14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

MODEL-BASED BIT ALLOCATION BETWEEN WAVELET SUBBANDS AND MOTION INFORMATION IN MCWT VIDEO CODERS

Marie Andrée Agostini, Marc Antonini and Michel Barlaud

CReATIVe Research Group - I3S Laboratory, CNRS - University of Nice-Sophia Antipolis
2000 route des Lucioles - BP121 - 06903 Sophia Antipolis Cedex - France
Phone: 33(0)4.92.94.27.20 - Fax: 33(0).92.94.28.98 - {agostini, am, barlaud}@i3s.unice.fr

ABSTRACT

A precise motion estimation is necessary, in motion-compensated wavelet based video coders (MCWT), in order to minimize the wavelet coefficients energy. Nevertheless, a motion vectors field of high precision is expensive in binary resources compared to wavelet subbands and it is thus necessary to optimize the rate-distortion trade-off between motion information and wavelet coefficients. To this end, we have proposed in previous works to quantize the motion vectors using a scalable and open-loop lossy coder [1] and, to evaluate the impact of this lossy motion coding on the decoded sequence, we have established a theoretical distortion model of the motion coding error [2]. We present in this paper an approach to realize an optimal model-based bit-rate allocation between motion and wavelet subbands. This method is based on the total distortion model of coding error on several decomposition levels, including both motion information and subbands quantization noise. Experimental validations are satisfactory.

1. INTRODUCTION

Video compression is an essential element in multimedia technologies, like high-definition television (HDTV), video conferences, digital cinema and internet-related applications. Recent years have seen the impressive growth in performance of video coding algorithms, such as the hybrid coders of the latest standards MPEG4 and H.264/AVC [3, 4]. Moreover, wavelet-based video coders [5, 6] with motion-compensated $t+2D$ lifting schemes [7] have been developed and have shown their efficiency, while allowing scalability [8], and while remaining close to the performances of the hybrid coders [9]. However, improvement is still possible. Especially, it is necessary to improve the motion vectors processing and to optimize the rate-distortion trade-off between motion information and wavelet coefficients in order to ameliorate the video coding efficiency at low bit-rates. The problem of the processing of motion vectors was explored in [10, 11]. But, today, in most of the video coders, the rate-distortion trade-off for a given rate is optimized by varying the motion estimation parameters, which is not well-adapted at low bit-rates, makes the scalability difficult to obtain, and remains expensive in terms of CPU time. We have presented in a previous work [1] a new scalable approach of motion information coding which consists in introducing losses on motion vectors of high subpixelic precision, while optimizing a rate-distortion criterion. This open-loop method of lossy motion coding allows to reduce the motion cost with good coding performances at low bit-rates. Obviously, the introduction of loss on the motion has an impact on the decoded sequence.

In [2], in order to evaluate analytically this impact, we established, for a MCWT video coder on one decomposition level, a theoretical input/output distortion model of motion coding error. In this paper, we improve this model by introducing the wavelet coefficients quantization error and by generalizing the model to several temporal decomposition levels.

The proposed input/output distortion model is then used to dispatch in an optimal way the binary resources between the motion vectors and the temporal wavelet coefficients. For that purpose, we introduce in this paper a model-based bit allocation process. The proposed approach is based on the search of the minimum of a model-based criterion J and allows to choose the optimal values for the motion and subbands bit-rates at a given target bit-rate.

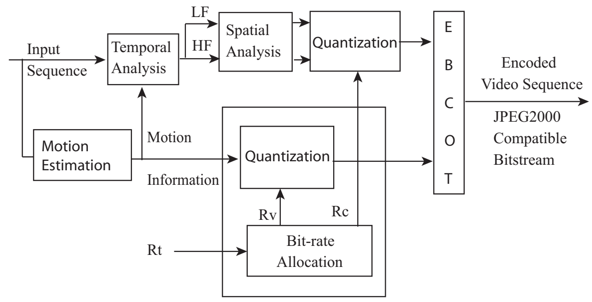

Figure 1: General structure of the encoder (R_c is the bit-rate of the subbands, R_v the one of the motion vectors and R_t the total bit-rate).

