LINEAR SCALAR QUANTIZATION OF WAVELET IMAGE DECOMPOSITION USING JOINT OPTIMIZATION

Riadh Abdelfattah[1], *Olivier Rioul*[2] *and Pierre Duhamel*[3]

[1] URISA, École Supérieure des Communications, Route de Raoued Km 3.5, 2083, El Ghazala-Ariana Tunisia, Email: riadh.abdelfattah@supcom.rnu.tn, Telephone: +216 71 857 000, Fax: +216 71 856 829
[2] COMELEC, École Nationale Supérieure des Télécommunications, 46 rue Barrault 75634 Cedex 13
[3] CNRS/LSS, École Supérieure d'Électricité, Plateau de moulon F-91192, GIF sur YVETTE

ABSTRACT

This paper proposes a new linear scalar quantization algorithm. A modified Lloyd's optimum algorithm, based on a linear quantization, is proposed to decompose a memoryless source into the bits. This is the case of the considered source which is the output of the wavelet transformation. The developed algorithm gives acceptable results for generalized Gaussian sources. Our approach is compared with the Lloyd max one's and a second one where the quantization intervals are equiprobable.

1. INTRODUCTION

For image transmission over a binary channel [3], a quantizer is used for the reduction of the image plane transform dynamics. In general, both scalar and vector quantization [6] could be used. A scalar quantizer is considered in this paper, providing a more simple bitwise decomposition of a still image.

In [7], Lloyd proposed two algorithms to design optimum quantizers in term of minimum distorsion [8]. It is based on a joint optimization of the encoder and the decoder. In [3], it is shown that for a uniform source, the optimum scalar quantization is obtained by a natural binary coding. However, this is not necessarily the case for other kinds of sources.

In this paper, this problem is treated for an arbitrary memoryless source and a modified version of Lloyd's algorithm, based on a linear quantization, is proposed. It is an optimization version of the algorithm proposed in [1, 5], where they consider equiprobable quantization intervals. The hypothesis to be considered here is to have a memory-less source. However, this is not a very restricting hypothesis. In fact, the considered source is the output of the wavelet transformation [2] which produce a memoryless output.

2. WAVELET IMAGE DECOMPOSITION

The wavelet transform allows the decorrelation of the image information by filtering the original image in to different frequency components using a filterbank. This filter bank divide each dimension of the image into a lowpass and a highpass region. Then only the lowpass region is divided into new lowpass- and highpass regions in the same way, and so on. The number of decomposition levels is the number of times the process is repeated. All the wavelet transformations considered in this paper use a minimum phase Daubechies low-pass filter for the image decomposition [4].

The lowpass-lowpass subimage must always be treated with special caution. This subimage contains a lot of essential information and do not have the same favorable statistical properties of the other subimages. Some extra processing is needed for this subimage.

3. JOINT OPTIMIZATION OF THE QUANTIZER

The goal of source coding or quantization is to reproduce the source by as few channel transmission symbols as possible, which means mapping a broad range of input values to a limited number of output values. In this paper, we choose the scalar quantization which is a case of lossy compression. The motivation behind lossy image compression is the limited channel bandwidth.

The scalar quantization is a function, Q: IR $\longrightarrow V$ where IR is the set of real numbers and $V = \{v_1 ... v_{2^N}\}$ the codebook (the output set). The thresholds of quantization are called decision levels, u_i, and the output levels are called the reconstruction levels or output levels, v_i. Thus, the compression done by the scalar quantization consists on selecting an output level v_i, and the only saved information is the index i. Two major functions are distinguished in scalar quantization: encoding: the inputs positioned in the i-th quantization interval, are encoded to i ; decoding: the encoded value i, is decoded to v_i.

