Available at
www.ComputerScienceWeb.com
POWERED BY SCIENCE DIRECT®

ACADEMIC PRESS

J. Vis. Commun. Image R. 14 (2003) 136–149

JOURNAL OF VISUAL Communication & IMAGE Representation

www.elsevier.com/locate/yjvci

A bit allocation scheme for a class of embedded wavelet video encoders

Rogério Caetano and Eduardo A.B. da Silva*

PEE/COPPE/DEL/EE, Universidade Federal do Rio de Janeiro, Cx. P. 68504, Rio de Janeiro, RJ 21945-970, Brazil

Received 7 May 2001; accepted 27 February 2003

Abstract

In this paper, we investigate bit allocation strategies for a class of embedded wavelet video encoders. They take advantage of the precise control that such coders have over the bit-rate of each frame. We first show that a piecewise-linear model suits the rate × distortion characteristics of these encoders better than an exponential model, specially in low bit-rate applications. Then, we use an effective iterative procedure for dealing with the problem of frame dependency which yields improved rate × distortion results. Two types of embedded wavelet coders, using scalar and vector quantization, are tested. The results are encouraging, showing that the adoption of an adequate rate-control strategy can improve both objective and subjective quality of video sequences encoded using such embedded wavelet video encoders.

Keywords: Rate control; Video coding; Wavelet transforms; Embedded encoders; Vector quantization; MPEG-4

1. Introduction

Embedded wavelet coders have been very successful in coding still images. They are part of a number of standards, including JPEG-2000 (ISO/IEC JTC1/SC29/WG1, 1999) and the texture coding part of MPEG-4 (ISO/IEC JTC1/SC29/WG11,

* Corresponding author. Fax: +55-21-290-6626.
E-mail addresses: caetano@lps.ufrj.br (R. Caetano), eduardo@lps.ufrj.br (E.A.B. da Silva).

1047-3203/03/$ - see front matter
doi:10.1016/S1047-3203(03)00020-8

1997). Nevertheless, for interframe video encoding, the classical DCT-based methods are in general preferred. However, the excellent performance of embedded wavelet coders for still images is a strong indication that it is worthy to further investigate their application to video coding. This assumes particular importance if one considers that MPEG-4 syntax allows the specification of alternative encoders for specific applications via its Systems Description Language (MSDL). In this paper, we investigate rate-control aspects related to the use of such encoders as replacements for the DCT encoding of the motion compensated frame difference in a MPEG-4 VM-8 encoder (ISO/IEC JTC1/SC29/WG11, 1997).

Among the most well-known embedded wavelet encoders, one can cite the EZW (Shapiro, 1993) and the SPIHT (Said and Pearlman, 1996) encoders. One of their main advantages is the absence of blocking effects due to the use of wavelet transforms instead of the DCT.

Another advantage of such coders is that they have precise control over the bitrate of every frame. This is unlike usual DCT-based methods, where the rate can only be indirectly controlled by the quantizer step-size, that is, a larger step-size implies smaller rate and signal-to-noise ratio, and vice versa.

For example, in the MPEG-4 VM-8 (ISO/IEC JTC1/SC29/WG11, 1997), the bitrate control strategy assumes a quadratic relation between the quantizer step-size and the rate spent in transmitting a frame. It is updated by a least-squares fitting as each new frame is encoded. Given the number of bits remaining to encode the sequence, one updates the quantizer step-size according to this relation. Although efficient, this is based on a number of ad hoc assumptions and is clearly suboptimal in a rate-distortion sense.

This paper investigates rate allocation strategies for a class of embedded wavelet encoders. This class can be characterized by the use of integer bit planes, like in the SPIHT (Said and Pearlman, 1996), EZW (Shapiro, 1993), and MGE (Lan and Tewfik, 1999) algorithms, as opposed to the use of fractional bit-planes, as in the EBCOT coder (ISO/IEC JTC1/SC29/WG1, 1999; Taubman, 1999). Unlike previous works (Li et al., 1997), which assume that the *R–D* curve for a single frame difference is exponential, we have shown that a piecewise-linear model provides a better approximation for the *R–D* characteristics of these embedded wavelet encoders than the exponential model. This enables us to perform rate allocation using Lagrangian optimization at low computational complexity. In addition, we use an iterative procedure similar to the one proposed in Hsu et al. (1997) to deal with the problem of frame dependency. We have tested two types of embedded wavelet encoders, the EZW (Shapiro, 1993) and its vector quantization counterpart, the SA-W-VQ coder (da Silva et al., 1996). We show that the proposed improvements generate embedded wavelet video encoders that outperform the MPEG-4 VM-8 by a comfortable margin.

The rest of the paper is divided as follows: Section 2 describes the modifications to the MPEG-4 VM-8 in order to use EZW and SA-W-VQ instead of the DCT. Section 3 analyses the rate-control problem, and proposes the new rate-control strategy. Section 4 presents experimental results and Section 5 states the conclusions.

