COMPRESSION OF 2-TONE MANUSCRIPT FOR MULTIMEDIA APPLICATION

by

BALAKRISHNAN SUBRAMANIAN

Thesis submitted in fulfilment of the requirements for the degree of Masters of Science

January 2008

ACKNOWLEDGEMENTS

In the name of GOD, I would like to thank and show my gratitude to few people whom has been a contributing factor for me to complete this research.

First and foremost, I would like to thank my thesis supervisor *Prof. Farid Ghani*, whom has been guiding and inspiring with so much of patience throughout. Our numerous meetings and discussions, led to many constructive comments which have greatly improved this work and me to be a better person. Thank you Sir.

I would like to thank my wife, *Ganti* for being very supportive throughout my research work.

My sons, *Vivek* and *Gautam*, thank you for being a great motivating factor for *Appa* to complete this research in timely manner.

Thank you *Dr. Ekram Khan*, for all your assistance and guidance though from a great distance.

Some close friends and relatives whom have been very understanding and helping either directly or indirectly for me to complete the research and thesis writing.

TABLE OF CONTENTS

CHAPTER 4 : MANUSCRIPT COMPRESSION USING WAVELET TRANSFORM

CHAPTER 5 : CONCLUSIONS AND FUTURE WORK

APPENDICES

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

$g(i, j)$	Original image
$\hat{g}(i, j)$	Reconstructed image
i	Number of rows
j	Number of Columns
$N * N$	Two dimensional image size
S	Maximum intensity value
dB	Decibel
N_c	Number of pixels changed from black to white or white to black
N_t	Total number of pixels in the original image (height x width)
ψ	Wavelet Function
τ	Translation (The location of the window)
s	Scale (s>1: dilate the signal, s<1: Compress the signal)
$\psi^*\left[\frac{t-\tau}{s}\right]$	Mother Wavelet
$x(t)$	Continuous signal in time domain
$x(n)$	Discrete signal
cA	Approximation Coefficients
cD	Diagonal Detail Coefficients
cH	Horizontal Detail Coefficients
cV	Vertical Detail Coefficients
F_x	Input sampling rate
F_y	Output sampling rate
D	Decimation Factor
L	Interpolation Factor

$h(n)$	Impulse response
$H_D(\omega)$	Frequency response
$X(\omega)$	Spectrum

LIST OF ABBREVIATION

MSE	Mean Square Error
PSNR	Peak Signal to Noise Ratio
bpp	Bits per pixel
FPD	Fractional Pixel Difference
IFS	Iterated Function System
HVS	Human Vision System
CODEC	Coder and Decoder
RLC	Run Length Coding
DPCM	Differential pulse code modulation
$K - L$	Karhunen-Loeve
DCT	Discrete Cosine Transform
JPEG	Joint Photographic Experts Group
JBIG	Joint Bi-Level Image Expert Group
DSP	Digital Signal Processing
CWT	Continuous Wavelet Transform
DWT	Discrete Wavelet Transform
IDWT	Inverse Discrete Wavelet Transform
LPF	Low Pass Filter
HPF	High Pass Filter

LIST OF APPENDICES

LIST OF PUBLICATIONS & SEMINARS

1.1 Two Tone Hand Written Text & Manuscript Compression for Multimedia Applications.
IES 2004 Politeknik Elektronika Negeri Surabaya, Indonesia

PEMAMPATAN MANUSKRIP 2-TON UNTUK APLIKASI MULTIMEDIA

ABSTRAK

Malaysia seperti negara lain kaya dengan dokumen lama berlandaskan unsur sejarah dan kebudayaan yang jarang ditemui. Secara amnya, kebanyakan dokumen ini adalah hitam putih dan didapati dalam berbagai tulisan dan disimpan di perpustakaan atau muzium di serata negara. Para sarjana sering merujuk manuskrip sebegini untuk tujuan penyelidikan. Bagi mengelakkan daripada sebarang kerosakan yang tidak boleh diperbaiki akibat kecuaian, adalah penting bagi dokumen-dokumen ini diimbas dan disimpan secara berdigit supaya dapat dicapai dengan mudah. Justeru, manuskrip ini harus dimampat bagi tujuan simpanan dan transmisi yang efisien. Pelbagai kajian telah dan masih dijalankan untuk pemampatan imej kelabu dan warna. Namun begitu, menurut laporan hanya sedikit kajian telah dilakukan untuk pemampatan yang efisien terhadap imej 2-ton atau hitam putih. Kajian ini tertumpu kepada membangunkan kaedah pemampatan yang ringkas dan efisien serta sesuai untuk teks 2-ton yang ditulis dalam tulisan Arab, Jawi dan Tamil. Beberapa *CODEC* berasaskan kombinasi efisien menggunakan kaedah pemampatan *lossless* dan *lossy* telah direkabentuk untuk mencapai nisbah pemampatan yang tinggi tanpa mengabaikan kualiti imej. *Run Length Coding (RLC)* telah digunakan bagi membina *CODEC* pemampatan *loseless*. Bagi pemampatan lossy, teknik *Multirate Digital Signal Processing* (DSP) dan penjelmaan *Wavelet* telah digunakan. *CODEC* yang telah direkabentuk daripada integrasi *RLC* dan teknik *Multirate DSP* berjalan lancar dan memberi nisbah pemampatan yang baik. Disamping itu, penjelmaan *Wavelet* memberi nisbah pemampatan yang tinggi terhadap imej 2-ton dan ini berpotensi untuk penyelidikan lebih lanjut. Parameter seperti *Peak Signal to Noise Ratio (PSNR)*, *Fractional Pixel Difference (FPD)*, dan *Bits per Pixel (bpp)* telah digunakan untuk menilai kualiti dan nisbah pemampatan untuk imej yang dibina semula.

