AIN SHAMS UNIVERSITY
FACULTY OF COMPUTER AND INFORMATION SCIENCES
COMPUTER SCIENCE DEPARTMENT

A COMPARATIVE STUDY OF IMAGE COMPRESSION TECHNIQUES

This thesis is submitted as a partial fulfillment of the requirements for the degree of **Master of Computers and Information**.

By

Applicant **Amr Abdul Mageed Gamgoum**
Computer Science Department
Faculty of Computer and Information Sciences
Ain Shams University

Under Supervision of

Prof. Dr. M. Hatem El-Ayadi
Professor of Signals and Systems,
Faculty of Computer and Information Sciences,
Ain Shams University.

Prof. Dr. Mustafa M. Syiam,
Professor of Computer Science,
Faculty of Computer and Information Sciences,
Ain Shams University.

Cairo ٢٠٠٧

ACKNOWLEDGMENTS

I acknowledge my deep gratitude to ALLAH the most beneficent and most merciful, who helped me complete this work on a level that I hope will please the reader.

I would also like to thank my dear parents whose kindness, care and never ending support and encouragement made me the person I am today.

I must also thank Prof. Dr. Mohamed Essam Khalifa, Dean of the faculty of Computer and Information Sciences, Ain shams University, for his support and concern.

Special thanks are due to Prof. Dr. Mohamed Hatem El-Ayadi, Professor of signals and systems, at the Faculty of Computer and Information Sciences, Ain Shams University, who constantly guided me in elaborating this thesis, assisted me in understanding and analyzing problems, and continuously provided support and valuable comments.

I would like also to thank others who have helped me out in a variety of ways including Prof. Dr. Taha Ibrahim El-Arif, Prof. Dr. Said El-Ghonaimy, at the Faculty of Computer and Information Sciences, Ain Shams University, and my colleagues Salma Hamdy and Mohamed Abd El-Megeed.

Finally I would like to dedicate this thesis in memory of the late Prof Dr. Mustafa Mahmoud Syiam, my co-supervisor, whose tragic departure has affected us all.

ABSTRACT

Image compression addresses the problem of reducing the amount of data required to represent a digital image. The underlying basis of the reduction process is the removal of redundant data, and sometimes least significant data as well. Thus compression techniques fall into two broad categories: Information preserving, and lossy. This thesis focuses on comparing the performance of various image compression techniques commonly used. This entails elaborate presentation, implementation, performance evaluation, and comparison of known image compression techniques. Suggestion of possible improvement, enhancements, or upgrades of studied techniques falls within the scope of interest of the thesis. Concepts from information theory related to data compression such as Shannon's noiseless and noisy coding theories and rate distortion theory are reviewed and used as guidelines in the presentation.

Lossless image compression techniques include variable-length coding schemes, LZW coding, bit-plane coding schemes, and lossless linear predictive coding. These techniques are reviewed and some modifications are suggested to improve their performance. A comparison based on the average compression ratio and the average compression time of nine test images of different sizes and different levels of detail, is presented for the original and modified schemes. It demonstrates the performance improvement due to suggested modifications. Because of their high time complexities, the studied schemes cannot compress image streams in real time. Thus we investigate vector/parallel implementations of such schemes and introduce two configurations that make the schemes suitable for real time image compression.

Lossy image compression includes both spatial domain and transform based compression techniques. Spatial domain methods include improved

gray-level scaling, lossy predictive coding, block truncation coding, and a new quantization technique based on the rate distortion function which is introduced here. Transform based coding techniques include discrete cosine transform with maximum variance selection, JPEG compression, and wavelet based compression. These techniques are reviewed and their performances are compared with respect to average compression ratio, the average compression time, average relative perceptual distortion and compression consistency over the same nine test images used in the evaluation of the lossless techniques. Lossy techniques can also be extended to include compression of color images. However some color models such as the YIQ represent color images for efficiently and are more appropriate for compression than the simple RGB model commonly used in graphics and image display devices. Finally the thesis is concluded the discussion of video streams which can be considered as a group of image frames. So in addition to the previously discussed compression techniques we can exploit the temporal redundancy found in the successive frames.

