IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

Gray Scale Image Compression

Mr. Neeraj Manglani[1], Ms. Savita Singh[2]

Assistant Professor, Department of Computer Science, Jagannath University, Jaipur, India[1]

M. Tech Scholar, Department of Computer Science, Jagannath University, Jaipur, India[2]

Abstract: The process of reducing the size of a data file is referred to as data compression. Compression is useful because it helps in resources usage reduction, such as data storage space or transmission capacity over internet. Image compression is a main technology in transmission and storage of images and scanned documents, because of large amount of data and information is associated with them. This paper proposes a compression technique for scanned documents and images, based compression method using discrete wavelet transformation (DWT), Discrete Cosine Transformation, (DCT) and Code Book Vector. The importance of the DCT in an image compression is that it takes correlated input image and concentrates its energy in just the first few transform coefficients. This feature of DCT makes it useful in data compression. The effectiveness of the algorithm has been justified over some real images and scanned documents. The performance of the proposed algorithm has been compared with other common compression algorithm.

Keywords: Discrete Wavelet Transformation (DWT), Discrete Cosine Transformation (DCT), Code Book Vector, Customized file format, JPEG images.

I. INTRODUCTION

Scanned documents in digital form play an important role in everyday life. It provides an effective way to store historically important books and data. Compression of the compound documents needs more work than the compression of the images because if the compressed image is having some lossy information, also the human visual system can't able to identify the lossy information and it does not affect the whole image content or information. But in the documents, if text having the minor lossy information then it affects the quality of whole document and can be easily identify by a human visual system. It cannot satisfy the users even if the data is compressed by high ratios.

In this paper we used DCT (Discrete Cosine Transform) [11], DWT (Discrete Wavelet Transform) [8] [10] and Code Vector Book to achieve maximum and effective compression ratio for scanned documents and gray scale images with keeping good quality of these images. In 2D-DCT, the image is broken into blocks of n × m pixels and then 2D- DCT is used to produce a block of n × m DCT coefficients for each block of pixels, the resultants coefficients are quantized, which must result in lossy but highly efficient compression ratio [11]. The wavelet-based compression scheme contains transformation, quantization, and lossless entropy coding[10].

II. TYPE OF IMAGE COMPRESSION

A. Lossless Compression

Image compression can be lossy or lossless. [1] [2] Lossless compression reduces bits by identifying and eliminating statistical redundancy, where redundancy in information theory is the number of bits used to transmit a message minus the number of bits of actual information in the message. In a lossless compression algorithm, compressed data is used to regenerate the original image again. No information is lost in lossless compression. This type of compression is includes entropy coding and Bit-plane coding [9]. Lossless compression is useful for exact recreation of original image, but lossless compression generally does not provide sufficiently high compression ratio which is must in image compression.

B. Lossy Compression

Lossy data compression does not allow reconstruction of exact original data from the compressed data. Lossy compression reduces bits by identifying unnecessary information and removing it. A lossy data compression algorithm is one in which compress the data and after decompress that data, the resulting data may be different from the original one. But it is almost similar to original one, in some useful way. The process of reducing the size of a data file is referred to as data compression. Lossy data compression is used frequently on the Internet. Repeatedly compressing and decompressing the file will cause the information (data) to progressively lose its quality.

III. DISCRETE COSINE TRANSFORMATION (DCT)

DCT expresses a finite sequence of data points in term of a sum of cosine functions oscillating at different frequencies. DCT is important to numerous applications in science and for compression of audio-video files and images, where small and high frequency components are discarded. The use of cosine rather than sine function is critical for compression, since it turn out the that fewer functions are needed to approximate a signal, whereas for different equations the cosines express a particular choice of boundary conditions.