COMPRESSION OF 2-TONE MANUSCRIPT FOR MULTIMEDIA APPLICATION

ABSTRACT

Malaysia like any other country has old and rare documents that depict its history and culture. These documents are generally in black and white and written in various scripts that are stored in libraries and museums throughout the country. These manuscripts are frequently accessed by scholars for research purpose. In order to preserve them from potential irreparable damages due to mishandling, there is a strong need to scan and store it digitally for better accessibility. Furthermore, for storage and transmission efficiency, compression is required. Considerable amount of work has been done and being done for compression of gray level and colored images. Unfortunately, little work has been reported for efficient compression of 2-tone or black and white images. This research is concerned with the development of simple and efficient compression methods suitable for compressing 2-tone texts that are written in Arabic, Jawi and Tamil. Several CODECs based on efficient combinations of lossless and lossy compression methods have been designed to obtain high compression ratios while not sacrificing much on the image quality. Run Length Coding (RLC) has been used to develop CODECs for the lossless compression. For lossy compression, Multirate Digital Signal Processing (DSP) techniques and Wavelet Transform are used. CODECs with the integration of RLC and Multirate DSP techniques perform well and give good compression ratio. On the other hand, Wavelet Transform provides higher compression of 2-tone images and shows a lot of potential for further development. Peak Signal to Noise Ratio (PSNR), Fractional Pixel Difference (FPD) and Bits per Pixel (bpp) are parameters that have been adopted to measure the quality and compression ratio of the reconstructed images.

CHAPTER 1
INTRODUCTION

1.0 Introduction

Image compression plays an important role in digital imaging technology. It produces an efficient and cost effective image which requires less storage space and smaller bandwidth for transmission. The advancement of imaging technology has been producing superior quality images that require large storage space and higher bandwidth for data transmission. By applying digital image compression techniques, the number of bits required to represent an image can be minimized primarily for achieving information transmission and storage efficiency (Ghafourian, 1995). Image compression addresses the problem of reducing the amount of data required to represent a digital image. In simple term, redundant data in the original image is removed during the compression process to produce the compressed image. Later, the compressed image can be decompressed to retrieve the original image or an approximation of it (Gonzalez, 2002). Various compression techniques have been developed to represent the images with fewer bits than required by the original images (Chanda et al., 2002). Test images such as scripts and texts are being used throughout the thesis.

Table 1.1 shows the requirement in terms of the disk space, transmission bandwidth and transmission time needed for various uncompressed data. This example clearly illustrates that uncompressed images and videos require big storage space, large transmission bandwidth and long transmission time. The only solution at the moment is to compress the multimedia data before its storage and transmission, and decompress it at the receiver for play back (Amhamed et al., 2001).

Table 1.1: Multimedia Data Types and Uncompressed Storage Space, Transmission Bandwidth, and Transmission Time Required

Multimedia Data	**Size/Duration**	**Bits per pixel**	**Uncompressed Size**	**Transmission Bandwidth**	**Transmission Time Using a 28.8k Modem**
Page of text	11” x 8.5”	Varying resolution	4-8 KB	32-64 Kb/page	1.1 – 2.2 sec
Telephone quality speech	10sec	8bps	80KB	64Kb/sec	22.2 sec
Gray scale Image	512x512	8bps	262KB	2.1Mb/image	1 min 13 sec
Color Image	512x512	24bps	786KB	6.29Mb/image	3 min 39 sec
Medical Image	2048x2048	12bps	5.16MB	41.3Mb/image	23 min 54 sec
Full-motion Video	640x640, 1min (30 frames/sec)	24bps	1.66GB	221Mb/sec	5 days 8 hrs

[Resources: (Amhamed et al., 2001)]

2-tone or hand written manuscripts contain a lot of useful information that needs to be converted into digital format for easy access to the users. Due to the limitations of the Optical Character Recognition (OCR) softwares, the hand written manuscripts in languages such as Jawi, Arabic and Sanskrit has to be stored as images, which often require huge storage memory. The cost and size of memories are reducing day by day according to Moore’s Law. However, due to ever increasing demands for new multimedia applications, the storage of the documents, textual information, books and handwritten documents in non-OCR languages with acceptable quality are still very memory demanding. Therefore, if these images need to be transmitted or stored, it is impractical to do so without any compression. The problem of transmitting or storing an image affects all of us daily in terms of cost and time. Therefore, effective image compression techniques, which are well suited for compression of such images, are essentially required. The purpose of manuscript compression is to reduce the number of bits to represent images, while maintaining the visual quality of the images for their readability.

