1-1-2015

Region of Interest Based Compression of Grayscale Images

Adhokshaja Achar Budihal Prasad
University of South Florida, adhokshajaac@mail.usf.edu

Region of Interest Based Compression of Grayscale Images

by

Adhokshaja Achar Budihal Prasad

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Electrical Engineering
Department of Electrical Engineering
College of Engineering
University of South Florida

Major Professor: Ravi Sankar, Ph.D.
Rangachar Kasturi, Ph.D.
Selçuk Köse, Ph.D.

Date of Approval:
February 26, 2015

Keywords: Tsallis Entropy, Image Segmentation, Wavelet Compression

DEDICATION

I dedicate this to my family. A special feeling of gratitude to my loving parents Prasad B. R. and Suma G.N and my brother Aniruddha Achar whose words of encouragement, endless support and constant love have sustained my throughout my life. I could not have done this without them.

ACKNOWLEDGMENTS

I wish to thank my committee members, who were more than generous with their expertise and precious time. A special thanks to Dr. Ravi Sankar, my major professor for his countless hours of reflecting, reading, encouraging, and most of all patience throughout the entire process. Thank you Dr. Rangachar Kasturi, and Dr. Selçuk Köse for agreeing to serve on my committee. Their excitement and willingness to provide feedback made the completion of this research an enjoyable experience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Image compression based on Region of Interest (ROI) has been one of the hot topics of interest in image processing. There is not a single widely accepted method for detecting the ROI automatically form an image. To reduce the transmission bandwidth and storage space requirements of gray scale images, an algorithm is suggested for detecting the ROI automatically based on Tsallis entropy method.

Tsallis entropy method is used to segment the image into two segments, the ROI and the background. These two segments can then be compressed at different rates, to avoid losing information in the ROI while achieving a good compression. Different approaches of compression based on wavelets and use of various compression methods are also discussed.

CHAPTER 1: INTRODUCTION

1.1 Background

In many applications, the compression of images is very necessary but there is always a trade-off between the extent of compression and the image quality. In most applications, due to bandwidth limitations, it becomes very essential to have higher compression rates, but still keep the details in the image in good condition and preserve the image quality. And for this reason, the areas of the image that hold the most information need to be compressed separately; for this reason, the Region of Interest (ROI) based compression of images becomes essential.

The ROI depends on the type of images that are being used. If the images under consideration were infrared images from a battlefield, the ROI would be the target area and the rest would be the background (BG). In case of medical images, the ROI determination is a little more complex. By coding the ROI and BG separately, the information in the ROI can be preserved and the information in the BG can be curtailed; thus, achieving higher compression rates. The general theme is to preserve quality in critical regions while allowing lossy compression of the other regions.

1.2 Current Methods

The compression of images based on an ROI involves two key steps. The first step is to identify the ROI automatically. This is done using a process called image segmentation. The second step of the process is image compression.

1.2.1 Image Segmentation

Image segmentation is a very important part in the process of detecting the ROI from the image. Over the years there have been various methods to perform image segmentation process. One such method that is widely used is the Ostu method for gray segmentation [1]. This method uses the image histogram to determine a threshold for segmentation process. This method may not be suitable for ROI detection as the segmentation process works best if the ROI is greater than 30% of the overall image [2]. The other methods that have been used include growing segment approach to image segmentation [3] and threshold determination by 2-D entropy [2]. The growing segment approach selects a pixel/ pixels that would be assumed to be the part of the ROI. A growing principle is applied to the pixels neighboring the said pixel and if the neighbors adhere to this principle, they are included in the ROI. The process is completed until all the pixels in the image are considered. This process is computationally very intensive and time consuming.

The use of 2-D entropy is better suited for ROI determination in images since it is relatively fast and accurate. The process of segmentation based on the 2-D entropy approach, calculated a threshold for segmentation based on the 2-D entropy of the 2-D histogram of the image. The point with the maximum entropy is determined to be the threshold point. The calculations of 2-D entropy can be computationally intensive as well. A modified version of Shannon's entropy is used in [2] to determine the threshold. The use of Tsallis entropy for threshold determination is introduced in [4], [5] and [6]. Tsallis entropy is simpler in terms of calculations and a lot more approximations can be made during calculations.

1.2.2 Image Compression

Image compression being one of the most researched topics in image processing, has a lot of different methods for reducing the image size. The most common compression methods are the JPEG, JPEG 2000 or other methods which involve Lempel–Ziv–Welch Encoding, DPCM or run length Encoding [7]. However, over the recent years a great deal of research is being done on progressive wavelet based compression schemes to reduce the image size. These methods are resolution based progressive encoding schemes where a lower resolution image is encoded first and progressively used to encode higher resolution of the image. This method of encoding is very efficient in terms of providing a good quality compressed image while keeping the size of the compressed file relatively small. JPEG 2000 uses a wavelet transformation and quantization for compression before encoding and truncating using methods such as EBCOT (Embedded Block Coding with Optimal Truncation) [2]. The use of a JPEG 2000 frame for compression of the images is found in [2] and [3]. The use of Haar wavelet based progressive compression method has been shown to be effective for image compression [8]. A method of coding the ROI in a medical image by compressing the image using wavelets and using a Max-shift algorithm to encode the ROI so as to minimize information loss in the ROI in introduced in [9]. This method is very similar to the JPEG2000 method. The other non-wavelet based methods that are being researched are compression of medical images using Discrete Cosine Transforms (DCT) [10] and Fractal image compression techniques [11].