2. Implementation of the video encoder

In this work, we have employed two types of embedded wavelet encoders. The EZW (Embedded Zero-Trees Wavelet) (Shapiro, 1993), which is based on successive approximation *scalar* quantization, and the SA-W-VQ (Successive Approximation Vector Quantization), which is based on successive approximation *vector* quantization (da Silva et al., 1996). In both coders, we have used a 1-stage biorthogonal wavelet transform as recommended in Sampson et al. (1995). In the SA-W-VQ case, the vectors were formed by dividing the wavelet coefficients into 2×2, 4×2, or 4×4 blocks. The orientation codebooks used were the first shells of the root lattices D_4, E_8, and Λ_{16}, related to the solution of the sphere packing problem (Conway and Sloane, 1988) at dimensions 4, 8, and 16, respectively, with $\alpha = 0.55$ (da Silva et al., 1996).

It is important to point out that both schemes (EZW and SA-W-VQ) are built on the EZW framework. It is true that we could have used a more efficient framework, like the one of the SPIHT coder (Said and Pearlman, 1996), which would certainly improve both the scalar and vector quantization results (Mukherjee and Mitra, 1998). However, this is of secondary importance, since the main emphasis of this work is on the rate-control strategy.

The video encoder used in this work is based on the MPEG-4 VM-8. Since, in this encoder, a block DCT is employed, its structure assumes that the images are divided into blocks. Therefore, in order to use a wavelet transform in such a framework, some modifications have to be made to the MPEG-4 VM-8 structure.

The main modification to the VM-8 is that the DCT, quantization, and run-length encoding followed by Huffman coding have been replaced by an embedded wavelet encoder followed by an adaptive arithmetic coder (Bell et al., 1990).

Other modification has to do with the fact that a wavelet transform is no longer applied to independent image blocks, but to the image as a whole. The image is divided in macroblocks only for the purposes of motion estimation and compensation. In this implementation, we are restricted interframe forward prediction. Therefore, in a coded frame all of its macroblocks are interframe.

We have also chosen the option of considering each frame composed of only one rectangular VOP (ISO/IEC JTC1/SC29/WG11, 1997), identical to the frame itself; also, the overlapped motion compensation option is always turned on, as is usually done in wavelet-based video encoders (Otha and Nogaki, 1993).

In our simulations, the first frame of a sequence was never encoded and the other frames were all interframe-coded. Therefore, any difference in the performance of the encoders is due only to the performance of the interframe encoders.

3. Rate control

Embedded wavelet encoders like EZW and SA-W-VQ have the capability to set up precisely the bit-rate of each frame. In constant bit-rate (CBR) systems, if one looks at this problem purely from the rate-control point of view, the trivial solution would be to divide equally the bit-rate among all frames. This solution would in principle have the advantage of obviating the need of a buffer in order to smooth out bit-rate variations. However, very often one needs to find out how many bits should be allocated to each frame in order to obtain the highest average signal-to-noise ratio (or, alternatively, subjective image quality) of the entire sequence for a given rate budget. This optimal bit allocation problem can be stated more formally as (Ramchandran and Ortega, 1998):

Given a set of N frames $\{f_i,\ i = 1, \ldots, N\}$ and an average target rate of R_{target} bits/frame, we have to encode frame f_i with rate R_i, yielding distortion D_i such that:

$$D = \frac{1}{N}\sum_{i=1}^{N} D_i \text{ is minimum,} \tag{1}$$

$$R = \frac{1}{N}\sum_{i=1}^{N} R_i \leqslant R_{\text{target}}, \tag{2}$$

where R and D are the average bit-rate and average distortion, respectively, for the N frames.

Considering that the rate–distortion (R–D) functions of each frame f_i are convex, this problem can be solved via Lagrangian optimization. In this case, it can be restated as:

$$\text{minimize } J = D + \lambda R, \tag{3}$$

where the function J is minimized for a given value of the Lagrange multiplier λ. Each solution of Eq. (3) for a value of λ corresponds to a solution of Eq. (2) for a particular R_{target} (Everett, 1963; Shoham and Gersho, 1988).

The advantage of this formulation is that, if λ is fixed, the rates that minimize Eq. (3), $\{R_i(\lambda),\ i = 1, \ldots, N\}$, can be found by solving only a non-restricted optimization problem, which is in general much easier to solve than the restricted optimization problem in Eqs. (1) and (2). Having the rates $R_i(\lambda)$, one has just to search in the λ space for the solution corresponding to a λ_0 such that:

$$\frac{1}{N}\sum_{i=1}^{N} R_i(\lambda_0) = R_{\text{target}}. \tag{4}$$

The Lagrangian optimization problem in Eq. (3) can be further simplified if we make the extra assumption that the R–D characteristics of frame f_i are independent of the particular point (R_j, D_j) in which frame j is being coded, $\forall j \neq i$. Then Eq. (4) is equivalent to solving:

$$\text{minimize } J_i = D_i + \lambda R_i, \quad i = 1, \ldots, N. \tag{5}$$

If we know an analytical expression for $D_i(R_i)$, $i = 1, \ldots, N$, then the minimization problem in either Eq. (3) or (5) can be first solved as a function of λ, and then one finds λ such that Eq. (4) is satisfied.