Image compression techniques are generally classified into three categories, namely pixel coding, predictive coding, and transform coding (Heidi et al., 1999). The idea behind pixel coding is to process or encode each pixel independently, ignoring the inter pixel dependencies. The pixel values that occur more frequently are assigned with shorter code words (fewer bits), and those pixel values that are rare are assigned with longer code words. This makes the average length of code word to decrease. Run Length Coding (Gonzalez et al., 2002) and Huffman Coding (Gonzalez et al., 2002) are the examples of this type of coding.

Predictive coding is based upon the principle that neighboring pixels in images are most likely are similar. That is if pixel “b” is physically close to pixel “a”, the value (intensity) of pixel “b” will be similar to that of pixel “a”. The compression is achieved by exploiting these inter-pixel redundancies in the image (Gonzalez et al., 2002). Redundancy is a characteristic related to factors such as predictability, randomness and smoothness in the data. The underlying principle is to remove mutual redundancy between successive pixels and encode only the new information. When compressing an image using predictive coding, quantized past values are used to predict future values, and only the new information (or more specifically, the error between the value of pixels “a” and “b”) is coded.

Transform coding plays an important role in image and audio compression. It uses a unitary matrix which has a fast algorithm that represents an excellent compromise in terms of computational complexity versus coding performance. Thus, transform coding outperforms more sophisticated schemes by a margin for a given cost (Donoho et al., 1998). A typical image's energy often varies significantly throughout the image. It makes the image compression in the spatial domain difficult. However, images tend to have a compact representation in the frequency domain packed around the low frequencies, which makes compression in the frequency domain more efficient and effective. In transform coding, first an image is transformed from spatial domain to the frequency domain, and then frequency domain coefficients are quantized and encoded to achieve the compression (Shapiro, 1993). The transform coefficients should be de-correlated to reduce redundancy and to have a maximum amount of information stored in the smaller number of coefficients. These coefficients are then coded as accurately as possible to retain the information. In this thesis, all the three techniques will be used to achieve the best possible compression for the manuscript images.

1.1 Importance of Data Compression

The main objective of data compression is to reduce the number of bits to represent the image and being able to reproduce the image with the best possible fidelity. The compression can be achieved by neglecting the redundant information found in the image. In addition to the redundancy, the high frequency components in the image can be ignored as well without degrading the quality of the decoded image significantly. Two fundamental components of image compression are the redundancy and irrelevancy reduction. By getting rid of these components from the image will not affect the image quality significantly (Ghafourian et al., 1995).

There are a lot of useful manuscripts in certain languages that can benefit the mankind. As mentioned earlier, the limitation of the OCR software is preventing the digitization of the valuable information. The only possible way at present is to scan the manuscripts and store them as images. Compressed image files can be stored and shared across the network easily. Usually the manuscripts are scanned at high resolution depending on the font size to ensure most of the information is captured correctly. These high resolution images contain higher pixel count which increases the file size of the images. The storage requirement for these images goes up tremendously and needs a good compression engine to reduce the bits representation of them.

High resolution images contain more detail information as compared to the lower resolution images. In most of the cases, the high frequency components that preserve the details of the images can be removed without degrading the image quality. In fact, some of the extra information presented in the images is not needed at all. By applying the right compression techniques, the redundant information can be discarded accordingly. Due to the increasing size and resolution of the images, it is necessary to apply compression methods to accommodate the storage and bandwidth limitation. On the other hand the cost can be reduced significantly by data compression.

1.2 Motivating Factors for Compression

The main motivation for data compression is conservation of the storage space. Uncompressed raw data file leads to larger inventories. Through data compression, we can store a lot more of the manuscripts in local computers and are easy to access anytime. Compressed images can be squeezed into smaller bandwidth to transmit more video channels on fiber networks. Limitation in the broadcast channels is another important factor, which drives the data compression. Satellite communication links with the remote sensors for weather and other earth-resource applications is more feasible with compressed data (Gonzalez et al., 2002). Relatively low capacity links that people wish to use, such as satellite communication links that cannot handle scientific data from remote sensors, is still bound by the bandwidth limitation.

Ultimately, enormous image databases can be browsed in reasonable time and good compression algorithms can provide efficient data structures, speed communication, and requires less memory. Furthermore, a lot of cost saving can be achieved (Mahapatra et al., 2003). Since, most of the available compression schemes are centered on the gray and colour images, it's necessary to focus on the compression of 2-tone manuscripts which gets less attention. Furthermore, most of the manuscripts are available in 2-tone image format. To the best of my knowledge there is not much work carried out in the area of 2-tone image compression.

There are still a lot of 2-tone manuscripts written in Jawi and Arabic scripts that can be found specifically in our libraries and museums. In order to preserve the originality of these documents, they are not accessible to us. For the common benefit, these manuscripts can be digitized and make it available for all. By doing so, it will definitely allow more people to have access to them without deteriorating its originality. On the other hand, it gives us a chance to appreciate and understand our ancestor's work.

