JPEG Picture Compression Using Discrete Cosine Transform

Nitesh Kumar More[1], Sipi Dubey[2]

[1]M.Tech Student,
RCET, Bhilai, India
nitesh.more7@gmail.com

[2]Professor
Department of Computer Science & Engineering
RCET, Bhilai
Chhattisgarh, Bhilai - India
drsipidubey@gmail.com

Abstract: ***For the past few years, a joint ISO/CCITT committee known as JPEG (Joint Photographic Experts Group) has been working to establish the first international compression standard for continuous tone still images, both grayscale and color. JPEG's proposed standard aims to be generic, to support a wide variety of applications for continuous tone images. To meet the differing needs of many applications, the JPEG standard includes two basic compression methods, each with various modes of operation. A DCT based method is specified for "lossy'' compression, and a predictive method for "lossless'' compression. JPEG features simple lossy technique known as the Baseline method, a subset of the other DCT based modes of operation. The discrete cosine transform (DCT) is a technique for converting a signal into elementary frequency components. It is widely used in image compression. Here we develop some simple functions to compute the DCT and to compress images. 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.***

Keywords: Image Compression, JPEG, RLE, Discrete cosine transform (DCT), 2-D DCT

1. Introduction

Image compression is very important for efficient transmission and storage of images. Demand for communication of multimedia data through the telecommunications network and accessing the multimedia data through Internet is growing explosively. With the use of digital cameras, requirements for storage, manipulation, and transfer of digital images, has grown explosively. These image files can be very large and can occupy a lot of memory. A gray scale image that is 256 x 256 pixels have 65, 536 elements to store and a typical 640 x 480 color image have nearly a million. Downloading of these files from internet can be very time consuming task. Image data comprise of a significant portion of the multimedia data and they occupy the major portion of the communication bandwidth for multimedia communication.

Therefore development of efficient techniques for image compression has become quite necessary [9]. For- tunately, there are several methods of image compression available today. These falls into two general categories: lossless and lossy image compression. JPEG [7] process is a widely used form of lossy image compression [10] that centers on the Discrete Cosine Transform. The DCT works by separating images into parts of differing frequencies. During a step called quantization, where part of compression actually occurs, the less important frequencies are discarded, hence the use of the term "lossy".

Then only the most important frequencies that remain are used to retrieve the image in the decompression process. They occupy the major portion of the communication bandwidth for multimedia communication. Therefore development of efficient techniques for image compression has become quite necessary [9]. For-tunately, there are

several methods of image compression available today. This fall into two general categories: lossless and lossy image compression. JPEG process is a widely used form of lossy image compression that centers on the Discrete Cosine Transform. The DCT [1] works by separating images into parts of differing frequencies. During a step called quantization [3], where part of compression actually occurs, the less important frequencies are discarded, hence the use of the term "lossy".

2. JPEG Process

a) Original image is divided into blocks of 8 x 8.
b) 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.
c) DCT is applied to each block by multiplying the modified block with DCT matrix on the left and transpose of DCT matrix [5] on its right.
d) Each block is then compressed through quantization.
e) Quantized matrixes are then entropy encoded.
f) Compressed image is reconstructed through reverse process.
g) Inverse DCT is used for decompression [11].

3. Backgrounds: Requirements and Selection Process

JPEG's goal has been to develop a method for Continuous tone image compression which meets the following requirements:

a) Be at or near the state of the art with regard to Compression rate and accompanying image Fidelity, over a wide range of image quality ratings, and especially in the range where visual fidelity to the original is characterized as “very good” to “Excellent”; also, the encoder should be parameterizable, so that the application (or user) can set the desired compression/quality tradeoff;

b) Be applicable to practically any kind of continuous tone digital source image (i.e. for most practical purposes not be restricted to images of certain dimensions, color spaces, pixel aspect ratios, etc.) and not be limited to classes of imagery with restrictions on scene content, such as complexity, range of colors, or statistical properties;

c) Have tractable computational complexity, to make feasible software implementations with viable performance on a range of CPU’s, as well as hardware implementations with viable cost for applications requiring high performance;

- Have the following modes of operation
  - Sequential encoding: each image component is encoded in a single left to right, Top to bottom scan.
  - Progressive encoding: the image is encoded in multiple scans for applications in which transmission time is long, and the viewer prefers to watch the image build up in multiple coarse to clear passes.
  - Lossless encoding: the image is encoded to guarantee exact recovery of every source image sample value (even though the result is low compression compared to the loss modes)
  - Hierarchical encoding: the image is encoded at multiple resolutions so that lower resolution Versions may be accessed without first having to decompress the image at its full resolution.

