Wavelet Based Image Compression System

[1]Bo Zhang, [2]Xiuwen Liu, [1]Li Shen
[1]Department of Electrical Engineering
[2]Department of Computer and Information Science
The Ohio State University
Columbus, OH 43210

I. Introduction

In this project, we developed a wavelet based image compression system primarily for gray-scale images. After wavelet transformation, subbands with similar statistical characteristics are grouped as a subblock. Each subblock is uniformly quantized with respect to an expected compression rate and then converted to a byte stream in such a way that most zeros are consecutive. The byte stream is encoded using run length encoding (RLE), mainly where the compression is achieved. The reconstruction is the inverse of the compression process. Fig. 1. shows an overview of the complete system.

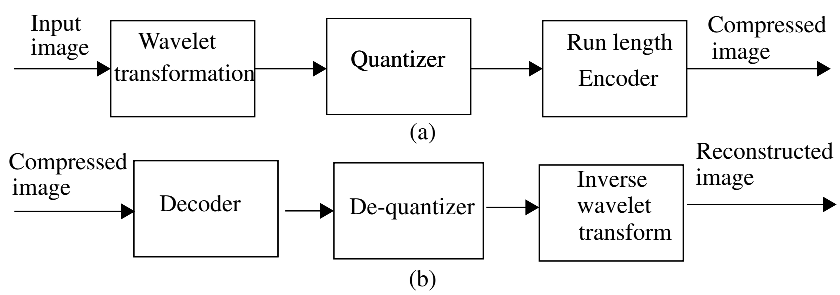

Fig. 1. Wavelet based compression system: (a). Compression process. (b). Reconstruction process.

Due to the properties of wavelet transform, we achieved high compression rate with small visual distortion. At a compression rate of about 8:1, the original and reconstructed images are visually indistinguishable. At a rate about 16:1, most of details such as textures are preserved with some noticeable blurring effects along long edges. Experimental results are provided for comparison and evaluation.

As to the computational complexity, our entire system is linear in term of the size of input, i.e., the number of pixels in the input image, and thus could be implemented very efficiently. Because most parts of the system are implemented in Matlab for this project, it is very difficult to evaluate the system complexity quantitatively in time.

II. Design method

1. Wavelet transformation

Wavelet decompostion is a multiresolution representation of the original image. The original image is decomposed into a set of biorthogonal subbands. The decomposition is along vertical and horizontal directions at different scales and is computed with a pyramidal algorithm using convolutions of quadrature mirror filter pairs, which is similar to a subband coding method.

In our system, (9,7) biorthogonal wavelets are used. A four stage wavelet transform is applied to the whole image. Fig. 2. shows a two-stage decomposition for illustration. The image is decomposed into seven subbands. Each subband is obtained by filtering the subband with lowest frequency components in the previous layer in the pyramid with two separable filters.

There are several advantages using biorthorgnal wavelet decompositions. First, high compression rate can be achieved with less visual distortion. Because human eyes are more sensitive to low frequency wavelet coefficients, weighting factors are used in the pyramidal algorithm so that those components are kept more. Second, wavelet transform is linear in term of the input image size. It can be implemented efficiently. Also because wavelet transform can be applied to the whole image as efficiently as breaking into small blocks, block artifacts are avoided.

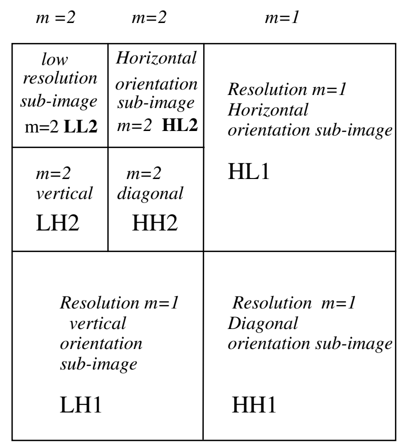

Fig. 2. Image decomposition for two stages. In this case, subblock 1 is LL2, subblock 2 consists of LH2, HL2, HH2, and subblock 3 consists of LH1, HL1, HH1.

2. Quantization

After wavelet decompostion, the coefficients are quantized with respect to an expected compression rate. As mentioned before, the pyramidal algorithm has a weight factor for each subband, and thus we used a global threshold for the entire image, which is calculated from the histogram of the coefficients. Then those coefficients with absolute values smaller than the threshold will be quantized to zero. To simplify the implementation, subbands that have similar statistical properties are grouped together as one subblock, as illustrated in Fig. 2. For a four-stage implementation, the wavelet coefficient matrix is grouped into five subblocks. A uniform quantizer is applied on each subblock.The quantization bit and quantization step are calculated based on the threshold and the statistical distribution of the coefficients. Due to the fact that the coefficients are bipolar and are not symmetric with respect to zero in general, an additional parameter is used to compensate the offset. Therefore, the smallest coefficient will be quantized to be zero, the largest coefficient will be quantized to the largest possible number with the given quantization bits, and the coefficients whose absolute values are smaller than global threshold will be quantized to the offset. For the four given images, we found that in most cases, the two subblocks with lowest frequencies are quantized to about eight bits, while other subblocks are quantized with less than eight bits

To make most zeros consecutive, each subband must be scanned in a certain way, which is determined by the filters used in the horizontal and vertical directions. For example, the subband HL1 shown in Fig. 2 is obtained by low-pass filtered vertically and high-pass filtered horizontally and thus it should be read vertically to produce more consecutive zeros. Similarly the scanning order for other subband is determined.

