A Novel Entropy-Constrained Adaptive Quantization Scheme for Wavelet Pyramid Image Coding

J.Q. Ni
jqni@hkueee.hku.hk

K.L. Ho
klho@hkueee.hku.hk

K.W. Tse
kwtse@hkueee.hku.hk

Department of Electrical and Electronic Engineering
The University of Hong Kong
Pokfulam Road , Hong Kong

Abstract - It is well known that orthogonal wavelet transform with filters of nonlinear phase gives poor visual results in low bit rate image coding. Biorthogonal wavelet is a good substitute, which is, however essentially nonorthogonal. A greedy steepest descend algorithm is proposed to design an adaptive quantization scheme based on the actual statistics of the input image. Since the L^2 norm of the quantization error is not preserved through the nonorthogonal transform, a quantization error estimation formula considering the characteristic value of reconstruction filters is derived to incorporate the adaptive quantization scheme. Computer simulation results demonstrate significant SNR gains over standard coding technique, and comparable visual improvements.

I. Introduction

Recently wavelet based image coding has become the domain of extensive research. For image coding application, it is desirable that an exact reconstruction subband coding scheme should correspond to an orthonormal basis with reasonable smooth mother wavelet. However, the resulting analysis and synthesis filters are in general not linear phase. On the other hand it is also desirable the FIR filter used be linear phase, since such filter can be easily cascaded in pyramidal filter structure without phase compensation [2]. Unfortunately, there are no nontrivial orthonormal linear phase FIR filters with the exact reconstruction property [5]. Relaxing the orthonormality requirement and using biorthogonal basis is a good alternative.

As is well known, the L^2 norm of quantization errors is not preserved through the nonorthogonal transform, so there does not exist an exact analytical formula for optimal bit allocation as in the case of orthogonal wavelet transform [3]. The goal of the paper is to improve the applicability of biorthogonal wavelet transform to image coding by proposing an adaptive quantization design scheme based on the statistics of the input image. The design procedure is based on a criterion which best describes the trade-off between distortion and rate, and hence results in superior performance.

II. Biorthogonal Wavelets

II.1 *Siganl Analysis with Biorthogonal Wavelet*

Biorthogonal wavelet bases were firstly introduced by Cohen [5] and were extensively used by numerous authors in image coding application. The basic idea of biorthogonal wavelet transform is to represent a signal $f(x)$ as a superposition of two biorthogonal bases

$$\psi_{m,n}(x) = 2^{-m/2} \psi(2^{-m}x - n)$$

and $\psi_{m,n}^*(x) = 2^{-m/2} \psi^*(2^{-m/2}x - n)$, where $\psi^*(x)$ is the dual basis of $\psi(x)$, m and n are scaling and translation parameters respectively. Therefore the signal $f(x)$ can be expressed as

$$f = \sum_{m,n} c_{m,n}(f) \psi_{m,n}^* \qquad (1)$$

and

$$c_{m,n}(f) = \left\langle \psi_{m,n}, f \right\rangle = \int_{-\infty}^{+\infty} \psi_{m,n}(x) f(x) dx$$

$$(2)$$

As in the orthonormal case, one can introduce multiresolution analysis. We then define two dual scaling functions: $\phi_{m,n}(x) = 2^{-m/2} \phi(2^{-m}x - n)$ and its dual $\phi_{m,n}^*$.

A pair of filter h and h^* have been introduced for decomposition and reconstruction of signal f to and from its wavelet representation respectively. Generally speaking, there are two pairs of filters: one pair h and g for decomposition and another pair h^* and g^* for reconstruction.

For signal decomposition , we have

$$c_{m,n} = \sum_k g_{2n-k} a_{m-1,k}$$

$$a_{m,n} = \sum_k h_{2n-k} a_{m-1,k} \qquad (3)$$

while for signal reconstruction,

$$a_{m-1,n} = 2 \times \sum_k (h_{2k-n}^* a_{m,k} + g_{2k-n}^* c_{m,k})$$

$$(4)$$

II.2 *Extension to the 2-D case*

The results described in the previous section can be

easily extended to 2-dimensional case by applying two 1-D wavelet transform separately [2].

In 2-D dyadic wavelet analysis, as in the 1-D case, we introduce a scaling function $\phi(x,y)$ such that:

$$\phi(x,y) = \phi(x)\phi(y) \qquad (5)$$

where $\phi(x)$ is a one-dimensional scaling function.