In June 1987, JPEG conducted a selection process based on a blind assessment of subjective picture quality, and narrowed 12 proposed methods to three. Three informal working groups formed to refine them, and in January 1988, a second, more rigorous selection process revealed that the “ADCT” proposal [11],

Based on the 8x8 DCT had produced the best picture Quality. At the time of its selection, the DCT based Method was only partially defined for some of the modes of Operation. From 1988 through 1990, JPEG undertook the sizable task of defining, documenting, simulating, testing, validating, and simply agreeing on the plethora of details necessary for genuine interoperability and Universality.

4. Processing Steps for DCT Based Coding

These figures illustrate the special case of single component (grayscale) image compression. The reader can grasp the essentials of DCT based compression [6] by thinking of it as essentially compression of a stream of 8x8 blocks of grayscale image samples. Color image compression can then be approximately regarded as compression of multiple grayscale images, which are either compressed entire lone at a time, or are compressedalternatelyinterleaving8x8sampleblocksfromeachinturn.

For DCT sequential decodes, which include the Baseline sequential codec, the simplified diagrams indicate how single component Compression works in a fairly complete way. Each8x8 block is input; makes its way through each processing step, and yields outputting compressed form into the data stream. For DCT progressive mode codec’s, an image buffer exists prior to the entropy coding step, so that an image can bestrode and then parceled out in multiple scans with successively improving quality. For the hierarchical mode of operation, the steps shown are used as building blocks within a larger framework

5. 8x8 FDCT and IDCT

At the input to the encoder, source image samples regrouped into 8x8 blocks, shifted from unsigned integers with range [0, 2P 1] to signed integers with range[2P1,2P11],and input to the Forward DCT (FDCT).At the output from the decoder, the Inverse DCT(IDCT) outputs 8x8 sample blocks to form the reconstructed image. The following equations are the idealized mathematical definitions of the 8x8 FDCT and 8x8 IDCT:

$$F(u,v) = \frac{1}{4}C(u)C(v)\left[\sum_{x=0}^{7}\sum_{y=0}^{7} f(x,y) * \cos\frac{(2x+1)u\pi}{16}\cos\frac{(2y+1)v\pi}{16}\right] \quad (1)$$

$$f(x,y) = \frac{1}{4}\left[\sum_{u=0}^{7}\sum_{v=0}^{7} C(u)C(v)F(u,v) * \cos\frac{(2x+1)u\pi}{16}\cos\frac{(2y+1)v\pi}{16}\right] \quad (2)$$

where: $C(u),\ C(v) = 1/\sqrt{2}$ for $u,v = 0$;

$C(u),\ C(v) = 1$ otherwise.

The DCT is related to the Discrete Fourier Transform (DFT). Some simple intuition for DCT based compression can be obtained by viewing the FDCT as a Harmonic analyzer and the IDCT as a harmonic Synthesizer. Each 8x8 block of source image samples is effectively a 64pointdiscrete signal which is a function of the two spatial dimensions x and y. The FDCT takes such a signal as its input and decomposes it into 64 orthogonal basis signals. Each contains one of the 64 unique two-dimensional (2D) “spatial frequencies’’ which comprise the input signal’s “spectrum.” The output of the FDCT is the set of 64basissignalamplitudes or “DCT coefficients” whose values are uniquely determined by the particular64pointinput signal.

The DCT coefficient values can thus be regarded as the relative amount of the 2D spatial frequencies contained in the 64pointinput signal. The coefficient with zero frequency in both dimensions is called the “Coefficient” and the remaining 63 coefficients are called the “AC coefficients.’’ Because sample values typically vary slowly from point to point across an image; the FDCT processing step lays the foundation for achieving data compression by concentrating most of the signal in the lower spatial frequencies. For atypical 8x8 sample block from a typical source image, most of the spatial frequencies have zero or near zero amplitude and need not be encoded.

