WAVELET CODING OF IMAGE USING QUADTREE REPRESENTATION AND BLOCK ENTROPY CODING

Wenxing Zheng Hua Nan

Laboratory of TV Technology
Beijing University of Posts and Telecommunications
Beijing 100088 , P.R.China

ABSTRACT

A new efficient image coding scheme , based on Quadtree Representation and Block Entropy Coding (QRBEC), for encoding the wavelet transform coefficients of images is presented. The property of HVS is also incorporated into the quantization process. In addition, how to flexibly control the quantization level as well as output bitrate of the coder is also investigated. The coding efficiency of the coder is quite competitive with the well-known EZW coder, and requires less computation burden. The proposed coding scheme can also be applied in image sequence coding, resulting in satisfactory performance.

1. INTRODUCTION

The main task of a wavelet codec for image compression is to code the quantized coefficients efficiently. In the literature , some good algorithms have been presented. In [1], a novel method known as EZW is shown to be promising. In [2], Lattice Vector Quantization (LVQ) is investigated. In this paper, a novel scheme better than above works is reported. In the proposed coder, wavelet coefficients are quantized and encoded within the framework of quadtree representation. The proposed coder owns the merit of high coding efficiency and easy implementation .

2. QUADTREE REPRESENTATION

The original fullband image is transformed into individual frequency bands using compactly supported biorthogonal wavelet filters [3]. In the proposed coder, separable lowpass and highpass filters are used to decompose the image in the horizontal and vertical directions. In this paper, the image is decomposed into 4 resolution levels to produce 13 subbands. The lowest frequency band is denoted as W_0 , others are denoted as W_1 - W_{12}. Subbands W_1 - W_{12} have the following properties: (1) After quantization, the zero coefficients are usually adjacent to one another, and appear as local regions of various size and form. This is due to the fact that natural scene has flat regions, the wavelet coefficients corresponding to these regions usually have small amplitudes, and are liable to be quantized to zero. (2) After quantization, non-zero coefficients usually cluster. The reason is that non-zero coefficients correspond to edges and textures in the original image, while these components are usually localized. Utilizing above properties, one can encode each subband efficiently .

Subband W_0 is DPCM coded as usual. For suband W_k , $k \geq 1$, it is split into many B_k by B_k blocks , where B_k is a preset constant, and takes the value of 16 ,8 or 4, depending on which subband is under consideration. Such a B_k by B_k coefficients block is denoted as $\bar{V}=[\ \bar{V}(m,n)]$, m,n = 1 - B_k. In the following , a *data block* is called a *zero block* if all its *elements* are zero, otherwise it is a *non-zero block* . $\bar{V}$ is quantized to V using a positive *stepsize* S_k as follows:

$$V(m,n) = \mathrm{Int}\{\ (2*|\bar{V}(m,n)| + S_k)/(2*S_k)\ \}* \mathrm{sign}(\bar{V}(m,n)) \quad (2.1)$$

where V(m,n) is the element of V , Int(x) is the integer part of x, and sign(x) is the symbol function. At the decoder, the reconstructed coefficients block is Y with $Y(m,n) = V(m,n)*S_k$, m , n = 1 - B_k .

When V is obtained, we apply a *quadtree decomposition* on it . If V is a zero block , 1-bit symbol “0” is assigned to it , and the quadtree decomposition is terminated. Otherwise , 1-bit symbol “1” is assigned to this non-zero block , and this block is *subdivided* into four ($B_k/2$) by ($B_k/2$) *smaller* sub-blocks . The subdividing process can be *repeated recursively* until any 2 by 2 block is encountered . For a 2 by 2 block , if it is a zero block , it is coded using 1-bit symbol "0", otherwise, it is coded using 1-bit symbol "1" followed with a *variable length code* "C" which is obtained using *block entropy coding* detailed in next section .

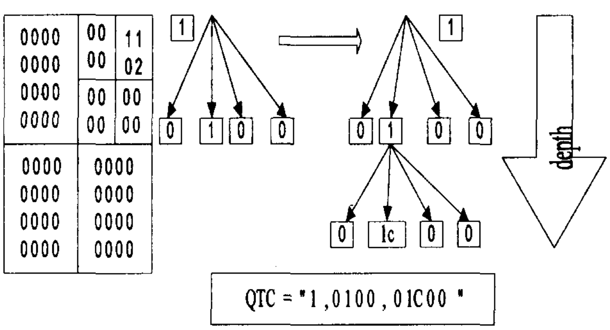

Fig.1 : Formation of a quadtree of depth 3 . The symbol C denotes the variable length code for encoding the 2 by 2 non-zero block located at the up-right corner .

