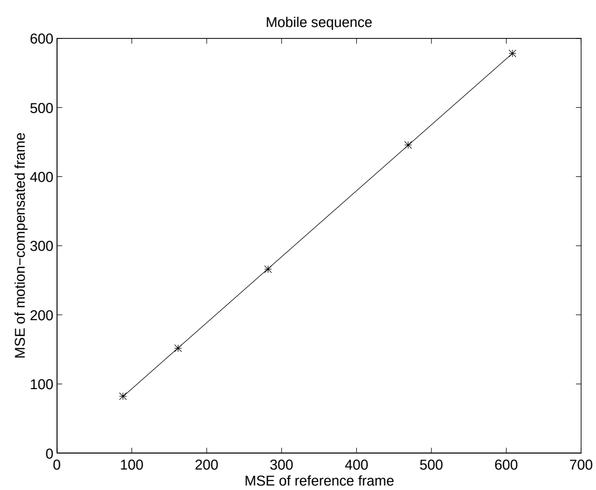


(d) Cheer

Figure 1: The R-D characteristics of the embedded wavelet video coder for I and P frames.



(a) Flower



(b) Mobile

Figure 2: The coding error of the motion-compensated reference frame versus the coding distortion of the reference frame.

3.2 Basic rate control scheme

By ignoring the interframe dependency resulted from motion compensation, we will derive a basic rate control scheme with the exponential R-D model (4) in this section. We first use the Lagrangian method to solve the constrained optimization problem (6), and the solution is of the form:

$$\frac{\partial D_i}{\partial R_i} = -\lambda \quad \text{(constant)}. \tag{7}$$

By assuming the exponential R-D model (4), we can obtain

$$\frac{\partial D_i}{\partial R_i} = -\sigma_i^2 2^{-\beta_i R_i} \beta_i \ln 2 = -D_i \beta_i \ln 2. \tag{8}$$

By substituting (8) into (7), we have

$$D_i = \frac{K}{\beta_i}, \qquad \text{with} \quad K = \frac{\lambda}{\ln 2}, \tag{9}$$

and

$$R_i = \frac{1}{\beta_i} \log_2 \frac{\beta_i \sigma_i^2}{K}. \tag{10}$$

With the bit rate constraint specified in (6), constant K can be calculated as:

$$\log_2 K = \frac{\sum_{i=1}^{N} \frac{1}{\beta_i} \log_2 \beta_i \sigma_i^2 - R_{\text{GOP}}}{\sum_{i=1}^{N} \frac{1}{\beta_i}}. \tag{11}$$

Equations (10) and (11) are the basic rate control scheme. However, as shown in (11), we have to estimate variance σ_i^2 and coding efficiency β_i to determine the allocated bit rate R_i for frame i.

For the I frame, variance σ_1^2 is the the mean square error of the wavelet decomposition. For the P frame, σ_i^2 is the mean square error of the motion prediction residue. We assume that the coding efficiency for all P frames are equal, i.e.

$$\beta_1 = \beta_I, \quad \text{and} \quad \beta_i = \beta_P, \qquad \text{for} \quad i = 2, \cdots, N. \tag{12}$$

Coding efficiency parameters β_I and β_P are estimated through the coding of the previous GOP.

3.3 Advanced rate control scheme

In the design of the basic rate control scheme, we assume that the frame distortion D_i, $1 \leq i \leq N$, in a GOP is independent. However, this is in general not true. Typically, an I frame of a better quality improves the motion predicted P frames in a GOP, thus reducing the coding bit rate required for following P frames. Consequently, an optimized rate control scheme should allocate more bits to the I frame. In this section, we first investigate the R-D relationship between a reference frame and its predictively coded frame, and then derive an advanced rate control scheme.

