

Combined DWT and DCT Image Compression Using Sliding RLE Technique

*Nushwan Y. Baithoon**

Received 25 , July, 2010

Accepted 23, October, 2010

Abstract:

A number of compression schemes were put forward to achieve high compression factors with high image quality at a low computational time. In this paper, a combined transform coding scheme is proposed which is based on discrete wavelet (DWT) and discrete cosine (DCT) transforms with an added new enhancement method, which is the sliding run length encoding (SRLE) technique, to further improve compression.

The advantages of the wavelet and the discrete cosine transforms were utilized to encode the image. This first step involves transforming the color components of the image from RGB to YUV planes to acquire the advantage of the existing spectral correlation and consequently gaining more compression.

DWT is then applied to the Y, U and V color space information giving the approximate and the detail coefficients. The detail coefficients are quantized, coded using run length encoding (RLE) and SRLE.

The approximate coefficients were coded using DCT, since DCT has superior compression performance when image information has poor power concentration in high frequency areas. This output is also quantized, coded using RLE and SRLE.

Test results showed that the proposed DWT DCT SRLE system proved to have encouraging results in terms of Peak Signal-to-Noise Ratio (PSNR), Compression Factor (CF) and execution time when compared with some DWT based image compressions.

Key words: Image Compression, Discrete Cosine Transform, Discrete Wavelet Transform, Run Length Encoding, Peak Signal-to-Noise Ratio, Compression Factor.

Introduction:

Image compression is important for the compact storing and fast transmission of images. To reach this target, numerous investigations were done to achieve this task, such as sub-block DCT and DWT coding. [1, 2] Although sub-block DCT coding is efficient for image compression, a block artifact sometimes appears when the low average bit rate is employed, which is an inherent deficiency, of sub-block DCT coding. A solution to such a problem might be with the use of DWT, although the resulting reconstructed image may not

necessarily be good for high picture quality coding. [3] In this paper, a combination of DWT and DCT with SRLE is introduced to improve image compression factor (CF), while keeping good quality level which is assessed using the conventional fidelity measure of peak signal-to-noise ratio (PSNR).

1. YUV Model

In order to invest in the high spectral correlation that is inherent in the YUV model and hence lower computational

*University of Baghdad, College of Science for Women, Department of Computer Science

complexities, the RGB input data is converted to YUV color space. In this model, Y denotes the luminance component; U and V are the two chrominance components. The luminance Y, equation (1), can be determined from the RGB model using the following relation:

$$Y=0.299R+0.587G+ 0.114B \quad (1)$$

From equation (1), it can be noted that the three weights associated with the three primary colors, R, G, and B, are not the same. Their different magnitudes reflect different responses of the human visual system (HVS) to different primary colors. Also, instead of being directly related to hue and saturation, the other two chrominance components, U and V, are defined as color differences, as shown in equations (2) and (3). [4, 5]

$$U=0.492(B-Y) \quad (2)$$

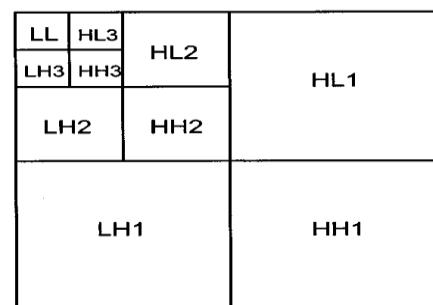
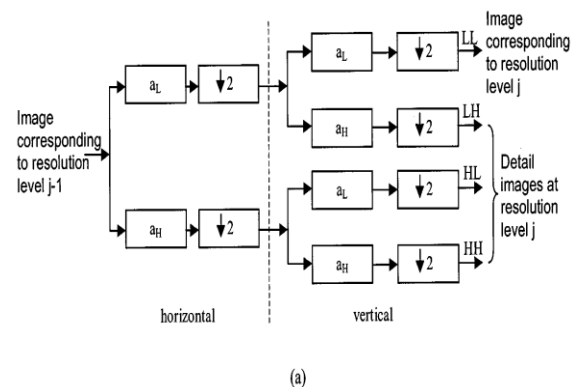
$$V=0.877(R-Y) \quad (3)$$