Section 2 briefly presents the general principle of the motion coding approach. Section 3 deals with the distortion model of the coding error on N temporal decomposition levels. We propose in section 4 an approach for the bit allocation between motion information and wavelet subbands. Finally, section 5 presents the results of the allocation process. Conclusion and further works are presented in section 6.

2. MOTION INFORMATION CODING IN A MCWT VIDEO CODER

Fully scalable, our video encoder is based on a lifted motion-compensated wavelet transform (figure 1). In this type of coder, the cost of the motion vectors can be very significant compared to the wavelet coefficients. To reduce this cost, we quantize with losses precise motion vectors, while controlling the rate-distortion trade-off between the original and the decoded sequence.

The vectors are encoded in open-loop [1], which allows full scalability and a good quality for the motion-compensated temporal analysis. Motion vectors are first quantized using an uniform scalar quantizer whose quantization step q controls the motion rate-distortion trade-off and, then, the quantized vectors are encoded using the MQ-Coder of an EBCOT encoder [9] and embedded in a JPEG2000-compliant bitstream.

This approach presents interesting performances on CIF and on SD sequences. Indeed, at low bit-rates, we show in previous work [1] that quantizing precise motion vectors with losses allows to obtain better coding performances than lossless motion coding.

3. DISTORTION MODEL OF MOTION AND SUBBANDS CODING ERROR

In a recent work [2], we established an expression for the distortion model on one temporal decomposition level including only the motion coding error. This result is briefly presented in section 3.1. Here, we include to this model the subbands quantization noise (section 3.2) and an extension to several decomposition levels is

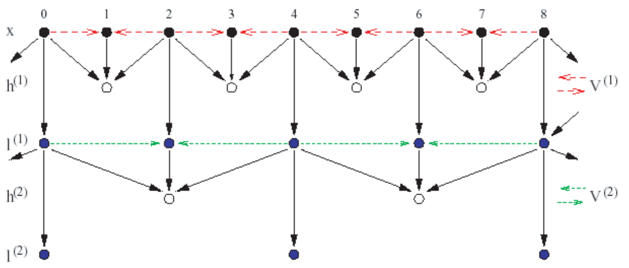

Figure 2: (2,0) lifting scheme for two temporal wavelet decomposition levels.

proposed in section 3.3. Finally, section 3.4 presents an experimental validation of the distortion model for the coding error.

3.1 Motion vectors distortion model

Our model is based on the computation of the input-output distortion D (MSE between input signal x and output signal $\widetilde{x}$) expressed by (with $\mathbf{Pn}$ the signal power and K the size of the sequence):

$$D = \frac{1}{K} \sum_{k=0}^{\frac{K}{2}-1} \left(\mathbf{Pn}\left(x_{2k+1} - \widetilde{x}_{2k+1}\right) + \mathbf{Pn}\left(x_{2k} - \widetilde{x}_{2k}\right) \right)$$

We assume that for one temporal wavelet decomposition level, the second part of this equation is equal to zero for a (2,0) lifting scheme (figure 2).

Thus, the distortion D, which expresses the reconstruction error due to the motion quantization on one decomposition level, can be written as [2]:

$$D = \frac{1}{2K} \sum_{k=0}^{\frac{K}{2}-1} \left[\mathbf{Pn}\left(x_{2k}\right) - \Gamma_{x_{2k}}\left(\eta_{B_{2k+1}}\right) + \mathbf{Pn}\left(x_{2k+2}\right) - \Gamma_{x_{2k+2}}\left(\eta_{F_{2k+1}}\right) \right]$$

where $\mathbf{Pn}\left(x_{2k}\right)$ and $\mathbf{Pn}\left(x_{2k+2}\right)$ are the power of images x_{2k} and x_{2k+2}; $\Gamma_{x_{2k}}\left(\eta_{B_{2k+1}}\right)$ and $\Gamma_{x_{2k+2}}\left(\eta_{F_{2k+1}}\right)$ the autocorrelation function of images x_{2k} and x_{2k+2}, with $\eta_{B_{2k+1}}$ and $\eta_{F_{2k+1}}$ the quantization error on the "Backward" and "Forward" motion vectors.

