High Performance Coding using a Model-based Bit Allocation with EBCOT

C. Parisot, M. Antonini and M. Barlaud

Laboratoire I3S - UMR 6070 (CNRS/University of Nice-Sophia Antipolis)
Les Algorithmes - 2000, route des Lucioles - BP 121 - F-06903 Sophia Antipolis Cedex - France
{parisot, am, barlaud}@i3s.unice.fr

ABSTRACT

The bit allocation procedure of JPEG 2000 is based on a rate-distortion curve directly computed from quantized and encoded wavelet coefficients. Thus, JPEG 2000 exploits at best the efficiency of its bit plane context-based arithmetic coder. However, encoding data which will not be saved in the final bit stream introduces complexity and JPEG 2000 bit allocation requires a lot of tests. We propose a new compression scheme using a low complexity model-based bit allocation followed by scalar quantizers with optimized deadzone sizes and the EBCOT bit plane coder to entropy encode the quantized subbands. The resulting compression scheme provides the same performances as JPEG 2000 with less complexity and simpler hardware implementation.

1 Introduction

Image compression has been under study for many years and the new JPEG 2000 standard [1] can be considered as the state of the art. The bit allocation procedure of JPEG 2000 is based on a rate-distortion curve directly computed from the image to compress. Thus, JPEG 2000 can exploit at best the efficiency of its bit plane context-based arithmetic coder. However, the method has two drawbacks. The first one is the need to encode data to get the rate-distortion curve used in the bit allocation process. Therefore, JPEG 2000 encodes more data than will be saved in the final codestream. The second one is that the set of possible quantization steps in a subband is limited to the product of a pre-defined quantization step times powers of two. When all the bit planes are encoded, the resulting quantizer is a uniform scalar quantizer with deadzone twice as wide as the quantization step. It has been proved that such a deadzone size is not optimal [2].

In this paper, we propose an alternative compression method using a model-based bit allocation procedure. The scalar quantizers are applied in each subband and EBCOT's bit plane arithmetic coder encodes the quantized subbands. Our bit allocation is computed from subband statistics. Once it is computed, both quantization and entropy coding can be performed concurrently for all subbands. Thus, quantizers are no more chosen from a finite set of possible ones and the deadzone width can be adapted to the subband bitrate. The complexity is reduced since only usefull data are encoded. The proposed method architecture is composed of independent procedures, which is welcome for hardware implementation.

In section 2, we present the structure of the proposed compression scheme and its different components. Section 3 deals with the model-based bit allocation method and section 4 shows some experimental results.

2 Compression scheme

Fig. 1 shows the flow chart for our new compression scheme. First, we perform a wavelet decomposition of the signal. The second step consists in using statistical information of the subbands to compute an optimized model-based bit allocation. In the third step, we perform the wavelet coefficients quantization. Finally, the quantized wavelet coefficients are lossless entropy encoded. The first difference with EBWIC [3] is that the bit allocation has been updated to take account of scalar quantizers with optimized deadzone sizes [2]. The second one is the use of EBCOT's [4] bit plane encoding rather than Run-Length and Huffman. Thus, the bit allocation and the quantization are performed before the encoding process. This avoids to encode data which will not be kept in the final bit stream and lots of tests can be saved since the encoding has to be performed for all the bit planes. An important aspect of our approach is that once the bit allocation has been computed, the quantization and encoding of subbands is done separately and can be performed simultaneously on separate ASICs for hardware implementations.

Let us now describe more precisely each block of this compression scheme.

2.1 Wavelet transform

We have selected the 9-7 biorthogonal wavelet transform [5] which is known to be almost orthogonal and gives the best results for dyadic sampled images [6]. We use a three level decomposition.

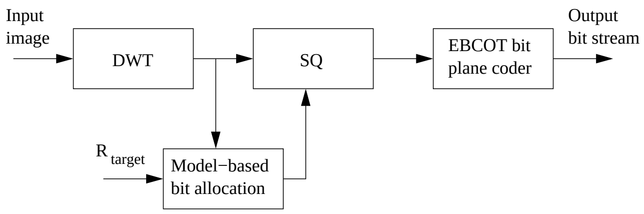

Figure 1: Flow chart of the compression scheme

2.2 Scalar quantization

The quantization is performed by uniform threshold quantizers with optimized deadzone sizes. The ratio $\frac{z}{q}$ between the zero quantization bin size z and the size q of all other quantization bins depends on the source probability density function and q. See [2] for more details on the computation of the optimized deadzone size. Uniform threshold quantizers with deadzone provide higher coding performance at the cost of a small increase of complexity. The decoded wavelet coefficient is the centroid of its quantization bin according to the probability density function model of the subband. The parameters z and q of each quantizer are computed in the bit allocation procedure.