We denote the pixel of the original reference frame, the encoded reference frame and its predictively coded frame by $g(i,j)$, $\hat{g}(i,j)$ and $f(i,j)$, respectively. The residue of motion compensation with respect to the original frame can be written as:

$$e(i,j) = f(i,j) - g(d(i,j)) \tag{13}$$

where $d(i,j)$ is a displacement function which connects pixel (i,j) in the predictively coded frame to its motion compensated counterpart in the reference frame. The variance of the motion compensated residue can be calculated as:

$$\sigma_g^2 = E[e^2(i,j)] = E\{[f(i,j) - g(d(i,j))]^2\}. \tag{14}$$

In video coding, motion compensation is actually based on the encoded reference frame $\hat{g}(i,j)$ since the original reference frame is not available at the decoder. Therefore, the actual residual error is:

$$\hat{e}(i,j) = f(i,j) - \hat{g}(d(i,j)) \tag{15}$$

with variance

$$\hat{\sigma}_g^2 = E[\hat{e}^2(i,j)] = E\{[f(i,j) - \hat{g}(d(i,j))]^2\}. \tag{16}$$

We can rewrite $\hat{e}(i,j)$ as

$$\hat{e}(i,j) = [f(i,j) - g(d(i,j))] + [g(d(i,j)) - \hat{g}(d(i,j))], \tag{17}$$

where the first part is the residue of motion compensation with respect to the original reference frame while the second part is the coding error of the motion compensated reference frame. We may reasonably assume that these two parts are uncorrelated so that

$$\hat{\sigma}_g^2 = \sigma_g^2 + E\{[g(d(i,j)) - \hat{g}(d(i,j))]^2\}. \tag{18}$$

Note that the second term of (18) is not the coding distortion of the original reference frame but that of the motion compensated reference frame. These two quantities are nevertheless very close to each other. Their relationship for two test videos (Flower and Mobile) is shown in Fig. 2, where the relationship is linear with a slope α approximately equal to 1. Thus, we have

$$E\{[g(d(i,j)) - \hat{g}(d(i,j))]^2\} \approx \alpha D_f, \qquad D_f = E\{[g(i,j) - \hat{g}(i,j)]^2\}. \tag{19}$$

Combining (19) and (18), we obtain

$$\hat{\sigma}_g^2 = \sigma_g^2 + \alpha D_f, \tag{20}$$

where D_f is the coding distortion of the reference frame. We see from (20) that the variance of the motion compensated residue grows linearly with the coding distortion of the reference frame offset by the residue of motion compensation with respect to the original frame. Practically speaking, if the previous frame is poorly encoded, the quality of the forward predictive frame with motion compensation is also poor regardless of the effort spent in motion compensation. The derived relationship (20) has been confirmed by extensive experiments. We plot the variance of the motion compensated residue versus the coding distortion of the reference frame for Flower, Mobile, Tennis, and Cheer sequences in Fig. 3. The affine relationship is observed in all experiments. By taking the dependency of the reference frame and the predictive frame as shown in (20) into consideration, we propose an advanced rate control scheme in this section.

Consider a group of pictures (GOP) consisting of N frames, where the first frame is the I frame and the remaining $N-1$ frames are P frames. To derive this scheme, we convert the constrained optimization problem (6) to an unconstrained optimization problem by using the Lagrangian method. That is, our objective is to minimize

$$J(R_1, \cdots, R_N) = \sum_{i=1}^{N} D_i + \lambda(\sum_{i=1}^{N} R_i - R_{\text{GOP}}), \tag{21}$$

where R_{GOP} is the total number of bits assigned to a group of pictures, R_i the bit rate to be allocated to frame i and

$$D_i = \hat{D}_{\max,i} 2^{-\beta_i R_i} = \hat{\sigma}_i^2 2^{-\beta_i R_i}, \tag{22}$$

and

$$\hat{\sigma}_i^2 = \sigma_i^2 + \alpha_i D_{i-1}, \tag{23}$$

are constraints from (4) and (20), respectively.