1.3 Motivation and Project Objectives

The motivation for this thesis is to achieve file size reduction without losing quality in the important regions of the image. Since some parts of the image are more important than the other parts, it is possible to compress these two separately. It is acceptable to lose information or quality in the non-important regions of the image. However, retaining the information or quality in the ROI part of the image is crucial.

The objectives of this thesis are to develop an algorithm to automatically detect the ROI of an image while minimizing the time required for computation of the ROI; compress the image's ROI and background such that the loss of information in the ROI is reduced; compare different compression methods based on perceptual and compression parameters and determine which best compression method. Study the effectiveness of the algorithm on medical, military and everyday gray scale images.

1.4 Document Organization

Chapter 2 of this document outlines the methodology proposed to extract the ROI automatically from the image, compression and reconstruction of the images based on this ROI. The proposed method for ROI determination is based on the modified Tsallis entropy method of segmentation, a global thresholding method for image segmentation. This method of segmentation allows extraction of the important regions from the image. The discussion of this method is done in Chapter 3.

The compression of the ROI and the background of the image, the reconstruction of the compressed image, the methodology involved, trade-offs are described in Chapter 4. Chapter 5 involves the discussion of results based on the different approaches for compression and comparison of results with reference papers. Chapter 6 provides conclusion and recommendations for future research.

CHAPTER 2: PROPOSED SCHEME

The scheme for compression is for gray-scale images with one or more regions of interest is shown in the form of a block diagram in Figure 1.

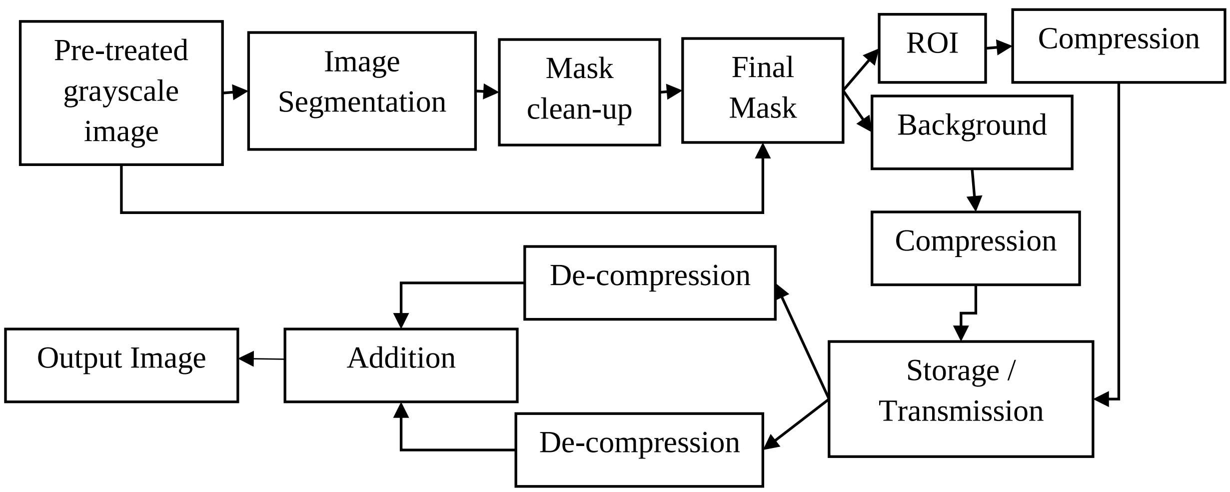


Figure 1 Block diagram of the proposed scheme

The input of the process is a pre-treated gray scale image. A pre-treated gray-scale image is any gray scale image that is ready to be compressed for storage or transmission. The pre-processing steps may involve cropping, de-noising or removing other undesirable artifacts that may exist in the image.

Once the image is preprocessed, it needs to be segmented. The segmentation of the image is done based on the global threshold. The regions of the image above the threshold are all included in the ROI and the rest of the image becomes part of the background. A mask is obtained by setting all pixels in the ROI to a binary value '1' and the pixels in the background to a binary value '0'. This is referred to as a binary mask or simply a mask. The obtained mask of the ROI, from this process needs to be continuous for the compression process. Hence the mask is cleaned, removing small artifacts and regions from the image. The regions of the background that are completely enclosed within an ROI segment are added to the mask by simply "filling the holes in the ROI". The obtained mask should consist of one or more closed segments. This will serve as a template to extract the ROI and background regions from the input image. This step is referred to as mask clean-up.

Using the mask on the input image, the background is separated from the ROI. This process gives two separate images, one with the ROI information and the other with the background information. These two images serve as the starting point for the compression process. The mask can now be discarded.

The compression process consists of using the two images obtained from the previous process and compressing them separately using wavelet based compression techniques. Since the images are separate, they can be separately compressed with different compression rates or even different compression methods to suit various needs. The compression process outputs two compressed images one each for the background and ROI parts of the image.

