Multi-Resolution Structure Color Image Compression Using DWT and VQ

Wen-Yuan Chen
Department of Electrical Engineering, National Cheng Kung University Tainan, 70101, Taiwan, R.O.C.
cwy@chinyi.ncit.edu.tw

Chin-Hsing Chen
Department of Electrical Engineering, National Cheng Kung University Tainan, 70101, Taiwan, R.O.C.
chench@eembox.ncku.edu.tw

Ghin-Tang Lin
Department of Electronic Engineering National Chin-Yi Institute of Technology, Taichung, Taiwan, R.O.C.
johnny@symbol.com.tw

Abstract

Quad-tree structure is an excellent algorithm for block based image processing, especially image compression. The structure divided an image block into homogenous blocks of size 16×16, 8×8, 4×4, and non-homogenous blocks of size 4×4. The bottom-up Quad-tree structure is very suitable for very low bit rate image compression of high performance using DWT and VQ. The multi-resolution decomposition by Discrete Wavelet Transform (DWT) is an efficient analysis tool for image processing. In general, 4-level decomposition is suitable for color image compression. In the paper, an adaptive clustering skill for generating multi-resolution codebooks by vector quantizing wavelet transformed sub-images of four level of decomposition is proposed to color image compressed. The quality of the recover is varying reconstructed from a multi-resolution codebook. Experiment results shown that the proposed scheme can achieve very low bit rate with high performance for color image.

Keyword: DWT, VQ, Multi-resolution codebook, bottom-up Quad-tree structure.

1. Introduction

Color images are everywhere high quality color images with higher resolution and huge memory spaces are the favorite of modern people. Color image compression is an important technique to reduce the image space and retain high image quality. The popularly used JPEG is the current color image compression standard which is based on Discrete Cosine Transform and run-length encoding procedures. But, the wavelet transform based image compression algorithm has excellent compression performance and is the core technique for the JPEG-2000 standard.

Shapiro's Zero-tree data structure of wavelet coefficients [11] has been widely used in image compression and data hiding fields. Rate-distortion optimization [3] and statistical characterization schemes [4] under Zero-tree structure were proposed. In addition, based on the test results of 24 proposals submitted by various companies and universities, the JPEG-2000 [5] committee decided on a framework using wavelet Sub-band coding for the new standard. In fact, the test results shown that wavelet coding is superior than DCT coding in both subjective and objective fidelity of image quality.

Color image compression using wavelet coding or hybrid of wavelet and other methods is extremely popular. L. T. Leung and L. M. Po [7] proposed a method to encode large size blocks instead of using conventional bilinear interpolation. In the encoding process, the image is first divided into 16x16 non-overlapped blocks and checked whether the reconstruction is sufficient good by calculating the mean square error (MSE), if not, the block will be subdivided into four 8x8 sub-blocks with equal size for further encoding. Based on the encoding strategies, the 16×16, 8×8, 4×4, and 2×2 blocks are compressed using different blocks size VQ coding. The method produces no significant blocking artifacts at the block boundaries.

M. J. Tsai [1] presented a low complexity intrasubband based coding method by only addressing the information within the sub-band. In the paper, the YIQ plane is used in place of the RGB plane to increase the efficiency and uniform scaler quantization. Data conversion is performed for individual sub-band and quantized coefficient. A "Stack-Run Coding" algorithm different from JPEG's run grouping is hired to achieve higher performance compression.

The neural-network model mixed with principle component analysis (PCA) is a suitable color image compression scheme. C. Clausen and H. Wechaler [6] presented a color image compression scheme using PCA and back-propagation learning. In the scheme, a neural-network model based upon Sanger's algorithm is used to represent an image in terms of principal components and back-propagation algorithm restore the original representation. The PCA method produces a black and white image with the same size as the original color image to replace the three vectors of the RGB component.

In our paper, a wavelet coding scheme in which VQ mixed with multi-resolution codebooks is proposed for a low bit rate high performance color image compression. The remainder of this paper is organized as follows. In Section 2, the encoding algorithm is presented. Empirical test is presented in Section 3. Finally, Section 4 concludes this paper.

