ARPN Journal of Science and Technology

http://www.ejournalofscience.org

Analysis of Various Image Compression Techniques

[1] G.M.Padmaja, [2] P.Nirupama
[1] Senior Assistant Professor in CSE Dept, BVRIT
[2] Associate Professor in CSE Dept, SIET
[1] padmaja.gmp@gmail.com, [2] nirupama.cse1@gmail.com

ABSTRACT

With the rapid development of digital technology in consumer electronics, the demand to preserve raw image data for further editing or repeated compression is increasing. Image compression is minimizing the size in bytes of an image without degrading the quality of the image to an unacceptable level. There are several different ways in which images can be compressed. This paper analyzes various image compression techniques. In addition, specific methods are presented illustrating the application of such techniques to the real-world images. We have presented various steps involved in the general procedure for compressing images. We provided the basics of image coding with a discussion of vector quantization and one of the main technique of wavelet compression under vector quantization. This analysis of various compression techniques provides knowledge in identifying the advantageous features and helps in choosing correct method for compression.

Keywords: *Compression, wavelet compression, SPIHT, global compression scheme, Kohonen's method*

1. INTRODUCTION

With the rapid development of digital technology in consumer electronics, the demand to preserve raw image data for further editing or repeated compression is increasing. **I**n the context of image processing, compression schemes are aimed to reduce the transmission rate for images, while maintaining a good level of visual quality. Compressing an image is significantly different than compressing raw binary data. General purpose compression programs can be used to compress images, but the result is less than optimal.

Image compression is a problem of reducing the amount of data required to represent a digital image. It is a process intended to yield a compact representation of an image, thereby reducing the image storage/transmission requirements. The reduction in image 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. Compression is achieved by the removal of one or more of the three basic data redundancies

1. Coding Redundancy
2. Inter pixel Redundancy
3. Psycho visual Redundancy

Coding redundancy is present when less than optimal code words are used. Inter pixel redundancy results from correlations between the pixels of an image. Psycho visual redundancy is due to data that is ignored by the human visual system (i.e. visually non essential information). Image compression techniques reduce the number of bits required to represent an image by taking advantage of these redundancies.

An inverse process called decompression (decoding) is applied to the compressed data to get the reconstructed image. The objective of compression is to reduce the number of bits as much as possible, while keeping the resolution and the visual quality of the reconstructed image as close to the original image as possible. Image compression systems are composed of two distinct structural blocks: an encoder and a decoder.

Lossless compression involves with compressing data which, when decompressed, will be an exact replica of the original data. This is the case when binary data such as executables, documents etc. are compressed. They need to be exactly reproduced when decompressed. On the other hand, images need not be reproduced 'exactly'. An approximation of the original image is enough for most purposes, as long as the error between the original and the compressed image is tolerable.

1.1 ERROR METRICS

Two of the error metrics used to compare the various image compression techniques are Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR). The MSE is the cumulative squared error between the compressed and the original image, whereas PSNR is a measure of the peak error. , if you find a compression scheme having a lower MSE (and a high PSNR), you can recognize that it is a better one.

ARPN Journal of Science and Technology

http://www.ejournalofscience.org

2. PROCEDURE FOR IMAGE COMPRESSION

The general steps involved in compressing an image are

1. Specifying the Rate (bits available) and Distortion (tolerable error) parameters for the target image.
2. Dividing the image data into various classes, based on their importance.
3. Dividing the available bit budget among these classes, such that the distortion is a minimum.
4. Quantize each class separately using the bit allocation information derived in step 3.
5. Encode each class separately using an entropy coder.

2.1 Classifying Image Data

An image is represented as a two-dimensional array of coefficients, each coefficient representing the brightness level in that point. Most natural images have smooth colour variations, with the fine details being represented as sharp edges in between the smooth variations. Technically, the smooth variations in colour can be termed as low frequency variations and the sharp variations as high frequency variations. The low frequency components (smooth variations) constitute the base of an image, and the high frequency components (the edges which give the detail) add upon them to refine the image, thereby giving a detailed image. Therefore, the smooth variations are demanding more importance than the details. Separating the smooth variations and details of the image can be done in many ways. One such way is the decomposition of the image using a Discrete Wavelet Transform (DWT).

a. The DWT of an Image

The procedure for decomposition of the image is as follows. A low pass filter and a high pass filter are chosen, such that they exactly halve the frequency range between themselves. This filter pair is called the Analysis Filter pair. First, the low pass filter is applied for each row of data, thereby getting the low frequency components of the row. But since the lpf is a half band filter, the output data contains frequencies only in the first half of the original frequency range. So, by Shannon's Sampling Theorem, they can be sub sampled by two, so that the output data now contains only half the original number of samples. Now, the high pass filter is applied for the same row of data, and similarly the high pass components are separated, and placed by the side of the low pass components. The same is repeated for all rows. Next, the filtering is done for each column of the intermediate data. The resulting two-dimensional array of coefficients contains four bands of data, each labeled as LL (low-low), HL (high-low), LH (low-high) and HH (high-high). The LL band can be decomposed once again in the same manner, thereby producing even more sub bands. This can be done up to any level, which results in a pyramidal decomposition shown in Figure 1.