At the optimum point, we have

$$\frac{\partial J}{\partial R_i} = 0, \qquad i = 1, 2, \cdots, N. \tag{24}$$

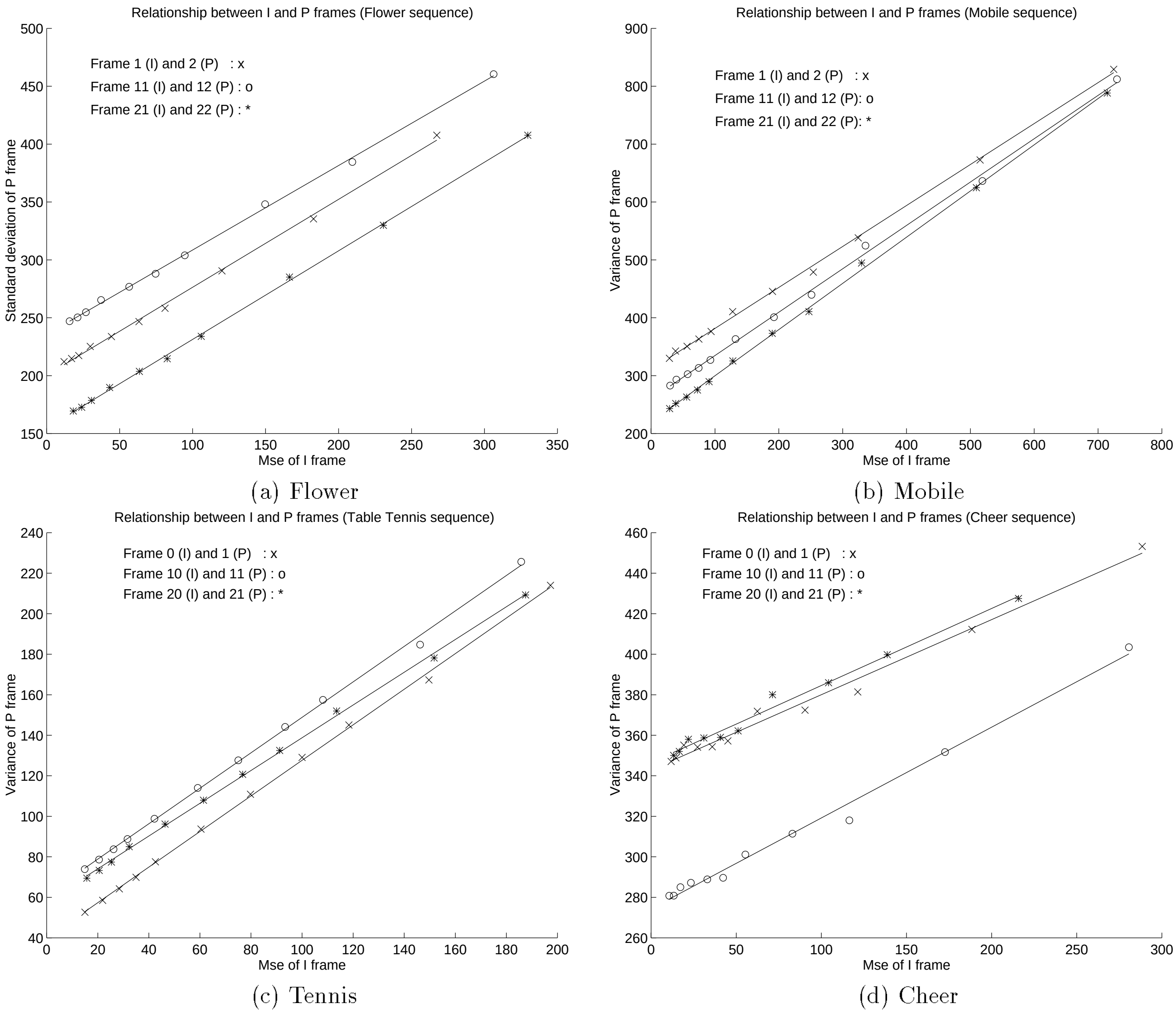


Figure 3: The relationship between the variance of the motion compensated residue and the coding error of the reference frame. For each sequence, we test three frame pairs: the 1st and the 2nd frames ('x'), the 11th and the 12th frames ('o'), and the 21st and the 22nd frames ('*').

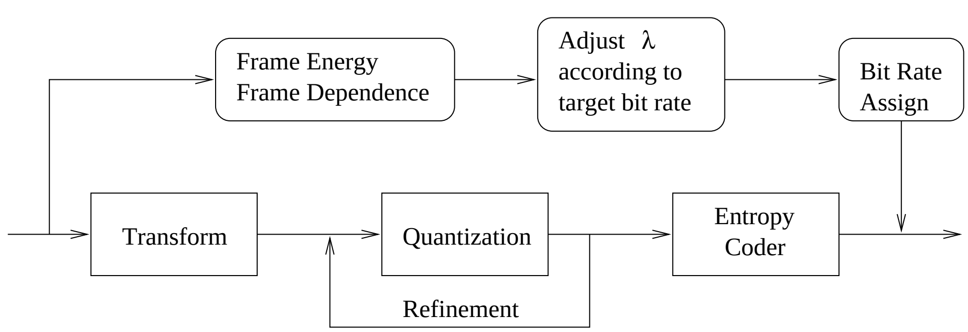


Figure 4: Block diagram of the advanced rate control scheme.

