Filter Evaluation for Zerotree Wavelet Image Coding

Michael Miloslavski and Yo-Sung Ho

Kwangju Institute of Science and Technology
572 Sangam-dong, Kwangsan-gu, Kwangju, 506-712, KOREA
Phone: +82-62-970-2258, Fax: +82-62-970-2204, E-mail: michael@gogh.kjist.ac.kr

Abstract - In this paper, we discuss the significance of certain transform characteristics for the efficiency of zerotree wavelet image coders and we demonstrate that regularity of the synthesis wavelet should not be too low. We also introduce a new measure of the coding performance, a distortion amplification factor, which can be considered as a measure of nonorthogonality because it reflects amplification of quantization errors in the synthesis stage of the filter bank. Distortion amplification is highly correlated with PSNR and it can be used as a measure of the wavelet suitability for image coding, when combined with a minimum regularity requirement.

I. INTRODUCTION

Recently, an efficient and computationally inexpensive image coding algorithm, called Embedded Zerotree Wavelet coder (EZW), has been proposed by J. M. Shapiro [1] and further developed by A. Said and W. A. Pearlman, as is called Set Partitioning In Hierarchical Trees (SPIHT) [2]. It is based on a wavelet decomposition with progressive transmission and successive refinement of subband coefficients. Locations of nonzero coefficients are encoded with a tree structure, which can exploit the self-similarity of the wavelet pyramid decomposition across different scales. Performance of this technique can be further improved by an appropriate choice of the filter bank. In this paper, we discuss what kind of properties should be considered for a proper choice of the wavelet filters.

To improve performance, an optimal wavelet basis can be used for a particular image. Evidently, in this case the wavelet must be constructed by considering the image being processed, and the scheme becomes data-dependent. Furthermore, construction of the optimal basis is, usually, a computationally intensive procedure, and its complexity grows with the filter size. Therefore, in this paper, we restrict ourselves for the choice of a wavelet with overall “good” characteristics.

Only linear phase filters are considered, since such filters can be easily cascaded without the need for the phase compensation. Linear phase filters permit symmetric extension of the input image, thus, reducing artifacts at the boundaries at each step of the reconstruction. A problem with this filter type is that such filters can not be orthogonal.

The choice of the filter for a compression scheme is determined by a set of filter’s characteristics. The most significant of these characteristics are regularity, filter length, number of vanishing moments, and coding gain. Importance of every particular characteristic depends on the compression scheme. In the next sections, we analyze influence of the transform characteristics on the performance of zerotree wavelet coder, and also introduce a measure of distortion amplification due to the nonorthogonality of the filter bank.

II. FILTER CHARACTERISTICS

Since natural images are mostly smooth, it is appropriate that an exact reconstruction subband coding scheme should correspond to a basis with a reasonably smooth wavelet. Being a measure of the wavelet smoothness, regularity has been suggested as a criterion for filter evaluation [3]. Smoother wavelets also make compression artifacts less objectionable.

High regularity requires long filters. However, very long filters are usually avoided because of their computational complexity and also because they tend to spread coding errors. Thus, if the filters are long, they cause an annoying artifact, known as ringing, around edges. Filter order also has an influence on the time-frequency localization. As the filter order increases, time-frequency localization is reduced up to some order. Beyond that, increasing the filter length has a negative effect. Therefore, short smooth filters are preferred in image compression.

Another measure of filter performance is the coding gain. It is a popular characteristic to determine energy compaction of a transform. Generalization of the notion of coding gain for nonorthogonal subband decomposition has been proposed by Katko and Yasuda [4]:

This work was supported in part by the Korea Science and Engineering Foundation (KOSEF) through the Ultra-Fast Fiber-Optic Networks Research Center at Kwangju Institute of Science and Technology.
\[ G_{\text{gen}} = \frac{1}{\prod_{k=0}^{N-1} \left( \frac{A_k B_k}{\alpha_k} \right)^{\alpha_k}}, \]  
(1)

where:
\[ \alpha_k = \frac{L_k}{L} \]  
(2)
\[ \sigma_k^2 = A_k \sigma_x^2 \]  
(3)
\[ \sigma^2 = \sum_{k=0}^{N-1} B_k \sigma_k^2 \]  
(4)

Here an input signal \( x \) with \( L \) elements is split into \( N \) subbands. Each subband has \( L_k \) elements, quantization distortion \( \sigma_k^2 \), and variance \( \sigma^2 \).

In the derivation of this formula it is assumed that each subband is quantized with a single quantizer. However, zero-tree coding technique violates an assumption of equal number of bits assigned to all coefficients in a subband. Number of bits used for each coefficient is determined uniquely by its magnitude, irrespective of the subband it belongs to. Another problem is that a significant part of the bit budget, especially at a low bitrate, is spent to encode the tree information, or locations of nonzero coefficients. Thus, the same reduction in distortion can be achieved at very different prices in terms of bitrate. In the next section, we introduce a filter characteristic more suitable for the use with zero-tree coding approach.