An important characteristic that we observed in the previous formulation is that a good estimate of the R–D characteristics is critical to the success of the rate-control method. Two popular estimates are: (a) when an analytical model is assumed for the *R–D*: in this case an exponential model is often used; (b) an operational *R–D* curve, where the points of the *R–D* curve are estimated by actually encoding the input.

In Li et al. (1997), it was investigated the use of an EZW-like encoder in interframe coding, and developed an optimum rate allocation strategy supposing that the *R–D* characteristics of each frame are exponential. This technique outperforms the fixed allocation rate control by 0.1–0.4 dB in high rate (≈1.2 Mb/s).

However, we argue here that the exponential model is neither the best nor the simplest model for such *R–D* characteristics, specially in low bit-rate applications. Our argument is supported by the fact that, in embedded wavelet encoders based on integer bit-planes (Lan and Tewfik, 1999; da Silva et al., 1996; Said and Pearlman, 1996; Shapiro, 1993), the wavelet coefficients are transmitted by bit-planes (in the SA-W-VQ algorithm one considers the set of codevectors in a given pass as a "vector bitplane" (da Silva et al., 1996)). Within a certain bit-plane, for each bit sent, there is a fixed reduction in the total distortion. This corresponds to the precision associated with that bit-plane, that is, $\Delta D/\Delta R = k$, and therefore the *R–D* characteristics are linear. Within the next bit-plane the precision is doubled, and $\Delta D/\Delta R = k/2$. Therefore, the *R–D* characteristics are linear with half the slope of the previous one. Such behavior gives rise to piecewise-linear characteristics instead of exponential ones.

For example, in EZW, each dominant and subordinate pass has only one bit-plane associated to it, and hence a single slope $\Delta D/\Delta R$ (one should note that, in the EZW coder, the subordinate pass $n-1$ correspond to the same bit-plane as the dominant pass n). The same applies to the SPIHT, MGE, and SA-W-VQ coders.

This can be confirmed in Fig. 1, where one can see the *R–D* characteristics of the motion compensated frame difference between frames 45 and 42 of the Mother sequence. It is encoded by a SA-W-VQ encoder using Λ_{16} as the orientation codebook (da Silva et al., 1996). It is important to note that an exponential model would only

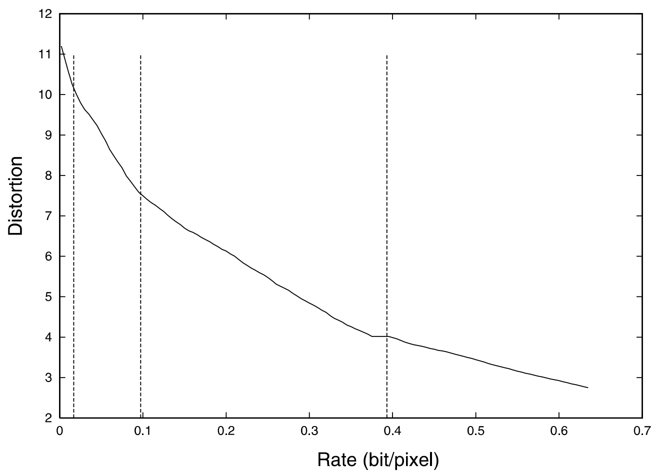

Fig. 1. *R–D* Characteristics of an interframe prediction error encoded by an embedded wavelet encoder.

fit to a curve as in Fig. 1 if the number of bit-planes was large, which seldom is the case in low bit-rate applications (in (Li et al., 1997), since the target rates are of the order of 1 Mb/s, then the exponential model is a good one). It is important to observe that, in the case of coders using fractional bit-planes, e.g., EBCOT (Taubman, 1999) in JPEG 2000, each bit-plane is coded by several passes. Then, each bit sent in same bit-plane is not likely to provide the same distortion reduction, and thus their *R–D* characteristics are not piecewise-linear.

An important advantage of the piecewise-linear model is that it allows a large reduction in the complexity of a Lagrangian optimization (Lin and Ortega, 1998). In addition, by closely examining Fig. 1, we see that the "breakpoints" (represented by the vertical lines) of the piecewise-linear curve lie at the boundaries of the bit-planes (da Silva et al., 1996; Shapiro, 1993). Therefore, in order to estimate the operational *R–D* characteristics in this case, the encoder only needs to encode the frame for a certain number of bit-planes and decode it just for the rates corresponding to the breakpoints.

Noting that Eq. (5) corresponds to a straight line in the plane $R_i \times D_i$ that intercepts the axis D_i at the point J_i, we see that to minimize J_i in this equation corresponds to finding the tangent to the *R–D* curve of frame f_i with inclination $-\lambda$ (Ramchandran and Ortega, 1998). This leads to the following simple algorithm to find the optimum rate allocation among N frames:

1. Given λ, find for each frame f_i the tangent to the *R–D* curve with inclination $-\lambda$; the tangency point is $(R_i(\lambda), D_i(\lambda))$;
2. Compute $R(\lambda)$ according to Eq. (2);
3. If $R(\lambda) = R_{\text{target}}$ then the optimal rate allocation is given by $R_i,\ i = 1, \ldots, N$; else, vary λ and go to step 1.