3.2 Including the subbands quantization noise

The high and low frequency temporal subbands are also quantized. We have thus to include the subbands quantization noise into the previous model.

If we denote by $\widehat{h}_k(\mathbf{p}) = h_k(\mathbf{p}) + \varepsilon_{h_k}(\mathbf{p})$ the quantized high frequency temporal subband, with $\varepsilon_{h_k}(\mathbf{p})$ the subband quantization noise, the (2,0) lifting scheme analysis equation for the high frequencies on one decomposition level becomes:

$$\widehat{h}_k\left(\mathbf{p}\right) = x_{2k+1}\left(\mathbf{p}\right) - \frac{1}{2}\left(x_{2k}^{B_{2k+1}} + x_{2k+2}^{F_{2k+1}}\right) + \varepsilon_{h_k}\left(\mathbf{p}\right) \qquad (1)$$

The synthesis equation is then given by:

$$\widetilde{\widehat{x}}_{2k+1}\left(\mathbf{p}\right) = \widehat{h}_k\left(\mathbf{p}\right) + \frac{1}{2}\left(\widetilde{\widehat{x}}_{2k}^{\widehat{B}_{2k+1}} + \widetilde{\widehat{x}}_{2k+2}^{\widehat{F}_{2k+1}}\right) \qquad (2)$$

with respectively $\widetilde{x}_k^{\widehat{B}_{k+1}} = \widetilde{x}_k(\mathbf{p} + \widehat{B}_{k+1}(\mathbf{p}))$ and $\widetilde{x}_k^{\widehat{F}_{k-1}} = \widetilde{x}_k(\mathbf{p} + \widehat{F}_{k-1}(\mathbf{p}))$ the "Backward" and "Forward" motion compensated pixels with quantized motion vectors.

We have to take also into account the low frequency temporal subbands coding noise introduced on the images x_{2k} and x_{2k+2} and denoted respectively by ε_{2k} and ε_{2k+2}:

$$\widehat{x}_{2k}(\mathbf{p}) = x_{2k}(\mathbf{p}) + \varepsilon_{2k}(\mathbf{p}) \text{ and } \widehat{x}_{2k+2}(\mathbf{p}) = x_{2k+2}(\mathbf{p}) + \varepsilon_{2k+2}(\mathbf{p})$$

Therefore, the synthesized pixel with quantized motion vectors ("Backward" and "Forward") and with coding error on the low frequency subbands can be written as:

$$\widetilde{\widehat{x}}_{2k}^{\widehat{B}_{2k+1}} = \widetilde{x}_{2k}^{\widehat{B}_{2k+1}} + \widetilde{\varepsilon}_{2k}^{\widehat{B}_{2k+1}} \text{ and } \widetilde{\widehat{x}}_{2k+2}^{\widehat{F}_{2k+1}} = \widetilde{x}_{2k+2}^{\widehat{F}_{2k+1}} + \widetilde{\varepsilon}_{2k+2}^{\widehat{F}_{2k+1}}$$

By using equations (1) and (2), we obtain the following relation:

$$x_{2k+1}(\mathbf{p}) - \widetilde{\widehat{x}}_{2k+1}(\mathbf{p}) = \frac{1}{2}\Big((x_{2k}^{B_{2k+1}} - \widetilde{x}_{2k}^{\widehat{B}_{2k+1}} - \widetilde{\varepsilon}_{2k}^{\widehat{B}_{2k+1}}) + (x_{2k+2}^{F_{2k+1}} - \widetilde{x}_{2k+2}^{\widehat{F}_{2k+1}} - \widetilde{\varepsilon}_{2k+2}^{\widehat{F}_{2k+1}}) \Big) - \varepsilon_{h_k}$$