CONTENTS

Chapter I: Introduction.

١,١ Importance of Image Compression

Over the years the amount of information stored and/or transferred has been increasing substantially due to the rapid technical upgrade of data sensors and the continuing growth of world wide communication systems and the Internet. For example, during the years ١٩٩٧ – ٢٠٠١ increase of scope and resolution of satellite imaging sensors led to an increase of their data rates from about ٥٠ Mbps, to almost ٤ Gbps [١]. The technology and hardware used in communication and data processing systems have also evolved significantly but with a rate slower than is needed (from ٢٠٠ MHz up to ٢,٦٦ GHz processor clock speed in the years ١٩٩٧ – ٢٠٠٣ [٢]). Thus data compression has become indispensable. In fact several academic communities and university research labs are now dedicated to developing new data compression schemes and/or improving, enhancing and upgrading existing systems [٣, ٤]. Data compression is a smart solution that helps existing communication and data processing systems to deal with the massive increase of information flow, or at least to survive until new systems and technologies are developed. This extends the useful life span of existing systems. For these reasons, developing techniques for data compression has been and will always be a hot topic that stimulates the interest of many scientists and engineers.

Because a great deal of the information being communicated or stored is of pictorial nature, image compression techniques are becoming very essential to accommodate the data increase due to enhanced spatial resolutions of modern imaging sensors and evolving broadcast television standards. Image compression is an ever-growing field and is recognized as an "*enabling technology*". Furthermore it plays a major role in many important and diverse applications including televideo-conferencing, remote sensing via satellite imaging, and the control of remotely piloted vehicles in military, space and hazardous waste management applications. In short, there is an ever-expanding need of efficient manipulation, storage and transmission of binary, gray-scale and color images.

Image compression has developed greatly over the years, from using statistical techniques and methods from information theory that are based on the theoretical work that began in the ١٩٤٠'s when C. E. Shannon and others first formulated the probabilistic view of information and its representation, transmission, and compression [٥]. Modern techniques have also been developed such as region based compression and recently compression using fractals and wavelets [٦, ٧].

١,٢ Introduction of Image Compression

Digital images are usually represented as ٢-D pixel arrays, a format which makes the requirements for storage and transmission very immense. Image compression addresses the problem of reducing the amount of data needed to represent a digital image. The underlying basis of the reduction process is the removal of redundant data. From a mathematical point of view, this amounts to transforming the ٢-D pixel array into a statistically uncorrelated data set. The transformation is performed prior to storage or transmission of the image. At some later time, the compressed image is decompressed to reconstruct the original image or an approximation of it [٨].

Data redundancy is a central issue in digital image compression, it is not an abstract concept but rather mathematically quantifiable. For instance if n and n denote the number of information carrying units in two representations of the same information, the *relative data redundancy* of the first representation with respect to the second one is measured by:

$$R_D = \;\; - \;\; /C_R \qquad (١ - ١)$$

where C_R is referred to as the *compression ratio*,

$$C_R = n \, / \, n \qquad (١ - ٢)$$

In digital image compression, three basic types of data redundancy can be identified and exploited [٨]:

) Coding redundancy:

This type of redundancy is present when the code assigned to represent the different pixel gray levels do not take full advantage of the probabilities of their occurrences. For example if a natural binary code is used, which assigns the same number of bits to both the most and least probable gray levels, it fails to minimize the overall average number of bits, and results in coding redundancy. A process known as *variable – length encoding* [٨], which assigns fewer bits to more probable gray levels than to less probable ones, reduces the overall number of bits needed to represent the image thus achieving data compression.

) Interpixel redundancy:

This type of redundancy is directly related to the interpixel correlations within an image. As such the value of any given pixel can be predicted from the values of its neighbors, making the information carried by individual pixels relatively small and so can be considered redundant. Exploiting this interpixel redundancy usually involves mapping the original ٢-D pixel array into a more efficient but usually non-visual format. One such mapping is the *Run Length Encoding* [٨], in which each row of pixels in the original image is mapped into a sequence of pairs (g_i, w_i), in which g_i denotes the ith gray level encountered along the row and w_i the number of consecutive repetitions of this gray level (i.e. the run length). Thresholding the gray level values, has the effect of increasing the run lengths and so decreasing the needed sequenced of pairs. This however leads to loss of information and thus an approximation of the original image is reconstructed instead of the exact image.