2. Encoding algorithm

2.1 RGB planes to YIQ plane mapping

The Red, Green, and Blue colors (RGB) are the three primary elements of a color image. From energy distribution analysis, the RGB component exist a lot of redundant information that can be reduced. Therefore, the RGB planes of a color image is usually mapped YIQ planes to reduce the redundant by formula (1) as follows:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

Where Y is the brightness, or the luminance, I is Hue, and Q is the saturation. Because brightness constitutes about 93% energy of a color image, so that color image processing is usually performed in the Y plane. The YIQ mapping in formula (1) has an inverse. Its inverse maps the YIQ signals back to the RGB planes, via the following formula.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & 0.956 & 0.621 \\ 1.000 & -0.272 & -0.647 \\ 1.000 & -1.106 & 1.703 \end{bmatrix} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix} \quad (2)$$

2.2 Multi-resolution by DWT

The Haar transform is a simple and efficient wavelet. It is performed in several stages, or levels. The first level is the mapping H_1 defined by $f \to (a_1 \mid d_1)$, from a discrete signal f to its first trend a_1 and first fluctuation d_1. The mapping H_1 has an inverse mapping back to signal f, via the formula

$$f = (\frac{a_1 + d_1}{2}, \frac{a_1 - d_1}{2}, \ldots \frac{a_{n/2} + d_{n/2}}{2}, \frac{a_{n/2} - d_{n/2}}{2}).$$

A discrete signal begins with a very lower resolution signal and successively adds on details to create higher resolution versions. It ends with a complete synthesis of the signal at the finest resolution. This is known as multi-resolution analysis (MRA) as shown in Fig. 1. MRA is the heart of wavelet analysis. In general, if the number N of signal values is divisible k times by 2, then a k-level MRA described as follows can be performed on the signal f.

$$f = A^k + D^k + \ldots + D^2 + D^1. \quad (3)$$

$$A^k = (f \cdot V_1^k)V_1^k + \cdots + (f \cdot V_{N_k}^k)V_{N_k}^k \quad (4)$$

$$D^k = (f \cdot W_1^k)W_1^k + \cdots + (f \cdot W_{N_k}^k)W_{N_k}^k \quad (5)$$

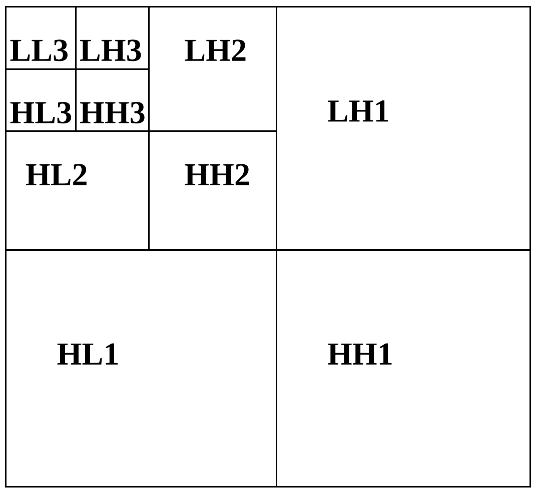


Fig. 1. The 3-level MRA of DWT.

In the paper, the original images will be transformed by DWT. Then except the LL band of the last level, the others bands will be compressed by VQ coding.

2.3 Bottom-up quad-tree partition and multi-resolution codebook

A block based coding system usually uses variable size image partition. The most popular partition mechanism is the well know Quad-tree segmentation which subdivides an image into four equal-size blocks and test a criterion, if met then continues to divide each subdivides until the criterion not met or minimum size block is reached. The criterion is a split-decision function in which a threshold value is compare with the calculated result of image blocks. The split-decision function and threshold value will affect the fidelity of the reconstruction and the encoding time.