To solve the above system, let us first examine the coding rate for the last frame, i.e. frame N. As before, we introduce a constant K which is equal to $\lambda/\ln 2$ and substitute (22) into (24). This leads to

$$D_N = \frac{K}{\beta_N}. \tag{25}$$

Next, we proceed to frame $N-1$. Since the coding distortion D_N of frame N is related to the coding rate R_{N-1}, we have

$$\frac{\partial J}{\partial R_{N-1}} = \frac{\partial D_{N-1}}{\partial R_{N-1}} + \frac{\partial D_N}{\partial R_{N-1}} + \lambda. \tag{26}$$

By substituting (22) and (23) into (26), we get

$$\frac{\partial D_{N-1}}{\partial R_{N-1}} F_{N-1} = -\lambda, \quad \text{where} \quad F_{N-1} = 1 + \alpha_N 2^{-\beta_N R_N}. \tag{27}$$

Similar to (8), we can calculate $\partial D_{N-1}/\partial R_{N-1}$ in above and obtain

$$D_{N-1} = \frac{K}{\beta_{N-1} F_{N-1}}. \tag{28}$$

Also, from (22), we have

$$2^{-\beta_N R_N} = \frac{D_N}{\hat{\sigma}_N^2} = \frac{D_N}{\sigma_N^2} = \frac{K}{\beta_N \sigma_N^2} \tag{29}$$

By combining (27) and (29), we get

$$F_{N-1} = 1 + \alpha_N \frac{K}{\beta_N \sigma_N^2}. \tag{30}$$

Similarly, we can calculate the partial derivatives of J with respective to R_i, $i < N-1$, as

$$\frac{\partial J}{\partial R_i} = \frac{\partial D_i}{\partial R_i} + \frac{\partial D_{i+1}}{\partial R_i} + \frac{\partial D_{i+2}}{\partial R_i} + \cdots + \frac{\partial D_N}{\partial R_i} + \lambda. \tag{31}$$

It is easy to see that frames $i+2$, $i+3$, $\cdots$, N are correlated with frame i through frame $i+1$. Consequently, D_{i+2}, $\cdots$, D_N are also correlated with R_i through D_{i+1}. It is not difficult to show that

$$\frac{\partial J}{\partial R_i} = \frac{\partial D_i}{\partial R_i} + \frac{\partial D_{i+1}}{\partial R_i} F_{i+1} + \lambda, \tag{32}$$

where

$$F_{i+1} = \frac{K}{\beta_{i+1} D_{i+1}}. \tag{33}$$

We substitute (22) into (32) and get

$$\frac{\partial J}{\partial R_i} = \frac{\partial D_i}{\partial R_i} F_i + \lambda, \quad \text{with} \quad F_i = 1 + \alpha_{i+1} 2^{-\beta_{i+1} R_{i+1}} F_{i+1}. \tag{34}$$

Furthermore, by using (22) and (33), we have

$$F_i = 1 + \frac{\alpha_{i+1} K}{\hat{\sigma}_{i+1}^2 \beta_{i+1}}. \tag{35}$$

By comparing the right-hand-sides of (34) and (24), we obtain

$$D_i = \frac{K}{\beta_i F_i}, \quad \text{with} \quad F_i = 1 + \frac{\alpha_{i+1} K}{\hat{\sigma}_{i+1}^2 \beta_{i+1}}.$$

	Flower	Mobile	Tennis	Cheer
Coder A(dB)	24.29	22.97	30.04	25.42
Coder B(dB)	24.93	23.55	31.29	25.99
Coder C(dB)	25.06	23.62	31.51	26.10

Table 2: Average PSNR for the three test video coders at 1.152Mb/sec.