2.3 Entropy coder

The lossless entropy coding of the quantized wavelet coefficients is supplied by EBCOT [4]. For each subband, EBCOT produces an embedded bit stream based on the following three passes for each bit plane:

1. The significance pass: This pass encodes the bits of the coefficients not yet found to be significant which are predicted to become significant during the processing of the current bit plane. If the bit to encode is one, the coefficient becomes significant and its sign is also encoded in this pass.

2. The refinement pass: This pass encodes the bits of the coefficients that were found to be significant in previous bit planes.

3. The cleanup pass: This pass encodes the bits of the coefficients not processed in the significance and refinement passes. If the bit to encode is one, the coefficient becomes significant and its sign is also encoded in this pass.

All the bits are encoded using a context based arithmetic coder. The values for the probability models in each context have been taken from JPEG 2000 [1]. We do not have split subbands into blocks in our implementation. As the rate-distortion optimization has been performed in the model-based bit allocation procedure, this coder is applied to all the bit planes.

2.4 Bit allocation

The bit allocation procedure is responsible for the final rate-distortion performance of the method. When it is computed before the quantization and encoding steps, it must use rate and distortion models that fit well with the quantizer and entropy coder performances. EBCOT uses a context based arithmetic coder. Thus, it is supposed to produce a bitrate lower than the entropy of the quantized wavelet coefficients. However, as the output bitrate depends on the spatial distribution of the wavelet coefficients, we do not know how to predict it efficiently. Therefore, the bit allocation proposed in the following section makes the assumption of an entropy coder.

3 Bit allocation

3.1 General purpose

The purpose of the bit allocation is to determine the quantizers in each subband which minimize the average mean squared error for a target bitrate [7]. The subband quantizers are scalar quantizers with deadzone. They are defined by the size of their zero quantization bin z and the size of all other quantization bins q. Therefore, the solution of the bit allocation problem is obtained by the minimization of the following criterion using Lagrange operators:

$$J = \sum_{i=1}^{\#SB} \Delta_i \pi_i \sigma_{Q_i}^2(z_i, q_i) + \lambda\left(\sum_{i=1}^{\#SB} a_i R_i(z_i, q_i) - R_T\right) \quad (1)$$

where $\sigma_{Q_i}^2(z_i, q_i)$ and $R_i(z_i, q_i)$ are respectively the mean squared error and the bitrate produced in the i^{th} subband; a_i denotes the weight of subband i in the total bitrate; $\{\pi_i\}$ are weights used to take account of the non-orthogonality of the filter bank and $\{\Delta_i\}$ are optional weights for frequency selection. $\#SB$ denotes the number of subbands while R_T is the target bitrate.

3.2 Rate and distortion models

The only way to compute an efficient bit allocation without pre-quantizing subbands is to accurately model the distribution $p(x)$ of the wavelet coefficients and use theoretical models for both distortion and bitrate. This distribution can be approximated with generalized gaussians [5]. We have

$$p(x) = ae^{-|bx|^{\alpha}}$$

with $b = \frac{1}{\sigma}\sqrt{\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)}}$ and $a = \frac{b\alpha}{2\Gamma(1/\alpha)}$. To compute the bitrate produced by the deadzone scalar quantizer $\{z, q\}$, we use the probability of the quantization level m. We have $\Pr(m) = \Pr(-m) = \int_{\frac{z}{2}+(|m|-1)q}^{\frac{z}{2}+|m|q} p(x)\, dx$ for $m \neq 0$ and $\Pr(0) = \int_{-\frac{z}{2}}^{+\frac{z}{2}} p(x)\, dx$.