Optimization of the scalar quantizer can be achieved by optimizing the encoder and decoder functions. The Lloyd-Max algorithm is today a common tool of joint optimization, and it can be used for different types of sources. The algorithm is based on iteratively optimizing the encoder and decoder. We can begin by optimizing the encoder assuming a known decoder, then assuming this optimized encoder, we proceed to optimize the decoder. We continue this iteration until convergence to an optimum so-

lution. Considering this methodology, we propose to develop an optimized linear decoder without the constraint of equiprobable quantization intervals. It can be viewed as a special adaptation of Lloyd's algorithm to optimize the distorsion due to the source and the channel. The quantizer optimization is then achieved by minimizing the mean square error (MSE) between the centroid, v_i, and the decision level u_i. Analytically, it is equivalent to minimize the expectation value of $(u-v)^2$,

$$E[(u-v)^2] = \int (u-v)^2 p(u)du = \sum_{j=1}^{2^N} \int_{V_j} (u-v_j)^2 p(u)du \quad (1)$$

where N is the number of quantization bits, u is the vector of decision levels known the probability density function $p(u)$ and v the corresponding centroid vector. v_j is the centroid of the quantization interval V_j. It is given by:

$$v_j = \sum_{i=1}^{N} \alpha_i b_{ij} \quad (2)$$

where the α vector is the reconstruction weights. As for all Lloyd-Max algorithm, the initialization of α's values is important for both the convergence speed and the final result. There are plenty of local minima solutions, but choosing initialization values for a with care, will most likely give the global minima solution or a value very close to that one. Equation 2 is equivalent with a matrix representation to

$$\underline{\mathrm{v}} = \alpha.\mathrm{B} \quad (3)$$

where B is the natural binary encoding matrix ($0 \leq j \leq 2^N - 1$) where 0 are replaced by -1. b_{ij} is then the i-th bit of the binary encoding (-1 or +1) of the integer j [9]. Thus, minimisation of (1) gives:

$$\frac{\partial E[(u-v)^2]}{\partial \alpha_i} = \sum_{j=1}^{2^N} \int_{V_j} \left(u - \sum_{i=1}^{N} \alpha_i b_{ij} \right) b_{ij} p(u)du = 0 \quad (4)$$

which is equivalent to

$$\sum_{j=1}^{2^N} \left(\int_{V_j} up(u)du \right) b_{ij} = \sum_{j=1}^{2^N} \left(\int_{V_j} p(u)du \right) \sum_{i'=1}^{N} \alpha_{i'} b_{i'j} b_{ij} \quad (5)$$

Putting $P_j = \int_{V_j} p(u)du$ and $E_j = \int_{V_j} up(u)du$, we obtain:

$$\sum_{i'=1}^{N} \alpha_i \underbrace{\left(\sum_{j=1}^{2^N} P_j b_{i'j} b_{ij} \right)}_{a_{ii'}} = \underbrace{\sum_{j=1}^{2^N} E_j b_{ij}}_{c_i} \quad (6)$$

which is equivalent with a matrix representation to

$$\mathrm{A}.\alpha = \mathrm{C} \quad (7)$$

As A is a non singular square matrix, α is deduced by inverting equation (7).

4. ALGORITHM DESCRIPTION

The quantization rule is defined as follows

$$u \in V_{\mathrm{w}} \qquad \Longleftrightarrow \qquad \mathrm{w} = [\mathrm{w}_k \ \cdots \ \mathrm{w}_2 \ \mathrm{w}_1] \quad (8)$$

where $\mathrm{w}_i \in \{0, 1\}$ and $[\mathrm{w}_k \ \cdots \ \mathrm{w}_2 \ \mathrm{w}_1]$ is the binary representation of w. In other words, we have to compute the decision levels and the corresponding centroids in order to define the quantizer. The developed quantizer is based on the Lloyd Max's rules:

- nearest neighbor rule: this rule states that for a given decoder, the optimum encoder is the one that encodes each source input to the nearest codevector,
- centroid rule: this rule states that for a given encoder, the optimum reconstruction level is the centroid of all input vectors encoded to this codevector, weighted by their probability.