Besides that, a lot of research activities can be carried out without meddling with the original manuscripts. By doing so, the originality of the documents can be preserved and at the same time the researchers will have the unlimited access to this information at any point of time. Due to efficient compression techniques, all the images can be saved in the user's workstation without worrying about the storage space. The researchers can easily share the images across the network for discussion with their counterparts around the world. This process can speed up the research timeline and might induce some breakthroughs in the research. Bottom line, everybody might get the benefit of the exposed manuscripts.

1.3 Objectives of the Research

This thesis focuses on the 2-tone manuscript compression using lossy and lossless compression techniques. There are 3 main objectives outlined for this work.

The objectives are as below.

- To specifically consider compression of text written in Jawi, Arabic and Tamil scripts as a fair amount of manuscripts are available in the libraries and museums of this country.
- To design and develop coders for the compression of 2-tone images including hand written manuscripts where OCR is not available.
- To implement the coders and compare their performances both qualitatively and quantitatively using some test images.

1.4 Organization of the Thesis

Detailed contents of the research has been put forth in the form of present thesis are organized into five chapters. Chapter 1 deals with the introduction of the thesis. Chapter 2 discusses about the data compression techniques and benefits. Various compression parameters and characteristics are discussed in detail. These include both lossy and lossless compression methods. The methodology adopted in this work is shown in this chapter.

Chapter 3 covers a new coding method introduced in this thesis using the Multirate digital signal processing techniques. The sequential coders are developed by using decimator and interpolator. The result and discussion are provided at the end of the chapter. Chapter 4 discusses on how to apply Wavelet theory to image compression. The integration of Wavelet with other coding techniques is discussed in detail and the results are provided as well. Chapter 5 contains the conclusions and provides suggestions for future work.

CHAPTER 2
DATA COMPRESSION TECHNIQUES & LITERATURE REVIEW

2.0 Introduction

Image compression has been studied for over 40 years as more and more digital images are being used in various applications. It is usually achieved by reducing the statistical redundancy presented in an image. In addition, the properties of Human Vision System (HVS) can be exploited further to increase the compression ratio. An early algorithm of data compression was the Morse code by Samuel Morse in the mid 19th century where letters were sent by telegraph. Since then, many approaches had been suggested including Vector Quantization (Klima et al., 2004), JPEG (Jackson, 1993), Wavelet (Mallat, 1999) and Fractal (Jackson, 1993) techniques.

Digital image compression can be applied to many areas such as, Video Conferencing, High Digital Definition TV (HDTV), Facsimile (FAX), Satellite remote sensing, network communications, internet, documents, sounds, graphics and medical imaging. The real world scene is represented in analog. The data is acquired by a sensor and is converted to digital form by the process of sampling and quantization. The image in the digital form is easily manipulated compared to its analog form.

A digital image is a two-dimensional array of pixels each with a value that corresponds to the image scene intensity. A single digital image may have millions of pixels. These images contain large amount of information. Although the capability of modern storage media and bandwidths of transmission systems are high, but still does not meet the user requirements. On the other hand, images are rapidly becoming so large that they cannot be adequately compressed for transmission or archival with existing techniques. Therefore, intensive research activities are still taking place to address this issue.

In this thesis, the manuscript compression will be focusing only on the 2-tone text images. To the best of my knowledge only JBIG2 (Joint Bi-Level Image Expert Group 2) has been developed exclusively for the 2-tone images. Colour and gray images will not be considered since there are a lot of optimized techniques already available such as “Set Partitioning In Hierarchical Trees (SPIHT)” (Said et al., 1996), “Set Partitioned Embedded bloCK coding (SPECK)” (Pearlman et al., 2004) and “Embedded Block Coding with Optimized Truncation (EBCOT)” (Taubman, 2000). It is my intention to apply some of the methods which are already in use for gray and colour images to the 2-tone image compression.

2.1 Overview of Compression

The recent growth in image data and intensive multimedia applications has not only sustained the need for more efficient ways to encode images but have made compression essential for such storage and communication technology. An early date of the image compression applications was on predictive coding used for linear prediction to television applications. A further landmark was the work of Huffman on constructing efficient variable length codes. In 1940s, C.E. Shannon and others provided the theoretical basis for efficient coding in general and later considered the application of fidelity criterion to his earlier results on coding (Gonzalez et al., 2002). This measure introduces degradation into the reconstructed image that leads to what is called the principle or fundamental intuitive image compression, which results in an image with acceptable quality using minimum number of bits for transmission or storage.

Vector quantization had been used in speech coding by Dudley in the late 1950s (Makhoul et al., 1985), activities increased with this work and this transferred to image compression in 1990s. The first scheme of transform coders was developed in 1960s, since then various methods were introduced until early 1980s when the wavelet concepts were introduced. The wavelet transform is in general a tool for multi resolution analysis of signal. In addition, there are many researchers who proposed algorithms for image data compression using wavelet transforms (Donoho et al., 1998). They provide a hierarchical frequency band split for requirements of image/video signals generation at various quality levels. M. Barnsley (Zhao et al., 2005) explored another technique for image compression or coding, which possess the self-similarity it may exist in the image. Then, Jacquin presents the first automated scheme for fractal compression (Saupe, 1996). In short, there is intensive work has been carried out in the field of image compression until today.

