



High Quality Image Compression

Corresponding Author:

Dr. Rash B Dubey,
Professor, ECE Dept, Hindu College of Engg, Sonapat, 121003 - India

Submitting Author:

Dr. Rash B Dubey,
Professor, ECE Dept, Hindu College of Engg, Sonapat, 121003 - India

Article ID: WMC001673

Article Type: Original Articles

Submitted on: 05-Mar-2011, 04:15:19 PM GMT **Published on:** 06-Mar-2011, 06:09:29 PM GMT

Article URL: http://www.webmedcentral.com/article_view/1673

Subject Categories: MISCELLANEOUS

Keywords: Image Compression, Discrete Cosine Transform, Quantization, Lossless Compression Technique.

How to cite the article: Dubey R B, Gupta R . High Quality Image Compression . WebmedCentral MISCELLANEOUS 2011;2(3):WMC001673

Source(s) of Funding:

NA

Competing Interests:

NA

High Quality Image Compression

Author(s): Dubey R B, Gupta R

Abstract

With the wide use of computers and consequently need for large scale storage and transmission of data, efficient ways of storing of data have become necessary. Dealing with such enormous information can often present difficulties. Image compression is minimizing the size in bytes of a graphics file without degrading the quality of the image to an acceptable level. The reduction in file size allows more images to be stored in a given amount of disk or memory space. Image compression addresses the problem by reducing the amount of data required to represent a digital image. Image compression is achieved by removing data redundancy while preserving information content. In this paper a simplified and more effective method is described and implemented. The proposed algorithm works on quantized coefficients of the discrete cosine transform (DCT) where there are a lot of coincident tokens. Experimental results show that the new approach attains competitive performance of JPEG image to a compression ratio of 40.6%.

Introduction

1. Introduction

Image compression is a technique used to reduce the storage and transmission costs [1]. The basic objective of image compression is to find an image representation in which pixels are less correlated. The two fundamental principles used in image compression are redundancy and irrelevancy. Redundancy removes redundancy from the signal source and irrelevancy omits pixel values which are not noticeable by human eye. Image compression techniques reduce the number of bits required to represent an image by taking advantage of these redundancies. An inverse process called decompression is applied to the compressed data to get the reconstructed image.

The rapid growth of digital imaging applications, including desktop publishing, multimedia, teleconferencing and high definition television (HDTV) has increased the need for effective and standardized image compression techniques. Among the emerging standards are JPEG, for compression of still images, MPEG, for compression of motion video [2-3]; and

CCITT H.261 for compression of video telephony and teleconferencing. All three of these standards employ a basic technique known as the discrete cosine transform (DCT) [4]. Its application to image compression is pioneered by Chen and Pratt [5].

The 2-D DCT is an invertible linear transform and is widely used in many practical image compression systems because of its compression performance and computational efficiency [6-7]. DCT converts data into sets of frequencies. The first frequencies in the set are the most meaningful; the latter, the least. The least meaningful frequencies can be stripped away based on allowable resolution loss. DCT-based image compression relies on two techniques to reduce data required to represent the image. The first is quantization of the image's DCT coefficients; the second is entropy coding of the quantized coefficients [9]. Quantization is the process of reducing the number of possible values of a quantity, thereby reducing the number of bits needed to represent it. Quantization is a lossy process and implies in a reduction of the colour information associated with each pixel in the image. Entropy coding is a technique for representing the quantized coefficients as compactly as possible [9]. The compression standard JPEG2000 [10] accepted quite recently is based on discrete wavelet transform (DWT) and it commonly provides considerably better quality of decoded images than JPEG.

The rest of this paper is organized as follows. A brief review of theoretical background and proposed methodologies is taken in Section 2. Finally, the Section 3 presents the experimental results and discussions.

Methods

1. Methodology

1.1. Background

Compressing an image is significantly different than compressing raw binary data. Of course, general purpose compression programs can be used to compress images, but the result is less than optimal. This is because images have certain statistical properties which can be exploited by encoders specifically designed for them. Also, some of the finer details in the image can be sacrificed for the sake of saving a little more bandwidth or storage space. The reduction in file size allows more images to be stored

in a given amount of disk or memory space. It also reduces the time required for images to be sent over the internet or downloaded from web pages.