The main difficulty with the above algorithm is that to find the value of λ which gives a rate allocation satisfying Eq. (4) can be a quite time-consuming task if the number of allowed values of λ is large. However, since in our case the *R–D* characteristics of the frames are piecewise linear, we just need to execute passes 1 and 2 of the algorithm for the values of λ which correspond to the negatives of the set of slopes of all the linear segments of the *R–D* curves of all frames.

This can be easily seen by supposing that the opposite is true, that is, that the optimum λ is such that it does not correspond to the negative slope of any linear segment of the *R–D* curves. It is equivalent to the straight lines touching all *R–D* curves at "corners." This is illustrated by the dashed straight lines in Fig. 2, which specify the allocation $(R_1, D_1), \ldots, (R_{k'}, D_{k'}), \ldots, (R_N, D_N)$. However, one can decrease λ up to the value represented by the solid lines without changing the rate allocation. Since this represents the slope of the linear segment $(R_{k'}, D_{k'}) - (R_{k''}, D_{k''})$, in curve k, we arrive at a contradiction. Therefore, we conclude that it suffices to test the values of λ corresponding to the negatives of the slopes of the segments of the *R–D* curves.

Using this property, in order to find the optimum point, we choose an inclination λ such that $R_1 + \cdots + R_{k'} + \cdots + R_N \leqslant NR_{\text{target}} < R_1 + \cdots + R_{k''} + \cdots + R_N$. Thus, the segment $(R_{k'}, D_{k'}) - (R_{k''}, D_{k''})$ is the one, among all the segments of all curves, with largest negative slope smaller than of equal to λ. Then, it suffices to choose $R_{k'} \leqslant R_k \leqslant R_{k''}$ such that $\sum_i R_i = NR_{\text{target}}$.

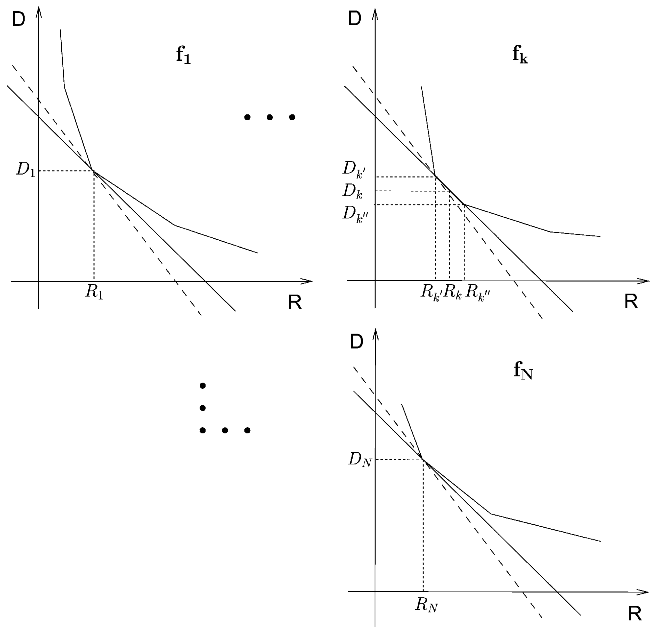

Fig. 2. Illustration of the rate allocation in the case of piecewise-linear *R–D* curves.

3.1. The frame dependency problem

A very important assumption made when deriving the rate allocation strategy above was that the *R–D* curves of each frame were independent of the operating point of the others. Unfortunately, in interframe video encoders this is not the case, because the prediction error corresponding to frame f_i depends on how frame f_{i-1} has been encoded, which completely violates the above assumption. In fact, for each operating point (R_j, D_j) in frame f_i, there will be a different *R–D function* for frame f_{i+1}. This is the so-called *dependency problem* (Ramchandran and Ortega, 1998).

The strategies to solve this problem are in general quite complex. In (Ramchandran et al., 1994), it was propose feasible solutions for the case when the number of points on each *R–D* curve is small enough. However, their solutions are unsuitable for embedded wavelet encoding schemes, that are usually able provide for each frame difference a very large number of points on the *R–D* curve. A different approach is the one employed in (Li et al., 1997), where the influence of frame f_i in frame f_{i+1} is restricted to a linear dependency of the variance of frame f_{i+1} on the coding error of frame f_i.

Here, we use a similar strategy to the one proposed in (Hsu et al., 1997) for dealing with the frame dependency problem. In our case, we apply the rate allocation algorithm described in this section to *N* consecutive frames (GOF) of a video sequence in several iterations. Having the reconstructed frames for iteration $n-1$, the *R–D* curves for iteration *n* are obtained and the rate allocation is computed, and so the frames for iteration n are reconstructed. This process proceeds until either the change in signal-to-noise ratio performance is below a threshold or a maximum number of iterations is exceeded. In our simulations, this process has led to an increase in signal-to-noise ratio performance in all cases considered. One should note that one advantage of this approach is that, instead of devising arbitrary models for the frame dependency, as in (Li et al., 1997), we have only resorted to the mild assumption that each iteration of the process leads in general to a signal-to-noise ratio improvement. Simulation results presented in Section 4 show that this process usually converges in less than 4 iterations.