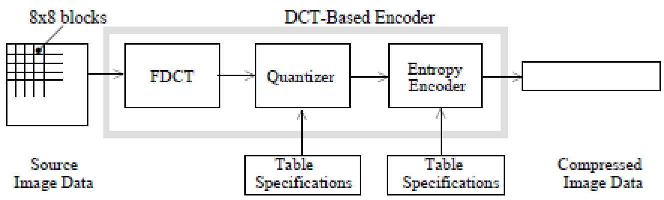

Figure 1: DCT Based Encoder Processing Steps

At the decoder the IDCT reverses this processing step. It takes the 64 DCT coefficients (which at that point have been quantized) and reconstructs a 64pointouputimage signal by summing the basis signals. Mathematically, the DCT is one to one mapping for 64point vectors between the image and the frequency Domains. If the FDCT and IDCT could be computed with perfect accuracy and if the DCT coefficients were not quantized as in the following description, the original 64pointsignal could be exactly recovered. In principle, the DCT introduces no loss to the source image samples; it merely transforms them to a domain in which they can be more efficiently encoded. Some properties of practical FDCT and IDCT implementations raise the issue of what precisely should be required by the JPEG standard. Fundamental property is that the FDCT and IDCT equations contain transcendental functions. Consequently, no physical implementation can compute them with perfect accuracy. Because of the DCT's application importance and its relationship to the DFT, many different algorithms by which the FDCT and IDCT may be approximately computed have been devised. Indeed, research in fast DCT algorithms is ongoing and no single algorithm is optimal for all implementations. What is optimal in software for a genera l purpose CPU is unlikely to be optimal in firmware for a programmable DSP and is certain to be suboptimal for dedicated VLSI.

Even in light of the finite precision of the DCT inputs and outputs, independently designed implementations of the very same FDCT or IDCT algorithm which differ even minutely in the precision by which they represent cosine terms or intermediate results, or in the way they sum and round fractional values, will eventually produce slightly different outputs from identical inputs.

To preserve freedom for innovation and customization with in implementations, JPEG has chosen to specify neither a unique FDCT algorithm nor a unique IDCT algorithm in its proposed standard. This makes compliance somewhat more difficult to confirm, because two compliant encoders (or decoders) generally will not produce identical outputs given identical inputs. The JPEG standard will address this issue by specifying an accuracy test as part of its compliance tests for all DCT based encoders and decoders; this is to ensure against crudely inaccurate cosine basis functions which would degrade image quality.

For each DCT based mode of operation, the JPEG Proposal specifies separate codec's for images with 8bitand 12bit (per component) source image samples. The12bitcodecs, needed to accommodate certain types of Medical and other images require greater computational resources to achieve the required FDCT or IDCT accuracy. Images with other sample precisions can usually be accommodated by either an8bitor 12bitcodec, but this must be done outside the JPEG standard. For example, it would be the responsibility of an application to decide how to fit or pad a 6bitsample into the 8bitencoder's input interface, how to unpack it at the decoder's output, and how to encode any necessary related information.

5.1 Quantization

After output from the FDCT, each of the 64 DCT Coefficients is uniformly quantized in conjunction with a 64elementQuantization Table, which must be specified by the application (or user) as an input to the encoder. Each element can be any integer value from 1to 255, which specifies the step size of the quantize forts corresponding DCT coefficient. The purpose of Quantization is to achieve further compression by representing DCT coefficients with no greater precision than is necessary to achieve the desired image quality. Stated another way, the goal of this processing step is to discard information which is not visually significant. Quantization is a many to one mapping, and therefore is fundamentally lossy. It is the principal source of lossyness in DCT based encoders. Quantization is defined as division of each DCT Coefficient by its corresponding quantizes step size, Followed by rounding to the nearest integer:

$$F^{Q}(u,v) = Integer\ Round\left(\frac{F(u,v)}{Q(u,v)}\right) \quad (3)$$

This output value is normalized by the quantize step size. Dequantization is the inverse function, which in this case means simply that the normalization is removed by multiplying by the step size, which returns the result to a representation appropriate for input to the IDCT:

$$F^{Q'}(u,v) = F^{Q}(u,v) * Q(u,v) \quad (4)$$

When the aim is to compress the image as much as possible without visible artifacts, each step size ideally should be chosen as the perceptual threshold or "just noticeable difference" for the visual contribution of its corresponding cosine basis function. These thresholds are also functions of the source image characteristics, display characteristics and viewing distance. For applications in which these variables can be reasonably well defined, psycho visual experiments can be performed to determine the best thresholds. The experiment described in has led to a set of Quantization Tables for CCIR601 [4] images and displays. These have been used experimentally by jpeg members and will appear in the ISO standard as a matter of information, but not as a requirement.