There are several different ways in which image files can be compressed. For internet use, the two most common compressed graphic image formats are the joint photographic experts group (JPEG) format and the graphics interchange format (GIF) format. The JPEG method is more often used for photographs, while the GIF method is commonly used for line art and other images in which geometric shapes are relatively simple. A text file or program can be compressed without the introduction of errors, but only up to a certain extent. This is called lossless compression. Beyond this point, errors are introduced. In text and program files, it is crucial that compression be lossless because a single error can seriously damage the meaning of a text file, or cause a program not to run. In image compression, a small loss in quality is usually not noticeable. There is no critical point up to which compression works perfectly, but beyond which it becomes impossible. When there is some tolerance for loss, the compression factor can be greater than no loss tolerance. For this reason, graphic images can be compressed more than text files or programs [8].

Today there are two distinct types of digital graphics, vector and bitmap. JPEG compression only works on bitmap images since vector graphics are not digital images and cannot be made any smaller. Bitmaps, on the other hand, can be compressed and the process is called lossy compression. Bitmap images are $m \times n$ matrices where every entry in the picture matrix corresponds to a small square in the picture. Bitmaps come in four main types. The first is binary (black and white), where the image is an $m \times n$ matrix with every element in the matrix being either a 0 (for black), or a 1 (for white), these images are very poor quality, and are almost never used for storing digital images. The second type is intensity images which store a number between 0 and 1, in every entry of the matrix. Intensity images, like binary images, have zero (0) as complete black, and one (1) as absolute white. However, unlike binary images, intensity images have many different shades of gray, corresponding to numbers between zero and one. The next type is indexed images, where there are 256 different colours, (or shades of gray) which are stored in the image file as the image index. Every element in image matrix has a number between 0-255 which corresponds to a certain colour (or shade of gray) in the index. The last type is truecolor, or RGB (Red, Green, Blue), where the image is composed of three layered $m \times n$ matrices, one for each red, green and blue. Today most pictures are truecolor, but there

are some indexed and intensity images in use. JPEG Compression is diverse enough to compress all three types of images [8].

There are two types of compression, relying on different approaches. Firstly, lossless or reversible compression, which can only be done if the original code was inefficient, for example by not taking into account the fact that some symbols are more frequently used than others, or by having unused bit patterns and secondly, lossy or irreversible compression, in which the original symbol, or its coded representation, cannot be reconstructed exactly, but instead the expander produces an approximation that is good enough.

JPEG usually have higher compression ratios for colour images than for gray-scale images. JPEG does not work very well for non-realistic images, such as line drawings and graphs, where a single wrong pixel is immediately obvious. Lossless data compression is used in many applications such as ZIP file format and in the UNIX tool. It is also often used as a component within lossy data compression technologies. Following techniques are included in lossless compression [8-10]: Run Length Encoding (RLE)

This is a very simple compression method used for sequential data. It is very useful in case of repetitive data. This technique replaces sequences of identical symbols, called runs by shorter symbols. The run length code for a gray scale image is represented by a sequence $\{V_i, R_i\}$ where V_i is the intensity of pixel and R_i refers to the number of consecutive pixels with the intensity V_i as shown in the figure. If both V_i and R_i are represented by one byte, this span of 12 pixels is coded using eight bytes yielding a compression ratio of 1: 5.

Huffman Encoding

This is a general technique for coding symbols based on their statistical occurrence frequencies. The pixels in the image are treated as symbols. The symbols that occur more frequently are assigned a smaller number of bits, while the symbols that occur less frequently are assigned a relatively larger number of bits. Huffman code is a prefix code. This means that the code of any symbol is not the prefix of the code of any other symbol. Most image coding standards use lossy techniques in the earlier stages of compression and use Huffman coding as the final step.

LZW Coding

LZW (Lempel- Ziv – Welch) is a dictionary based coding. Dictionary based coding can be static or dynamic. In static dictionary coding, dictionary is fixed during the encoding and decoding processes. In dynamic dictionary coding, the dictionary is updated on fly. LZW is widely used in computer industry and is

implemented as compress command on UNIX.

Area Coding

Area coding is an enhanced form of run length coding, reflecting the two dimensional character of images. This is a significant advance over the other lossless methods. For coding an image it does not make too much sense to interpret it as a sequential stream, as it is in fact an array of sequences, building up a two dimensional object. The algorithms for area coding try to find rectangular regions with the same characteristics. These regions are coded in a descriptive form as an element with two points and a certain structure. This type of coding can be highly effective but it bears the problem of a nonlinear method, which cannot be implemented in hardware.