### III. DISTORTION AMPLIFICATION FACTOR

Since we use a nonorthogonal transform, the energy preservation property is violated. To evaluate how much quantization error, introduced in the transform domain, is amplified by the filter bank, let us consider an equivalent scheme corresponding to the subband decomposition (see Figure 1). The equivalent filters can be expressed in the following form:

\[ h_k = h_L(n) \ast h_L(n/2) \ast \ldots \ast h_L(n/2^{k-1}) \ast h_H(n/2^{k-1}) \]  
for \( k = 1, \ldots, N - 1 \)
\[ h_k = h_L(n) \ast h_L(n/2) \ast \ldots \ast h_L(n/2^{k-2}) \ast h_L(n/2^{k-1}) \]  
for \( k = N, \)  
(5)

where \( h_L \) and \( h_H \) are the low-pass and high-pass filters of the original iterated filter bank. \( h_k \) is called the \( k \)-th iteration of the filter. The output of every channel is given by:

\[ y_k(i) = h_k * x = \sum_{j=0}^{L(H_k)-1} h_k(j) x(i-j) \]  
(6)

Encoding introduces errors in the transform domain. To describe these errors, we can utilize an additive noise model. In this case, the subband signals after the decoding are given by:

\[ \hat{y}_k(n) = y_k(n) + \varepsilon_k(n), \]  
(7)

where \( \varepsilon_k \) is the white noise. Each channel after the synthesis stage has the energy:

\[
E_k \left[ \sum_{j=0}^{L(H_k)-1} g_k(j) \left[ y_k(i-j) + \varepsilon_k(i-j) \right]^2 \right] \\
= \sum_{u=0}^{L(H_k)-1} \sum_{v=0}^{L(H_k)-1} g_k(u) g_k(v) E_k \left[ y_k(i-u) y_k(i-v) \right] \\
+ \sum_{v=0}^{L(H_k)-1} g_k^2(v) E_k [\varepsilon_k^2] \\
= \sum_{u=0}^{L(H_k)-1} \sum_{v=0}^{L(H_k)-1} g_k(u) g_k(v) r_{\varepsilon_k}(u-v) + \sigma_{\varepsilon_k}^2 \sum_{v=0}^{L(H_k)-1} g_k^2(v)
\]  
(8)

where \( g_k \) is the iterated synthesis filter, defined in the same way as the iterated analysis filter; \( r_{\varepsilon_k} \) (\( n \)) is the correlation coefficient for the subband signal. The second term on the last line of (8) corresponds to the energy of the amplified errors. Errors contribute to the overall distortion according to the number of samples in the subband, thus the amplification of errors for every channel, \( B_k \), is given by:

\[ B_k = \alpha_k \sum_{v=0}^{L(H_k)-1} g_k^2(v) \]  
(9)

where \( \alpha_k = N_k/N \), is the degree of channel downsampling. We can use this amplification factor to evaluate errors of the zero-tree coding scheme. At any point of iteration, the absolute value of error can not be greater than the half of the current threshold:

\[ \sigma_{\varepsilon_k}^2 \leq \sigma_q^2 = \left( \frac{T}{2} \right)^2 \]  
(10)

Therefore, the total error has an upper limit:

\[ \sigma^2 = \sum_k B_k \sigma_{\varepsilon_k}^2 \leq \sum_k B_k \sigma_q^2 = \sigma_q^2 \sum_k B_k \]  
(11)
\[ \sum B_k \] is the distortion amplification factor. It shows how much the quantization errors are amplified by the synthesis stage of the filter bank. It can also be considered as some measure of nonorthogonality (for the orthogonal transform this factor is always one). The distortion amplification factor can predict filter performance, when compared filters have similar energy compaction.

\[ \sum B_k \] is the distortion amplification factor. It shows how much the quantization errors are amplified by the synthesis stage of the filter bank. It can also be considered as some measure of nonorthogonality (for the orthogonal transform this factor is always one). The distortion amplification factor can predict filter performance, when compared filters have similar energy compaction.

IV. EXPERIMENTAL RESULTS

To estimate the analyzed transform characteristics, a series of experiments was conducted. We have used several filters, which have been recently proposed in the literature [5,6,7,8]. Figure 5.2(a) shows relationship between the regularity and the PSNR for Lenna image, encoded at 0.3 bpp.

We can see, that there is a strong influence of the regularity on the performance when regularity is less than two. Two groups are indicated as being inconsistent with the usual behavior. Filters in these groups have high values of the distortion amplification factor and are far from being orthogonal. Figure 5.3 shows the perceptual artifacts of filters with low regularity. Irregular filters results in low PSNR and poor perceptual quality. One could easily notice blocking effects. Increase of the filter regularity above two does not give significant improvement of the performance.

Figure 2(b) demonstrates that the distortion amplification factor can be a reliable measure of the filter performance, especially, when combined with the minimum regularity criteria (synthesis wavelet regularity should not be less than one). The correlation between the distortion amplification and PSNR is quite high.

V. CONCLUSIONS

In this paper we demonstrated that filters with low regularity provide low values of PSNR. Therefore, some regularity of the synthesis wavelet is necessary, and if it is less than one, performance is poor. However, when regularity is higher than two, it does not have significant influence on the image quality.

Another measure of filter performance is the distortion amplification factor, which can be considered as a measure of nonorthogonality, since it reflects the amplification of the quantization errors by the synthesis stage of the filter bank. Experiments proved that the distortion amplification is highly correlated with PSNR. When combined with the minimum regularity requirement, it can be used to choose a suitable wavelet for zerotree wavelet coder.

REFERENCES

(a) Fig. 2. PSNR for Lenna image encoded at 0.3 bpp as a function of (a) regularity, and (b) distortion amplification.

(a) Fig. 3. Lenna image encoded at 0.3 bpp (a) with filter of regularity 0, PSNR 30.692, and (b) filter of regularity 1, PSNR 33.516.