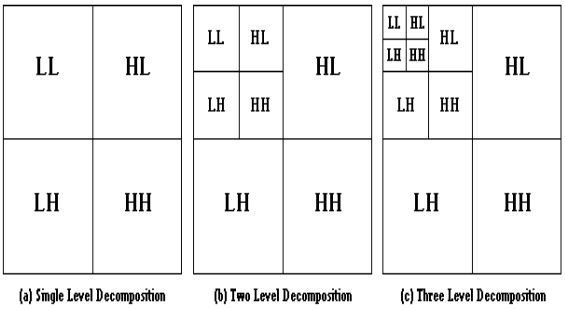

Fig 1: Pyramidal Decomposition of an Image

The LL band at the highest level can be classified as most important, and the other 'detail' bands can be classified as of lesser importance, with the degree of importance decreasing from the top of the pyramid to the bands at the bottom. A three-layer decomposition of an image is shown in figure 2.

Fig 2: The three layer decomposition of the image

2.2 Quantization

Quantization refers to the process of approximating the continuous set of values in the image data with a finite (preferably small) set of values. The input to a quantizer is the original data, and the output is always one among a finite number of levels. The quantizer is a function whose set of output values are discrete, and usually finite. This is a process of approximation, and a good quantizer is one which represents the original signal with minimum loss or distortion. There are two types of quantization -Scalar Quantization and Vector Quantization. In scalar quantization, each input symbol is treated separately in producing the output, while in vector quantization the input symbols are clubbed together in groups called vectors, and processed to give the output. This clubbing of data and treating them as a single unit increases the

optimality of the vector quantizer, but at the cost of increased computational complexity.

a. Scalar Quantization

A quantizer can be specified by its input partitions and output levels (also called reproduction points). If the input range is divided into levels of equal spacing, then the quantizer is termed as a Uniform Quantizer, and if not, it is termed as a Non-Uniform Quantizer. A uniform quantizer can be easily specified by its lower bound and the step size. Also, implementing a uniform quantizer is easier than a non-uniform quantizer. Figure 3 shows a sample uniform quantizer . If the input falls between n*r and (n+1)*r, the quantizer outputs the symbol n.

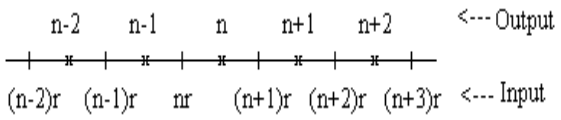

Fig 3: A uniform quantizer

In a similar manner ,a quantizer partitions its input and outputs discrete levels, a Dequantizer is one which receives the output levels of a quantizer and converts them into normal data, by translating each level into a 'reproduction point' in the actual range of data. An optimum quantizer (encoder) and optimum dequantizer (decoder) must satisfy the following conditions.

- Given the output levels or partitions of the encoder, the best decoder is one that puts the reproduction points **x'** on the centers of mass of the partitions. This is known as centroid condition.
- Given the reproduction points of the decoder, the best encoder is one that puts the partition boundaries exactly in the middle of the reproduction points, i.e. each **x** is translated to its nearest reproduction point. This is known as nearest neighbor condition.

The quantization error **(x - x')** is used as a measure of the optimality of the quantizer and dequantizer.

b. Vector Quantization

One of the main common methods to compress images is to code them through vector quantization (VQ) techniques. The principle of the VQ techniques is simple. At first, the image is splitted into square blocks of n×n pixels, for example 4× 4 or 8× 8; each block is considered as a vector in a 16- or 64-dimensional space, respectively. Second, a limited number (l) of vectors (code words) in this space is selected in order to approximate as much as possible the distribution of the initial vectors extracted from the image; in other words, more code words will be placed in the region of the space where there are more points in the initial distribution (image), and vice versa. Third, each vector from the original image is replaced by its nearest codeword (usually according to a second-order distance measure). Finally, in a transmission scheme, the index of the codeword is transmitted instead of the codeword itself; the compression is achieved if the number of bits used to transmit this index ($\log_2 l$) is less than the number of initial bits of the block (n× n ×m if m is the resolution of each pixel).