2.2 The Proposed Scheme

A 2-D JPEG image is a representation of finite set of values, called pixels. Typically, the pixels are stored in computer memory as a two dimensional array of small integers and these values are transmitted or stored in a compressed form. Images are very large in uncompressed form. Compression techniques are used to reduce image file size for storage, processing and transmission. Compression is a reversible conversion of data to a format that requires fewer bits, usually performed so that the data can be stored or transmitted more efficiently. It can be lossy or lossless. We use the DCT lossless compression technique. First we convert the dimensions of the image from unit 8 to double array by using the level shifting technique. Then convert the image into 8-by-8 blocks, take DCT of each block and then quantized the each block. After quantization arrange each 8-by-8 block into zigzag order and we get a matrix in zigzag form. In next step we distinguish the ac and dc coefficients. After that apply RLE on ac coefficients and differential pulse code modulation (DPCM) on dc coefficients. The RLE is a compression technique that replaces consecutive occurrences of a symbol with the symbol followed by the number of times it is repeated. Clearly this compression technique is most useful where symbols appear in long runs, and thus can sometimes be useful for images that have areas where the pixels all have the same value, cartoons for example. The DPCM is a relative encoding method, a transmission technique that attempts to improve efficiency by transmitting the difference between each value and its predecessor, in place of the value itself. After that the result of DPCM and RLE combined to get finally compressed data of the image. A schematic block diagram of proposed scheme is shown in Fig 1 and other associated stages are described as follows.

2.2.1 Algorithm Design Step

i) Original image is divided into blocks of 8 x 8.

ii) Pixel values of a black and white image range from 0-255 but DCT is designed to work on pixel values ranging from -128 to 127. Therefore each block is modified to work in the range. Calculate DCT Matrix.

iii) DCT is applied to each block by multiplying the modified block with DCT matrix on the left and transpose of DCT matrix on its right.

iv) Each block is then compressed through quantization.

v) Quantized matrix is then gone through zigzag scanning.

vi) After that, Huffman coding is applied.

vii) Compressed image is reconstructed through reverse process.

viii) Inverse DCT is used for decompression.

Fig 1: Schematic block diagram of proposed scheme.

2.2.2 Discrete Cosine Transform

The discrete cosines transform (DCT) is a technique for converting a signal into elementary frequency components. It is widely used in image compression. In particular, a DCT is a Fourier-related transform similar to the discrete Fourier transform (DFT), but using only real numbers. DCTs are equivalent to DFTs of roughly twice the length, operating on real data with even symmetry (since the Fourier transform of a real and even function is real and even), where in some variants the input and/or output data are shifted by half a sample. There are eight standard DCT variants, of which four are common. The most common variant of discrete cosine transform is the type-II DCT, which is often called simply the DCT; its inverse, the type-III DCT, is correspondingly often called simply the inverse DCT.

The DCT is one of many transforms that takes its input and transforms it into a linear combination of weighted basis functions. These basis functions are commonly the frequency, like sine waves. The DCT is a separable transform, implying that a 2-D DCT can be calculated by applying 1-D DCT on both the dimensions of the 2-D signal. The 2-D DCT is just a one dimensional DCT applied twice, once in the x direction, and again in the y direction. One can imagine the computational complexity of doing so for a large image. Thus, many algorithms, such as the Fast Fourier Transform (FFT), have been created to speed the computation [4].

Equation (1) Please refer illustration

The DCT-II is probably the most commonly used form, and is often simply referred to as the DCT. This transform is exactly equivalent to a DFT of $4N$ real inputs of even symmetry where the even-indexed elements are zero. That is, it is half of the DFT of the $4N$ inputs y_n , where $y_{2n} = 0$, $y_{2n+1} = x_n$ for n , and $y_{4N-n} = y_n$ for $0 < n < 2N$.

The DCT-II implies the boundary conditions: x_n is even around $n=-1/2$ and even around $n=N-1/2$; X_k is even around $k=0$ and odd around $k=N$.

2.2.3 Quantization

A typical code works by breaking the image into discreet blocks. These blocks can then be subjected to DCT to separate out the low frequency and high frequency components in both the horizontal and vertical direction. The resulting block is then divided by the quantization matrix, and each entry rounded. The coefficients of quantization matrices are often specifically designed to keep certain frequencies in the source to avoid losing image quality. Alternatively, the extent of the reduction may be varied by multiplying the quantizer matrix by a scaling factor, the quantizer scale code and prior to performing the division. After applying quantization on DCT Matrix, the resulting image is shown in Fig. 2.