The distortion D becomes therefore:

$$D = \frac{1}{K} \sum_{k=0}^{\frac{K}{2}-1} \left[\frac{1}{4}(\mathbf{Pn}(x_{2k}^{B_{2k+1}} - \widetilde{x}_{2k}^{\widehat{B}_{2k+1}} - \widetilde{\varepsilon}_{2k}^{\widehat{B}_{2k+1}}) + \mathbf{Pn}(x_{2k+2}^{F_{2k+1}} - \widetilde{x}_{2k+2}^{\widehat{F}_{2k+1}} - \widetilde{\varepsilon}_{2k+2}^{\widehat{F}_{2k+1}})) + \mathbf{Pn}(\varepsilon_{h_k}) \right]$$

By introducing the scalar products defined for the "Backward" vectors as (with similar notation for the "Forward" vectors):

$$\left\langle x_{2k}^{B_{2k+1}}, \widetilde{x}_{2k}^{\widehat{B}_{2k+1}} \right\rangle = \frac{1}{NM} \sum_{\mathbf{p}} x_{2k}^{B_{2k+1}} \times \widetilde{x}_{2k}^{\widehat{B}_{2k+1}}$$

and by developing, we obtain:

$$\begin{aligned} D = \frac{1}{2K} \sum_{k=0}^{\frac{K}{2}-1} [& \frac{1}{2}\mathbf{Pn}(x_{2k}^{B_{2k+1}}) + \frac{1}{2}\mathbf{Pn}(\widetilde{x}_{2k}^{\widehat{B}_{2k+1}}) + \frac{1}{2}\mathbf{Pn}(\widetilde{\varepsilon}_{2k}^{\widehat{B}_{2k+1}}) \\ & - \langle x_{2k}^{B_{2k+1}}, \widetilde{x}_{2k}^{\widehat{B}_{2k+1}} \rangle - \langle x_{2k}^{B_{2k+1}}, \widetilde{\varepsilon}_{2k}^{\widehat{B}_{2k+1}} \rangle - \langle \widetilde{x}_{2k}^{B_{2k+1}}, \widetilde{\varepsilon}_{2k}^{\widehat{B}_{2k+1}} \rangle \\ & + \frac{1}{2}\mathbf{Pn}(x_{2k+2}^{F_{2k+1}}) + \frac{1}{2}\mathbf{Pn}(\widetilde{x}_{2k+2}^{\widehat{F}_{2k+1}}) + \frac{1}{2}\mathbf{Pn}(\widetilde{\varepsilon}_{2k+2}^{\widehat{F}_{2k+1}}) + 2\mathbf{Pn}(\varepsilon_{h_k}) \\ & - \langle x_{2k+2}^{F_{2k+1}}, \widetilde{x}_{2k+2}^{\widehat{F}_{2k+1}} \rangle - \langle x_{2k+2}^{F_{2k+1}}, \widetilde{\varepsilon}_{2k+2}^{\widehat{F}_{2k+1}} \rangle - \langle \widetilde{x}_{2k+2}^{F_{2k+1}}, \widetilde{\varepsilon}_{2k+2}^{\widehat{F}_{2k+1}} \rangle \Big]. \end{aligned}$$

Assuming that the signal and the different quantization noises are mutually decorrelated, the crossed scalar products $\left\langle x_{2k}^{B_{2k+1}}, \widetilde{\varepsilon}_{2k}^{\widehat{B}_{2k+1}} \right\rangle$, $\left\langle \widetilde{x}_{2k}^{B_{2k+1}}, \widetilde{\varepsilon}_{2k}^{\widehat{B}_{2k+1}} \right\rangle$, $\left\langle x_{2k+2}^{F_{2k+1}}, \widetilde{\varepsilon}_{2k+2}^{\widehat{F}_{2k+1}} \right\rangle$ and $\left\langle \widetilde{x}_{2k+2}^{F_{2k+1}}, \widetilde{\varepsilon}_{2k+2}^{\widehat{F}_{2k+1}} \right\rangle$ are equal to zero. Then, high bit-rate assumptions involve:

$$\begin{aligned} \mathbf{Pn}\left(\widetilde{x}_{2k}^{\widehat{B}_{2k+1}}\right) &\approx \mathbf{Pn}\left(x_{2k}^{B_{2k+1}}\right) \approx \mathbf{Pn}\left(x_{2k}\right) \\ \mathbf{Pn}\left(\widetilde{x}_{2k+2}^{\widehat{F}_{2k+1}}\right) &\approx \mathbf{Pn}\left(x_{2k+2}^{F_{2k+1}}\right) \approx \mathbf{Pn}\left(x_{2k+2}\right), \end{aligned}$$