2.3 Bit Allocation

The first step in compressing an image is to segregate the image data into different classes. Depending on the importance of the data it contains, each class is allocated a portion of the total bit budget, such that the compressed image has the minimum possible distortion. This procedure is called Bit Allocation. The Rate-Distortion theory is used for solving the problem of allocating bits to a set of classes, or for bit rate control in general. The theory aims at reducing the distortion for a given target bit rate, by optimally allocating bits to the various classes of data.

One approach to solve the problem of Optimal Bit Allocation using the Rate-Distortion theory is given below.

1. Initially, all classes are allocated a predefined maximum number of bits.
2. For each class, one bit is reduced from its quota of allocated bits, and the distortion due to the reduction of that 1 bit is calculated.
3. Of all the classes, the class with minimum distortion for a reduction of 1 bit is noted, and 1 bit is reduced from its quota of bits.
4. The total distortion for all classes D is calculated.
5. The total rate for all the classes is calculated as R = p(i) * B(i), where p is the probability and B is the bit allocation for each class.
6. Compare the target rate and distortion specifications with the values obtained above. If not optimal, go to step 2.

In the approach explained above, we keep on reducing one bit at a time till we achieve optimality either in distortion or target rate, or both. An alternate approach is to initially start with zero bits allocated for all classes, and to find the class which is most 'benefitted' by getting an additional bit. The 'benefit' of a class is defined as the decrease in distortion for that class.

ARPN Journal of Science and Technology

http://www.ejournalofscience.org

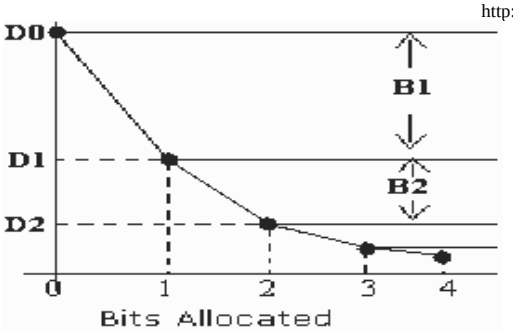

Fig 4: 'Benefit' of a bit is the decrease in distortion due to receiving that bit.

As shown above, the benefit of a bit is a decreasing function of the number of bits allocated previously to the same class. Both approaches mentioned above can be used to the Bit Allocation problem.

2.4 Entropy Coding

After the data has been quantized into a finite set of values, it can be encoded using an Entropy Coder to give additional compression. By entropy, we mean the amount of information present in the data, and an entropy coder encodes the given set of symbols with the minimum number of bits required to represent them.

3. THE GLOBAL COMPRESSION SCHEME

The global compression scheme for lossy compression is shown in Fig. 5. After victimization (transformation of image blocks into vectors), a DCT and a low-pass filter first reduce the quantity of information by keeping only the low-frequency coefficients. Then, the vector quantization is performed, with another loss of information. Finally, the indexes of the code words found by the vector quantizer are transformed by a differential coding, and the results are compressed by an entropic coder ; these two last steps do not introduce any loss in the information. The decompression scheme performs the same operations in the opposite way.

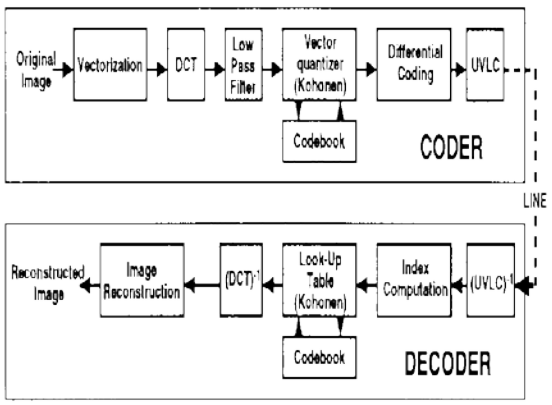

Fig. 5: Global compression scheme for lossy compression.

3.1 Image Decomposition

The image is first decomposed into blocks (4× 4 or 8×8 pixels as usual); the DCT transform is applied on each block, in order to eliminate a part of the information contained in the image, that is, high frequencies not visible to human eyes. The DCT transform of and by n by pixels block is again a n by n block. However, in the transformed block, low-frequency coefficients are grouped in the upper-left corner, while high frequency ones are grouped in the lower-right corner. The low-pass filter on the transformed block will keep only the coefficients nearest from the upper left corner, with $c \leq n^2$; the remaining coefficients $n^2 - c$ are discarded.