The compressed images can be transmitted or stored until further use. Optionally the images can be enclosed in a frame to form a single image. To obtain the output image which is the input image after compression, the compressed files need to be reconstructed.

For the reconstruction process, the two compressed image files obtained are uncompressed. The process of uncompressing the compressed files is simple and involves undoing the steps done while compressing the image. The process of reconstruction yields two separate images, the reconstructed background and the reconstructed ROI parts of the original image. To obtain the final over all image, these two need to be superimposed together. Since the ROI and the background have mutually exclusive parts of the original image, a simple mathematical addition operation of the two images will provide the full reconstructed image.

Since the aim of this proposed scheme is to obtain a compressed image with a minimal loss of quality in the ROI while keeping the file sizes low, various approaches to the compression process need to be studied. The way of comparing the compression processes is done using two perceptual parameters - MSE (mean square error), PSNR (Peak Signal-to-Noise Ratio), and two compression parameters - CR (compression Ratio) and BPP (Bits-Per-Pixel). Various compression methods such as Spatial Partitioning in Hierarchical Trees (SPIHT) [2] [3], Embedded Zerotree Wavelets (EZW) [2] [3], and Wavelet Bases, e.g. Haar [8], Bi-Orthogonal (BIOR) Wavelets [2] will have to be evaluated for different compression parameters in order to determine which one would be suitable for image compression. The effect of the proposed scheme also needs to be studied on various types of images including medical, military and images from day-to-day consumer use.

CHAPTER 3: IMAGE SEGMENTATION

3.1 Introduction

The technology of automatic ROI detection is based on image segmentation. The standards available are not suitable for practice because of their unreliability. In practical environments, where the image can be contaminated by various noise sources including sensor noises, environment and radiation and hence most of the methods currently available for automatic segmentation are not likely to be able to match the requirements.

Segmentation involves dividing the image into two regions (ROI and background) based on a global or a local threshold. Using a local threshold is obviously computationally intensive; hence for the sake of ROI detection, a global threshold approach for segmentation is ideal.

Threshold segmentation method is a widely used and effective method for segmentation grayscale image. The segmentation threshold value determines the segmentation results. Therefore it becomes very important to choose an optimal threshold in such a way that the ROI pixels and background pixels of image can be clearly distinguished by their gray level values. Segmentation by thresholding used only the gray levels or histogram of the image is used while ignoring other image information such as spatial information. By choosing an adequate threshold value, ROI can be extracted from grayscale image. Selection criterion of segmentation threshold value plays an important role in threshold segmentation method.

Various Global thresholding methods currently exist including Ostu [1], maximal entropic thresholding [4], minimum error threshold value method. At present the most common segmentation method is Ostu method which would be satisfactory if the ROI is greater than 30% of the image [2] . However for smaller targets or multiple targets, the approach has to be different or at least modified.

A growing segment technique may be used for selecting an initial seed from the target region based on histogram data and using a growing principle to determine which segments of the image that are to be added to the ROI cluster. This process is computation intensive and slow.

Since 1980, various methods that use entropy to determine the thresholding based on entropy have been suggested [3]. However, most of these utilize only one-dimensional entropy, which solely depended on the gray-level distribution. When the signal-to-noise of image is low and the background of image is complex, segmentation results with one-dimensional entropy segmentation method are poor [4].

Various research papers suggest the use of a 2-D entropy approach to determine a global threshold and segment the image based on that threshold [4] [5] [6]. This process involves calculating a 2-D histogram of the image and calculating the point corresponding to the maximal entropy. This maxima of the 2-D histogram serves as the threshold for the segmentation process. The calculation of 2-D histogram is not very computationally intensive, however the algorithms for calculating the entropy from the 2-D histogram are. Multiple algorithms to calculate the entropy in a short duration are suggested in [4], [5], [6] and [12]. One such approach is using the Tsallis entropy method. Tsallis entropy when modified can serve as a fast method for calculating the global threshold for a given image based on its 2-D Histogram.

3.2 2-D Histogram

3.2.1 Definition

A 2-D histogram is a plot of number of pixels with a given gray level and a gray average. The gray average is the average gray level in a 3x3 neighborhood of a pixel.

Essentially a 2-D histogram is a three dimensional plot with Gray Levels from 0 to L-1 and gray average also from 0 to L-1 as the two base axes and the number of pixels (often normalized) corresponding to each Gray level and Gray average pair as the third.

3.2.2 Generation of the 2-D Histogram

For the generation of the 2-D histogram, the average gray level (and Gray Average) of each pixel in the image needs to be calculated. If the pixel P_{xy} has a gray level s_{xy}, obtained using the value I (P_{xy}), where I is the gray level operator. The Gray Average t_{xy}, is computed using Eqn. (1) [2].

$$t_{xy} = \sum_{m=x-3}^{x+3} \sum_{n=y-3}^{y+3} I(P_{mn}) \qquad (1)$$

The use of an averaging filter that averages the gray values in a 3x3 neighborhood would be suitable for achieving the desired result. Now once the Gray Average for all pixels is calculated, the construction of 2-D histogram can be done. This is done by calculating the (s_{xy}, t_{xy}) pair for each pixel and using this data to plot the number of pixels r_{st} which correspond to each (s, t) pair vs. the Gray value s on one axis and the Gray Average value t on the other axis. This is the third axis or represented by color in the 2-D histogram. The 2-D Histogram need to be normalized before doing any further calculations and for this the value p_{st} is defined in Eqn. (2).

$$p_{st} = \frac{r_{st}}{M \times N} \qquad (2)$$

where M and N are the dimensions of the source image in pixels.