The developed algorithm is described in three main steps:

1. **Initialization of the algorithm**
   - Initialize the reconstruction weighs, α, using a uniform source, only for the first step of initialization;
   - Calculate the centroids for each interval

     $$v_j = \sum_{i=1}^{N} \alpha_i b_{ij} \qquad \text{for} \qquad 1 \leq j \leq 2^N \quad (9)$$

     An increasing order of the centroids v_j must be considered.
2. **Lloyds iterations**
   - With α and V, calculate the reconstruction levels vector U;
   - Calculate P_j and E_j;
   - Calculate A using (6);
   - Compute the new value of α with (7).
3. **Termination of the algorithm**
   - α and V are the solution to the optimization problem. α represents the optimum dequantizer while V represents the optimum quantizer.

As exemple, let's consider a scalar quantization using the developed algorithm with $N = 4$ bits: the corresponding optimized reconstruction weight vector is given by $\alpha = [1.50 \ \ 0.84 \ \ 0.45 \ \ 0.25]$, where the weight is proportional to the importance of the bit in the reconstruction process. As mentioned earlier, initializing the α vector is extremely important both for the convergence speed and for the final solution (because of local minima). Several initializing strategies can be used. One possibility is to use a set of more or less random values which statistically perform well. In [5], a more systematic approach based on a

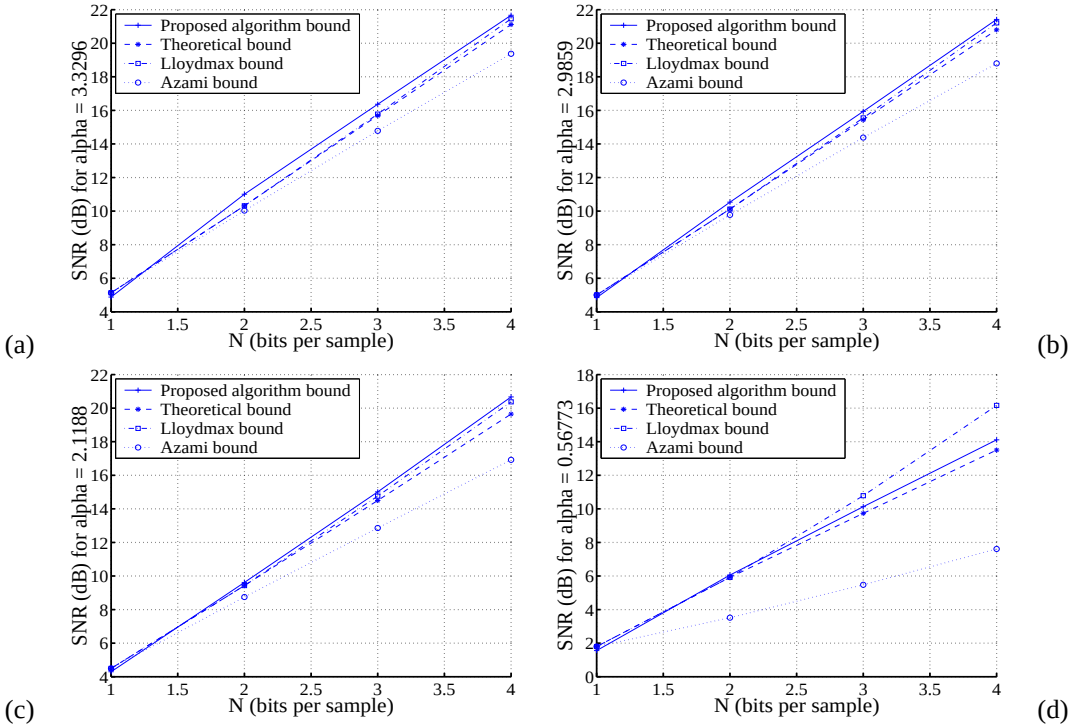

Figure 1. Rate SNR plots of Lena for Low pass - Low pass subimage with (a) 2, (b) 3, (c) 4 decomposition levels and (d) the rate SNR plot of Lena for the second subimage (High pass - High pass).

theoretical fundament is given. In [1] the optimal α-values are computed through exhaustive search. This is done for the most common perfect sources like uniform, Gaussian and Laplacian. Most subimages have a pdf that are close to Laplacian when normalized (except the low-low subimage which is more Gaussian-like).