Note $\hat{\sigma}_{i+1}^2$ depends on D_i with the interframe dependency relation (23), we thus further solve D_i as:

$$\begin{aligned} D_i &= \frac{-B_i + \sqrt{B_i^2 - 4A_iC_i}}{2A_i}, \qquad i = 1, 2, \cdots, N-2, \\ A_i &= \beta_i \beta_{i+1} \alpha_{i+1}, \\ B_i &= \beta_i \beta_{i+1} \sigma_{i+1}^2 + K\alpha_{i+1}(\beta_i - \beta_{i+1}), \\ C_i &= -K\beta_{i+1}\sigma_{i+1}^2. \end{aligned} \tag{36}$$

Finally, with the coding distortion for all frames specified by (25), (28) and (36) in a GOP, we can solve the allocated rate for each frame i by (22) as

$$R_i = \begin{cases} \frac{1}{\beta_1} \log_2 \frac{\sigma_1^2}{D_1}, & i = 1, \\ \frac{1}{\beta_i} \log_2 \frac{\alpha_i D_{i-1} + \sigma_i^2}{D_i} & i = 2, 3, \cdots, N. \end{cases} \tag{37}$$

In the implementation of the advanced rate control scheme (37), variables β_i and σ_i^2 are obtained in the same way as the basic rate control scheme. We assume the interframe correlation factor α_i to be equal for all frames, and it is estimated from the coding of the previous GOP. K is an adjustable control parameter so that the bit rate constraint in (6) can be met. It is adjusted through simple bisection iteration consisting of the following three steps:

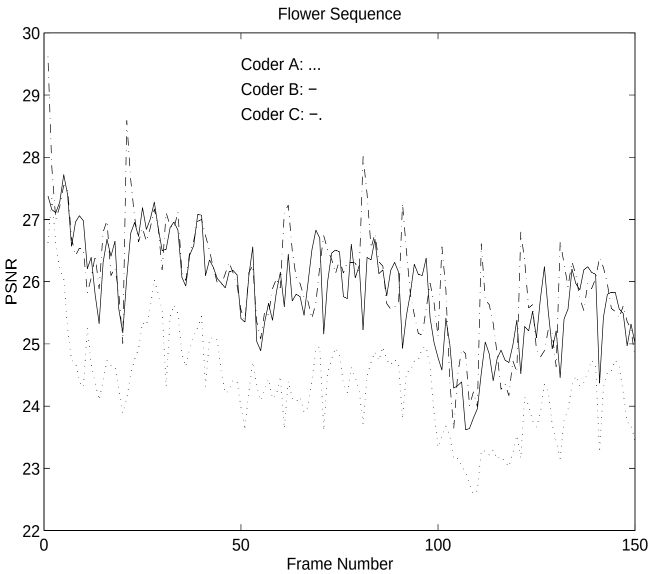
1. Select two initial parameters K_1 and K_2 so that bit rates obtained from (37) satisfy $R_{\text{GOP}}(K_1) > R_{\text{GOP}} > R_{\text{GOP}}(K_2)$.
2. Let $K_t = (K_1 + K_2)/2$. If $R_{\text{GOP}}(K_t) > R_{\text{GOP}}$, K_1 is replaced by K_t, otherwise K_2 is replaced by K_t.
3. Repeat Step 2 until $\mid R_{\text{GOP}}(K_t) - R_{\text{GOP}} \mid < \epsilon$, where ϵ is a nonnegative number close to zero. (It is chosen to be 10^{-6} in the experiments in Section 4.)

The block diagram of the advanced rate control scheme is shown in Fig. 4. Note that each iteration of the bisection algorithm only involves the evaluation of D_i by (25), (28) and (36) and R_i by (37) with frame parameters α_i, β_i and σ_i^2. The complexity of the algorithm is $O(N)$. Thus, the advanced rate control scheme is also computationally efficient.