A typical 2-D histogram for an image with 256 gray levels is shown in Figure 2. Coordinate y is gray average of region while coordinate x is gray level of the pixel. In the histogram, r_{ij} is the number of pixel in the image, whose gray level is i and gray average of region is j. For any threshold vector is (s, t) the 2-D histogram can be divided into four quadrants – A, B, C and D. The gray level and the gray average of the ROI or the background have little difference due to the slow change about the gray value of pixel in the ROI or the background region. So the diagonal areas in 2-D histogram are chiefly occupied by the ROI and the background pixels. The region A has the background pixels and D has the ROI. The quadrants B and C are mostly noise and edge pixels and are not considered since they do not contain important information [2], [4]. The threshold (s*, t*) that optimally divides the histogram into four quadrants has to be found out using the maximum entropy.

3.3 Entropy and Threshold Calculation

The maximum entropy principle states that, for a given amount of information, the probability distribution which best describes our knowledge is the one that maximizes the Shannon entropy subjected to the given evidence as constraints. The 2-D entropy thresholding segmentation method separates an image with the thresholding based on 2-D histogram. [2], [4].

The entropy of a discrete source is often obtained from the probability distribution p = {p_i}; therefore $0 \leq p_i \leq 1$ and $\sum_1^k p_i = 1$.

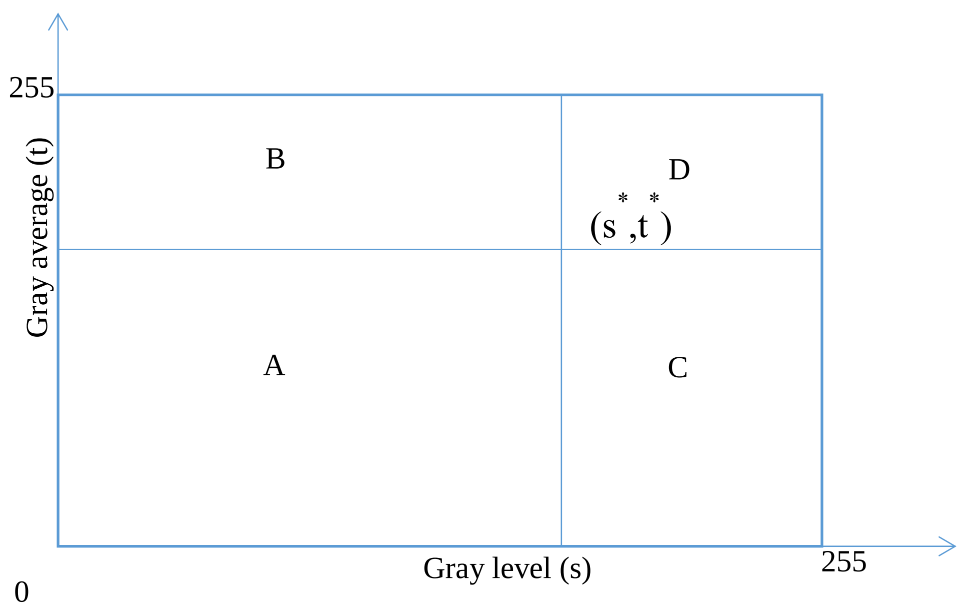


Figure 2 2-D gray histogram

3.3.1 Tsallis Entropy

Tsallis Entropy [13] is based on the extension of Shannon entropy principle which is described in Eqn. (3).

$$S = -\sum_{i=1}^{k} p_i \ln(p_i) \quad (3)$$

where k is the total number of states.

If we consider that a physical system can be decomposed in two statistical independent subsystems A and B, the Shannon entropy has the additivity property S(A+ B) = S(A) + S(B). This formalism has been shown to be restricted to the Boltzmann-Gibbs-Shannon (BGS) statistics [12]. However, for non-extensive physical systems, some kind of extension appears to become necessary. The Tsallis entropy of order q for two dimensions is defined in Eqn. (4) [1] [4].

$$S_q = \frac{1}{1-q}\left(1 - \sum_{i=0}^{255}\sum_{j=0}^{255} p(i,j)^q\right) \quad (4)$$

where $0 < q < 1$.

To speed up the process of calculations, the following assumptions are made. Since the quadrants B and C are not necessary, they can be ignored; and since the other two quadrants consist of the background and target, they both can be considered independent distribution. The normalization of these quadrants can be accomplished by calculating the posteriori probability, $P_A(s,t)$, $P_D(s,t)$ as defined in Eqns. (5) and (6).

$$P_A = \sum_{i=0}^{s}\sum_{j=0}^{t} p(i,j) \quad (5)$$

[1] All calculations of entropy assume the image has a total of 256 unique gray levels.

$$P_D = \sum_{i=s+1}^{255} \sum_{j=t+1}^{255} p(i,j) \qquad (6)$$

Now to simplify the calculations further, assuming the off-diagonal probabilities are zero, the following approximation is made, $P_D \approx 1 - P_A$.