and the distortion D can be simplified as:

$$\begin{aligned} D = \frac{1}{2K} \sum_{k=0}^{\frac{K}{2}-1} \Big[& \mathbf{Pn}\left(x_{2k}\right) - \Gamma_{x_{2k}}\left(\eta_{B_{2k+1}}\right) + \frac{1}{2}\mathbf{Pn}\left(\widetilde{\varepsilon}_{2k}^{\widehat{B}_{2k+1}}\right) \\ & + \mathbf{Pn}\left(x_{2k+2}\right) - \Gamma_{x_{2k+2}}\left(\eta_{F_{2k+1}}\right) + \frac{1}{2}\mathbf{Pn}\left(\widetilde{\varepsilon}_{2k+2}^{\widehat{F}_{2k+1}}\right) + 2\mathbf{Pn}(\varepsilon_{h_k}) \Big] \end{aligned}$$

with, $\Gamma_{x_{2k}}(\eta_{B_{2k+1}}) = \left\langle x_{2k}^{B_{2k+1}}, \widetilde{x}_{2k}^{\widehat{B}_{2k+1}} \right\rangle$ (similar notations for image x_{2k+2}).

Moreover, thanks to the high-rate assumption, we also have:

$$\begin{aligned} \mathbf{Pn}\left(\widetilde{\varepsilon}_{2k}^{\widehat{B}_{2k+1}}\right) &\approx \mathbf{Pn}\left(\varepsilon_{2k}^{B_{2k+1}}\right) \approx \mathbf{Pn}\left(\varepsilon_{2k}\right) \\ \mathbf{Pn}\left(\widetilde{\varepsilon}_{2k+2}^{\widehat{F}_{2k+1}}\right) &\approx \mathbf{Pn}\left(\varepsilon_{2k+2}^{F_{2k+1}}\right) \approx \mathbf{Pn}\left(\varepsilon_{2k+2}\right) \end{aligned}$$

The distortion D, which represents the reconstruction error due to the motion and the subbands quantization on one decomposition level, can thus be expressed as:

$$D = \frac{1}{2K}\sum_{k=0}^{\frac{K}{2}-1}\Big[\mathbf{Pn}(x_{2k}) - \Gamma_{x_{2k}}(\eta_{B_{2k+1}}) + \frac{1}{2}\mathbf{Pn}(\varepsilon_{2k})$$
$$+\mathbf{Pn}(x_{2k+2}) - \Gamma_{x_{2k+2}}(\eta_{F_{2k+1}}) + \frac{1}{2}\mathbf{Pn}(\varepsilon_{2k+2}) + 2\mathbf{Pn}(\varepsilon_{h_k})\Big]$$

3.3 Generalization to N decomposition levels

The expression of the input/output distortion D can be easily generalized to several temporal decomposition levels (with N the number of levels):

$$D = \frac{1}{2K}\sum_{n=0}^{N-1}\sum_{k=0}^{\frac{K}{2^{N-n}}-1}\Big[\mathbf{Pn}(x_{2^{N-n}k}) - \Gamma_{x_{2^{N-n}k}}(\eta_{B_{2^{N-n}k+2^{N-n-1}}})$$
$$+\mathbf{Pn}(x_{2^{N-n}k+2^{N-n}}) - \Gamma_{x_{2^{N-n}k+2^{N-n}}}(\eta_{F_{2^{N-n}k+2^{N-n-1}}})\Big]$$
$$+\frac{1}{K}\Big(\frac{1}{2^N}\mathbf{Pn}(\varepsilon_{l_k}^{(N)}) + \sum_{i=1}^{N}\frac{1}{2^i}\mathbf{Pn}(\varepsilon_{h_k}^{(i)})\Big) \quad (3)$$

with $l_k^{(N)}$ the low frequency subband and $h_k^{(i)}$ the high frequency subband at the i^{th} decomposition level.