٣) ***Psycho-visual redundancy:***

Due to the fact that the human eye does not respond with equal sensitivity to all visual information, some information simply has less relative importance than other information in normal visual processing [٨]. This information is said to be psycho-visually redundant, and can be eliminated without significantly reducing the perceived quality of the image.

Image compression techniques fall into two main categories:

١) ***Information preserving (lossless) techniques:***

These methods allow the image to be compressed and then decompressed without losing information. They are used in image archiving applications such as storage of legal and medical data in which the smallest details are of extreme importance and so must be preserved. E-government is a large field for this category of image compressing techniques [٨]. Examples of such techniques are the variable length encoding techniques [٩, ١٠, ١١], Lempel-Ziv-Welch(LZW) coding [١٢, ١٣], and lossless predictive coding [٨].

٢) ***Lossy techniques:***

These methods allow for higher levels of data compression. However the decompressed image is an approximation of the original rather than an exact representation [٤]. This type of compression is useful in applications such as broadcast television and facsimile transmission in which a certain amount of error is an acceptable trade-off for enhanced compression performance. Digital cinema is becoming a major field for lossy image compression techniques. Examples of such techniques are lossy predictive coding [٨], Discrete Cosine Transform coding [١٤, ١٥], and wavelet coding [١٦].

١,٣ Fidelity Criteria

Because of the fact that compression may lead to losing information of interest, a means of quantifying the nature and amount of this loss is of extreme importance for the evaluation of the performance of the compression scheme. Such means are referred to as *Fidelity Criteria* [٨]. Two general classes of criteria are used:

٠) Objective Fidelity Criteria:

This type of fidelity criteria expresses the amount of information loss (distortion) as a function of the original input image and the compressed and subsequently decompressed output image. One such criterion is based on the difference (distance) between the input image $f(x, y)$ and output image $f'(x, y)$ pixel values.

$$e(x, y) = f(x, y) - f'(x, y), \quad ٠ \le x \le M, ٠ \le y \le N, \qquad (١ - ٣)$$

where $e(x, y)$ is the distortion or error between the two images.

The objective in this case is to minimize a certain function of the error. Examples of such functions are the *Norms* of the error, which can be defined in several ways. Stacking the rows of the double array $e(x, y)$ yields a vector $\boldsymbol{e}$. Then we can make use of the class of L_p – *vector norms*

$$\| \boldsymbol{e} \|_p = \left(\sum_{i=1}^{MN} |e_i|^p \right)^{\frac{1}{p}} \qquad (١ - ٤)$$

We can also use L_p – *matrix norms* of the double array $\mathbf{E}$ whose elements are $e(x, y)$ with $٠ \le x \le M$ and $٠ \le y \le N$,

$$\| \boldsymbol{E} \|_p = \max_{\| \boldsymbol{x} \|_p = } \| \boldsymbol{E x} \|_p \qquad (١ - ٥)$$

The matrix norms most commonly used, are:

L١-norm which is the maximum absolute column sum value of **E**.

L٢-norm which is the maximum singular value of **E**.

L_∞-norm which is the maximum absolute row sum value of **E**.

Here in the text the symbol $||e(x, y)||$ will always refer to the norm either according to (١ – ٤) or (١ – ٥).

Another objective fidelity criterion is the *Signal to Distortion Ratio* (SDR), which is measured by:

$$SDR = \log \left[\frac{\| f'(x, y) \|}{\| e(x, y) \|} \right]^2 \quad [\mathrm{dB}] \qquad (١ – ٦)$$

The higher this ratio, the better the quality of the output image is.