DCT takes related input data information and concentrates its energy in just the first few transform coefficients. If the input data consists of correlated quantities, then most of the n transform coefficients produced by the DCT are

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

zeros or small numbers, and only a few are large as shown below [11]:

12	10	8	10	12	10	8	11
11	12	10	8	10	12	10	8
8	11	12	10	8	10	12	10
10	8	11	12	10	8	10	12
12	10	8	11	12	10	8	10
10	12	10	8	11	12	10	8
8	10	12	10	8	11	12	10
10	8	10	12	10	8	11	12

a. Original Data

81	0	0	0	0	0	0	0
0	1.57	0.61	1.90	0.38	1.81	0.20	0.32
0	0.61	0.71	0.35	0	0.07	0	0.02
0	1.90	0.35	4.76	0.77	3.39	0.25	0.54
0	0.38	0	0.77	8.00	0.51	0	0.07
0	1.81	0.07	3.39	0.51	1.57	0.56	0.25
0	0.20	0	0.25	0	0.56	0.71	0.29
0	0.32	0.02	0.54	0.07	0.25	0.29	0.90

b. DCT Coefficients

81	0	0	0	0	0	0	0
0	2	1	2	0	2	0	0
0	1	1	0	0	0	0	0
0	2	0	5	1	3	0	1
0	0	0	1	8	1	0	0
0	2	0	3	1	2	1	0
0	0	0	0	0	1	1	0
0	0	0	1	0	0	0	1

c. Quantized

12.29	10.26	7.92	9.93	11.51	9.94	8.18	10.97
10.90	12.06	10.07	7.68	10.30	11.64	10.17	8.18
7.83	11.39	12.19	9.62	8.28	10.10	11.64	9.94
10.15	7.74	11.16	11.96	9.90	8.28	10.30	11.51
12.21	10.08	8.15	11.38	11.96	9.62	7.68	9.93
10.09	12.10	9.30	8.15	11.16	12.19	10.07	7.92
7.87	9.50	12.10	10.08	7.74	11.39	12.06	10.26
9.66	7.87	10.09	12.21	10.15	7.83	10.90	12.29

d. Reconstructed data

Figure1. 2-D DCT of a Block of Correlated Data

IV. DISCRETEWAVELET TRANSFORMATION DWT

DWT is a frequency domain tools, which is used in proposed compression technique with pruning proposal based on discrete wavelet transform (DWT). The proposed technique first decomposes an image into coefficients called sub-bands and then the resulting coefficients are further applied for DWT to achieve more compression [8]. MATLAB provides an in-built function dwt2() for wavelet transformation for 2-D matrix which gives four coefficients. DWT2 performs single-level 2-D wavelet decomposition with respect to either a particular wavelet.

[CA,CH,CV,CD] = dwt2(X,'wname')

It computes the approximation coefficients matrix CA and details coefficients matrices CH, CV, CD, obtained by a wavelet decomposition of the input matrix X and 'wname' is the wavelet name.

Wavelet compression method for both CT and MRI images is better than DCT compression method [11].

V. CODE BOOK VECTOR

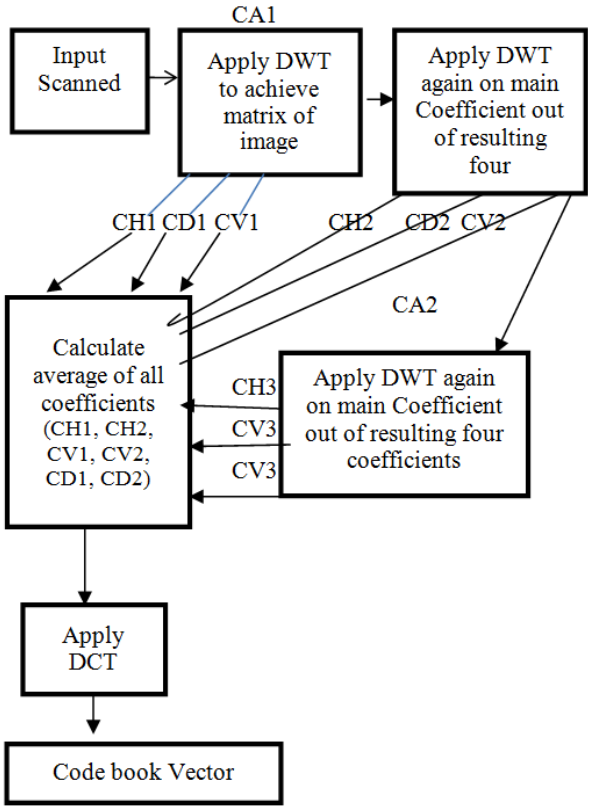
After applying DCT and DWT to data, we are creating a code book, in which the coefficients having approximate similar values, are collect them separate. Out of all coefficients some major coefficient are collected separately providing some index to distinguish between them. By this image data is codded in much compressed format. This code book is used for decoding the data to achieve image again.