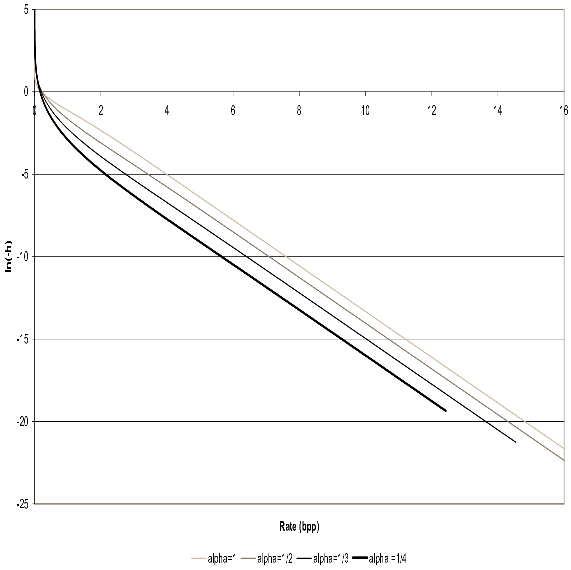

Figure 2: $ln(-h)$ versus bitrate

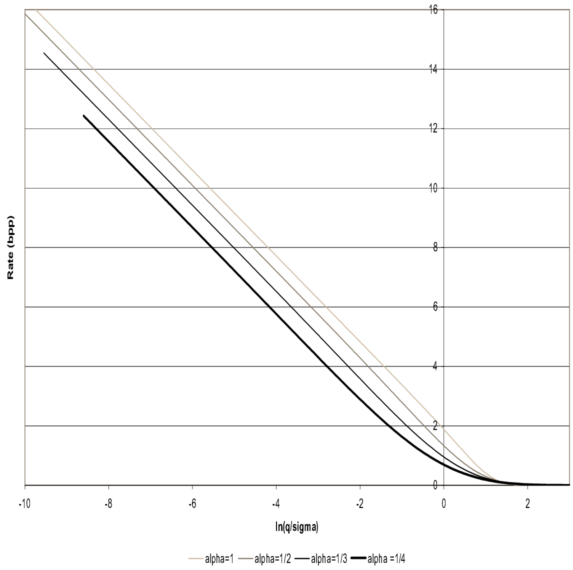

Figure 3: Bitrate versus $\ln\left(\frac{q}{\sigma}\right)$

Then, the bitrate is approximated by the entropy of the output quantization levels:

$$R = -\sum_{m=-\infty}^{+\infty} \Pr(m) \log_2 \Pr(m)$$

According to [8], the best decoding value for the quantization level m is the centroid of its quantization bin $\widehat{x_m} = \frac{\int_{\frac{z}{2}+(|m|-1)q}^{\frac{z}{2}+|m|q} xp(x)dx}{\Pr(m)}$ for $m \neq 0$ and $\widehat{x_0} = 0$. Then, the mean squared error is

$$\begin{aligned} \sigma_Q^2 &= \int_{-\frac{z}{2}}^{+\frac{z}{2}} x^2 p(x)\, dx \\ &\quad + 2\sum_{m=1}^{+\infty} \int_{\frac{z}{2}+(|m|-1)q}^{\frac{z}{2}+|m|q} \left(x-\widehat{x_m}\right)^2 p(x)\, dx \end{aligned}$$

3.3 Solution

For generalized gaussian distributions, (1) can be written as

$$J = \sum_{i=1}^{\#SB} \Delta_i \pi_i \sigma_i^2 D_i\left(\frac{z_i}{\sigma_i}, \frac{q_i}{\sigma_i}\right) + \lambda\left(\sum_{i=1}^{\#SB} a_i R_i\left(\frac{z_i}{\sigma_i}, \frac{q_i}{\sigma_i}\right) - R_T\right) \tag{2}$$

where D_i and R_i depend only on α and the ratios $\frac{z}{\sigma}$ and $\frac{q}{\sigma}$ (see [2] for more details).