The Tsallis entropies associated with the background and the object can now be calculated using Eqns. (7) and (8). S_q^A is the Tsallis entropy associated with the quadrant A and S_q^D is the entropy associated with the quadrant D.

$$S_q^A = \frac{1}{1-q}\left(1 - \sum_{i=0}^{s}\sum_{j=0}^{t}\left(\frac{p(i,j)}{P_A}\right)^q\right) \qquad (7)$$

$$S_q^D = \frac{1}{1-q}\left(1 - \sum_{i=s+1}^{255}\sum_{j=t+1}^{255}\left(\frac{p(i,j)}{1-P_A}\right)^q\right) \qquad (8)$$

3.3.2 Threshold Calculation

The threshold pair (s*, t*) which divides the histogram optimally into four quadrants is luminance pair (s, t) associated with the maximum entropy *max* (S_q^{A+D}). This is described in Eqn. (9) [4].

$$(s^*, t^*) = Arg\ Max\ S_q^{A+D}(s,t)$$

$$= Arg\ Max\left[S_q^D(s,t) + S_q^A(s,t) + (1-q)S_q^D(s,t)S_q^A(s,t)\right] \qquad (9)$$

The process of finding the threshold is a long iterative process since every value of (s, t) needs to be potentially checked for the entropy calculation. However, by restricting the threshold point to the diagonal of the 2-D histogram, the computations are reduced significantly. So instead of having the threshold at (s*, t*), the point (t*, t*) will now be considered.

The usual approach is to start at (0, 0) and go up to (255,255) and find the maxima of the 2-D entropy which will serve as the threshold. This process can also be significantly sped up by using a known range of threshold values for different set of images. For instance, if a set of MRI images would most likely have their threshold point at (127, 127) then the range (120,120) to (135,135) can be considered for calculation process. The known range of threshold values for a set of images is decided based on thresholds that were obtained for similar images in the past.

Once the threshold is determined, a mask needs to be created from the threshold that will serve as a template to distinguish the ROI from the background part of the image. This shall be referred to as "Binary ROI mask" or simply mask. The mask is created by assigning a value '1' to all pixels in the input image whose gray value and gray average exceed or equal to the threshold. The rest of the pixels are given a value '0'; yielding a black and white image of the mask.

The value of q in Eqns. 7 and 8 above also need to be varied for an optimal threshold determination. Again, for a specific type of images, the q value is most likely going to be the same. For calculations, a 256 × 256 pixels image shown in Figure 3 is chosen. This image has 256 gray levels. The image is an MRI scan of the human brain reproduced with permission from [14]. The 2-D histogram of the MRI scan is as shown in Figure 4. The effect of different values of q can be seen in the Table 1. All the images are masks generated by different q values based on the image shown in Figure 3.