For an image block of size $n \times n$, expressed as $X = \{x_1, x_2, ...x_m; m = n^2\}$, let $\overline{X}$, $\overline{X}_H$, $\overline{X}_L$ represent the block mean, high mean, low mean, respectively, which are defined as.

$$\overline{X}_H = \frac{1}{q} \sum_{x_i \geq \overline{X}} x_i, \quad (6)$$

$$\overline{X}_L = \frac{1}{m-q} \sum_{x_i \prec \overline{X}} x_i, \quad (7)$$

$$S_{th} = \left| \overline{X}_H - \overline{X}_L \right|. \quad (8)$$

Where q is the number of pixels whose intensity is equal to or greater than the block mean, $\overline{X}_H$ represents the mean of pixel intensity equal or greater then $\overline{X}$, $\overline{X}_L$ represents the mean of pixel intensities less then $\overline{X}$, m is the number of pixels of an image block, and S_{th} represents the threshold value in the split-decision function. If ($\overline{X}_H$ - $\overline{X}_L$) of the block is greater than the value of we set usually is 20, the block is called non-homogenous otherwise is called homogenous.

In the paper, the bottom-up Quad-tree partition scheme shown in Fig. 2 was used to create the multi-resolution codebook. The bottom-up Quad-tree partition structure, first, divides the original input image into several large blocks of size 16×16, then, further subdivides each large block into 16 small blocks of size 4×4. We test the small block using the split-decision function, if it is non-homogenous a small block is formed otherwise combined 4 small blocks to construct a block of middle size 8×8 and test if homogenous. By the rule, block size from small to large will be determined. Furthermore, a multi-resolution codebook was created using LBG algorithm by training these image blocks. In the paper, we use threshold value S_{th} =20 and test each block by equations (6)-(8) to decide what block size will be formed.

A color image of RGB components is mapped to YIQ components firstly. The IQ components are then sub-sampled to $\frac{1}{4}$ of the original size. Finally Y and $\frac{1}{4}$ IQ planes go through bottom-up Quad-tree segmentation to obtain four classes images array that are the homogenous blocks of 16×16, 8×8, 4×4, and non-homogenous block of size 4×4. The LBG algorithm is used to generate the four kinds of codebooks by training the four kinds of image blocks respectively.

3. Empirical tests

According to the human visual system, some amount of distortion between the reconstructed image and the original-image is allowed. In this paper we employ the $PSNR$ to indicate the performance of the method. The $PSNR$ is given by

$$PSNR = 10 \log_{10} \frac{255 \times 255}{\frac{1}{N \times N} \sum_{i=1}^{N} \sum_{j=1}^{N} (x_{i,j} - \hat{x}_{i,j})^2} \quad (9)$$

Where $x_{i,j}$, $\hat{x}_{i,j}$, are the gray-scale values of the original and the reconstructed images of size $N \times N$ respectively.

Color images “F16l”, “Lena”, “Splash”, and “Tiffany” with size 512×512, 8 bits for each color component were used as the test images. These images as shown in Fig. 3 are used as the multi-resolution codebook generation training images. The PSNR of the RGB component and its average of the reconstructed image are listed in tables 1. According to Tab. 1, we found the PSNR fall in the range from 29.61 to 31.39 and the compression ratios are all about 40. The Tab. 2 is the PSNR of the YIQ component of test images. The number of blocks of the decomposition of test images is shown in Tab. 3. The compression ratio is the most important parameter for color image compression. Although, difference images has difference reconstructed quality and compression ratio, but they all vender good performance in the bottom-up Quad-tree segmentation and VQ coding with the multi-resolution codebook algorithm.

4. Conclusion.

In this paper, we map a color image of RGB components to YIQ components to reduce redundancy. Further, according to energy proportion, we sub-sample the IQ components to decrease their weighting. The Haar Wavelet is hired to decompose the YIQ components into low-low (LL), low-high (LH), high-low (HL), high-high (HH) sub-bands. Because the LH, HL, HH bands are from edges of the image, which include only a small proportion energy that a lot of redundancy can be removed by using the bottom-up Quad-tree partition VQ coding structure and multi-resolution codebooks. Experimental results shown that our method is feasible for color image compression.