4. Experimental results

The MPEG-4 VM-8 adaptation described in Section 2 has been simulated using the rate control strategy proposed in Section 3. We have used the test sequences Mother-and-daughter, Silent-voice, Hall-monitor, Children and Weather with 300 frames in QCIF format at 30 frames/s. The bit-rates used in all experiments were in the range 14–64 kbit/s. The sequences were sub-sampled in time by a factor of 3 to generate 10 frames/s sequences.

Table 1 and Fig. 3 compare the average peak signal-to-noise ratio (PSNR) of the different MPEG-4 adaptations when coding the test sequences using 64 kbit/s. "DCT" refers to MPEG-4 VM-8, "EZW" refers to the EZW encoder and "Λ_{16}", "E_8", and "D_4" to the SA-W-VQ encoder using the referred codebooks. In the case of the MPEG-4 VM-8, the rate control strategy allows the skipping of some frames. This obliges the decoder to repeat frames in order to maintain the frame-rate. Hence, the PSNR calculation must take the repeated frames into consideration. The columns labelled "Constant Rate" refer to the use of the same bit-rate for every frame of the wavelet encoders, while the columns labelled "Rate Control" refer to the wavelet encoders employing the rate control strategy proposed in Section 3, using 4 iterations of the rate-control algorithm in a group of 40 frames (GOF = 40). One can see from this table that the use of the proposed rate control increases the rate $\times$ distortion performance of both the EZW-based and SA-W-VQ-based encoders. It is interesting to note that, even without rate control, the performance of the wavelet-based encoders is in general superior to the one of the MPEG-4 VM-8.

Table 1
Comparison between the average PSNR (dB) of the different MPEG-4 versions at a bit-rate of 64 kbit/s

	Constant rate				Rate control				
	Λ_{16}	E_8	D_4	EZW	Λ_{16}	E_8	D_4	EZW	DCT
Mother	39.77	39.53	39.15	38.21	40.06	39.88	39.60	38.83	37.38
Silent	37.62	37.39	36.96	35.88	37.90	37.71	37.35	36.44	34.31
Hall	40.18	40.17	40.18	38.90	40.36	40.39	40.37	39.54	39.62
Children	29.46	29.12	28.74	27.76	29.56	29.40	28.98	28.20	25.08
Weather	36.47	36.28	35.77	34.31	37.25	36.98	36.65	35.66	30.40