Fig 2: Quantized Image of Lena 256 Fig 3: Zigzag Scanning

2.2.4 Zigzag Scanning

Zigzag scanning is done in order to convert group non-zero components at the top left corner of the block into a sequence of non-zero components and to remove all the zeros. For the encoding case, the function receives a pointer to a block of DCT coefficients. As displayed above, the block is reordered in a zigzag pattern. One can think of this pattern as a mapping of locations from one buffer to another. A zigzag mapping table is predetermined and stored in memory. Following the mapping the next step is to remove the trailing zeros from the block, and placing a 255 to denote the end of the block. This is a lossless compression method because the blocks are zero-padded in the reverse transform in order to achieve the original block of DCT coefficients.

In the decoding case, the length of zigzag block is computed by searching for the value of 255. This is necessary so that the system knows how many zeros it must append to the end of the block. Next the forward transformation is reversed, and then zeros are added in the appropriate memory locations.

2.2.5 Run-Length Encoding (RLE)

RLE is a simple technique to compress digital data by representing successive runs of the same value in the data as the value followed by the count, rather than the original run of values. The goal is to reduce the amount of data needed to be stored or transmitted. The need to use run length encoding often arises in various applications in DSP, especially image processing and compression. The RLE Representation is shown in Fig.4.

Fig 4: RLE Representation

The advantage of RLE is a lossless compression technique. That is unlike some other compression techniques, one can obtain exactly the original RLE that can reduce data to merely two numbers if all the values in the original data are exactly the same, regardless of the size of the input. However, in the worst case, if there are no repeating values in the data, RLE could actually double the amount of numbers compared with the original data. Thus RLE should only be used in cases where runs of the same value are expected. After applying RLE on the AC coefficients, the image obtained is shown in Fig.5.

Fig 5: RLE on AC Coefficients

2.2.6 Differential Pulse Code Modulation (DPCM)

DPCM conducted on signals with correlation between successive samples leads to good compression ratios. Images and video signals are examples of signal which have above mentioned correlation. In images this means that there is a correlation between the neighbouring pixels, in video signals correlation is between the same pixels in consecutive frames and inside a frame. DPCM compression method can be conducted for intra-frame coding and inter-frame coding. Intra-frame coding exploits spatial redundancy and inter-frame coding exploits temporal redundancy [8-11].

In the intra-frame coding the difference is formed between the neighbouring pixels of the same frame, while in the inter-frame coding it is formed between the values of the same value in two consecutive frames. In both coding intra- and inter- frame the value of target pixel is predicted using the previously-coded neighbouring pixels.

DPCM is usually used with lossy compression techniques, like coarser quantization of differences can be used, which leads to shorter code words. This is used in JPEG and in adaptive DPCM (ADPCM), a common audio compression method. ADPCM can be watched as a superset of DPCM. Resulting image of Lena 256 after applying DPCM on DC coefficients is shown in Fig. 6.

Fig 6: DPCM on DC Coefficients

Results and Discussions

The algorithm is implemented on personal computer (1.8GHz CPU, 2GB RAM). The algorithm has been tested on 6 different types of images. Fig. 7 shows various original images. The results obtained after compressed images are shown in Fig. 8.

Fig. 7: Original images

Fig. 8: Compressed images

Image compression is very important for efficient

transmission and storage of images. The basic objective of image compression is to find an image representation in which pixels are less correlated. The carried out studies show that the method proposed and described in this paper provides better quality of decoded images than JPEG2000 in most of practical situations and its superiority for complex textured images in some cases. The proposed method is obtained by simple modifications of JPEG, in which DCT serves as its core. This indicates that DCT is at least not worse transformation for use in image compression than DWT used as the basis of JPEG2000. The algorithm designed for image compression has reduced the size of a 256x256 JPEG image to a compression ratio of 40.6%. Further work is in progress to test algorithm on larger sets of biomedical data to improve the accuracy of compression results.