A simple example of quadtree decomposition is illustrated in Fig.1. On the left is a 8*8 quantized coefficients block, and on the most right is the corresponding quadtree representation. The tree is scanned in a *breadth-first* order, i.e., from top to bottom, from left to right, which produces a *variable length code* called QTC (QuadTree Code) *to be transmitted to the decoder* . An example of QTC is shown in Fig.1 . From this figure, one can see that *various size zero blocks* (e.g. 2 by 2, 4 by 4) are coded as *zero leaf-nodes* of a quadtree in *various depth* using a few bits. Obviously, using quadtree representation can encode zero blocks very efficiently. Moreover, for a natural image, the majority of its coefficients are quantized to zero , so one can see that the coding efficiency of the coder is high .

3. BLOCK ENTROPY CODING

In the proposed coder, in order to further improve coding efficiency, we encode a 2 by 2 non-zero block using variable length code rather than encode four quantized coefficients *separately*. Experimental results show that, to encode the same 2 by 2 non-zero block the proposed algorithm costs much less bits as compared with encoding four elements separately using traditional Huffman codes.

The block entropy coding algorithm is described below. First, let T=2, and $M=T^2$. Consider a T by T wavelet coefficients block $\bar{X}=[\ \bar{x}(p,q)\ ]$ which will be quantized to X = [x(p,q)] using formula (2.1). X is called a point lying on pyramid P(M,R) if it satisfies $\sum_{p=1}^{T}\sum_{q=1}^{T}|x(p,q)| = R$, where R, the *pyramid radius*, is a non-negative integer. Furthermore, let N(M,R) be the number of points on P(M,R).

As well known, the probability density function (PDF) of $\bar{X}$ can be modeled well by i.i.d. Laplacian distribution , and **(i)** the equal probability surface of $\bar{X}$ is so-called Laplacian pyramids, **(ii)** the PDF of pyramid radius is biased towards zero, see [4]. Based on (i) and (ii) , X which is the *quantized version* of $\bar{X}$ has the corresponding properties: **(I)** the points on P(M,R) have *approximately equal* occurrence probability. **(II)** the PDF of integer random variable R is *biased towards zero* strongly, especially for higher frequency subbands. Based on property (I) and (II), one can encode X efficiently.

For each point X on P(M,R), an *unique address* can be assigned to it according to a *fast enumeration* method [4], and from this address , the components of X will be easily retrieved by the decoder. To explain in detail how to code a point X on P(M,R), $R \geq 1$, the following steps are required :

step-1: Compute pyramid radius $R = \sum_{p=1}^{T}\sum_{q=1}^{T}|x(p,q)|$, and then, R is *entropy coded* using a Huffman codebook H. Based on property (II), the entropy code for smaller R is shorter, while that for larger R is longer. Note that each subband has its own Huffman codebook in order to suit its own statistical characteristics. Denote the Huffman code for R as C(R).

step-2: An unique address which gives the position of X on P(M,R) is obtained by the fast enumeration algorithm [4]. This address is in the integer interval [0, N(M,R)-1]. Based on property (I), this address can be represented using $\lceil \log_2(N(M,R)) \rceil$-bits *fixed length code* without loss of coding efficiency. Denote this fixed length code as C(A).

Finally , a hybrid code "C" for encoding a 2 by 2 non-zero block is the combination of C(R) and C(A), with C(R) being the prefix, C(A) being the suffix.

For more details about the fast enumeration algorithm and the calculation of N(M,R) mentioned above, see [4].