5.2 DC Coding and Zig Zag Sequence

After quantization, the DC coefficient is treated separately from the 63 AC coefficients. The Coefficient is a measure of the average value of the 64image samples. Because there is usually strong correlation between the DC coefficients of adjacent 8x8blocks, the quantized DC coefficient is encoded as the difference from the DC term of the previous block in the encoding order (defined in the following), as shown in Figure 2. This special treatment is worthwhile, as DC coefficients frequently contain a significant fraction of the total image energy.

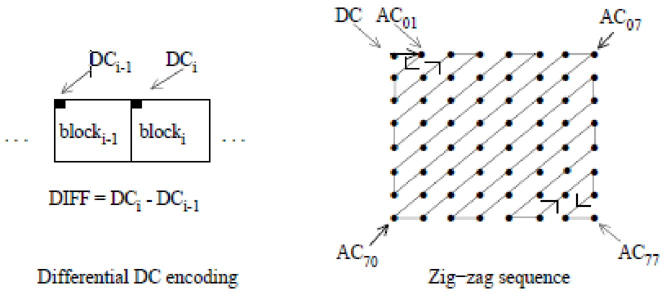

Figure 2: Preparation of Quantized Coefficients for Entropy Coding

Finally, all of the quantized coefficients are ordered in the "zigzag "sequence, also shown in Figure 2.This ordering helps to facilitate entropy coding by placing low frequency coefficients (which are more likely to be nonzero) before high frequency coefficients.

5.3 Entropy Coding

The final DCT based encoder processing step is entropy coding. This step achieves additional compression lossless by encoding the quantized DCTcoefficients more compactly based on their statistical characteristics. The JPEG proposal specifies two entropy coding methods Huffman coding [8] and arithmetic coding. The Baseline sequential codec uses Huffman coding, but codec's with both methods are specified for all modes of operation. It is useful to consider entropy coding as a 2stepprocess. The first step converts the zigzag sequence of quantized coefficients into an intermediate sequence of symbols. The second step converts the symbols to a data stream in which the symbols no longer have externally identifiable boundaries. The form and definition of the intermediate symbols is dependent on both the DCT based mode of operation and the entropy coding method. Huffman coding requires that one or more sets of Huffman code tables be specified by the application. The same tables used to compress an image are needed to decompress it. Huffman tables may be predefined and used within an application as defaults, or computed specifically for a given image in an initial statistics gathering passes prior to compression. Such choices are the business of the applications which use JPEG; the JPEG proposal specifies no required Huffman tables. Huffman coding for the Base line sequential encoder is described in detail in section 7.By contrast, the particular arithmetic coding method specified in the JPEG proposal [2] requires no tables to be externally input, because it is able to adapt to the image statistics as it encodes the image. (If desired, statistical conditioning tables can be used as inputs for slightly better efficiency, but this is not required.)Arithmetic coding has produced 510%bettercompression than Huffman for many of the images which JPEG members have tested. However, some feel it is more complex than Huffman coding for certain implementations, for example, the highest speed hardware implementations. (Throughout JPEG'shistory, "complexity" has proved to be most elusive a practical metric for comparing compression methods.)

For color images with moderately complex scenes, all DCT based modes of operation typically produce the following levels of picture quality for the indicated ranges of compression. These levels are only guideline quality and compression can vary significantly according to source image characteristics and scene content. (The units "bits/pixel" here menthe total number of bits in the compressed image including the chrominance components divided by the number of samples in the luminance component.)

- 0.250.5bits/pixel: moderate to good quality, Sufficient for some applications;
- 0.50.75bits/pixel: good to very good quality, sufficient for many applications;
- 0.751/5 bits/pixel: excellent quality, sufficient for most applications;
- 1.52.0bits/pixel: usually indistinguishable from the original, sufficient for the most demanding applications.