2.2 Compression Benefits

When data compression is used to reduce storage requirements and transmission time, it provides potential cost saving associated with sending less data over the communication channels. First, compression reduces the number of erroneous bits transmitted over the given channel as compared to the non compressed data since fewer bits are stored or transmitted when data is compressed. Secondly, it increases the efficiency of the equipment used and reduces transmission time. Finally, compression may provide a level of security against illicit monitoring, as it is useful for cryptology. The privacy and authentication provide protection against unauthorized intrusions in communication systems. Table 1.1 shows some examples of storage and data rate requirements for uncompressed digital images for various applications.

2.3 Compression Parameters

There are three main characteristics, by which one can judge image compression algorithms. Compression ratio describes the storage reduction achieved by the particular technique. Image quality describes the closeness of the decompressed data to the original image. Finally, compression speed or cost that refers to the computational effort required for the encoding and decoding processes.

The above mentioned parameters are used to determine the suitability of the compression techniques to different applications according to acceptable reconstructions and range of compression ratios available and implementation costs. The following sections discuss each of these attributes in more details.

2.3.1 Compression Ratio

Data compression is the process of reducing the amount of data required to represent a given amount of information. Differing amount of data may represent the same amount of information. The compression ratio is simply defined as the size of the original image divided by the size of the compressed image. This ratio gives an indication of how much compression is achieved by a given compression algorithm for a particular image.

$$Compression\ Ratio = \frac{Size\ of\ Original\ image}{Size\ of\ Compressed\ image} \qquad (2.3.1\text{-}1)$$

This ratio usually indicates the picture quality, since most of the compression techniques operate over a range of compression rate and decompression quality. Generally, the higher the compression ratio, the poorer the quality of resulting images. The trade-off between compression ratio and picture quality is an important metric to consider when image compression takes place.

2.3.2 Image Quality

Image quality is one of the most important measures for any image compression system. The quality of the reconstructed images depends very much on the application used, since an image might be degraded during the process of compression and decompression. Measures of image fidelity can be used to assess the degree of degradation. Image quality described as the fidelity that an image compression scheme recreates the source image data. The choice of distortion measure depends on various factors. Firstly, it should be easy to compute. Secondly, it should adapt to the human vision system's characteristics. There are two methods for determining image quality. Subjective image quality is determined by statistically processing the fidelity rating given by a group of human viewers, whereas the objective image quality is determined by a computational process that does not require human interventions.

There are two commonly used measures for objective quality, which 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. The MSE between two images with a size $N \, x \, N$ where, $g(i, j)$ is original image and $\hat{g}(i, j)$ is reconstructed image are:

$$MSE = \frac{1}{N \times N} \sum_{i=1}^{N} \sum_{j=1}^{N} (g_{ij} - \hat{g}_{ij})^2 \tag{2.3.2-1}$$

One problem with MSE is that it depends strongly on the image intensity scaling. However, PSNR avoids this problem by scaling the MSE according to the image. It is determined as follows.

$$PSNR = 10 \log_{10}\left(\frac{S^2}{MSE}\right) \tag{2.3.2-2}$$

Here, S is the maximum intensity value. PSNR is measured in decibel (dB). This measure (PSNR) is not ideal, but commonly used. Its main drawback is that the signal strength is estimated as $(S)^2$ (Value square), rather than the actual signal strength for the image. A lower value for MSE means lesser error, and as seen from the inverse relation between the MSE and PSNR, this translates to a high value of PSNR. Logically, a higher value of PSNR is good because it means that the ratio of Signal to Noise is higher.

Besides that, there is another measure called the bits per pixel (bpp). In this measurement, the number of bits required to store the information per pixel is calculated. The more bits there are the more memory is required to store or display the image. For a 2-tone image, each pixel is represented by either “0” or “1”. In this case, the bpp is 1. The bpp value is calculated by taking the average value for each pixel.

Objective measures such as PSNR do not indicate subjective quality at low PSNR values and high compression ratios (Grgic et. al., 2001). Subjective evaluation of a reconstructed image requires visual inspection of the image for artifacts, and measures perceptual quality of the reconstructed image.

2.3.3 Fractional Pixel Difference (FPD)

PSNR is not a good metric to measure the quality of 2-tone or bi-level images (Ye et.al., 2003). Therefore, this thesis introduces a new quality measurement metric specifically for 2-tone documents. The parameter is called Fractional Pixel Difference (FPD) and used to evaluate the quality of the reconstructed document. FPD is the ratio of the total number of pixels that are changed (from black to white and white to black) in the reconstructed image as compared to the original image to the total number of pixels. This method is very suitable for 2-tone images since every pixel can be evaluated individually. The FPD values are same as the MSE values for 2-tone images but does not yield the same results for gray and colour images.

Let,

N_c = Number of pixels changed from black to white or white to black.