2. DWT Process

In DWT, an image can be analyzed by passing it through an analysis filter bank followed by a decimation process. This analysis filter bank consists of a low pass filter and a high pass filter at each decomposition stage. When a signal passes through these filters, it is split into two bands. The low pass filter, which corresponds to an averaging operation, extracts the coarse information of the signal. The high pass filter, which acts as a differencing operation, extracts the detail information of the signal. The output of the filtering operations is then decimated by two. [6, 7]

The 2-D wavelet transform can be accomplished by performing two separate 1-D transform, as shown in figure (1a). First, the image is filtered

along the x- axis and decimated by 2. Then, it is followed by filtering the sub-image along the y-axis and decimated by 2. Finally, the image data contents are split into four bands (denoted by LL, HL, LH and HH) after one level decomposition with further decompositions that can be achieved by acting upon the LL subband successively, and then the resultant image is split into multiple bands. [8] In this paper, Haar wavelet transform was applied. One way of accomplishing such decomposition is by the use of pyramid wavelet decomposition, as shown in figure (1b). This technique is a very efficient way of transforming significant visual data to the detail coefficients. Its computational complexity is about 30% higher than other decomposition methods, such as the quincunx method, but its image reconstruction abilities are higher. [9]



(b)
Fig.(1): (a) One Filter Stage in 2-D DWT
(b) Pyramid Wavelet Decomposition

3. DCT Process

In DCT, the input image is first partitioned into sub-blocks of size 8*8 pixels in dimension. Then each block is transformed from spatial to frequency domain using 2-D DCT basis function. The resulting frequency coefficients are quantized and finally output to a lossless entropy coder. DCT is an efficient image compression method because it can de-correlate pixels in the image, since the cosine bases are orthogonal, and condense most of the image energy into few transform coefficients. [10, 11]

Broadly speaking, compressing a set of correlated pixel values using DCT may be done by:

- Compute DCT coefficients of the image.
- Quantize the coefficients.
- Entropy encodes the quantized coefficients either by variable-length coding or arithmetic coding.

In DCT image compression, it is common to use two-dimensional DCT which can be described by an n*n transform matrix.

The most common 2D DCT is as given in equation (4).

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} I(x, y) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad (4)$$

Where
I(x, y) = Pixel value

$$\alpha(k) = \begin{cases} \sqrt{\frac{1}{N}} & \text{for } k = 0 \\ \sqrt{\frac{2}{N}} & \text{for } k = 1, 2, \dots, N - 1 \end{cases}$$

With, N being the image size.

For matrix representation of equation (4), equation (5) may be used, giving equation (6).

$$T_{i,j} = \begin{cases} \sqrt{\frac{1}{N}} & \text{for } i = 0 \\ \sqrt{\frac{2}{N}} \cos\left[\frac{(2j+1)i\pi}{2N}\right] & \text{for } i \neq 0 \end{cases} \quad (5)$$

$$C = T I T' \quad (6)$$

While the inverse transform (I), equation (7), is

$$I = T' C T \quad (7)$$

4. Quantization

DCT-based image compression relies on the quantization of the image's DCT coefficients. Quantization is the process of reducing the number of possible values of a quantity, thereby reducing the number of bits needed to represent it. [2]

A simple example of quantization is the rounding of real into integers. To represent a real number between 0 and 7 to some specified precision takes many bits. Rounding the number to the nearest integer gives a quantity that can be represented by just three bits.

In this process, a reduction of the number of possible values of the quantity (and thus the number of bits needed to represent it) at the cost of losing information. A "finer" quantization, that allows more values and loses less information, can be obtained by dividing the number by a weight factor before rounding. Taking a larger value for the weight gives a "coarser" quantization. De-quantization, which maps the quantized value back into its original range, but not its original precision, is achieved by multiplying the value by the weight. [12]

5. SRLE Technique

Sliding run length encoding (SRLE) combines the goodness of RLE mechanism with the effectiveness of SRLE technique. The word sliding

comes about from the process of interrogating (sliding over) the randomness of the entire RLE output which will be at the encoder side.

Since, RLE output stream is a combinational pairs of a value (coefficient **C**) and its run (repetition **R**); we will then be interested on how well such a pair is represented as compactly as possible at the transmitter side so as to further decrease the amount of bits to be sent thus enhancing overall bandwidth requirements by increasing CF, without affecting PSNR.