3.4 Validation of the distortion model

To validate the proposed model, we compare the results obtained experimentaly for the input-output distortion of the coder with those obtained by applying the theoretical distortion formula (3). In figure 3, we present the results of the distortion for the CIF sequence "Foreman" on two decomposition levels with a (2,0) lifting scheme at 500 Kbps, for different motion bit-rates R_v, where both quarter-pixel motion vectors and wavelet subbands are quantized. Beside, the theoretical curve is approximated using Smoothing-B splines.

These results show that the theoretical and experimental curves are close and follow the same progression. We observe no more than 5 % on average for the errors between theory and experimentation. Therefore, the proposed theoretical distortion model for the coding error provides a good approximation.

4. MODEL-BASED BIT-RATE ALLOCATION

The input/output distortion model being defined by the previous equation, the problem is now to find the optimal bit-rates for the motion vectors and for the temporal wavelet coefficients in order to minimize the total distortion. The proposed method is described in the following section.

4.1 Optimization problem

The problem (P) of bit allocation between motion information and wavelet coefficients can be formulated as:

$$(P)\begin{cases}\min\limits_{R_v,R_c} D(R_v,R_c)\\ \text{under constraint } R_v + R_c = R_t\end{cases}$$

where R_v is the motion bit-rate, $R_c = \sum_i \frac{1}{2^i}R_{c_i}$ the wavelet coefficients bit-rate, and R_t the target bit-rate. In order to solve problem (P), we use the Lagrange multipliers and we introduce the following convex criterion:

$$J_\lambda(R_v,R_c) = D(R_v,R_c) + \lambda(R_v + R_c - R_t).$$

The optimal solution (R_v^*,R_c^*) is then obtained by minimizing $J_\lambda(R_v,R_c)$. For fixed λ, we can find the optimal solution by solving

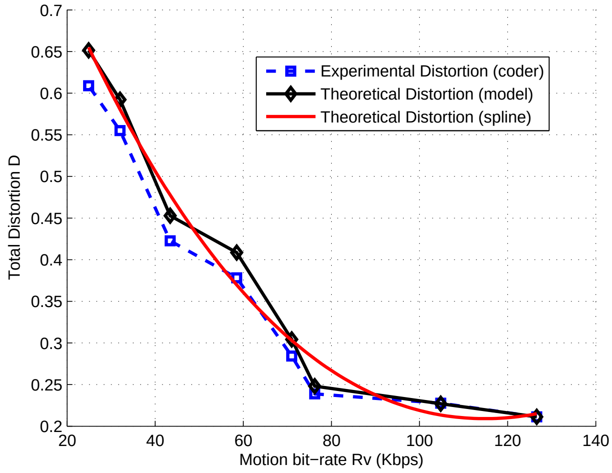

Figure 3: Validation of the theoretical distortion model (both quarter-pixel vectors and wavelet subbands are quantized) on the sequence "Foreman" at 500 Kbps, on two decomposition levels.

the following system:

$$\begin{cases}R_v^*(\lambda,R_c) = \arg\min\limits_{R_v} J_\lambda(R_v,R_c) \text{ for all } R_c\\ R_c^* = \arg\min\limits_{R_c} J_\lambda(R_v,R_c)|_{R_v=R_v^*(\lambda,R_c)}\,.\end{cases}$$

Varying the value of λ in an efficient way (for example by using dichotomy) permits to meet the constraint.

4.2 Bit allocation algorithm

The proposed allocation algorithm works as follows:

1. $\lambda = \lambda_{init}$
2. For each value of R_c, find the value $R_v^*(\lambda,R_c)$ that minimizes the criterion $J_\lambda(R_v,R_c)$
3. Find the value R_c^* that minimizes the criterion $J_\lambda(R_v^*(\lambda,R_c),R_c)$
4. If $R_v^* + R_c^* = R_t$ then stop, else change the value of λ and go to step 2.