References

1. M. I. Khalil, "Image compression using new entropy coder", *International Journal of Computer Theory and Engineering*, vol. 2, no. 1, pp. 39-41, February, 2010
2. A. Puri, "Video coding using the MPEG-1 compression standard", *Society for Information Display Digest of Technical Papers*, vol. 23, pp. 123-126, 1992.
3. G. Wallace, "The JPEG still picture compression standard," *Communications of the ACM*, vol. 34, no. 4, pp. 30-44, 1991.
4. N. Ahmed, T. Natarajan and K. R. Rao, "Discrete cosine transform", *IEEE Transactions on Computers*, vol.-C-17, no. 1, pp. 693-694, February 21, 1973.
5. W. H. Chen and W. K. Pratt., " Scene adaptive coder", *IEEE Transactions on Communications COM*, vol. 32, pp. 225-232, 1984.
6. T. S. Reddy, K. Ramani, S. Varadarajan and B. C. Jinaga, "Image Compression Using Transform Coding Methods", *IJCSNS International Journal of Computer Science and Network Security*, vol. 7, no. 7, July 2007.
7. Ken Cabeen and Peter Gent, *Image Compression and the Discrete Cosine Transform*, Math 45, College of the Redwoods.
8. A. B. Waston *Mathematica Journal*, vol. 4, no. 1, pp. 81-88, 1994. [http://vision.arc.nasa.gov/Image Compression Using Discrete Cosine Transform, Andrew B. Waston publications/mathjournal94.pdf](http://vision.arc.nasa.gov/Image%20Compression%20Using%20Discrete%20Cosine%20Transform,%20Andrew%20B.%20Waston%20publications/mathjournal94.pdf)
9. Kesavan, Hareesh. *Choosing a DCT Quantization Matrix for JPEG Encoding.*, <http://www-se.Stanford.EDU/class/ee392c/demos/kesvan/>
10. D. Taubman and M. Marcellin, "JPEG 2000: Image

Compression Fundamentals, Standards and Practice. In: Boston: Kluwer, 2002.

11. D. M. Reavy and C. G. Boncelet, "BACIC: a New Method for Lossless I-Level and Gray scale Image", *Electrical Engineering Department, University of Delaware, Newark.*

Illustrations

Illustration 1

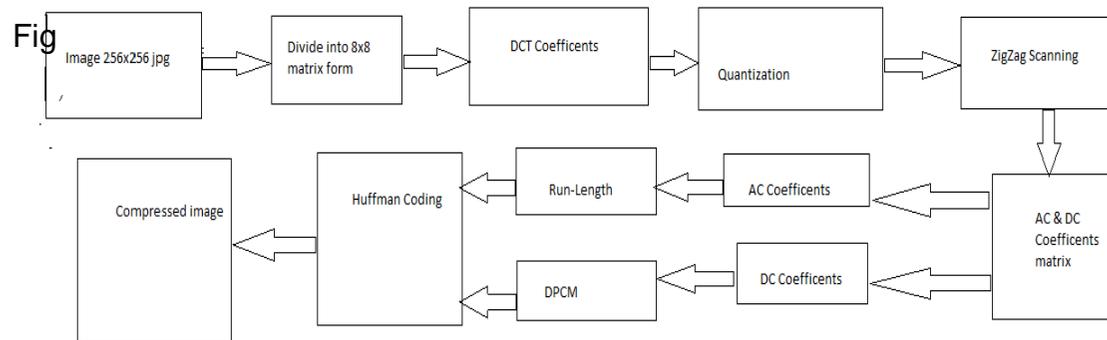


Fig 1: Schematic block diagram of proposed scheme.



Fig 2: Quantized Image of Lena 256

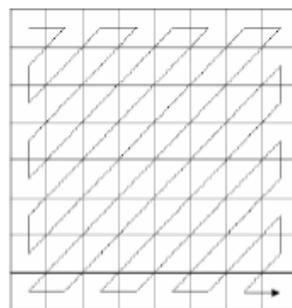


Fig 3: Zigzag Scanning

Original

1	1	1	0	0	1	1	1
---	---	---	---	---	---	---	---

WebmedCentral > Original Articles
RLE Representation

1	3	0	2	1	3
---	---	---	---	---	---

Fig 4: RLE Representation

$$X_k = \sum_{n=0}^{N-1} x_n \cos \left[\frac{\pi}{N} \left(n + \frac{1}{2} \right) k \right] \quad k = 0, \dots, N - 1. \quad \dots\dots(1)$$

Fig 6: DPCM on DC Coefficients



Fig. 7: Original images



Fig. 8: Compressed images

Disclaimer

This article has been downloaded from WebmedCentral. With our unique author driven post publication peer review, contents posted on this web portal do not undergo any prepublication peer or editorial review. It is completely the responsibility of the authors to ensure not only scientific and ethical standards of the manuscript but also its grammatical accuracy. Authors must ensure that they obtain all the necessary permissions before submitting any information that requires obtaining a consent or approval from a third party. Authors should also ensure not to submit any information which they do not have the copyright of or of which they have transferred the copyrights to a third party.

Contents on WebmedCentral are purely for biomedical researchers and scientists. They are not meant to cater to the needs of an individual patient. The web portal or any content(s) therein is neither designed to support, nor replace, the relationship that exists between a patient/site visitor and his/her physician. Your use of the WebmedCentral site and its contents is entirely at your own risk. We do not take any responsibility for any harm that you may suffer or inflict on a third person by following the contents of this website.