Let $\psi(x)$ denotes the one-dimensional wavelet associated with the scaling functionm $\phi(x)$. Then, we have three two-dimensional wavelets defined as follows:

$$\begin{aligned}
\psi^H(x,y) &= \phi(x)\psi(y) \\
\psi^V(x,y) &= \psi(x)\phi(y) \qquad (6) \\
\psi^D(x,y) &= \psi(x)\psi(y)
\end{aligned}$$

II.3 *Image Coding with Biorthogonal Wavelets*

A wide variety of wavelet-based image compression schemes have been reported in previous literatures. The system diagram reflecting the structure employed in this paper is shown in Fig.1. The principle of wavelet based image coding can be described as follows: the input image P_0 is fed into the first BWT unit, the "smoothed" version P_1 is again fed into next BWT unit for further decomposition while the wavelet details D_1^1, D_1^2 and D_1^3 are sent to a buffer for quantizing and encoding. This procedure is repeated J times, forming a J layer pyramid structure $\left\{ P_J, (D_j^1)_{1\leq j\leq J}, (D_j^2)_{1\leq j\leq J}, (D_j^3)_{1\leq j\leq J} \right\}$, which is called the wavelet pyramid. The "information packing" property of wavelet transform makes the image in wavelet domain have a lower entropy than the original, thereby increasing the possibility of compression.

III. Adaptive Quantization Scheme for Image Pyramid Coding

III.1 *Problem Formulation*

The proposed adaptive quantization scheme is a scalar quantizer. If a variable-length entropy encoder is used, uniform quantization is optimal. To achieve the best results, a separate quantizer should be designed for each subband, which is equivalent to determine the optimal step size Δ for each subband based on the statistics of the input image. If we represent the original image with J-level wavelet pyramid, for a predefined target bit rate R_{Target}, the problem is formulated as to find a set of optimal step size $\left\{ \Delta_J, (\Delta_j^k)_{1\leq j\leq J, 1\leq k\leq 3} \right\}$ to minimize the overall distortion

$$\varepsilon = D(\Delta_J, \Delta_j^k) \qquad (7)$$

subject to the target bit rate constraint:

$$(\frac{1}{4^J} R_J + \sum_{j=1}^{J} \frac{1}{4^j} \sum_{k=1}^{3} R_j^k) \leq R_{Target} \qquad (8)$$

Since L^2 norm is not preserved through biorthorgonal wavelet transform, the ε has not the simple form of direct summation of quantization error of each subband as in the orthogonal case. To estimate the overall distortion ε, it is not necessary to perform the complete inverse wavelet transform each time the set of quantization step is updated. In subband image coding, one can assume that quantization error generated in one subband will not influence the error generation in other subband during reconstruction [4]. Therefore the overall distortion depends upon the frequency characteristic of reconstruction filters. Then the average distortion per sample after reconstruction can be expressed as [4] :

$$\sigma_n^2 = \gamma_l \sigma_{nl}^2 + \gamma_h \sigma_{nh}^2 \qquad (9)$$

where $\sigma_{nl}^2, \sigma_{nh}^2$ and γ_l and γ_h are the quantization error and characteristic value of reconstruction filter corresponding to low-pass and high-pass subbands, respectively. And γ_l, γ_h have the form (for normalized filter pair h and h^*):

$$\gamma_l = \frac{1}{\pi} \int_{-\pi}^{\pi} |H^*(\omega)|^2 \, d\omega \qquad (10)$$

$$\gamma_h = \frac{1}{\pi} \int_{-\pi}^{\pi} |G^*(\omega)|^2 \, d\omega \qquad (11)$$

The extension of above results to 2-dimensional case is straightforward. Let γ_0 denote the characteristic value of reconstruction filter for LL band, γ_1 for LH band, γ_2 for HL band and γ_3 for HH band, for a J-layer pyramid image decomposition, we have

$$\varepsilon = \gamma_0^J \varepsilon_J + \sum_{j=1}^{J} \sum_{k=1}^{3} \gamma_0^{j-1} \gamma_k \varepsilon_j^k \qquad (12)$$

where $\gamma_0 = \gamma_l \gamma_l$, $\gamma_1 = \gamma_l \gamma_h$, $\gamma_2 = \gamma_h \gamma_l$ and $\gamma_3 = \gamma_h \gamma_h$.

Rate estimation is another key issue involved in the quantizer design procedure. We do not restrict ourselves to a particular entropy encoder for the sake of generality, instead we adopt an empirical entropy R_T introduced by Antonini [2]:

$$R_T = \frac{1}{4^J} R_J + \sum_{j=1}^{J} \frac{1}{4^j} \sum_{k=1}^{3} R_j^k \qquad (13)$$

where R_J and R_j^k represent the bit rates assigned to sub-image P_J and P_j^k, respectively. For sub-image P_j^k, we can estimate the average entropy as follows:

$$\Re_j^k = -\frac{1}{4^j} \sum_{i=1}^{L} p(w_i) \log_2 p(w_i) \qquad (14)$$

w_i in above expression denotes the codeword corresponding to P_j^k.