5. EXPERIMENTAL RESULTS

We present the results of the bit allocation algorithm for the sequence CIF "Foreman" with quarter-pixel motion vectors and the sequence SD "City" with pixel motion vectors, on two decomposition levels with a (2,0) lifting scheme. Figure 4 shows typical curves for J_λ, for sequence "City". Curve 4(a) represents, for a fixed λ, the criterion $J_\lambda(R_v,R_c)$ as a function of the motion bit-rate R_v for one subbands bit-rate R_c: when J_λ is minimum, we deduce the optimal motion bit-rate R_v^* for this R_c; whereas curve 4(b) represents, also for a fixed λ, the criterion $J_\lambda(R_v^*(\lambda,R_c),R_c)$ as a function of the subbands bit-rate R_c. According to the value of λ, we have thus the optimal couple (R_v^*,R_c^*).

Figure 5presents, for the two sequences, the performances comparison between the curve obtained by applying this bit allocation algorithm (with optimal values R_c^* and R_v^* presented in table 1 for each total rate R_t given by the value of λ, for "City") and the curve without allocation (triangular markers: for same target bit-rates R_t, R_v = 126.6 Kbps for "Foreman" and R_v = 160.4 Kbps for "City", lossless motion coding). These curves represent the input-output distortion as a function of the target bit-rate R_t. It appears that using the optimal bit-rates allows to improve the quality of the decoded sequence. These results show that the proposed approach of optimal bit-rate allocation gives satisfactory results.

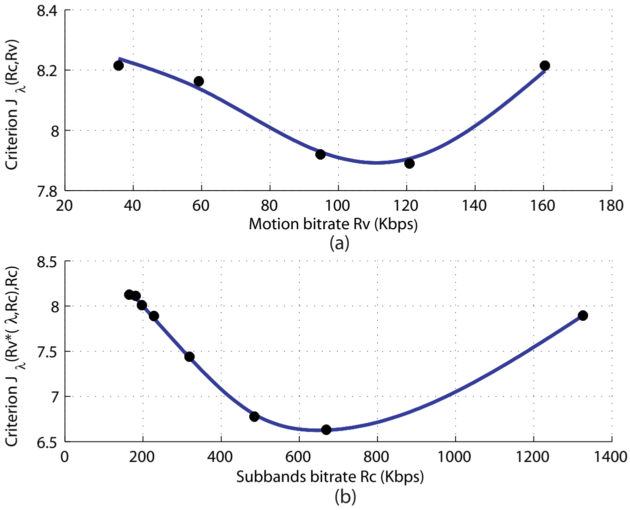

Figure 4: Bit-rate allocation on “City”, pixel motion vectors, two temporal decomposition levels at $\lambda = 0.004$: (a) Criterion $J_\lambda(R_v, R_c)$ function of R_v with $R_c = 230$ Kbps; (b) Criterion $J_\lambda(R_v^*(\lambda, R_c), R_c)$ function of R_c.

R_t	220	270	300	350	480	640	830
R_v^*	59.2	89	94.8	120.8	160.4	160.4	160.4
R_c^*	160	180	200	230	320	480	670

Table 1: Optimal bit-rates R_v^* and R_c^* for different target bit-rates R_t for “City” on two decomposition levels, with pixel motion vectors (bit-rates in Kbps).

6. CONCLUSION

In this paper, we have proposed a theoretical model for the input/output quantization distortion where both motion vectors and temporal wavelet coefficients are lossy quantized in the framework of MCWT video coder. This model has been established on several temporal decomposition levels and takes into account the motion quantization error as well as the wavelet coefficients quantization error.

Furthermore, using this model, we have derived an efficient bit allocation algorithm to dispatch the binary resources between motion and wavelet coefficients. Indeed, it is well known that optimizing the rate-distortion trade-off between motion information and wavelet coefficients is a crucial problem. The proposed model-based approach permits to find analytically, for a target bit-rate, which optimal rates to choose for the motion vectors and the wavelet coefficients in order to have a minimal distortion at decoding.

This feature decreases the computational complexity and cost of the coding and improves the coder performances. Experimental results show an improvement in term of MSE for the decoded video compared to a standard approach without optimal bit allocation.