The entire DCT procedure is shown in Figure.

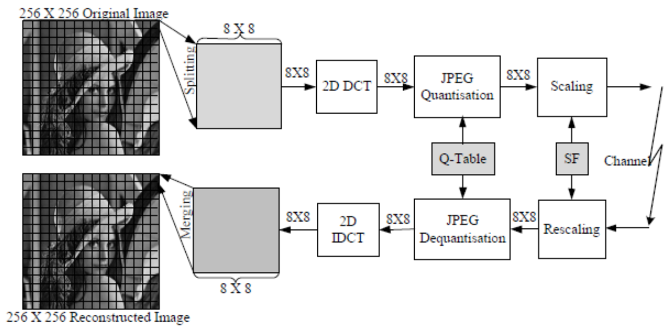

Figure 3: Block diagram of the JPEG based DCT

6. Result and Conclusion

JPEG continuous to image compression standard is not a panacea that will solve the myriad issues which must be addressed before digital images will be fully integrated within all the applications that will ultimately benefit from them. For example, if two applications cannot exchange uncompressed images because they use incompatible color spaces, aspect ratios, dimensions, etc. then a common compression method will not help. However, a great many applications are "stuck" because of storage or transmission costs, Argument over which (nonstandard) compression method to use, or because VLSI codec's are too expensive due to low volumes. For these applications, the thorough technical evaluation, testing, selection, validation, and documentation work which JPEG committee members have performed is expected to soon yield an approved international standard that will withstand the tests of quality and time. As diverse imaging applications become increasingly implemented on open networked computing systems, the ultimate measure of the committee's success will be when JPEG compressed digital images come to be regarded and even taken for granted as "just another data type," as text and graphics are today. Figure (a) and (b) shows the original image and reconstructed images at two different compression level using 2- D DCT. For the lower compression ratio, the distortion is unnoticed by human visual perception, which can be seen in Figure © In order to achieve higher compression it is required to apply quantization followed by scaling to the transformed coefficient.

Figure 4: Illustration of compression using DCT: (a) Original Image1, CR at (b) 88%, (c) 96%

References

[1] Andrew B. Watson, NASA Ames Research, "Image Compression Using the Discrete Cosine Transform", Mathematica Journal, 4(1),1994, p. 81-88.

[2] Nageswara Rao Thota, and Srinivasa Kumar Devireddy, "Image Compression Using Discrete Cosine Transform", Georgian

[3] Electronic Scientific Journal: Computer Science and Telecommunications 2008|No.3 (17).

[4] Swastik Das and Rashmi Ranjan Sethy, "A Thesis on Image Compression using Discrete Cosine Transform and Discrete Wavelet Transform", Guided By: Prof. R. Baliarsingh, dept of Computer Science & Engineering, National Institute of Rourkela.

[5] F. Feing, S. Winograd, "Fast algorithms for the discrete cosine transform", IEEE Transactions on Signal Processing, vol. 40, no. 9,September, 1992.

[6] N. Ahmed, T. Natarajan, and K.R. Rao," Discrete Cosine Transform", IEEE Trans. Computers, 90-93, Jan 1974.

[7] Wallace, G. 1991. The JPEG still picture compression standard communications of the ACM 34(4): 30-44.

[8] Chih-chang chen, Oscal T-C. Chen "A Low complexity computation scheme of discrete cosine transform and quantization withadaptation to block content", Dept of Electrical Engineering, 2000 IEEE.

[9] Chih-chang chen, Oscal T-C. Chen "signal and Media Laboratories", Dept of Electrical Engineering, 2000 IEEE.

[10] Yung-Gi Wu, "Medical Image compression by sampling DCT coefficient", 2002 IEEE.

[11] W. Yan, P. Hao, C. Xu, "Matrix factorization for fast DCT algorithms", IEEE International Conference on Acoustics, Speech, andSignal Processing, vol. 3, 2006.

Author Profile

Nitesh Kumar More is perusing M. Tech in Computer Science and Engineering from RCET, Bhilai, and Under CSVTU. He is doing his project under the guidance of Dr Sipi Dubey.