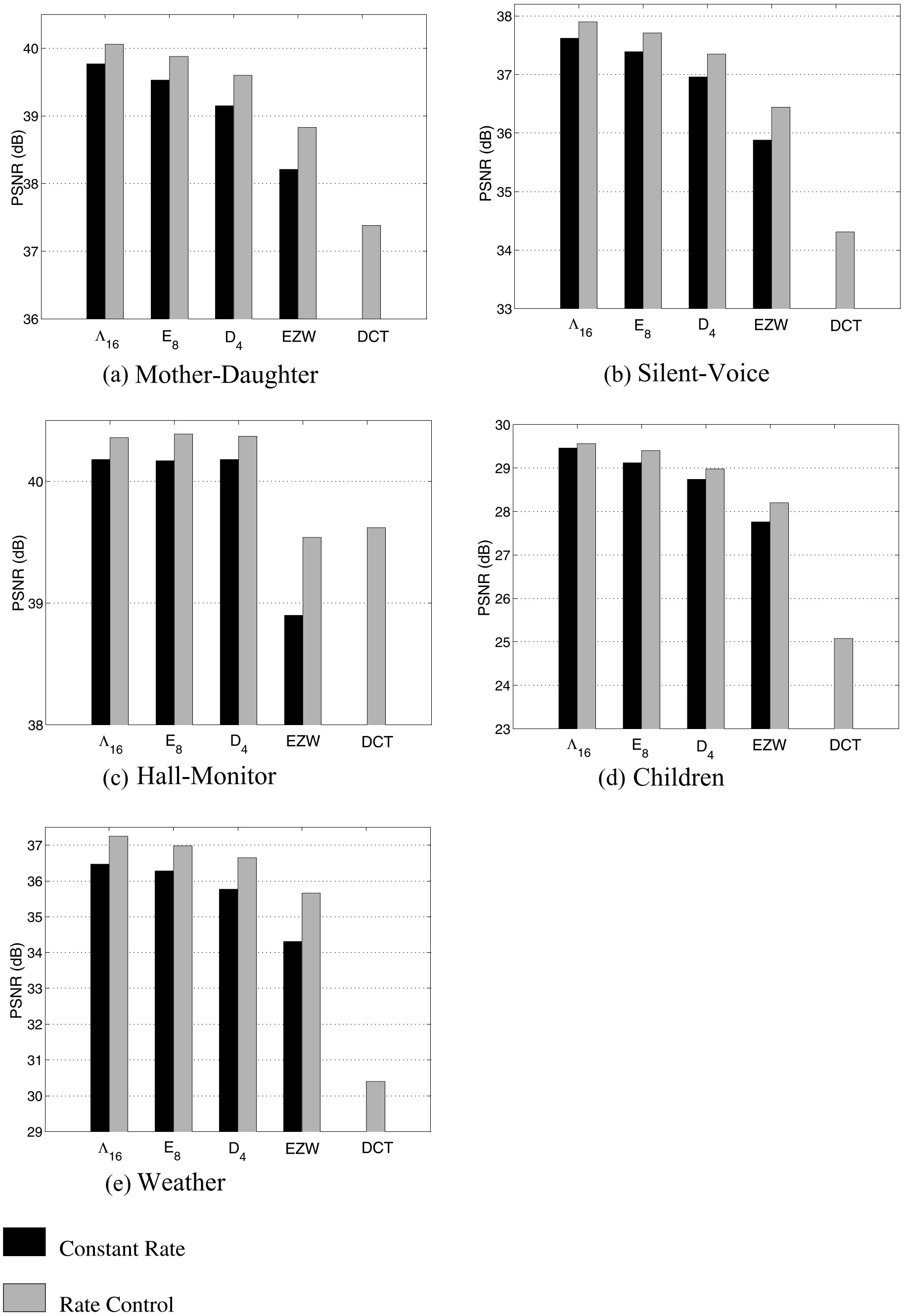

Fig. 3. Comparison between the average PSNR of the different MPEG-4 versions at a bit-rate of 64 kbit/s.

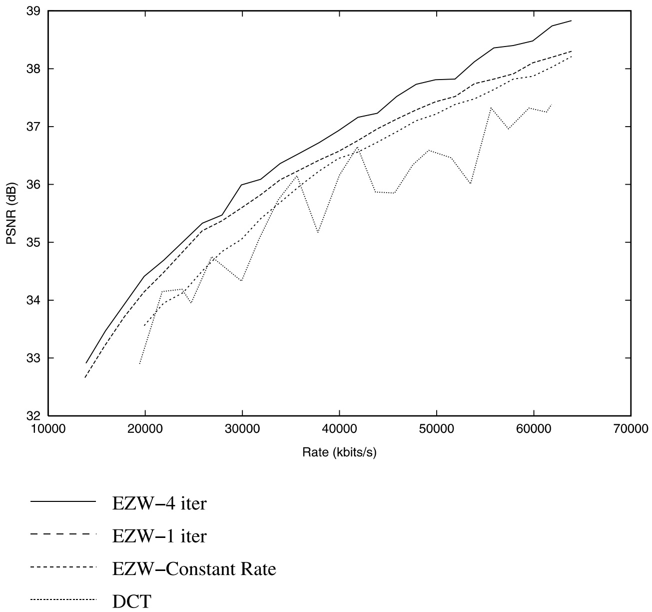

Fig. 4. Plot of PSNR versus rate for the Mother sequence.

In Fig. 4 there is a plot of the average PSNR of the Mother sequence for bit-rates ranging from 14 to 64 kbit/s. The coders employed are the MPEG-4 VM-8, the EZW coder with constant rate, and EZW using rate allocation with 1 and 4 iterations in a group of 40 frames (GOF = 40). From it, one can see that the rate-control strategy produces a consistent improvement for all rates. In fact, even the use of just one

Table 2
Average PSNR for different wavelet encoders and different rate control parameters for the sequence Mother at a bit-rate of 64 kbit/s

Constant rate		EZW	Λ_{16}	E_8	D_4
		38.21	39.77	39.53	39.15
GOF 5	1 iter	38.41	39.83	39.57	39.20
	2 iter	38.71	39.89	39.67	39.34
	3 iter	38.76	39.88	39.71	39.36
	4 iter	38.78	39.90	39.73	39.41
GOF 10	1 iter	38.35	39.84	39.58	39.25
	2 iter	38.69	39.93	39.66	39.37
	3 iter	38.69	39.97	39.78	39.40
	4 iter	38.85	39.96	39.70	39.43
GOF 40	1 iter	38.30	39.94	39.69	39.35
	2 iter	38.63	40.05	39.81	39.48
	3 iter	38.78	40.07	39.83	39.57
	4 iter	38.83	40.06	39.88	39.60

iteration, which does not take into consideration the frame dependency, produces a consistent improvement over the constant rate case. This is an indication of the adequacy of the piecewise-linear model for the *R–D* function of the frames. In addition, one can see that the case with 4 iterations performs consistently better than the one iteration case, confirming the effectiveness of the strategy used to deal with the frame dependency.

Table 2 and Fig. 5 show results for the proposed rate-control strategy using the embedded wavelet video encoders, applied to the sequence Mother. The rate control has been carried out in groups of 5, 10, and 40 frames (GOFs), and with 1, 2, 3, and 4