5. EXPERIMENTAL RESULTS

When dealing with image quantization and dequantization, we need some fidelity criteria to evaluate these techniques. In fact, they introduce distorsion to the reconstructed image. It is common to separate between objective fidelity criteria and subjective fidelity criteria. All though objective fidelity criteria provide a simple and convenient way to compare different systems, subjective fidelity criteria should not be neglected.

It is not necessarily a direct link between an objective fidelity criterion and the subjective human perception of the image. Small objective changes in the image quality can have a considerable impact on the perceived quality, and objective changes can have no effect on the observed image. Subjective fidelity criteria are therefore always necessary as a control / evaluation of the objective results. A common objective fidelity criteria in image quantization is the mean-square signal-to-noise ratio SNR_{ms}. The SNR (in dB) is defined as $SNR = 10\log_{10}(\sigma^2/D)$, where σ^2 is the variance of the source pdf. The distorsion D is given by the minimization of the mean square error

$$\begin{aligned} D &= \sum_{j=1}^{2^N} \int_{V_j} (u - v_j)^2 p(u) du \\ &= \underbrace{\int u^2 p(u) du}_{1} + \sum_{j=1}^{2^N} \underbrace{\left(\int_{V_j} p(u) du \right)}_{P_j} v_j^2 \\ &\quad -2 \sum_{j=1}^{2^N} \underbrace{\int_{V_j} u p(u) du}_{E_j} \end{aligned} \tag{10}$$

Thus, the distorsion is computed as follows

$$D = 1 + \sum_{j=1}^{2^N} v_j^2 P_j - 2 \sum_{j=1}^{2^N} v_j E_j \tag{11}$$

Three different quantizers / dequantizers have been simulated and compared:

1. The Lloyd max algorithm.
2. Lloyd-Max algorithm adaptation to optimize the α-values using equiprobable quantization intervals.
3. The proposed algorithm.

Figure 1 and 2 show the rate SNR results of Lena for Low pass - Low pass subimage with different decomposition levels (2, 3 or 4). The histograms of these subimages