Let f be any function and x_k its k^{th} variable and define $\frac{\partial f}{\partial x_k}$ as the derivative of f with respect to its k^{th} variable. Differentiating (2) with respect to z_i, q_i and λ provides the following system:

$$\frac{\frac{\partial D_i}{\partial x_1}\left(\frac{z_i}{\sigma_i}, \frac{q_i}{\sigma_i}\right)}{\frac{\partial R_i}{\partial x_1}\left(\frac{z_i}{\sigma_i}, \frac{q_i}{\sigma_i}\right)} = \frac{\frac{\partial D_i}{\partial x_2}\left(\frac{z_i}{\sigma_i}, \frac{q_i}{\sigma_i}\right)}{\frac{\partial R_i}{\partial x_2}\left(\frac{z_i}{\sigma_i}, \frac{q_i}{\sigma_i}\right)} \tag{3}$$

$$\frac{\frac{\partial D_i}{\partial x_2}\left(\frac{z_i}{\sigma_i}, \frac{q_i}{\sigma_i}\right)}{\frac{\partial R_i}{\partial x_2}\left(\frac{z_i}{\sigma_i}, \frac{q_i}{\sigma_i}\right)} = -\lambda \frac{a_i}{\Delta_i \pi_i \sigma_i^2} \tag{4}$$

$$\sum_{i=1}^{\#SB} a_i R_i\left(\frac{z_i}{\sigma_i}, \frac{q_i}{\sigma_i}\right) - R_T = 0 \tag{5}$$

The solution of Eq. (3) provides the optimal relationship between z and q for a given shape parameter α. We get $\frac{z_i}{\sigma_i} = g_i\left(\frac{q_i}{\sigma_i}\right)$ with g_i the function that provides the solution of (3). Inserting this in Eq. (4) and (5) gives

$$h_i\left(\frac{q_i}{\sigma_i}\right) = \frac{\frac{\partial D_i}{\partial x_2}\left(g_i\left(\frac{q_i}{\sigma_i}\right), \frac{q_i}{\sigma_i}\right)}{\frac{\partial R_i}{\partial x_2}\left(g_i\left(\frac{q_i}{\sigma_i}\right), \frac{q_i}{\sigma_i}\right)} = -\lambda \frac{a_i}{\Delta_i \pi_i \sigma_i^2} \tag{6}$$

$$\sum_{i=1}^{\#SB} a_i R_i\left(g_i\left(\frac{q_i}{\sigma_i}\right), \frac{q_i}{\sigma_i}\right) - R_T = 0 \tag{7}$$

where h_i is defined in Eq. (6) to simplify the notations.

The bit allocation process is the following:

1. Lambda is given. For each subband i, compute $\ln\left(\lambda \frac{a_i}{\Delta_i \pi_i \sigma_i^2}\right) = \ln(-h)$ and read the corresponding bitrate R_i from the curves shown in Fig. 2.
2. Compute $\left|\sum_{i=1}^{\#SB} a_i R_i - R_T\right|$. If it is lower than a given threshold, the constraint (7) is verified. Else, compute a new λ and go back to step 1.
3. For each subband i, read $\frac{q_i}{\sigma_i}$ from the curves shown in Fig. 3. q_i is the optimal quantization step for subband i.
4. Then, for each subband i, read z_i from the curves shown in Fig. 4.