III.2 *The Adaptive Quantization Scheme*

The proposed adaptive quantization scheme follows its spirit from Wu's algorithm [1] for JPEG baseline coders with some modification. It is felt that the proposed scheme is more suitable for wavelet-based image coding application. Compared to the quantization table in JPEG baseline coders, the number of entry in the set of quantization step, which equals to the number of subbands in wavelet pyramid, is greatly reduced, therefore the complexity of the algorithm is greatly reduced.

The basic concept of the scheme is described as follows: starting from a set of initial large quantization step $\left\{\Delta_J,(\Delta_j^k)_{1\leq j\leq J,1\leq k\leq 3}\right\}$ for each subband, corresponding to low bit rates and high reconstruction distortion, the algorithm decreases the step size of one entry of the quantization set at a time until a predefined target bit rate is reached. A reasonable quantity that best describes the rate-distortion trade-off is $\dfrac{-\Delta\varepsilon}{\Delta R}$, i.e., the ratio between the decrease in reconstruction error and the increase in bit rates. We re-lable the components of quantization table

$$\Delta = \left\{\Delta_J,(\Delta_j^k)_{1\leq j\leq J,1\leq k\leq 3}\right\}$$ to give

$$\Delta = \left\{\Delta_0,\Delta_1,\cdots,\Delta_{3J}\right\}$$ for convenient description of the scheme. Then we can describe the procedure of quantizer design as the following optimization problem:

$$\max_k \max_{\overline{\Delta}_k} \frac{-\Delta\varepsilon}{\Delta R}\bigg|\Delta_k\to\overline{\Delta}_k \qquad (15)$$

The optimization problem of (15) can be divided into two parts; the first part calculates

$$\lambda_k = \max_{\overline{\Delta}_k} \frac{-\Delta\varepsilon}{\Delta R}\bigg|\Delta_k\to\overline{\Delta}_k \qquad (16\text{-}1)$$

for all k and the second part finds

$$\lambda = \max_k \lambda_k \qquad (16\text{-}2)$$

We set the initial value of Δ_0 in quantization table $\left\{\Delta_k\right\}$ as some small value, e.g. 16, to preserve the visual quality of image. The algorithm also maintains two extra tables $\left\{\lambda_k\right\}$ and $\left\{\Delta_k^*\right\}$ where the Δ_k^* is the solution for (16-1). The design algorithm is iterative, and is given as follows:

1) Initially set the quantization table $\left\{\Delta_k\right\}$ by $\Delta_0 =$ some small step; $\Delta_k = \angle_{\max}$ for $k = 1,\cdots,3J$.

2) Initialize the tables of $\left\{\lambda_k\right\}$ and $\left\{\Delta_k^*\right\}$ by searching Δ_k within domain $\left\{1,\cdots,\Delta_{\max}-1\right\}$ for $k = 1,\cdots,3J$ and $\left\{1,\cdots,\Delta_0-1\right\}$ for k=0 so that (16-1) has the maximum.

3) Search k in $\left\{0,1,\cdots,3J\right\}$ for k^* to solve (16-2).

4) Update k^*-th entry of the quantization table $\left\{\Delta_k\right\}$ by setting $\Delta_{k^*} = \Delta_{k^*}^*$.

5) Update λ_{k^*} and $\Delta_{k^*}^*$ by searching Δ in $\left\{1,\cdots,\Delta_{k^*}^*-1\right\}$ to solve (16-1) for $k = k^*$.

6) if $R_T \geq R_{T\,\mathrm{arget}}$ then stop,else goto 3.

The design algorithm terminates when the bit rate equal or exceeds the target bit rate $R_{T\,\mathrm{arget}}$. For normalized filter pair $h(n)$ and $h^*(n)$, the absolute value of WT coefficient ranges from 0 to 255 and the $\Delta_{\max}$ can be set to 255. However, computer simulation suggests that, as stated in Wu'paper, for practical image coding application at a reasonable bit rate, initializing the quantization step table with some relatively small step size leads almost the same result as with the maximum allowable step size 255, but the number of iteration required to arrive at the target bit rate is greatly reduced.

IV Simulation Result

Computer simulation was carried out to verify the feasibility and effectiveness of the proposed adaptive quantization design scheme for wavelet pyramid image coding. We compare the performance between the conventional and adaptive quantization scheme with the standard test image "Lena" of size 512×512. In both cases the test image is decomposed with the same nine and seven taps biorthogonal wavelet [2]. The conventional quantization scheme uses the same step size for each band at the same resolution, decreasing by a factor of 2 when going to next coarser resolution, which is optimal for orthogonal wavelet decomposition [3]. We obtain PSNR of 34.5db at 0.4bbp and 32.3db at 0.2bbp with the proposed adaptive quantization scheme, compared to 32.2 at 0.4 bbp and 30.1db at 0.2bbp with conventional quantization scheme. Fig.2 gives the original test image and the reconstructed images with the proposed scheme at different bit rates. Significant improvement over conventional quantization scheme is observed both in PSNR perfomance and overall visual quality of coded image at the same bit rate.