SRLE will calculate the probability distribution of **C** for a given RLE output, hence, the repetition probabilities of the **C** values are found. The **C** values with the highest probability will then be assigned the required amount of bits needed for its representation; also, this bit assignment is used to find the highest value that can represent other **C**'s. A search of other possible values in **C**, **SS**, which can be represented with the same number of bits, is then sought.

The rest of the **C** values that does not fit the above criterion, are assigned with a variable bit length, needed for the highest value, or **HV**, of **C**, which will adequately represent them.

As an example, suppose a sequence of **C** coefficients is given as 100, 70, 60, 30, 7, 4, 4, 4 and 3. Normally, the 100 value will require 7 bits for its representation, thus 7 bits multiplied by the number of coefficients 9 giving 63 bits total number of bits required to convey the whole bit stream of the **C** values.

Whereas with SRLE, since the value of 4 has the highest probability of occurrence then the number of bits needed for its representation is found as 3 bits. The 3 bits also will cover the values of 3 and 7; hence, five out of the nine values will be represented with 3 bits giving 15 bits. The other

values of **C** will be represented with 7 bits i.e.7 multiplied by 4 giving 28 bits. For now, the total number of bits **T** will be 28 bits plus 15 bits giving 43 bits.

The **T** bits if transmitted will be meaningless to the decoder and thus it cannot reconstruct the original **C** values. Therefore a further technique is devised, to correctly symbolize the **T** string, by the use of tags.

Tagging will be used in such a way that the **SS** values may be given a 0 tag bit each, while the **HV** values are given a 1 tag bit each. In essence, the tags will act as a header to the trail of actual data values of **C**.

Therefore, these tags will be used here to differentiate the 4 bit assignment values from that of the 7 bits, when processed by the decoder. Hence, and since there are 9 **C** values, 9 tag bits will be required. The total number of bits required to convey the whole bit stream of the **C** values will now be 43 bits plus 9 tag bits giving 52 bits.

Comparing the above two results of 63 bits with that of 52 bits, it can thus be seen that there is a reduction of $((63-52)/63)\%$ which is about 17.4% of the total amount of data intended for transmission. It is therefore true to say that in a more practical condition with **C** being much longer and highly correlated; the results will be even more superior.

6. System Model

To implement the proposed system, the following algorithm has been devised:

- A. Encoder Side;
 1. A BMP RGB image "F" is loaded.
 2. Convert "F" to Y, U and V color space information "I".
 3. DWT "I" to get approximate and the details coefficients matrices.
 4. Prepare an 8*8 mask for the approximate coefficients matrix. This

mask will serve as a filter so that, during the DCT process, only a number of DCT coefficients are allowed for further processing. A light filter is used for the approximate coefficients matrix. The tailoring of such a filter is image dependent.

5. 2-D 8-by-8 blocks DCT is computed for step 4.

6. Quantize and RLE, outputs of the details coefficients matrix and that of step 5.

7. SRLE process for step 6.

Figure (2) shows the block diagram of the proposed encoder system.

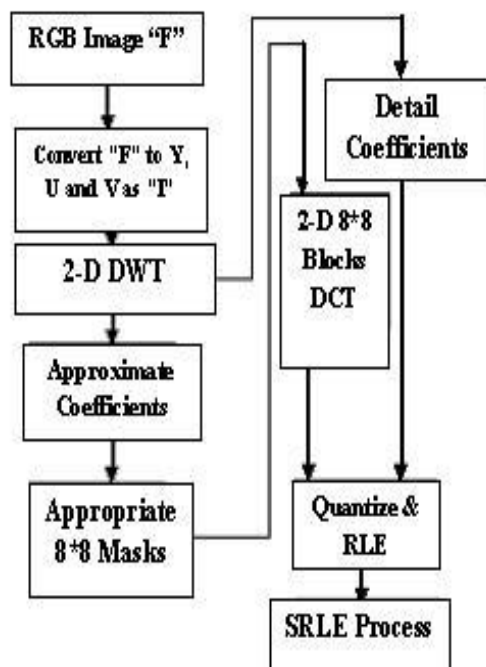


Fig. (2): Encoder Block Diagram

B. Decoder Side;

A reverse procedure of the above encoding process is done.