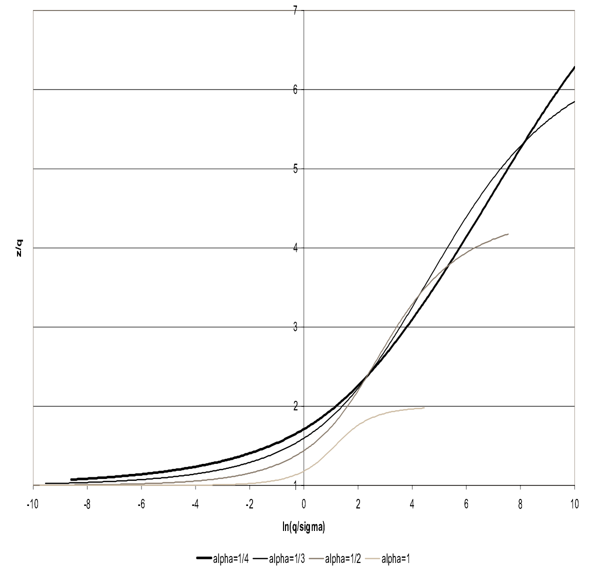

Figure 4: Optimal $\frac{z}{q}$ ratio versus $\ln\left(\frac{q}{\sigma}\right)$

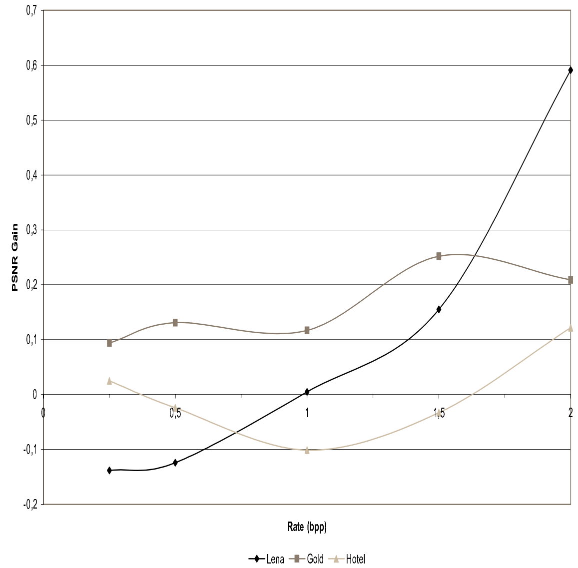

Figure 5: PSNR of the proposed method minus PSNR of JPEG 2000

	Lena		Gold		Hotel	
bpp	JPEG	Proposed	JPEG	Proposed	JPEG	Proposed
0.25	34.10	33.96	31.56	31.65	30.17	30.20
0.5	37.25	37.13	34.20	34.33	34.06	34.04
1	40.35	40.36	37.75	37.86	38.45	38.34
1.5	42.74	42.89	40.43	40.69	41.11	41.07
2	44.77	45.37	43.13	43.34	43.52	43.64

Table 1: PSNR versus bitrate comparisons between JPEG 2000 (VM8.6) and the proposed method

4 Experimental results

Table 1 compares the PSNR of our method and JPEG 2000 (Verification Model 8.6) for three test images when both coders use a three level 9-7 wavelet decomposition. The difference between the PSNR of our method and the PSNR of JPEG 2000 is plotted in Fig. 5. We can see that the performances of our approach are close to JPEG 2000 except for Lena at 2 bpp for which we have a gain of about 0.6 dB. Therefore, the method proposed in this paper provides state of the art performances with reduced complexity compared to JPEG 2000. Actually, the proposed compression scheme is composed of a low complexity model-based bit allocation [3] and independent quantization and entropy coding procedures while JPEG 2000 computes the bit allocation during the encoding step. A possible hardware architecture is to implement the DWT, subband statistics computation, scalar quantization and EBCOT bit plane coding procedures on ASICs and the bit allocation on a DSP.

5 Conclusion

We have proposed a new compression scheme using a model-based bit allocation to determine the quantizers to apply in each subband and EBCOT to entropy encode the quantized subbands. Our algorithm provides the same performances as JPEG 2000 with less complexity and simpler hardware implementation. Finally, the proposed method can be easily adapted to provide a scalable JPEG 2000-like file format syntax.

References

[1] "Jpeg-2000 part-1 standard," *ISO/IEC 15444-1.*

[2] C. Parisot M. Antonini M. Barlaud, "Optimal nearly uniform scalar quantizer design for wavelet coding," *Proc. of SPIE Visual Communications and Image Processing Conference*, Jan. 2002.

[3] C. Parisot M. Antonini M. Barlaud, "EBWIC: A low complexity and efficient rate constrained wavelet image coder," *Proc. of IEEE International Conference on Image Processing*, Sept. 2000.

[4] D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Transactions on Image Processing*, vol. 9, no. 7, July 2000.

[5] M. Antonini M. Barlaud P. Mathieu I. Daubechies, "Image coding using wavelet transform," *IEEE Transactions on Image Processing*, vol. 1, no. 2, pp. 205–220, 1992.

[6] J.D. Villasenor B. Bellzer B. Liao, "Wavelet filter evaluation for image compression," *IEEE Transactions on Image Processing*, vol. 4, no. 8, pp. 1053–1060, Aug. 1995.

[7] K. Ramchandran M. Vetterli, "Best wavelet packet bases in a rate-distorsion sense," *IEEE Transactions on Image Processing*, vol. 1, no. 2, pp. 160–176, Apr. 1993.

[8] A. Gersho R. M. Gray, "Vector quantization and signal compression," *Kluwer Academic Publishers*, 1992.