N_t = Total number of pixels in the original image (height x width)

Then,

$$FPD = \frac{N_c}{N_t} \qquad (2.3.3\text{-}1)$$

Smaller is the value of FPD, better will be the quality of reconstructed image. That shows the number of pixels with contradicting values as compared to the original image are less. For all the methods used in this thesis at the same bits per pixel, the FPD will vary with the type of coder. In this way we can compare the effect of various coding techniques and its implementation.

2.3.4 Compression Speed

Compression and decompression periods are defined as the amount of time required for compressing and decompressing a picture. Their values depend on the following considerations:

- The complexity of the compression algorithm, where a complex compression technique can produce better quality images, but it could be time consuming which it is not suitable for some real time applications (Said et. al., 1996).
- The efficiency of the software or hardware implementation of the algorithm (Pearlman et. al., 2004).
- The speed of the utilized processor or auxiliary hardware (Jackson et.al., 1993).

Generally, the faster the compression or decompression can be performed, the better it is. Fast compression time increases the speed with which resulting compressed image can be coded. Fast decompression time increase the speed with which the user can display and interact with the reconstructed images.

Speed of compression usually matters much more if the data is to be transmitted rather than stored. The decompression speed is quite important for storage and retrieval and is vital for reception of transmitted data. Some compression techniques show symmetry for compression and decompression speeds. Others, such as Fractal coding, are extremely asymmetric.

2.4 Image Redundancy and Human Perception

Image compression takes the advantage of the fact that there is much redundant information contained in the images. Basically there are three types of redundancies:

- Spatial redundancy or inter-pixel redundancy: the images almost have strong neighbouring pixel correlation. This kind of redundancy may lead to better compression.
- Spectral redundancy in images composed of more than one spectral band; the spectral values for the same pixel location are often correlated.
- Statistical redundancies: some pixel intensities are more likely to occur than the others.

Videos have an additional redundancy in addition to the above three.

- Temporal redundancy, adjacent frames in a video sequence often show very little changes.

Image compression research aims at reducing the bits needed to represent an image by removing these redundancies as much as possible. Since, the focus is on 2-tone still image compression, so temporal redundancy will not be considered in this work. The removal of spatial and spectral redundancies is often accomplished by transform coding, which uses some reversible transform to the de-correlated image data. The statistical redundancies are removed by efficient entropy coding such as Huffman or Arithmetic Coding.

Images often have similar structure at various scales. Describing the image through contracted parts of the same image (self transformability) performs redundancy reduction. Often, these transformed parts do not fit together to form an exact copy of the original image but an approximation scheme attempt to eliminate the redundancy induced by this restricted self-similarities. It is achieved by the use of mathematical theory of Iterated Function System (IFS) (Hoskins et al., 1992) through self-similarity of the image.

The human eye does not respond with equal sensitivity to all image signals. Some information has less relative importance than others in our visual system. Therefore, eliminating some information may be acceptable.

2.5 Compression Types

Compression can be studied according to the losses of the decoded image compared to the original image. There are two types of image compression, the lossy and lossless compression. Based on the characteristics of the compression method, it can be easily segregated into lossy or lossless compression. Both of the compression types will be discussed in detail in the subsequent sections.

2.5.1 Lossless Compression

Lossless compression methods accurately preserve all of the original information. It removes the redundancy without degradation. The data after decompression exactly matches the original uncompressed data bit by bit. It is guaranteed that the reconstructed image is identical to the original image. Data compression is achieved in lossless techniques by representing more frequently occurring values by shorter codes and less frequently occurring values by longer codes (Chanda et al., 2002).

The main advantage of this type of compression is that all image details are preserved in the reconstructed image. This is an important requirement for some applications where high quality data is in demand such as text, executable code programs and medical images. For example, compressing digital radiographs with lossy techniques may result in misdiagnosis and incur other unwanted problems. Lossless compression takes the advantage of allocating different number of bits to different pixels in images or different characters or sequence of characters or patterns to reduce the data to a compressed form. It should be noted that since lossless algorithms operate within the perfect reconstruction limit, the amount of compression should not be expected to be very high; generally it does not exceed 2-3 times for typical images. If this loss restriction does not exist, lossy image compression can be used.

2.5.2 Lossy Compression

By using lossy compression techniques, it is possible to achieve a higher degree of compression (in the range of 50 ~ 150 times), where there is a trade-off between compression ratio and quality: good quality is achieved at the expense of compression ratio, and vice versa. These techniques allow the image quality degradation and introduce errors into the data, so that the original image cannot be recovered perfectly. The efficiency of a compression method is measured not only by the ability of data compression, but also by other metrics such as the image quality, speed, and complexity in implementation.

The achievable compression rate is bounded by the tolerable degradation in image fidelity. Similarly, lossy image compression also has many different types of application areas. One of the examples can be found in daily life that is digital camera. These applications often require less image fidelity in the reconstructed image. Thus, they can afford some loss of information during the coding process. On the other hand, as long as humans can recognize the reconstructed images it is considered to be a success. Further compression can be achieved by manipulating the human visual characteristics. Therefore, due to the human-based nature of lossy image compression techniques, its compression performance usually outperforms that of lossless compression. This is because it has more factors that can be exploited to achieve higher compression performance in terms of both compression ratio and visual quality of reconstructed image.