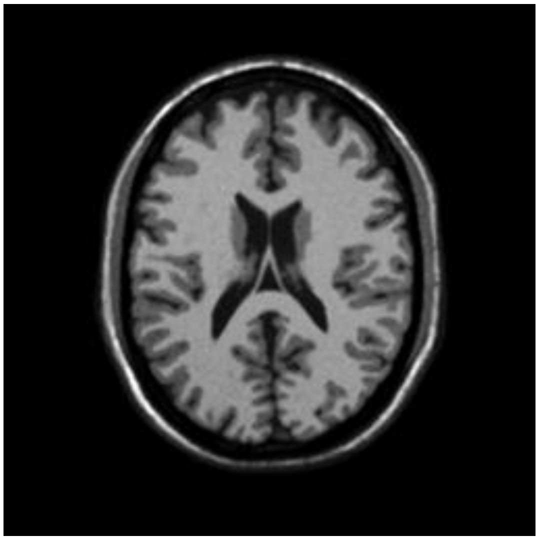


Figure 3 MRI scan of brain, input image

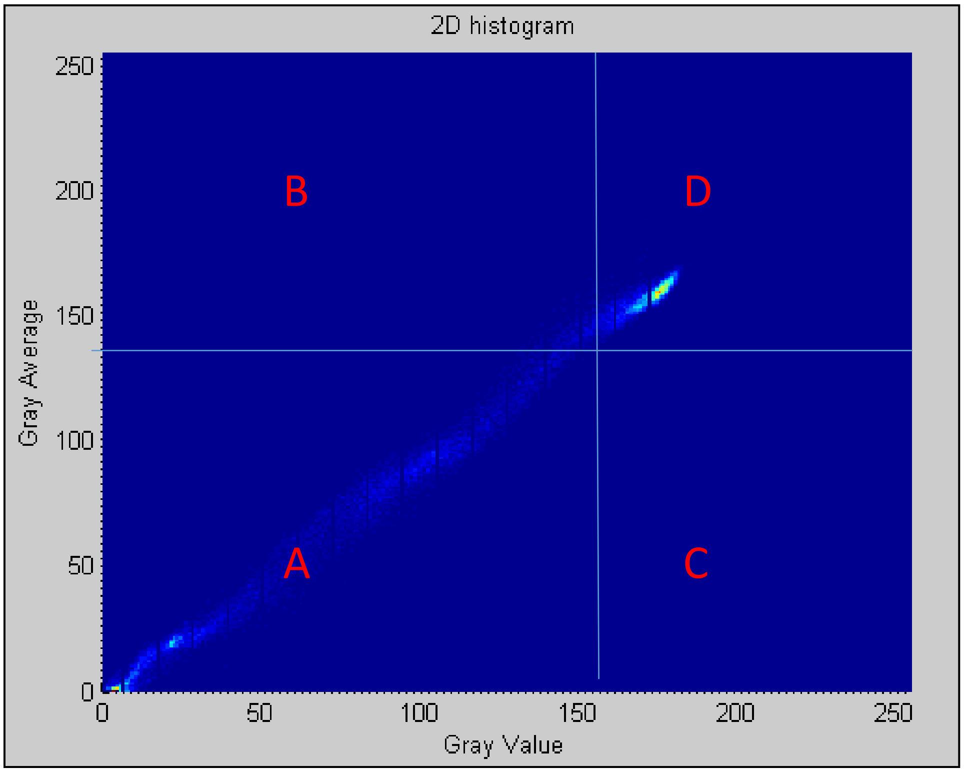


Figure 4 2-D histogram of MRI image

Table 1 Effect of different "*q*" values on segmentation process

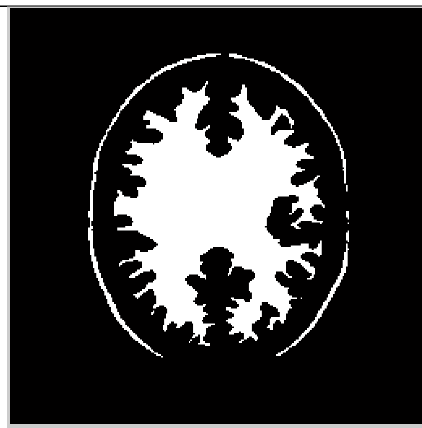
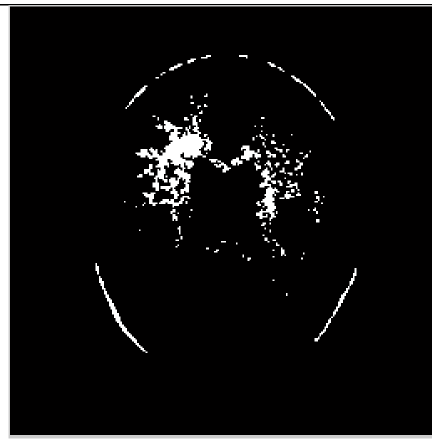
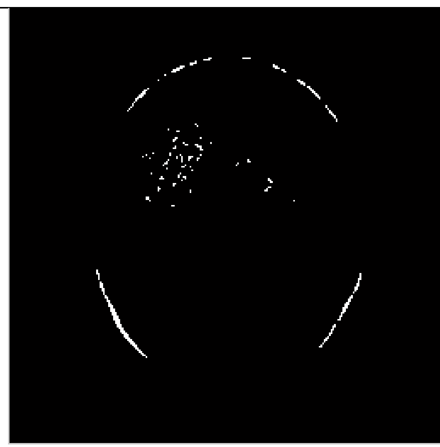
q value	Segmented mask of ROI without clean-up	Threshold pair (s,t)
0.1		(168 ,118)
0.25		(170, 115)
0.4		(178 , 150)

Table 1 (Continued)

q value	Segmented mask of ROI without clean-up	Threshold pair (s,t)
0.5		(185,183)
0.6		(190, 185)
0.7		(192,187)

As observed, higher q values have drastic effect on the segmentation process reducing the ROI to a very small region. At values of $q > 0.6$, the ROI is virtually absent. Hence the use of lower q values is preferred. A q value in the range of 0.25 to 0.4 is suggested for different applications such as military or medical images.

The mask needs to be continuous for improved efficiency of the compression algorithms, hence the pixels of the background completely enclosed within the ROI pixels, are included in to the ROI. Once the mask is generated, it is cleaned of different artifacts (small unconnected regions) using image erosion or any other noise reduction morphological processing. This step referred to as the mask clean-up is optional, however, highly recommended so as to improve performance in compression.

The mask clean-up is done using morphological process of erosion which is the removal of pixels from the object boundaries. The number of pixels removed depends on the structure element. The erosion is done by considering the minimum value of the pixel in the neighborhood defined by a diamond structure. Using a q value of 0.25, after erosion of the mask using a diamond morphological structure, the image is segmented and the ROI and background are extracted and the images are shown in Figures 5 and 6.

It may seem like a very bad segmentation of the image, however, the portions of the image with highest information content such as tumours or lesions seen on the MRI would be of a higher gray-level (close to white) and thus would be included in the ROI automatically by this algorithm.

Once the ROI is separated from the background, they can be compressed separately to obtain their compressed forms.