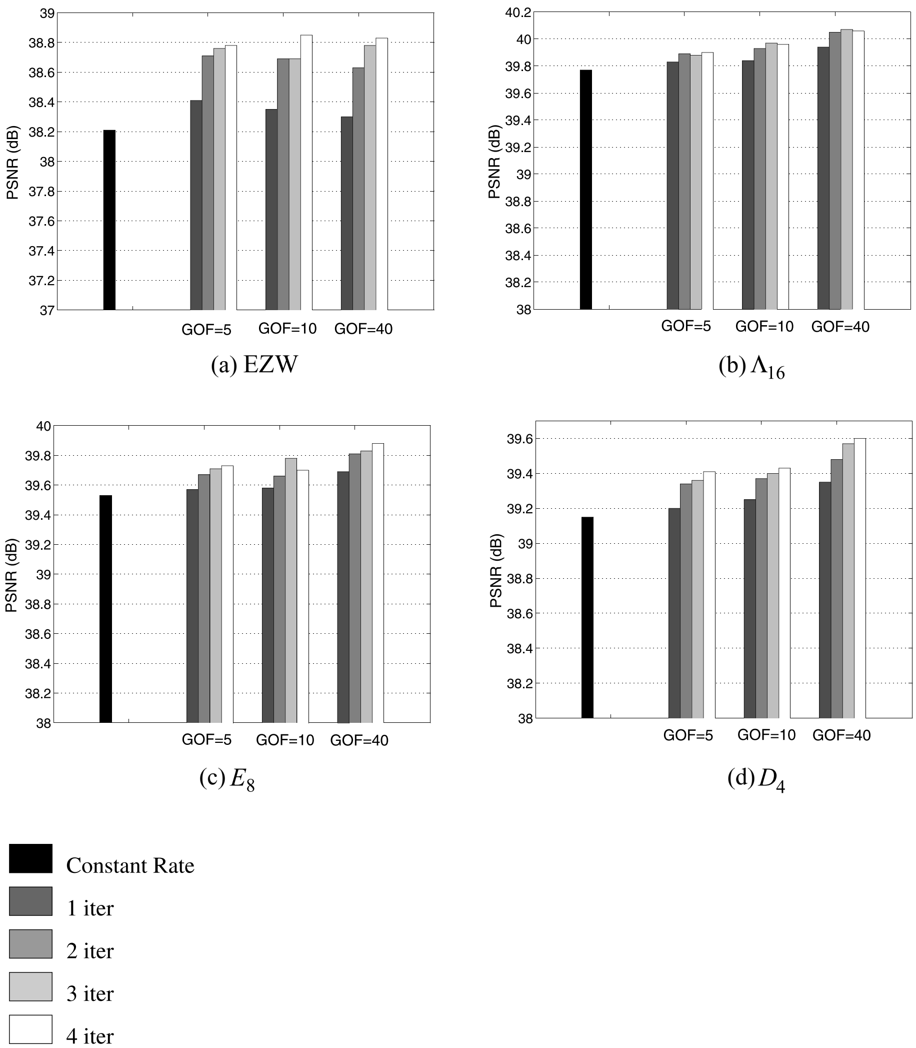

Fig. 5. Average PSNR for different wavelet encoders and different rate control parameters for the sequence Mother at a bit-rate of 64 kbit/s.

iterations (see Section 3). These results show that in most cases there is little improvement in going from 3 to 4 iterations of the algorithm. Also, one can observe that there is a small tendency for the average PSNR to grow with the GOF size. It is important to point out that, although a larger GOF size implies a large encoding/decoding delay, it does not imply an increase in complexity. It is also important to note that the complexity grows linearly with the number of iterations of the rate-control algorithm. In addition, one can see from Table 2 that the proposed rate-control strategy provides gains in PSNR in the vast majority of cases.

Table 3
The average PSNR of different sequences at different bit-rates for the encoders Λ_{16} and MPEG-4 VM-8

Sequence	Bit-rate	PSNR (dB)				
		RC	CR	DCT	RC–CR	RC–DCT
Mother–Daughter	20 K	35.27	33.98	32.90	1.29	2.37
Hall-Monitor	20 K	35.65	35.07	34.43	0.58	1.22
Silent-Voice	20 K	31.12	30.64	29.58	0.48	1.54
Weather	20 K	29.85	29.14	28.27	0.71	1.58
Mother–Daughter	48 K	38.85	38.50	36.34	0.35	2.51
Silent-Voice	48 K	36.22	36.00	32.41	0.22	3.81
Weather	48 K	35.14	34.58	29.78	0.56	5.36
Mother–Daughter	64 K	40.06	39.77	37.38	0.29	2.68
Hall-Monitor	64 K	40.36	40.18	39.62	0.18	0.74
Silent-Voice	64 K	37.90	37.62	34.31	0.28	3.59
Weather	64 K	37.25	36.47	30.40	0.78	6.85

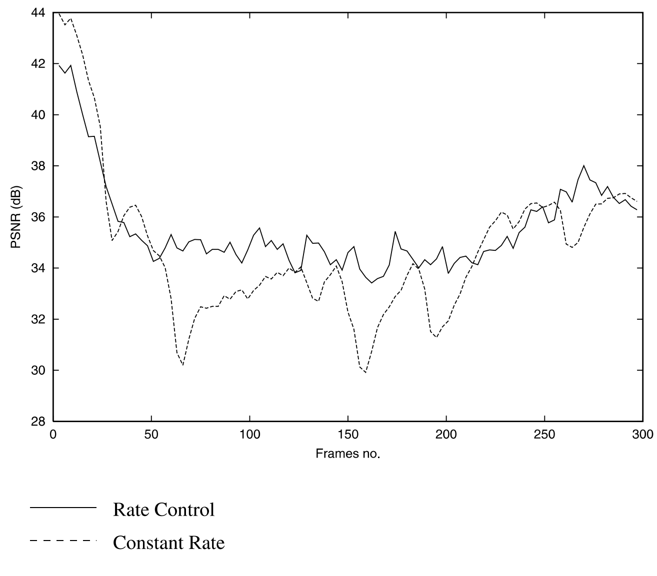

Fig. 6. Plot of PSNR versus frame number for the Mother sequence, coded using the Λ_{16} encoder.