3. Coding

Here based on the assumption that most zeros are consecutive, and nonzeros are randomly distributed, we chose run-length encoding as the encoding method. The encoder converts the byte stream of each subblock produced by the quantization process to a bit stream and encodes long repeated bytes using counters. Specifically, our algorithm works as follows. 1). Find the number that occurs most, which will be encoded most effectively using counters. Find also the number that occurs least. This number will be used as a special flag to indicate the following will be an encoding of a long sequence. If this number occurs in the byte stream, a bit immediately following it will indicate whether it is a special flag or the number itself. 2). To be more effective for long repeated sequences of nonzeros, similar to zero sequence encoding, another special flag may be used to indicate that the following is an encoding of a repeated sequence. Here the repeated value itself will also be encoded along with the number of occurrence. 3). Several different encoding parameters are tried first and the optimal one is used for actual encoding.

The major advantage of this method is that it is very computationally effective for both encoding and decoding. We can see that it is linear in term of the input size. This part is implemented using C language. For a 512 x 512 image, the decoding can be done less than 0.2 second on a HP machine. Also it achieves a very high compression rate. As shown in next section, the loss because of encoding is small.

III. Experimental results

1. Experimental results

The four given images are compressed and reconstructed using our system. Fig. 3 shows the original and reconstructed images along with difference image and the error distribution. The compression rate is about 7.5:1. The error distribution is good and the original and reconstructed images are visually equivalent.

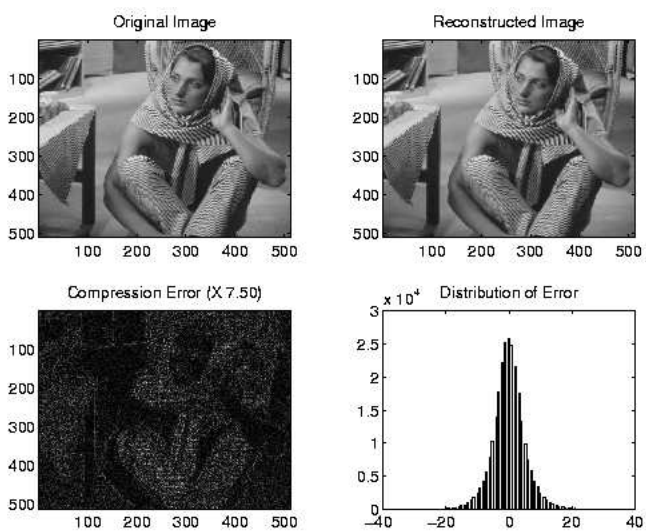

Fig. 3. The reconstructed image and compression error.

Fig. 4 shows the original images and reconstructed images for an expected compression rate of 8:1 and 16:1. For all the four images, the reconstructed images are very close to the original at a compression rate of about 8:1. When compressed by a factor of 16:1, some noticeable blurring effects are generated along especially long edges.

2. Objective fidelity criteria

Two different and yet related objective criteria are used. *Error per pixel* is defined as the average of pixel difference between the original image and reconstructed one. *Peak signal noise ratio*

(PSNR) is defined by the following formula, where f(x, y) is the original image and $\bar{f}(x, y)$ is the reconstructed version.

$$PSNR = 20Log_{10}\left(255/\sqrt{\left(\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}[\hat{f}(x,y)-f(x,y)]^2\right)/MN}\right)$$

The two criteria are calculated for each image at two compression rates and the results are shown in the following tables.

Criteria	pepper	barb	tiger	leopard
compression rate	6.6388:1	7.5237:1	6.5312:1	6.9817:1
PSNR(dB)	36.85	33.99	32.86	32.80
error per pixel	2.8458	3.7703	4.3798	4.3231

Criteria	pepper	barb	tiger	leopard
compression rate	14.4972:1	16.2947:1	13.0966:1	13.006:1
PSNR(dB)	34.65	29.50	29.85	29.33
error per pixel	3.723	6.067	6.060	6.239

3. Subjective assessment

All of us agree that we can hardly see the difference between the original and reconstructed images at a compression rate of about 7:1. Even small details such as tiny spots are also preserved without any noticeable difference. When compressed at a rate of about 15:1, the textures and most of the details are still preserved very well. But along long edges such stripes in the tiger image we can see the blurring effects because the transition now is not as sharp as in the original image. This is consistent with the objective criteria. In general, when PSNR is more than 30 dB, the reconstructed image should be visually very close to the original image.

Because we applied a wavelet transformation through the entire image, there is no block artifact, which JPEG and other small block-based methods may suffer.

IV. Conclusion

A wavelet based compression system was presented. At a compression rate of about 8:1, the PSNR for all testing images is above 30 dB. At that rate, the reconstructed image is visually equivalent to the original image. At a compression rate of 16:1, most of details of the image are preserved very well only with some blurring effects along long edges.

Conceptually, the system can be implemented very efficiently because the entire system is linear in term of the input image size. By using more sophisticated quantization method and encoding schema, the compression rate could be improved even more. Because the wavelet coefficients have certain distribution, an optimal quantization can be developed. Also Huffman coding might be able to give some additional compression when applied on the run-length coded stream.

References

[1]. M. Antonini, and et al, “Image coding using wavelet transform”, *IEEE Transactions on Image Processing,* vol. 1, no. 2, pp. 205-220, 1992.

Fig. 4. The original image and reconstructed ones. Left column: original image, middle: compression rate about 7:1, and right: compression rate about 15:1. Please refer to the tables for accurate compression rate and objective assessment.