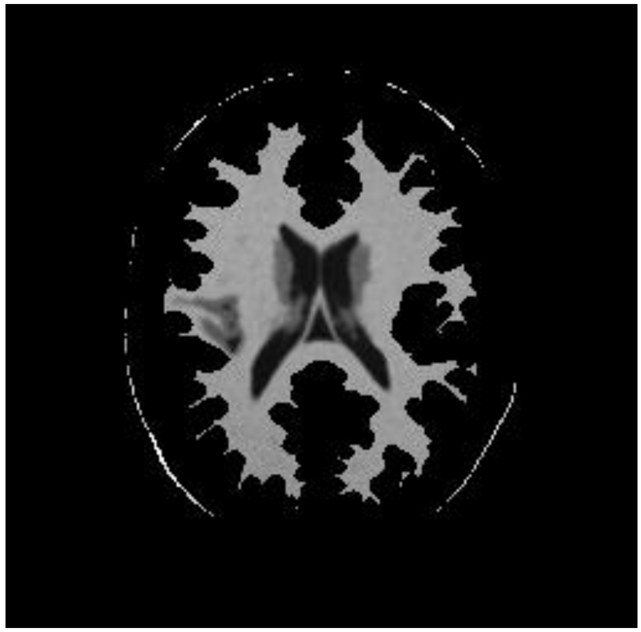


Figure 5 ROI of the input image (q = 0.25)

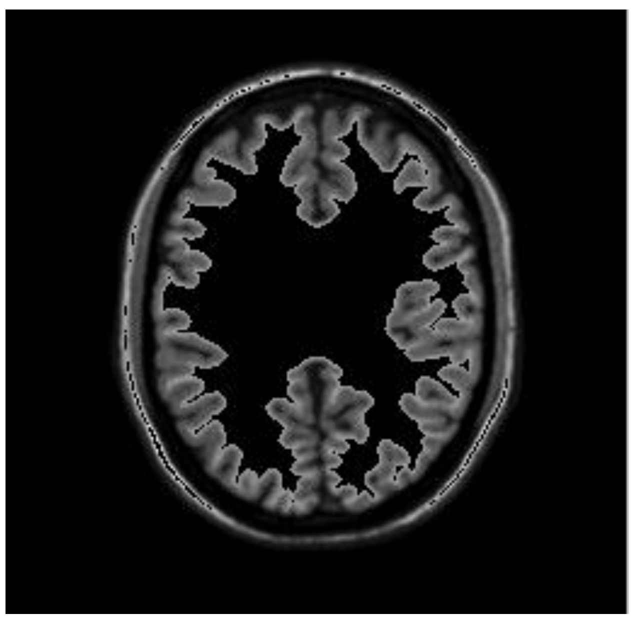


Figure 6 Background of the input image (q = 0.25)

CHAPTER 4: COMPRESSION

Compression and decompression of the images is the last step in the process. These are aimed at reducing the file size of the image without compromising the visual quality of the image. The methods used are lossy compression methods.

4.1 Wavelet Based Image Compression and Decompression

One of the most successful applications of wavelet methods is transform-based image compression (also called coding) [15]. The compression is done by a series of steps involving wavelet transform, thresholding and quantization before being encoded for storage or transmission. While quantization reduces the number of bits required for each pixel by reducing the number of gray levels from 256 to 16. However this quantization is done after the image has been decomposed by wavelet. The coefficient thresholding can be either global or local, and Fixed or Huffman coding can be used for the encoding process. The full process of compression is shown in the block diagram in Figure 7. Decompression of the image is done by reversing the above process. This is shown in Figure 8.

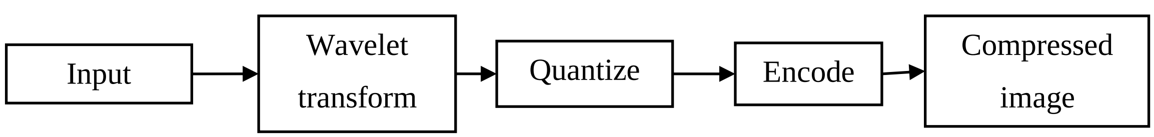


Figure 7 Image compression using wavelets.

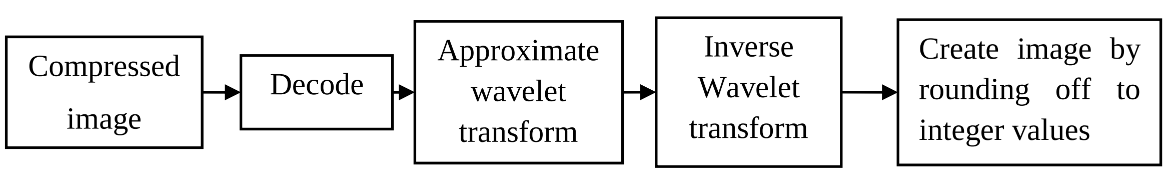


Figure 8 Decompression process

4.1.1 Progressive Methods of Image Compression

More sophisticated methods are available which combine wavelet decomposition and quantization. These are progressive methods. The methods are most commonly used include EZW (Embedded Zerotree Wavelet) algorithm, SPIHT (set partitioning in hierarchical trees) algorithm, ASWDR (adaptively scanned wavelet difference reduction) algorithm which achieve some of the lowest errors per compression rate and highest perceptual quality [15]. More in-depth description and analysis of these methods can be found in [15].

4.1.2 Selection of Wavelet Bases

There are a various number of wavelet bases available to choose from for the decomposition of the image, these include Haar, Daubechies, Bi-orthogonal, and Symlets. Bi-orthogonal wavelets, especially Bior3.5 and Bior3.7 are among the best wavelets for compression based on PSNR and compression performance [2]. Haar wavelets can also provide similar compression performance as other wavelets.