References

[1] M. J. Tsai, “Very Low Bit Rate Color Image Compression by Using Stack-Run-End Coding”, *IEEE trans. Consumer Electronics*, Vol. 46, No. 2, pp. 368-374, 2000.

[2] I. K. Kim and R. H. Park, “Still Image Coding Based on Vector Quantization and Fractal Approximation,” *IEEE Trans. Image Processing*, Vol. 5, No. 4, pp.587-597, 1996.

[3] J. Li and S. Lei, “An Embedded Still Image Coder with Rate-Distortion Optimization,” *IEEE Trans. Image Processing*, Vol. 8, pp. 913-924, July 1999.

[4] R. W. Buccigrossi and E. P. Simoncelli, “Image Compression via Joint Statical Characterization in the Wavelet Domain,” *IEEE Trans. Image Processing*, Vol. 8, pp. 1688-1701, December 1999.

[5] O. K. Al-Shaykh, I. Moccagatta, and H. Chen, “JPEG-2000:A New Still Image Compression Standard,” *IEEE conference* , pp. 99-103 1998.

[6] C. Clausen and H. Wechsler, “Color Image compression using PCA and back-propagation learning,” *Pattern Recognition*, Vol. 33, pp.1555-1560, 2000

[7] L. T. Leung and L. M. Po, “Hierarchical Predictive VQ for Color Image Compression,” *International Symposium on Signal Processing and Applications*, ISSPA, Gold Coast, Australia, pp.25-30, August 1996.

[8] C. C. Wang and C. H. Chen, “low-complexity fractal-based image compression using two-stage search strategy,” *Optical Engineering*, Vol. 38, No. 6, pp.1006-1013, 1999.

[9] S. C. Pei and Y. S. Lo, “Color Image Compression and Limited Display Using Self-Organization Kohonen Map,” *IEEE Trans. Circuit and System for Video Technology*, Vol. 8, No. 2, April 1998.

[10] C. K. Yang, J. C. Lin, and W. H. Tsai, “Color Image Compression by Moment-Preseverving and Block Truncation Coding Techniques,” *IEEE Trans. Communications*, Vol. 45, No. 12, December 1997.

[11] J. M. Shapiro, “Embedded Image Coding Using Zerotrees of Wavelet Coefficients,” *IEEE Trans. Signal Processing*, Vol. 41, No. 12, December 1993.