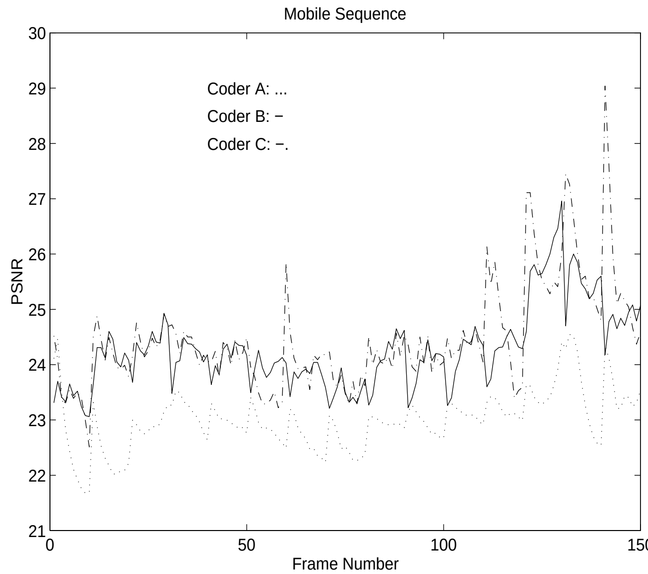
4 EXPERIMENTAL RESULTS

In this section, we compare the proposed embedded wavelet video coder with the basic and advanced rate control schemes with MPEG test model 5 (TM 5). MPEG TM 5 and the wavelet video coder with basic and advanced rate control schemes are denoted by coders A, B and C, respectively. The target bit rate is 1.152 Mb/sec. Four test image sequences are used in the experiments: Flower, Mobile, Tennis and Cheer. The sequences are of the CIF format with frame size 352×240 and 30 frames per second. Each sequence runs for 5 seconds, or 150 frames. The structure for each GOP is chosen to be one I frame followed by nine P frames so that there are 3 GOPs in one second.

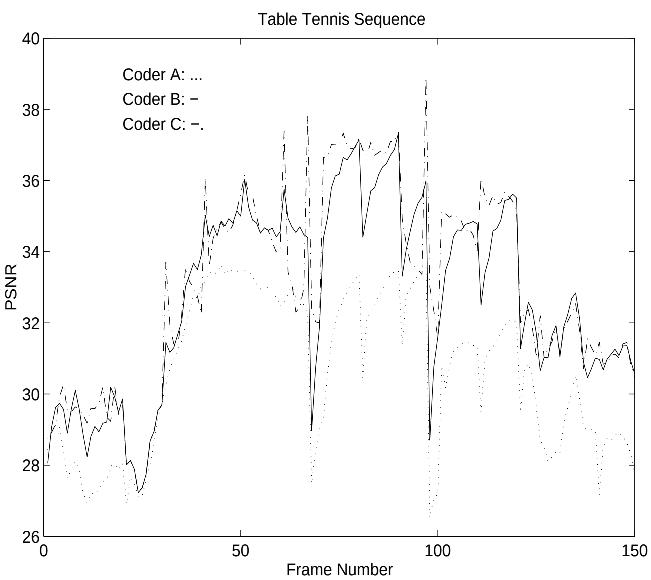
We plot the PSNR as a function of the frame number in Fig. 5. The average PSNR is also shown in Table 2. In the average, the embedded wavelet coder with the basic rate control scheme offers an average PSNR gain of 0.6-1.3 dB over MPEG TM5, the advanced rate control scheme offers an additional PSNR gain of 0.07-0.22dB.