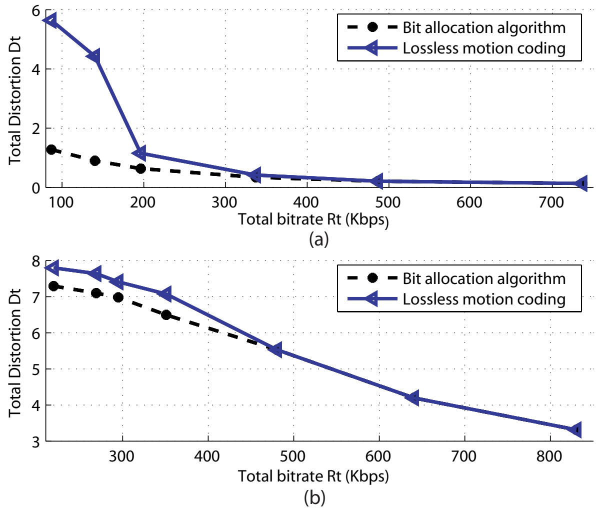

Figure 5: Performances comparison, (a) Sequence ‘Foreman”, quarter-pixel motion vectors, two temporal decomposition levels; (b) Sequence ‘City”, pixel motion vectors, two temporal decomposition levels: Approach with bit allocation algorithm compared to approach without allocation and with lossless motion coding.

REFERENCES

[1] M. A. Agostini, T. André, M. Antonini, and M. Barlaud, “Scalable motion coding for video coders with lifted MCWT,” in *Proc. of International Workshop on Very Low Bit-rate Video-coding (VLBV)*, Costa Rei, Sardinia - Italy, Sept. 2005.

[2] M. A. Agostini, T. André, M. Antonini, and M. Barlaud, “Modeling the motion coding error for MCWT video coders,” in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toulouse, France, May 2006.

[3] MPEG4 Video Group, *Coding of audio-visual objects: Video, ISO/IEC JTC1/SC29/WG11 N2202*, Mar. 1998.

[4] Gary J. Sullivan, Pankaj Topiwala, and Ajay Luthra, “The h264 / avc advanced video coding standard : Overview and introduction to the fidelity range extensions,” in *SPIE Conference on Applications of Digital Image Processing XXVII, special Session on Advances in the New Emerging Standard : H264 / AVC*, Aug. 2004.

[5] S. J. Choi and J.W. Woods, “Motion-compensated 3-D subband coding of video,” *IEEE Transaction on Image Processing*, vol. 8, no. 2, pp. 155–167, Feb. 1999.

[6] S. Cho and W.A. Pearlman, “A full-featured, error-resilient; scalable wavelet video codec based on the set partitioning in hierarchical trees (SPIHT) algorithm,” *IEEE Transaction on Circuits and Systems for Video Technology*, vol. 12, pp. 170–171, Mar. 2002.

[7] J. Viéron, C. Guillemot, and S. Pateux, “Motion compensated 2D+t wavelet analysis for low rate fgs video compression,” in *Proc. of Tyrrhenian Intern. Workshop on Digital Comm.*, Capri, Italy, Sept. 2002.

[8] G. Pau, C. Tillier, B. Pesquet-Popescu, and H. Heijmans, “Motion compensation and scalability in lifting-based video coding,” *EURASIP Signal Processing: Image Communication, special issue on Wavelet Video Coding*, pp. 577–600, Aug. 2004.

[9] A. Secker and D. Taubman, “Lifting-based invertible motion adaptive transform (LIMAT) framework for highly scalable video compression,” *IEEE Transaction on Image Processing*, vol. 12, no. 12, pp. 1530–1542, Dec. 2003.

[10] L. A. Da Silva Cruz and J. W. Woods, “Adaptive motion vector quantization for video coding,” in *IEEE Intern. Conf. on Image Processing*, Vancouver, Canada, Sept. 2000, pp. 867–870, vol 2.

[11] R. L. Joshi, T. R. Fischer, and R. H. Bamberger, “Lossy encoding of motion vectors using entropy-constrained vector quantization,” in *IEEE Intern. Conf. on Image Processing*, Washington, USA, 1995, pp. 3109, vol 3.