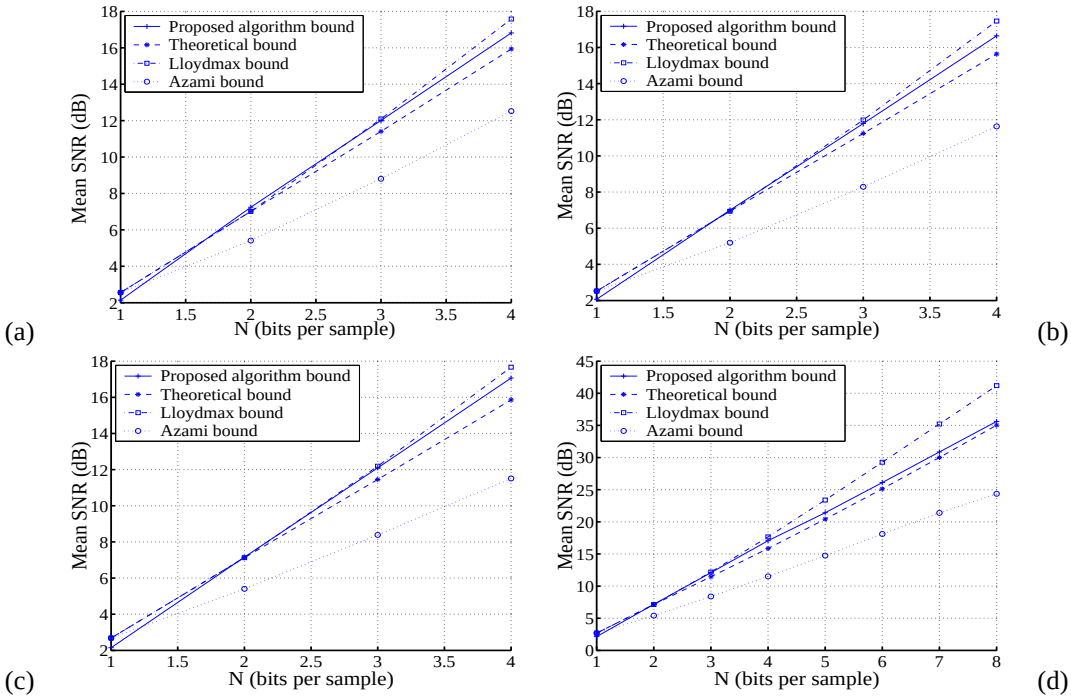

Figure 2. Rate mean SNR plots of Lena with (a) 2, (b) 3, (c) 4 decomposition levels and (d) untill 8 quantification bits.

(Low pass - Low pass) have a non zero mean values. Thus, for the equivalent distorsion computation of these subimages, a normalization parameter (maximum plus minimum values divided by two) is removed from the processed subimage. The SNR is used as a quality criteria of the reconstructed images. This does not necessarily mean that a higher SNR-value gives a superior visually percepted image, but generally we accept this assumption. The quantization with the constraint in equation (2) gives an increasing quantization loss with increasing bit rates compared with Lloyd max quantization. All the curves confirm this result. However, we remark the significant improvement of the proposed algorithm against the one proposed in [1]. For low rates, the difference between the proposed system and the Lloyd max quantization is negligible. For high rates, the difference becomes quite large.

6. CONCLUSION

In this paper, we have presented a linear scalar quantizer of wavelet image decomposition (memoryless source) using joint quantizer/dequantizer optimization. The developed algorithm is based on the Lloyd Max rules. The illustrated rate SNR on this paper show a significant improvement of the proposed algorithm. To have a more objective idea on the robustness of the developed algorithm, the quantizer has to be tested in a transfer system with different noisy channels.

7. REFERENCES

[1] S.B.Z. Azami, *Codage conjoint source/canal, protection hiérarchique*, Thèse de doctorat de l'Ecole Nationale Supèrieure des Télécommunications, ENST99E007, 1999.

[2] R.W. Buccigrossi and E.P. Simoncelli, *Image Compression via joint statistical characterization in the wavelet domain*, IEEE Trans. on Image Processing, vol. 8, No. 12, pp. 1688-1701, 1999.

[3] R.J. Clarke, *Digital compression of still images and video*, Academic Press, 1995.

[4] I. Daubechies, *Orthonormal bases of compactly supported wavelets*, Com. on Pure Appl. Math., vol. 41, pp. 909-996, nov. 1988.

[5] S.I. Evj, *Bite-Plane image coding using joint source-channel optimization*, Rapport technique, Département COMELEC, Ecole Nationale Supèrieure des Télécommunications, 2001.

[6] A. Gersho and R.M. Gray, *Vector quantization and signal compression*, Kluwer Academic Publishers, 1997.

[7] S.P. Lloyd, *Least squares quantization in PCM*, IEEE Transaction on Information Theory, vol. 28, No. 2, 1982.

[8] J. Max, *Quantization for minimum distorsion*, IEEE Transaction on Information Theory, vol. 6, No. 1, 1960.

[9] COSOCATI, *Codage conjoint source canal pour la transmission d'images*, Rapport technique, Réseau National de Recherche en Télécommunications, ENST-Paris, 2003.