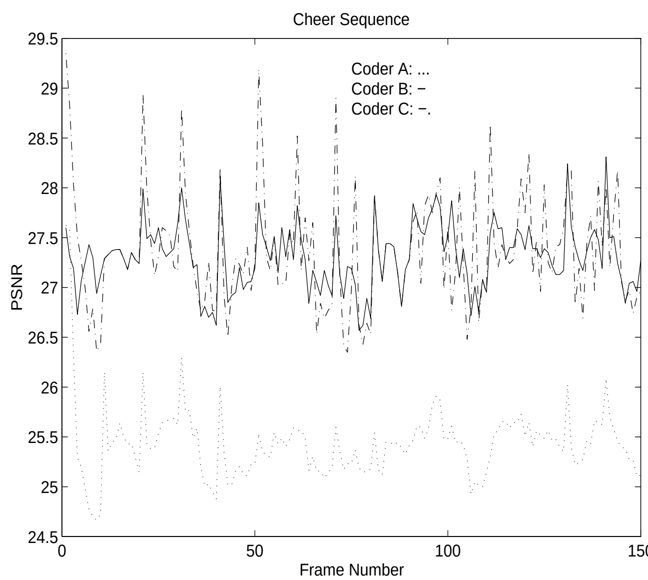
(a) Flower



(b) Mobile

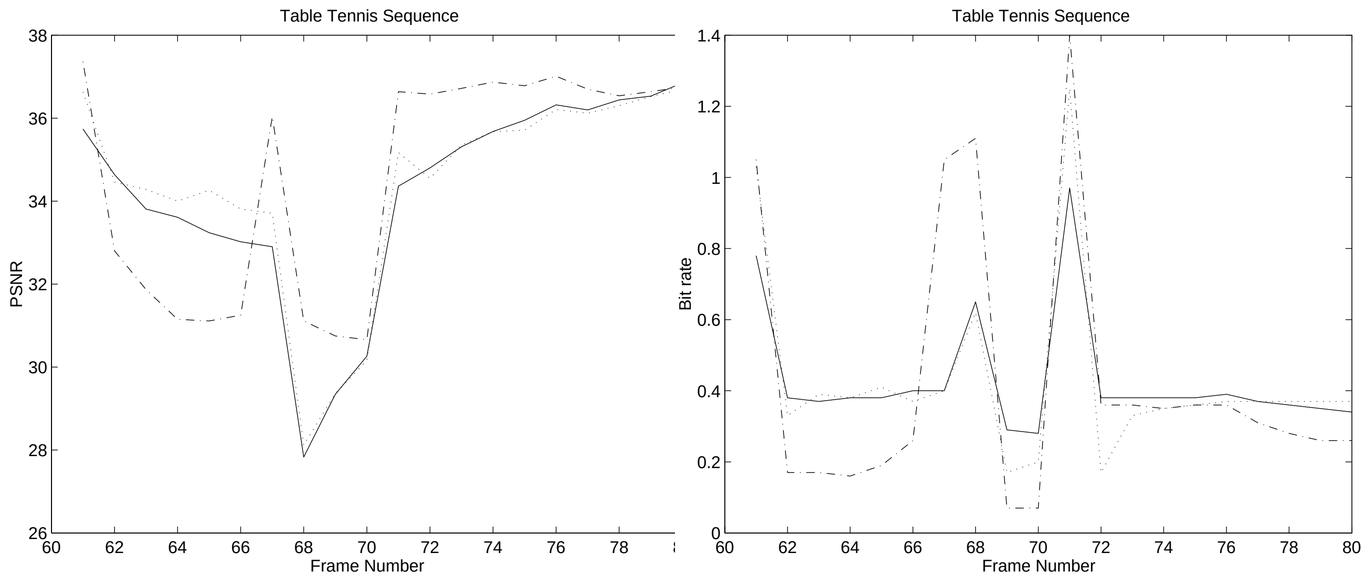


(c) Tennis



(d) Cheer

Figure 5: The PSNR performance as a function of the frame number by using MPEG TM5 (dotted line), the embedded wavelet coder with the basic rate control scheme (solid line) and the advanced rate control scheme (dash dotted line).



(a) PSNR (b) Bit rate

Figure 6: The (a) PSNR values and (b) coding bit rates in the neighborhood of scene change (frame 60-80 of Tennis) for an embedded wavelet coder using the bit rate determined by MPEG TM5 (dotted line), the basic rate control scheme (solid line), and the advanced rate control scheme (dash dotted line).

	MPEG TM5	Wavelet (MPEG TM5)	Basic Scheme	Advanced Scheme
Frames 61-80 (dB)	31.33	33.19	33.30	33.72
Frames 91-100 (dB)	30.14	32.57	32.41	33.18

Table 3: Average PSNR in the neighborhood of scene change.

It is observed that the advanced rate control scheme performs especially well around scene change. To demonstrate this, we focus on the PSNR and bit rate variation in the neighborhood of a scene change (Frame 60-80 of Tennis), where the scene change occurs at the 67th frame. We compare MPEG TM5 and three different wavelet-based coders with different bit rates, where the first wavelet coder follows the exact coding bit rate assigned by MPEG TM5 and the second and third coders use the basic and advanced rate control schemes, respectively. The average PSNR for this interval is shown in Table 3, The advanced rate control scheme is about 0.4-0.8 dB better than the other two wavelet coders, and 2.4-3.0 dB better than MPEG TM 5. We also plot the PSNR and bit rates as functions of the frame number in Fig. 6 for the three wavelet coders. For the wavelet coder with the MPEG TM 5 rate and with the basic rate control scheme, the bit rate assigned to the frame immediately after the scene change is not large enough (see Fig. 6 (b)) so that the quality of the scene change frame becomes very poor. It also affects the quality of frames thereafter as shown in Fig. 6 (a).

Note also that the bit rate for each GOP of the proposed wavelet video coder with basic and advanced rate control schemes is exactly the same as the allocated bit rate $R_{\rm GOP}$ for one GOP. Such a property provides an additional advantage in video editing since we can modify or replace a whole GOP without affecting other GOPs.