2.5.3 Lossless versus Lossy Compression

The theoretical type of image compression, whether lossless or lossy, to be used for a given application is primarily based on the redundancy inside the image. Normally, it is assumed that the neighbouring pixels exhibit higher degree of correlation. The redundancy is due to the correlation between neighbouring pixel data. This assumption is true for most of the images. However, occasional exceptions occur at the edges of objects in image. Based on this assumption, prediction models can be constructed easily for correlating pixels. In other words, if one pixel is given accurately, its surrounding pixels are nearly known. This correlation points out, that there is little useful information in pixel values of an image. By representing only this useful information, we should be able to compress the original image to a fraction of its size.

In addition to this redundancy in image, lossy image compression can also utilize the characteristics of the Human Vision System (HVS) as mentioned above. These characteristics divide the information in an image into two types: relevant to HVS and irrelevant to HVS. By excluding information irrelevant to HVS, a better compression performance can be expected. Some HVS characteristics such as spatial and spectral redundancies help to predict what information among pixels is relevant to our vision (Grgic et al., 2001). Table 2.1 shows some comparison between lossless and lossy image data compression.

Table 2.1: Comparison between Lossless and Lossy Compression

Lossless Compression	**Lossy Compression**
A compression scheme in which no bits of information are permanently lost.	A compression scheme in which some bits of information are permanently lost during compression and decompression of an image.
Reconstructed image is exactly identical to the original.	The loss is usually only minimal and hardly detectable. Reconstructed image is not identical to the original.
High image quality in demand.	Image quality is trade-off with image compression.
For most types of data, lossless compression techniques can reduce the space needed by only about 50%.	Lossy techniques have the ability to achieve much greater compression. It can reduce file sizes by 95%.

As a consequence, an image compression algorithm should have considerations concerning the above arguments. In order to be efficient for practical applications, a good image compression algorithm should possess the following properties:

1. Good image fidelity in reconstructed image in the decoding phase gives the compression algorithm efficient performance in particular if the other factors such as complexity and execution time are good.

2. High compression and acceptable ratio in lossy and lossless coding respectively are one of the important measures of the compression scheme.
3. Adequate statistical coding. Huffman or arithmetic coding techniques are good coding techniques that can help us to get near-optimal or optimal average code length.

In modern image compression formats, these properties are often present. Different techniques approach these properties in different ways. More images can fit into a floppy disk or hard disk because they have been compressed and take up less space. More importantly, the smaller file size also means that it can be sent over the transmission channels much faster.

2.6 Compression Techniques

There are various compression techniques has been developed to achieve a very optimal compression. Some of the available methods have been selected and customized for the purpose of this thesis. Overview of each technique is elaborated in the following sections.

2.6.1 Run Length Coding (CODEC1)

Run Length Coding (RLC) is a lossless coding method. CODEC1 is designed based on the RLC method. The original image can be reconstructed without losing any information (Gonzalez et al., 2002). RLC is very easy to implement and most popular technique for 2-tone images. The basic concept of RLC is to replace the sequence of the same data values within a file by a count number and a single value (Chanda et al., 2002; Gonzalez et al., 2002). For 2-tone images, where the pixels only have 2 values, “0” and “1”, there is no need to store the code for every run except the first run. Based on the first run the consecutive pixel values can be determined. Figure 2.1 and 2.2 shows the implementation of the run length encoding and decoding systems respectively. This coding engine is very flexible and can be deployed in any 2-tone imaging system easily.

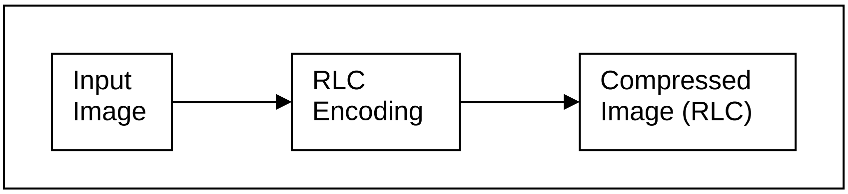

Figure 2.1: Run Length Encoder System

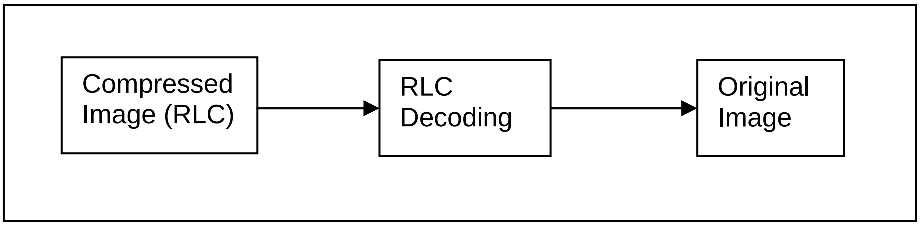

Figure 2.2: Run Length Decoder System

Figure 2.3 shows the structure of CODEC1 which consist of the RLC encoder and decoder. In chapter 3 and 4, RLC will be integrated with other compression techniques to enhance the compression ratio.