Table 3 compares the results for the proposed rate-control strategy (RC) with the constant rate (CR), both using the Λ_{16} codebook, and the encoder MPEG-4 VM-8 (DCT) of different sequences at different bit-rates. In case of the rate control we use a GOF = 40 with 4 iterations.

Fig. 6 shows the PSNR plotted against frame number for the sequence Mother coded with the Λ_{16} encoder, both without rate control and with a rate control using a GOF of 40 and 4 iterations at a bit-rate of 20 kbit/s. It can be seen that the larger average PSNR achieved using the proposed rate control strategy (see Table 3) has not been obtained at the expense of a larger PSNR variation.

5. Conclusions

In this paper, we have proposed a novel rate-control strategy based on an MPEG-4 framework, for use in a class of embedded wavelet video encoders. We have shown that the *R–D* characteristics of the frame differences are better described by a piecewise-linear model than by an exponential model in such embedded wavelet encoders, and propose an algorithm for rate allocation. The frame dependency problem has been tackled by applying iteratively the proposed algorithm to a group of frames.

Simulation results have shown that the proposed rate-control strategy is effective, with the advantage of not requiring any arbitrary assumption about neither the rate $\times$ distortion characteristics of the encoders nor the frame dependency. Also, the results show that the use of a proper rate allocation strategy in embedded wavelet coders can produce results consistently superior to the ones of the MPEG-4 VM-8.

References

Bell, T.C., Cleary, J.G., Witten, I.H., 1990. Text Compression. Prentice-Hall PTR, Englewood Cliffs, NJ.

Conway, J.H., Sloane, N.J.A., 1988. Sphere Packings, Lattices and Groups. Springer, New York.

da Silva, E.A.B., Sampson, D.G., Ghanbari, M., 1996. A successive approximation vector quantizer for wavelet transform image coding. IEEE Transactions on Image Processing Special Issue on Vector Quantization 5, 299–310.

Everett III, H., 1963. Generalized Lagrange multiplier method for solving problems of optimum allocation of resources. Operations Research 11, 399–417.

Hsu, C.Y., Ortega, A., Reibman, A.R., 1997. Joint selection of source and channel rate for VBR video transmission under ATM policing constraints. IEEE Journal on Selected Areas in Communications 15, 1016–1028.

ISO/IEC JTC1/SC29/WG11, MPEG-4 video verification model version 8.0, July 1997.

ISO/IEC JTC1/SC29/WG1 (ITU/T SG28), JPEG2000 verification model 5.3, 1999.

Lan, T.-H., Tewfik, A.H., 1999. Multigrid embedding (MGE) image coding. In: Proceedings of the 1999 International Conference on Image Processing, Kobe.

Li, J., Cheng, P.-Y., Kuo, C.-C.J., 1997. Rate control for an embedded wavelet video coder. IEEE Transactions on Circuits and Systems for Video Technology 7, 696–702.

Lin, L.-J., Ortega, A., 1998. Bit-rate control using piecewise approximated rate–distortion characteristics. IEEE Transactions on Circuits and Systems for Video Technology 8, 446–459.

Mukherjee, D., Mitra, S.K., 1998. Vector set-partitioning with successive refinement Voronoi lattice VQ for embedded wavelet image coding. In: 1998 IEEE International Conference on Image Processing, Chicago, IL, October.

Otha, M., Nogaki, S., 1993. Hybrid picture coding with wavelet transform and overlapped motion compensated interframe prediction coding. IEEE Transactions on Signal Processing 41, 3416–3423.

Ramchandran, K., Ortega, A., 1998. Rate–distortion methods for image and video compression. IEEE Signal Processing 15, 23–50.

Ramchandran, K., Ortega, A., Vetterli, M., 1994. Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders. IEEE Transactions on Image Processing 3, 533–545.

Said, A., Pearlman, W.A., 1996. A new, fast and efficient image codec based on set partitioning in hierarchical trees. IEEE Transactions on Circuits and Systems for Video Technology 6, 243–250.

Sampson, D.G., da Silva, E.A.B., Ghanbari, M., 1995. Low bit rate video coding using wavelet vector quantization. IEE Proceedings Part-I: Vision, Image and Signal Processing 142, 141–148.

Shapiro, J.M., 1993. Embedded image coding using zero trees of wavelet coefficients. IEEE Transactions on Acoustics, Speech and Signal Processing 41, 3445–3462.

Shoham, Y., Gersho, A., 1988. Efficient bit allocation for an arbitrary set of quantizes. IEEE Transactions on Acoustics, Speech and Signal Processing 36, 1445–1453.

Taubman, D., 1999. High performance scalable image compression with EBCOT. In: 1999 IEEE International Conference on Image Processing, Kobe, October.