3.2 Kohonen's Algorithm

The goal of this algorithm is to create a correspondence between the input space of stimuli and the output space constituted of the codebook elements, the code words. Then these last ones have to approximate the vectors in the input space in the best possible way. All code words are physically arranged on a square grid; it is thus possible to define k-neighborhoods on the grid, which include all code words whose distance (on the grid) from one (central) code word is less or equal to k.

3.3 Differential Coding

Since most parts of the image are smooth, a differential coding applied to the code words after vector quantization will lead to "small" codes in average. The use of an entropic coder, which encodes these differences into variable-length words (i.e., words which will use fewer bits if the differences themselves are small), will thus lead to further compression.

3.4 Entropic Coder

Run length coding (RLC) and variable length coding (VLC) are widely used techniques for lossless data compression. By combining these two techniques, we can achieve a higher compression ratio. Because of the entropic coder, the compression ratio will be higher if the difference between code words is low. A simple differential scheme (zeroth-order predictor) where each codeword is subtracted from the codeword corresponding to the previously encoded block in the image (i.e., the one at the left of the current block).

ARPN Journal of Science and Technology

http://www.ejournalofscience.org

4. TECHNIQUES AVAILABLE

4.1 Vector Quantization Scheme

There are many ways to achieve the vector quantization process of image compression. Kohonen's algorithm is a reliable and efficient way to achieve VQ, and has shown to be usually faster than other algorithms and to avoid the problem of "dead units" that can arise for example with the LBG algorithm.

Kohonen's algorithm also realizes a mapping between an input and an output space that preserves topology; in other words, if vectors are near from each other in the input space, their projection in the output space will be close too.

Topology-preserving property is used for progressive transmission of the image. It can also be used for further differential coding, but on images without discrete cosine transform (DCT) transform, and with a less-performant zeroth-order predictor (instead of first-order).The topology-preserving property is also used to minimize the effect of transmission errors in noisy channels.

4.2 Entropy Coding Schemes

Two of the most popular entropy coding schemes are Huffman coding and Arithmetic coding.

Huffman coding is an entropy encoding algorithm used for lossless data compression. The term refers to the use of a variable-length code table for encoding a source symbol (such as a character in a file) where the variable-length code table has been derived in a particular way based on the estimated probability of occurrence for each possible value of the source symbol. The Huffman coding algorithm is an optimal compression algorithm when only the frequency of individual letters is used to compress the data. The basic idea is that the letters that are more frequent than others uses fewer bits to encode those letters than to encode the less frequent letters.

Arithmetic coding completely bypasses the idea of replacing an input symbol with a specific code. Instead, it takes a stream of input symbols and replaces it with a single floating point output number. The longer (and more complex) the message, the more bits are needed in the output number. The output from an arithmetic coding process is a single number less than 1 and greater than or equal to 0. This single number can be uniquely decoded to create the exact stream of symbols that went into its construction. In order to construct the output number, the symbols being encoded have to have a set probabilities assigned to them.

4.3 Spiht

One of the important algorithm using a wavelet-based image compression coder is Set Partitioning in Hierarchical Trees (SPIHT) . It first converts the image into its wavelet transform and then transmits information about the wavelet coefficients. The decoder uses the received signal to reconstruct the wavelet and performs an inverse transform to recover the image. SPIHT is significant breakthroughs in still image compression in that it offered significantly improved quality over vector quantization, JPEG, and wavelets combined with quantization, while not requiring training and producing an embedded bit stream. SPIHT displays exceptional characteristics over several properties all at once including:

- Good image quality with a high PSNR (Peak Signal to Noise Ratio)
- Fast coding and decoding
- A fully progressive bit-stream
- Can be used for lossless compression
- May be combined with error protection
- Ability to code for exact bit rate or PSNR

The Discrete Wavelet Transform (DWT) runs a high and low-pass filter over the signal in one dimension. The result is a new image comprising of a high and low-pass sub band. This procedure is then repeated in the other dimension yielding four sub bands, three high-pass components and one low pass component. The next wavelet level is calculated by repeating the horizontal and vertical transformations on the low-pass sub band from the previous level. The DWT repeats this procedure for however many levels are required. Each procedure is fully reversible (within the limits of fixed precision) so that the original image can be reconstructed from the wavelet transformed image.