VI. CUSTOMIZE FILE HEADER FORMAT

As in our proposed work we are created a code book after eliminating zero and applying DCT and DWT (three times) we are also creating file format in which compressed data are arranged so that we can use that code to decode and achieve image again. The compressed file created after compression, is also provides our own file extension that is .sdc.

VII. IMPLEMENTATION OF PROPOSED METHOD

In our methodology, Scanned document (gray scale) is input for our proposed algorithm. Flow chart of proposed methodology is given below:



Now apply DWT (a frequency domain tool) which commonly used for image compression. DWT is used to decompose a signal into component wavelets. Wavelets have the great advantage of being able to separate the fine details in a signal. MATLAB provides some inbuilt functions to apply DWT, like dwt(), dwt2() etc. We are using dwt2() for 2-D matrix. This will result into four coefficient that are main coefficient (CA1), horizontal coefficient (CH1), vertical coefficient (CV1) and diagonal coefficient (CD1) from the given matrix. Apply DWT again on main coefficient (CA1), which again results into four coefficients that is main coefficient (CA2), horizontal coefficient (CH2), vertical coefficient (CV2) and diagonal coefficient (CD2).

Then apply third time DWT on main coefficient (CA2) to achieve much compressed data, which again results into four coefficients that is main coefficient (CA3), horizontal coefficient (CH3), vertical coefficient (CV3) and diagonal coefficient (CD3).

After applying three times DWT, calculate average of all coefficients resulting from all DWT except main coefficients. Then eliminate zeros from the code. Then finally apply DCT on average of all coefficients and last main Coefficient. In our purposed work we are creating a separate file for header format. For this we are creating a Code Book Vector which stores very less entropy of

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

original scanned document (gray scale) but provide much compressed file and this compressed file is save with .sdc file extension. Code Book Vector tool is important tool to decode the compressed file into image which is achieve with some lossy entropy of original image. After applying inverse DCT and inverse DWT this Code Book Vector (.sdc file) is decoded into .jpeg file again.

VIII. RESULT OF GRAY SCALE IMAGE

Table 1.Compression of Original Image (.jpeg) to .scd file.

Original Image Size(JPEG) (800x545)	Compress Factor	Compress Image Size (.sdc)	Compress Ratio
158 kb	1	45.6 kb	3.4
158 kb	2	49.5 kb	3.1
158 kb	3	55.0 kb	2.8
158 kb	4	63.1 kb	2.5
158 kb)	5	78.0 kb	2.0

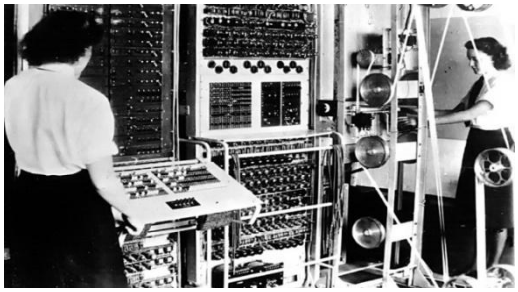


Figure 3.Original JPEG image Size: 158 kb -800x545

After compression by proposed algorithm, a .scd file is created having size 45.6 kb which is almost more than 3 times compressed. To achieve original image again, decompressed this .sdc file to JPEG format. Reconstruction from compressed file (.sdc) results in minor poor quality, but this cannot recognize by human eyes, as shown below figure:

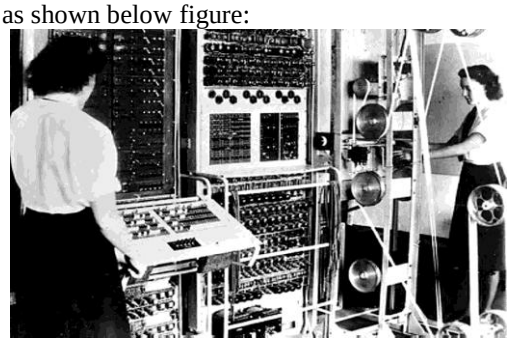


Figure4. Decompressed JPEG image Size: 85.7 kb

B. Decompression of Compressed file (.scd) to .jpeg file

Compressed Image Size(.sdc)	Compressed Ratio	Decompressed Image Size(.JPEG)
45.6 kb	3.4	85.7 kb
49.5 kb	3.1	85.0 kb
55.0 kb	2.8	84.6 kb
63.1 kb	2.5	84.3 kb
78.0 kb	2.0	83.1 kb

IX. CONCLUSION

A new image compression scheme based on discrete wavelet transform and discrete cosine transform proposed in this research which provides sufficient high compression ratios with little degradation of image quality. The effectiveness and robustness of this approach has been justified using a set of real images. The images are taken can compressed in five levels according to requirement. From the experimental results it is evident that, the proposed compression technique gives better performance compared to other traditional techniques. Wavelets are better suited to time-limited data and wavelet based compression technique maintains better image quality by reducing errors. The future direction of this research is to implement a compression and decompression for colour image.

REFERENCES

[1]. M. J. Weinberger, G. Seroussi and G. Sapiro, "The LOCO-I lossless image compression algorithm: Principles and standardization into JPEG-LS", IEEE Trans. on Image Processing, Vol. 2, pp. 1309-1324, Aug. 2000.

[2]. Madhuri A. Joshi, "Digital Image Processing, An Algorithmic Approach", PHI, New Delhi, pp. 175-217, 2006.

[3]. Digital image processing using MATLAB by Rafael C. Gonzalez ,Richard E. Wood Steven L.Eddins,2004, Pearson Prentice Hall.

[4]. A. Alice Blessie1, J. Nalini and S.C.Ramesh, "Image Compression Using Wavelet Transform Based on the Lifting Scheme and its Implementation", JCSI International Journal of Computer Science Issues, Vol. 8, Issue 3, No. 1, May 2011.

[5]. AlexandreZaghetto, Member, IEEE, and Ricardo L. de Queiroz, Senior Member, IEEE. "Scanned Document Compression Using Block-Based Hybrid Video Codec." IEEE Transactions On Image Processing, Vol. 22, No. 6, June 2013

[6]. S. Vimala, M. Sathya , K. Kowsalya Devi ," Enhanced AMBTC for Image Compression using Block Classification and Interpolation" International Journal of Computer Applications (0975 – 8887) Vol 51– No.20, August 2012.

[7]. Prabhakar, Telagrapu, V .Jagan Naveen etal. Image compression using DCT and wavelet transformations, international journal of signal processing, image processing and pattern recognition, vol.4,No.3, September ,2011.

[8]. M. MozammelHoqueChowdhury and AminaKhatun "Image Compression Using Discrete Wavelet Transform" Proc. Of IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 1, July 2012.

[9]. FathimaNajeeb , Rahul R Nair "Hybrid Compression Scheme for Scanned Documents" in Proc of International Journal of Engineering Trends and Technology (IJETT) – Volume17 Number 10–Nov2014.

[10]. K Gopi1, Dr. T. Rama Shri "Medical Image Compression Using Wavelets" in Proc. IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) Vol 2, Issue 4 May. – Jun. 2013

[11]. Gullanar M. Hadi "Medical Image Compression using DCT and DWT Techniques." Advances in Image and Video Processing, Vol 2 No 6, Dec (2014). http://scholarpublishing.org/index.php/AIVP/article/view/771