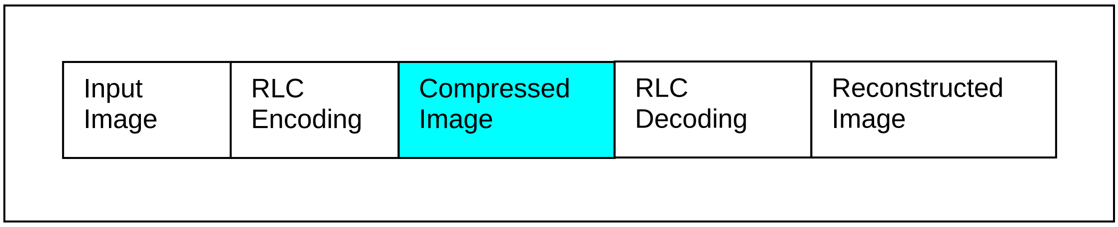

Figure 2.3: CODEC1 Pipeline

It is important to note that there are many different ways to implement the run length coding schemes. The method proposed in (Moinuddin et al., 1997) seems to be very much applicable for the 2-tone manuscripts. The runs in the image are divided into two categories. Category 1 has run length less than 255 and category 2 has the run length of equal to or greater than 255. Single byte is used to code the runs between 0 to 254 and three bytes for runs greater than 254. For category 2, the first byte must be 255 that indicate the coded run is greater than or equal to 255. The following second byte represents the multiplication factor of the first byte. Finally, the third byte contains the remainder. The maximum run length can be achieved with this implementation is 65535 (Moinuddin et al., 1997).

The following steps explain the implementation of the run length encoder for a 2- tone image.

1. Identification of the first bit. If “0”, store “0” in the first byte of the runs. Otherwise, can be ignored.
2. The size of the input image is determined and the column width is stored for the reconstruction of the image.
3. Accumulate the runs of 0’s and 1’s.
4. Evaluate each of the accumulated numbers and segregate to category 1 or category 2. Less than 255, can be left as it is in a single byte. More or equal to 255, has to be separated into 3 bytes following the pseudo codes below. All the implementation work has been carried out using the MATLAB functions.

For X==255,

$$\begin{aligned} &First\ Byte = 255 \\ &Second\ Byte = 0 \\ &Third\ Byte = 0 \end{aligned}$$

For X>255,

$$\begin{aligned} &First\ Byte = 255 \\ &Second\ Byte = floor[(X-255)/255] \\ &Third\ Byte = remainder[(X-255)/255] \end{aligned}$$

Example: X=750,

$$\begin{aligned}
&First\ Byte = 255 \\
&Second\ Byte = floor[(750-255)/255] \\
&\qquad\qquad\quad = floor[495/255] = 1 \\
&Third\ Byte = remainder[(750-255)/255] \\
&\qquad\qquad = remainder[495/255] = 240
\end{aligned}$$

5. Finally, the column width of the original image is stored in the last 3 bytes for reconstruction purpose.

The reverse process of the encoding produces the decoded image. The implementation of the run length decoder for a 2-tone image is explained in the following steps.

1. Identify the starting run. If the first byte is “0”, the run starts with “0”. Otherwise it starts with 1.
2. Last three bytes are read to determine the column width for the reconstructed image.
3. Each byte is walked through to determine if the data belong to category 1 or category 2. If the value is less than 255, no further processing is required. If the value is more or equal to 255, that byte and the consecutive 2 bytes of data need to be processed to obtain a single value. The following explains the calculation of the 3 bytes from the previous example.

For X>=255,

First Byte = 255

Second Byte = 1

Third Byte = 240

Computed Value = First Byte + (255 x Second Byte) + Third Byte

= 255 + (255 x 1) + 240

= 750

4. Expand the matrix to binary format based on the calculated magnitude. The first byte determines the run starts with “0” or “1”.
5. Based on the computed column width, the 2 D image is created.

Further integrating it with other techniques can deliver more compression without sacrificing the image quality. The flip side of this implementation is that it is not suitable for images with short or extremely high run lengths. Minor modification is necessary to cater such needs.

2.6.2 Entropy Coding

Entropy coding is also a lossless compression that is equally applicable to both binary and gray scale images. The key idea of entropy coding is to reduce coding redundancy by identifying most common symbols and representing it with the fewest number of bits. Huffman coding is the most efficient technique to assign binary words of unequal length to the gray levels (Chanda et al., 2002). Huffman code, constructs unique codes that represent frequent symbols with short bit patterns, and infrequent symbols with long bit patterns. Thus it is also a form of variable length coding. Huffman coding can be incorporated with any coding algorithms to remove the statistical redundancies (Gonzalez et al., 2002).

The first step in Huffman coding is to identify the symbols and their probabilities of occurrences. Based on this information, Huffman coder will create a series of source reductions by ordering the probabilities of the symbols under consideration. It combines the lowest probability symbols into a single symbol that replaces them in the next source reduction.

Figure 2.4 explains the steps in the Huffman algorithm. The first 2 columns in the table list the source symbols and their probabilities in descending probability order