5 CONCLUSIONS

In this research, we proposed two new rate control schemes for an embedded wavelet video coder. They do not only meet the constant bit rate (CBR) transmission constraint but also minimize the coding distortion. The basic rate control scheme ignores the frame dependency. In deriving the advanced rate control scheme, we investigated the relationship between the reference and predictive frames, and showed that the variance of the motion compensated residue depends linearly on the coding distortion of the reference frame. By incorporating this observation with an exponential decaying R-D codec model, we solved the optimal bit allocation problem by using the Lagrangian method and demonstrated that the allocated bit rate for each frame R_i can be calculated as a function of several image dependent parameters such as the variance of each frame (σ_i^2), the correlation parameter between the reference and the predictive frame (α_i) and the coding efficiency parameter (β_i). Both basic and advanced rate control scheme have a relatively low computational complexity. Experimental results demonstrated that the proposed rate control schemes achieved a better PSNR performance. The advanced rate control scheme performs especially well around the scene change.

6 ACKNOWLEDGEMENT

This work was supported by the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, and the National Science Foundation Presidential Faculty Fellow (PFF) Award ASC-9350309.

7 REFERENCES

[1] J.-J. Chen and H. M. Hang, "A transform video coder source model and its application," in *Proc. of ICIP'94*, vol. II, (Austin, Texas), pp. 962–966, 1994.

[2] J. Choi and D. Park, "A stable feedback control of the buffer state using the controlled Lagrange multiplier method," *IEEE Trans. Image Proc.*, Vol. 3, pp. 546–558, Sep. 1994.

[3] W. Ding and B. Liu, "Rate control of MPEG video coding and recording by rate-quantization modeling," *IEEE Trans. on Image Processing*, pp. 12–20, Feb. 1996.

[4] D. L. Duttweiler and C. Chamzas, "Probability estimation in arithmetic and adaptive huffman entropy coders," *IEEE Trans. on Image Processing*, Vol. 4, pp. 237–246, Mar. 1995.

[5] E. Frimout, J. Biemond, and R. L. Lagendijk, "Forward rate control for MPEG recording," in *Proc. of SPIE Visual Communications and Image Processing '93*, (Cambridge, MA), pp. 184–194, Nov. 1993.

[6] J. Li, P.-Y. Cheng, and C.-C. J. Kuo, "On the Improvements of Embedded Zerotree Wavelet (EZW) Coding," in *SPIE: Visual Communication and Image Processing'95*, vol. 2501, (Taipei, Taiwan), pp. 1490–1501, May 1995.

[7] D. W. Lin, M.-H. Wang, and J.-J. Chen, "Optimal delayed-coding of video sequences subject to a buffer-size constraint," in *Proc. of SPIE Visual Communications and Image Processing '93*, (Cambridge, MA), pp. 223–234, Nov. 1993.

[8] L.-J. Lin, A. Ortega, and C.-C. J. Kuo, "A gradient-based rate control algorithm with applications to MPEG video," in *Proc. of ICIP'95*, vol. III, (Washington, D.C.), pp. 392–395, 1995.

[9] MPEG-2, *Test Model 5 (TM5) Doc. ISO/IEC JTC1/SC29/WG11/93-225b*, Test Model Editing Committee, Apr. 1993.

[10] A. Ortega, K. Ramchandran, and M. Vetterli, "Optimal trellis-based buffered compression and fast approximation," *IEEE Trans. Image Proc.*, Vol. 3, pp. 26–40, Jan. 1994.

[11] K. Ramchandran, Z. Xiong, K. Asai, and M. Vetterli, "Adaptive transforms for image coding using spatially-varying wavelet packets," *Submitted to the IEEE Trans. on Image Processing*, 1995.

[12] A. Said and W. A. Pearlman, "A new, fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 6, pp. 243–250, Jun 1996.

[13] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. on Signal Processing*, Vol. 41, No. 12, pp. 3445–3462, Dec. 1993.

[14] D. Taubman and A. Zakhor, "Multirate 3-D subband coding of video," *IEEE Trans. on Image Processing*, Vol. 3, No. 5, pp. 572–588, Sep. 1994.

[15] S.-W. Wu and A. Gersho, "Rate-constrained optimal block-adaptive coding for digital tape recording of HDTV," *IEEE Trans. Circuits and Sys. for Video Tech.*, Vol. 1, pp. 100–112, Mar. 1991.