V Conclusion

In this paper, we have presented a entropy-constrained adaptive quantization scheme for wavelet based pyramid image coding. The scheme is based on a criterion which gives the best trade-off between distortion and bit rate and adapts to the actual statistics of the input image. It gives the best coding efficiency when followed by entropy

coding. The use of biorthogonal wavelet leads to better visual quality of coded image, which is, however, nonorthogonal in nature and the L^2 norm of the quantization error is not preserved through nonorthogonal transform. To incorporate the adaptive quantization scheme, we have derived the formula of quantization error estimation, considering the characteristic value of reconstruction filter, for biorthogonal wavelet transform under the assumption of cross-uncorrelated quantization error among subbands. Empirical entropy is also used to evaluate the bit rate of the coded image when the adaptive quantization scheme is applied to the decomposed image. The immediate estimation of both reconstruction error and bit rate makes the proposed adaptive scheme feasible and fast in implementation. Computer simulation shows considerable SNR gains over conventional scheme and comparable visual improvement.

Reference

[1] S.W. Wu and A. Gersho, "Rate-Constrainted Picture-Adaptive Quantization for JPEG Baseline Coders," in Proc. Inter. Conf. on ASSP, pp. 365-368, April,1993.

[2] M. Antonini, M. Barlaud and Daubechies, "Image Compression Using Wavelet Transform," IEEE Trans. Image Processing, Vol.1, no.2, pp. 205-220,1992.

[3] J. Chen, S. Itoh and T. Hashimoto, "Scalar Quantization Noise Analysis and Optimal Bit Allocation for Wavelet Pyramid Image Coding," IEICE Trans. Fundamentals, Vol. E76-A, pp. 1502-1514, Sep. 1993.

[4] C.K. Cheong, K. Aizawa, T. Saito and M. Hatori, "Subband Image Coding with Biorthogonal Wavelets," IEICE Trans. Fundamentals, Vol. E75-A, pp. 871-881, July 1992.

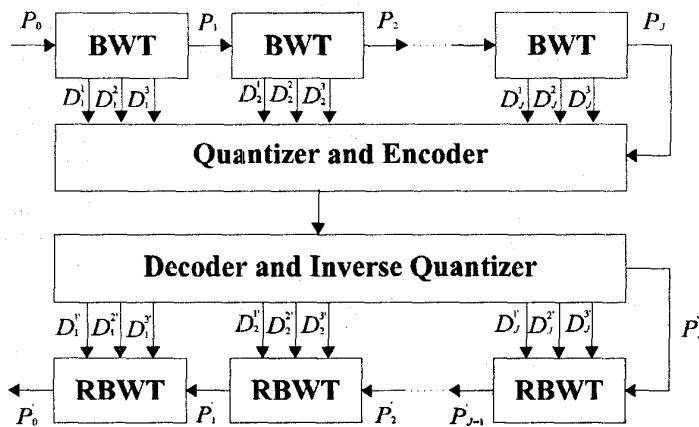[5] Charles K. Chui, Wavelets: A Tutorial in Theory and applications, Academic Press,Inc.,New York,1992.
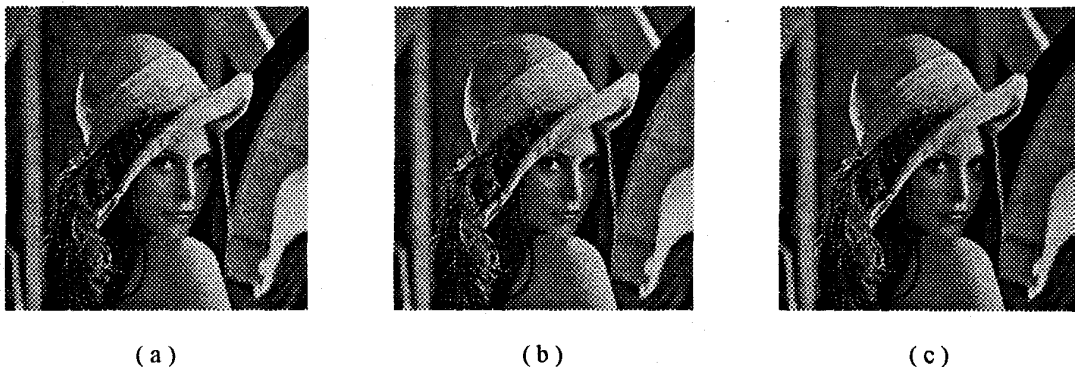
Fig.1 Block diagram of wavelet-based image coders

(a)   (b)   (c)

Fig.2 Reconstructed results for 512 × 512 image Lena
(a) original;
(b) PSNR: 34.5 at 0.4bbp;
(c) PSNR: 32.3 at 0.2 bbp