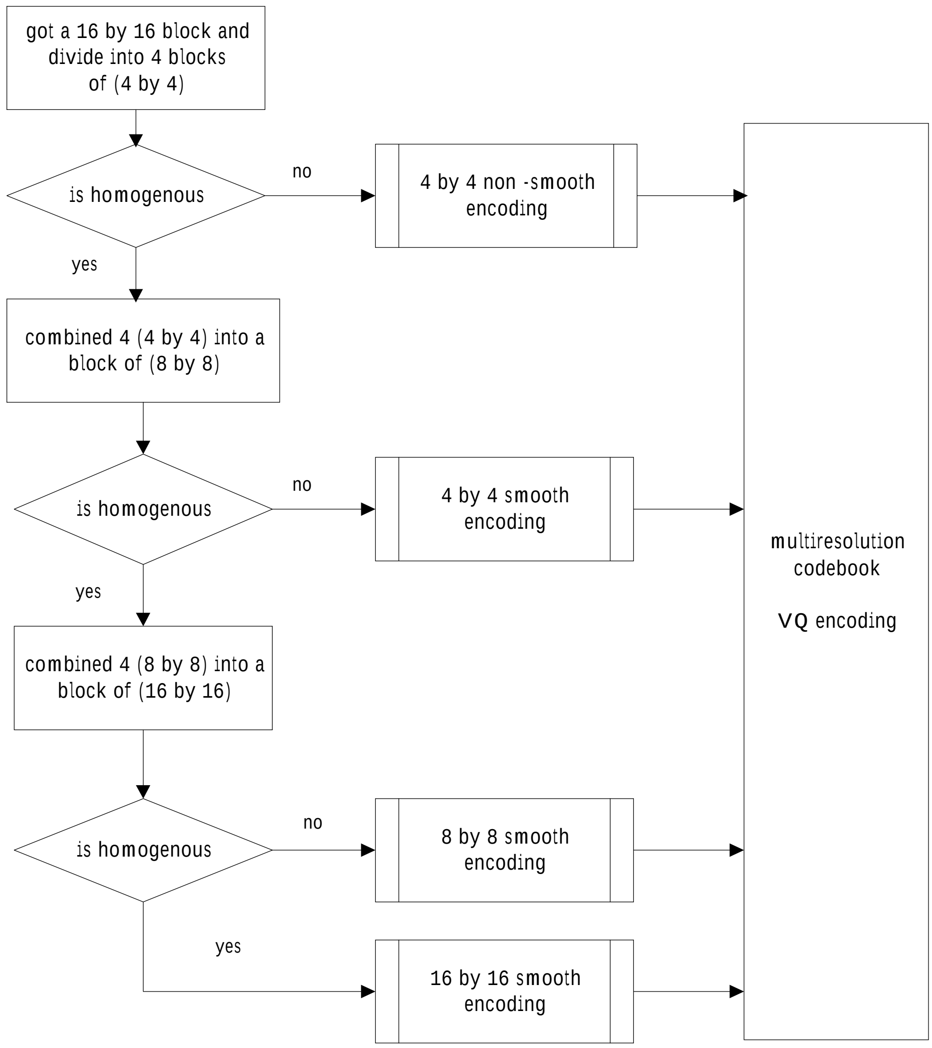


Fig. 2 The VQ encoding with bottom-up structure and the multi-resolution codebooks.

Tab. 1 Color compression results for image size 512×512.

Picture	R	G	B	Average	CR
F16	28.07	28.01	28.01	28.03	43.97
Lena	29.40	30.05	30.05	29.84	38.47
Splash	29.96	32.10	32.11	31.39	44.75
Tiffany	30.44	30.04	30.04	30.17	42.17

Tab. 2 The PSNR of the YIQ components.

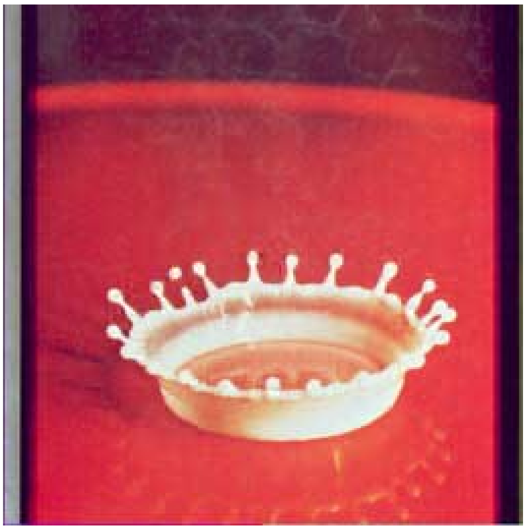
Picture	Y	I	Q	Average	CR
F16	28.97	35.72	34.75	33.15	43.97
Lena	30.88	35.15	37.75	34.59	38.47
Splash	34.20	32.77	36.60	34.53	44.75
Tiffany	32.72	35.15	32.96	33.61	42.17

Tab. 3 The block numbers of test images.

The size of block		16×16	8×8	4×4	non-homogenous 4×4
F16	Y	769	471	1068	872
	I	156	157	391	325
	Q	146	166	416	424
	Total	1071	794	1875	1621
Lena	Y	677	618	18453	1371
	I	94	230	742	674
	Q	69	251	829	903
	Total	840	1099	3024	2948
Splash	Y	834	326	783	697
	I	159	160	353	303
	Q	150	174	372	372
	Total	1143	660	1508	1372
Tiffany	Y	787	496	911	641
	I	58	300	839	873
	Q	180	97	309	263
	Total	1025	893	2059	1777

(a)

(b)



(c)

(d)

Fig. 3 the reconstruct Images. (a)F16 (512×512), (b) Lena (512×512), (c) Splash (512×512), (d) Tiffany (512×512).