SPIHT is a method of coding and decoding the wavelet transform of an image. By coding and transmitting information about the wavelet coefficients, it is possible for a decoder to perform an inverse transformation on the wavelet and reconstruct the original image. The entire wavelet transform does not need to be transmitted in order to recover the image. Instead, as the decoder receives more information about the original wavelet transform, the inverse-transformation will yield a better quality reconstruction (i.e. higher peak signal to noise ratio) of the original image. SPIHT generates excellent image quality and performance due to several properties of the coding algorithm.

SPIHT divides the wavelet into Spatial Orientation Trees. Each node in the tree corresponds to an individual pixel. The offspring of a pixel are the four

pixels in the same spatial location of the same sub band at the next finer scale of the wavelet. Pixels at the finest scale of the wavelet are the leaves of the tree and have no children.

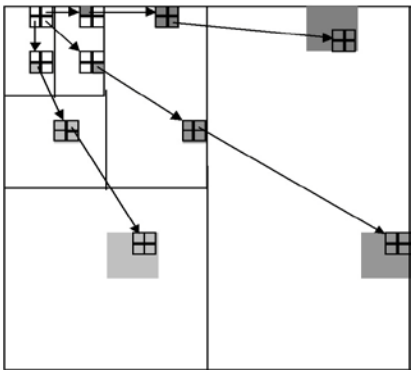

Fig 6: Spatial-orientation trees

Every pixel is part of a 2 x 2 block with its adjacent pixels. Blocks are a natural result of the hierarchical trees because every pixel in a block shares the same parent. Also, the upper left pixel of each 2 x 2 block at the root of the tree has no children since there only 3 sub bands at each scale and not four. Figure 6 shows how the pyramid is defined. Arrows point to the offspring of an individual pixel, and the grayed blocks show all of the descendents for a specific pixel at every scale.

SPIHT codes a wavelet by transmitting information about the significance of a pixel. By stating whether or not a pixel is above some threshold, information about that pixel's value is implied. Furthermore, SPIHT transmits information stating whether a pixel or any of its descendants are above a threshold. If the statement proves false, then all of its descendants are known to be below that threshold level and they do not need to be considered during the rest of the current pass. At the end of each pass the threshold is divided by two and the algorithm continues.

By proceeding in this manner, information about the most significant bits of the wavelet coefficients will always precede information on lower order significant bits, which is referred to as bit plane ordering. Within each bit plane data is transmitted in three lists: the list of insignificant pixels (LIP), the list of insignificant sets (LIS) and the list of significant pixels (LSP).

5. CONCLUSION

Image compression is governed by the general laws of information theory and specifically rate-distortion theory. However, these general laws are non constructive and the more specific techniques of quantization theory are needed for the actual development of compression algorithms. Vector quantization can theoretically attain the maximum achievable coding efficiency. Transform coding techniques, in conjunction with entropy coding, capture important gains of VQ, while avoiding most of its difficulties. We analyzed various techniques of image compression.

REFERENCES

[1] T. Cover and J. Thomas, Elements of Information Theory. New York: John Wiley & Sons, Inc., 1991.

A. Gersho, "Asymptotically optimal block quantization," IEEE Transactions on Information Theory, vol. IT-25, pp. 373–380, July 1979.

[2] C. M. Chakrabarti, M. Vishwanath, Owens R.M, Architectures for Wavelet Transforms:A Survey," Journal of VLSI Signal Processing,Vol. 14, pp 171-192, 1996.

[3] T. Cormen, C. Leiserson, R. Rivest,Introduction to Algorithms, The MIT Press,Cambridge, Massachusetts, 1997.

[4] T. W. Fry, Hyper Spectral Image Compression on Reconfigurable Platforms, Master Thesis, University of Washington, Seattle, Washington, 2001.

[5] K. K. Parhi, T. Nishitani, "VLSI Architectures for Discrete Wavelet Transforms," IEEE Transactions on VLSI Systems, pp 191 – 201, June 1993.

AUTHOR PROFILES:

G.M.Padmaja received the Master of Technology in Information Technology from Sathyabama University, Chennai in 2010. Currently, she is a Senior Assistant Professor at Padmasri Dr.B.V.Raju Institute of Technology, Narsapur. Her interests are in networks, image processing and data mining.

P.Nirupama received the Master of Technology in Computer Science & Engineering from Sathyabama University, Chennai, in 2006. Currently, she is an Associate Professor at Siddharth Institute of Engineering & Technology, Puttur. Her interests are in mobile wireless networks, image processing and data warehousing.