Results and Discussion:

The proposed scheme was implemented using MATLAB Ver. 7 and tested on a Pentium IV 2.2 MHz Core 2 Due CPU. In all of these investigations, a Lena image (256x256, 24 bits/pixel) has been used as a test

material. Different tests were obtained for better insight to the actual gains that can be accomplished with the usage of the SRLE technique. Note that, in the development process of the proposed technique, a sustained peak signal-to-noise ratio (PSNR) is maintained, to give emphasis to compression factor (CF) and execution time (ET) gains. PSNR and CF, equations (8) and (9), are calculated as follows. [3]

$$PSNR = 10 \log_{10} \frac{(L-1)^2}{RMS}$$

(8)

Where,

$$RMS = \frac{1}{m \cdot n} \sum_{i=1}^m \sum_{j=1}^n (x_{i,j} - x'_{i,j})^2$$

L is the number of gray levels, X is the original image, X' is the compressed image, m and n are the image width and height respectively.

$$CF = \frac{\text{Image size before compression}}{\text{Image size after compression}}$$

(9)

Table (1) shows the outcome of the gained results.

Table (1): Test Outputs Obtained from Comparing Non-SRLE with SRLE Schemes

#	Scheme	PSNR (dBs)	CF	ET (Sec)
1.	DWT	30.018	6.17	1.50
2.	DWT SRLE	30.023	6.962	1.62
3.	DWT DCT	30.015	4.60	1.67
4.	DWT DCT SRLE	30.020	9.06	1.70

From table (1), the following observations were noted:

- The results in the first row are due to the direct application of 2-D DWT on "F". The approximate and the details coefficients are quantized and then run length encoded at the encoder side.

- As shown in the second row, when SRLE is in action, a further boost to CF is sustained, while maintaining approximately the same PSNR output with an increase in ET, of result 1.
- The results listed in the third row are due to the direct application of 2-D DWT on "I" giving the approximate and the details coefficients then DCT, with an 8*8 mask, is applied on the approximate coefficients, since DCT has superior compression performance when image information has poor power concentration in high frequency areas. The decrease in CF was due to the fact that RLE not being able to properly symbolize the randomness of the available data for effective encoding.
- The results of the fourth row indicate that when SRLE is in action a better CF is gained, since SRLE is able to appropriately symbolize the randomness of the available RLE data for effective encoding, while

maintaining approximately the same PSNR output with fine ET when compared to the above three ET results.

Figure (3) shows the original image with output images of table (1) results.

Ibraheem [13] used DWT for image compression. His test results indicated a PSNR of (24 to 40) dBs with CF of (2 to 4) respectively. Saha [14], in his comparison of wavelet compression methods, showed that using DWT with fixed length coders can achieve a PSNR of just lower than 30db at a CF of about 5.

When comparing these results with that of table (1), compression enhancement gained by the usage of the SRLE technique, and in particular DWT DCT SRLE case, achieved good improvements to CF.



Fig. (3): Original Image with output Images of Table (1)

Conclusions:

A range of image compression techniques have been proposed to achieve high image qualities and good compression factors with low computational complexities. In this paper a DWT DCT SRLE scheme, which is a double transform technique with enhanced encoding, was proposed. Such a technique showed low distortions with affirmative PSNR, encouraging CF and good ET, when compared to other various schemes. The gains attained using the proposed technique was due to the fact that SRLE improved the performance of the conventional RLE method by further increasing its probabilistic randomness.

References:

1. Mano, J. 1999. Compressed Image File Formats JPEG, PNG, GIF, XBM, BMP. Addison Wesley Longman, Inc. pp. 121-148.
2. Salomon, D. 2008. A Concise Introduction to Data Compression. Springer-Verlag London Limited. pp. 144-155.
3. Pu, I. 2006. Fundamental Data Compression. Butterworth-Heinemann Publications. pp. 163-165.
4. Archarya, T., Tsai, P. 2005. JPEG2000 Standard for Image Compression: Concepts, Algorithms and VLSI Architectures. Wiley, Hoboken. pp. 60-64.
5. Salomon, D. 2007. Data Compression the Complete Reference. Springer-Verlag London Limited. pp. 337-353.
6. Shi, Y. and Sun, H. 2000. Image and Video Compression for Multimedia Engineering. CRC Press. pp. 36-38.
7. Graps, A. 1995. An Introduction to wavelets. IEEE Computational Sciences and Engineering, 2(2):50-61.
8. Hilton, M., Jawerth, B., Sengupta, A. 1994. Compressing Still and Moving Images with wavelets. Multimedia Systems, Vol. 2, No. 3. pp. 191-198.
9. Truchetet, F., Joanne, B., Perot, F. and Laligant, O. 2000. High-Quality Still Color Image Compression. Society of Photo-Optical Instrumentation Engineers. pp. 409-411.
10. Nelson, M. and Gally, J. 1996. The Data Compression Book. New York NY: M&T Books. pp. 326-344.
11. Salomon, D. 2007. Variable-Length Codes for Data Compression. Springer-Verlag London Limited. pp. 9-12.
12. Watson, A. 1994. Image Compression Using the Discrete Cosine Transform. NASA Ames Research Center, pp. 81-88.
13. Ibraheem, N., I. 2004. Image Compression Using Wavelet Transform. M.Sc. thesis, Baghdad University. pp. 80-85.
14. Saha, S. 2004. Image Compression – from DCT to Wavelets: A Review. The Association for Computing Machinery. Inc. pp. 15-16.

ضغط الصور المزدوج بطريقتي التحويل المويجي المتقطع و تحويل الجيب تمام المتقطع بأستخدام تقنية الانزلاق التشفيري بحساب طول الخطوة

نشوان يوسف بيثون*

*جامعة بغداد/ كلية العلوم للبنات/ قسم علوم الحاسبات

الخلاصة:

لقد اقترحت طرق عدة لضغط الصور والتي تهدف الى تحقيق نسب ضغط عالية مع الحفاظ على جودة عالية للصور على ان يتم اجراء عملية الضغط في زمن قصير. في هذا البحث تم اقتراح طريقة تشفير ذات تحويل مزدوج بأستخدام التحويل المويجي المتقطع (DWT) و تحويل الجيب تمام المتقطع (DCT) اضافة الى طريقة جديدة هي الانزلاق التشفيري بحساب طول الخطوة (SRLE) لغرض تحسين الاداء الكلي لعملية الضغط . و بالتالي فأن فوائد الاستخاد المزدوج للتحويلين قد وجهت للاستفادة من الخصائص المنفردة لكليهما . في البداية تحول معلومات الصورة الى مركبات لونية اذ يتم التحويل من النظام اللوني (احمر ، اخضر ، ازرق (RGB)) الى نظام الشدة (Y) و مركبتي اللون (U,V) ومن ثم يتم اعادة اعتبار المركبتان اللونيتان U و V الى مستوى دقة اقل بسبب افتقارهما للمعالم المكانية و بالتالي الحصول على معدلات ضغط اعلى . بعد ذلك يتم تطبيق طريقة التحويل المويجي المتقطع (DWT) على المركبات (YUV) و التي ينتج عنها مصفوفة المعاملات التقريبية (approximate coefficients) و المعاملات الدقيقة (details coefficients). يتم تكميم مصفوفة المعاملات الدقيقة و تشفيرها بطريقة حساب طول الخطوة (RLE) و بعدها بطريقة الانزلاق التشفيري بحساب طول الخطوة (SRLE) . اما بالنسبة الى حزمة المعاملات التقريبية فيتم تشفيرها بأستخدام طريقة تحويل الجيب تمام المتقطع (DCT) و ذلك لاداء ضغطها العالي عندما تحتوي المعلومات الصورية على مجاميع قليلة من الطاقات المركزة للمعاليم في ظل مناطق عالية التردد. ايضا، يتم تكميم المعلومات الناتجة و تشفيرها بطريقة حساب طول الخطوة (RLE) و بعدها بطريقة الانزلاق التشفيري بحساب طول الخطوة (SRLE) . اشارت نتائج الطريقة المقترحة، للتحويل المزدوج بأستخدام تقنية الانزلاق التشفيري بحساب طول الخطوة ، اعطاء نسب ضغط (CF) مشجعة مع الحفاظ على عامل نسبة قمة الاشارة الى الضوضاء (PSNR) و في زمن تنفيذ جيد مقارنة مع الطرق الاخرى التي تستخدم طريقة التحويل المويجي المتقطع (DWT).