4. OPTIMAL BITRATE CONTROL

In formula (2.1), there is a *stepsize* S_k for each subband W_k. Obviously, the output bitrate of the coder is controlled by $\{S_k\}$. So the problem is how to get the *optimal* $\{S_k\}$ for the image to be coded under a target bit rate ? The rate distortion (R-D) performance of the coder for subband W_k can be reasonably modeled by

$$D_k = \alpha * \sigma_k^2 * 2^{-\beta * R_k} \quad (4.1)$$

where D_k , σ_k^2 and R_k are quantization error , variance and encoding bitrate for subband W_k , respectively . α and β are parameters that depend on the particular coder structure and the PDF of the data. In another point of view , D_k can also be approximated over a range of bitrate by

$$D_k = (1/12) * S_k^2 \quad (4.2)$$

This is due to the fact that formula (2.1) is used to *uniformly* quantize each coefficient . Now, one can write that

$$(1/12) * S_k^2 = \alpha * \sigma_k^2 * 2^{-\beta * R_k} \quad (4.3)$$

Using (4.1) through (4.3), one can addresses the following optimization problem:

$$\text{Min}\left(\sum_k \omega_k * D_k * \lambda_k \right) \quad (4.4)$$

$$\text{subject to : } \sum_k R_k * \lambda_k = R \quad (4.5)$$

where R is the target bit rate available for encoding the image . ω_k which is determined empirically , is the visual weighting factor for subband W_k . λ_k is in proportion to the number of samples in subband W_k , i.e. λ_{10} - λ_{12} =1/4 , λ_7 - λ_9 =1/16 , and so on.

Using (4.2) and (4.3), one can rewrite (4.4) and (4.5) *in terms of* S_k , and the *optimal solution* for S_k is obtained by applying the method of *Lagrange multiplier*, and takes the *form* of

$$S_k = \Delta / \sqrt{\omega_k} \quad (4.6)$$

where Δ can be derived from a function with input variables $\{\sigma_k^2\}$ and R . It is through *varying* Δ (*consequently* , $\{S_k\}$) that one can flexibly control both the *quantization level* and the coder's *output bitrate* . In practice, Δ is quantized using 8-bits, and sent to the decoder as side information to inform the decoder of the correct stepsizes .

5. EXPERIMENTS

The proposed wavelet codec is extensively tested. Here, the results obtained with the 8-bit grey level image "Lena" are reported. In Table-1 , Psnr obtained by QRBEC and DCT at different bitrates are compared. Note that QRBEC significantly outperforms DCT, with Psnr gains ranging from 4.13 dB to 1.31 dB. For a bitrate of 0.2 bpp, the visual quality obtained by QRBEC is much better than that obtained by DCT, see Fig.2-3. When bitrate is 0.25 bpp, Psnr obtained by EZW coder[1] is 33.17dB, whereas that obtained by QRBEC is 33.28 dB, slightly better than the former. When output bitrate is 0.167 bpp, Psnr from Barlaud's LVQ[2] is 30.2dB, whereas that from QRBEC is 31.69 dB, with a 1.49 dB gain.

bitrate (bpp)	0.2	0.25	0.3	0.35	0.4	0.45	0.5
QRBEC	32.24	33.28	34.02	34.78	35.52	36.01	36.39
DCT	28.11	30.10	31.56	32.82	33.79	34.57	35.08

Table - 1 Psnr (dB) performance comparison between QRBEC and DCT

Fig.2 : Lena coded at 0.2 bpp using QRBEC.

Fig.3 : Lena coded at 0.2 bpp using DCT.

In addition, the proposed coder has been applied in video coding. This coding algorithm is applied on the motion compensated prediction error signal. A *buffer* is placed between the coder and the channel which *regulates* the output bitstream. The problem of buffer overflow and underflow is avoided by adjusting the parameter Δ stated above on a *picture by picture* basis . When the channel rate is set to 4Mb/s , the average Psnr for Y signal of sequence "flower garden" (720*576 , 50 fields/sec , 4:2:0 format) is 30.8 dB , and that of "mobile" is 29.4 dB. Moreover, the perceptual quality of the reconstructed video is quite satisfactory.

6. REFERENCES

[1] J.M.Shapiro."Embedded image coding using zerotrees of wavelet coefficients". IEEE Trans on SP. Vol 41,pp 3445 − 3462. Dec,1993.

[2]M.Barlaud. "Pyramid lattice vector quantization for multiscale image coding" . IEEE Trans on IP vol.3 No.4. pp367 − 380. July 1994.

[3] A.Cohen. I.Daubechies. J.C. Feauveau ."Biorthoganal bases of compactly supported wavelets". Commu on Pure and Applied Mathematics, Vol XLV, pp 485 − 560. 1992.

[4] T.R.Fisher "A pyramid vector quantizer " .IEEE Trans on Information Theory, Vol.32 No.4. pp568 − 583.July 1986.