We can also define a relative distortion measure as

$$\rho = \frac{\|e(x, y)\|}{\|f'(x, y)\|}. \qquad (١ – ٧)$$

In particular we can define a relative perceptual distortion measure as

$$\rho = \frac{\|\tilde{e}\|_4}{\|\tilde{f}'\|_4}, \qquad (١ – ٨)$$

where $\tilde{e}$ and $\tilde{f}'$ are the DFTs of **e** and **f** ' respectively. We generally require that $\rho \leq$ ٠,٠٥, i.e. SDR $\geq$ ٢٦ dB, so that $f'(x, y)$ approximates $f(x, y)$ to at least one significant digit.

To illustrate the use of the objective fidelity criteria, we demonstrate a form of image compression that exploits psycho-visual redundancy in an image. This is performed by means of *quantization* (mapping the wide range of input values to a limited number of output values), which results in a loss of quantitative information.

Consider the images shown in figure ١,١. Figure ١,١(a) shows a monochromatic image with ٢٥٦ possible gray levels i.e. ٨ bits per pixel. Figure ١,١(b) shows the same image after we quantized to ١٦ gray levels or four bits per pixel. The resulting compression ratio is ٢ : ١. However false contours are present in the previously smooth regions of the original image.

Figure ١,١(c) is the version we obtained after applying an improved ١٦ gray level quantization of the same image, which takes advantage of the human visual system. The compression ratio is also ٢ : ١, however, the false contouring is greatly reduced at the expense of additional but less objectionable graininess. This method is known as *improved gray – scale* (IGS) *quantization* [٨] and will be explained in detail later in section ٤,٢,١

(a) (b) (c)

Figure ١,١: Uniform vs IGS quantization. (a) The Original Image; (b) Uniform quantization to ١٦ levels; (c) IGS quantization to ١٦ levels.

Table ١.١: Comparison between uniform gray – scale quantization, and IGS quantization according to the given fidelity criteria.

Fidelity Measure	Uniform Quantization	IGS Quantization
$L_{١}$ norm of **E**	٢٤٠٨	١٦٥٧
$L_{٢}$ norm of **E**	١٩٥٨.٨	٢٤٣,١٣٥٧
L_{∞} norm of **E**	٢٣٠٩	١٦١١
SDR (Using $L_{٢}$ norm)	٢٣,٨٤١ dB	٢٧,٥٢٨٨ dB
Relative perceptual distortion (ρ)	٠,٠٦٧	٠,٠٠٤٣
Relative Perceptual SDR	٢٣,٤٧٥٢ dB	٤٧,٢٧٥٥ dB

Table ١.١ shows a comparison we made between both compression schemes using figures ١,١(a), ١.١(b), and ١,١(c) as examples, in terms of the previously discussed fidelity criteria where **E** is the error matrix between the original image and the decompressed image. It can be seen that the IGS quantization yields better results that the uniform gray – scale quantization with a very close margin. However, the relative perceptual distortion ratio emphasizes greatly the quality of the IGS quantization over uniform

quantization from a perceptual point of view indicating that the IGS quantization produces visually better images, and thus is more preferable.

) Subjective Fidelity Criteria:

Because most of the visual information is viewed by human beings, having the output image evaluated subjectively by a number of human observers may be more adequate [٨]. This can be accomplished by showing the evaluators a number of sample decompressed image and averaging their absolute evaluations. Evaluation scale can be based on the quality of the image as perceived by the evaluators with a value ranging from ١ denoting an "excellent" image in which no distortion is present, to ٦ for "unusable" images which are so bad that they are of no use.

Evaluations may also be performed comparatively by means of side – by – side comparisons between the decompressed images and the original uncompressed image. Comparisons can be done with a scale such as {- ,- ,- , , , , } to represent the subjective evaluations {*much worse, worse, slightly worse, the same as, slightly better, better, much better*}, respectively.

١,٤ General Model of Image Compression System

In section ١,٢ we discussed the major categories of image compression techniques. However a practical image compression system is usually formed by combining these techniques together. A general compression system consists of two distinct blocks, an *encoder* and a *decoder* (figure ١,٢). The input image *f(x, y)* is fed into the encoder which generates a set of symbols representing that image in a more compact form. After transmission over the *channel*, the encoded representation is fed into the decoder, which in turn reconstructs the output image *f'(x, y)*. This output image is either an exact representation or an approximation of the original image, depending on whether the system is noise-free (information preserving), or noisy.