4.2 Compressing the ROI and Background Separately

In the first approach, Global thresholding of coefficients and Huffman encoding is used for compression of the ROI and background separately. The ROI is specifically set to have a lower rate of compression and the background is set to a higher rate of compression. The results of compression with other methods such as SPHIT, EZW, and ASWDR along with quality measurement comparisons are discussed in Chapter 5.

CHAPTER 5: QUALITY MEASUREMENTS AND COMPARISIONS

This chapter defines four quantitative measures, two each for compression performance and perceptual quality. The results of using different compression methods and comparison of their quality are also included.

5.1 Definitions of Measurements

5.1.1 Compression Performance

The effectiveness of compression is measured usually by two parameters. These two parameters define how well the input image is compressed. There will be four total measurements for each input image, two each for the ROI and background.

5.1.1.1 Compression Ratio (CR)

The Compression Ratio (CR) is a ratio of the final compressed file size to the original file size expressed in percentage. Lower the CR, better is the compression.

$$CR\% = \frac{Final\ Fize\ Size}{Initial\ File\ Size} \times 100 \tag{10}$$

5.1.1.2 Bit-per-pixel Ratio (BPP)

The Bits-Per-Pixel ratio (BPP) is a measure of number of bits required to store an image pixel. Lower the BPP, better is the compression.

5.1.2 Perceptual Quality

The two common measurement of perceptual quality are Mean Square Error (MSE) and Peak Signal-to-Noise Ratio (PSNR). These describe how well the reconstructed image looks to a human observer.

5.1.2.1 Mean Square Error (MSE)

Mean Square Error (MSE) between two images P and P_c with dimensions M×N is defined as in Eqn. 11. Lower the value of MSE, lower the error.

$$MSE = \frac{1}{M \times N}\sum_{i=1}^{M}\sum_{j=1}^{N}(P[i,j] - P_c[i,j])^2 \tag{11}$$

5.1.2.2 Peak Signal-to-Noise Ratio (PSNR)

PSNR tends to be cited more often since it is a logarithmic measure, and our brains seem to respond logarithmically to intensity. It represents a measure of the peak error and is expressed in decibels (dB) as defined in Eqn. 12.

$$PSNR = 10 \times \log_{10}\left(\frac{255^2}{MSE}\right) dB \tag{12}$$

Increasing PSNR represents increasing fidelity of compression; which means, higher the PSNR, better the quality of the compressed or reconstructed image. Generally, when the PSNR is 40 dB or higher, the two images are virtually indistinguishable by human observers.

5.2 Comparison of Performance of Different Compression Schemes

5.2.1 Effect of Compression Methods

The image shown in Figure 3 after segmentation and separating the ROI from the background was compressed using various schemes including Global thresholding and Huffman encoding, SPHIT, EZW and using various wavelet bases. Table 2 shows the various images after the compression process and compares the performance based on the parameters defined above. There are two sets of compression performance measurements one each for Background (BG) and the ROI. The PSNR is calculated over the whole image rather than just the ROI.

Table 2 Comparison of different compression methods

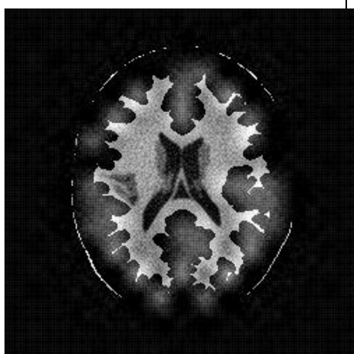
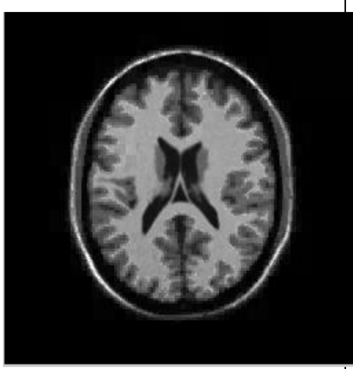
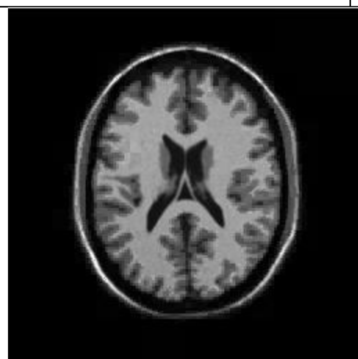
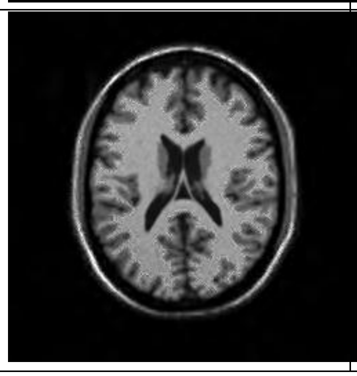
Compression Method	Image Output	Compression Performance		Perceptual Quality	
		BPP	CR (%)	MSE	PSNR (dB)
Global thresholding of coefficients and Huffman encoding		BG : 0.03 ROI : 1.04	BG : 0.33 ROI : 12.96	1573.6	16.16
SPIHT with Haar wavelet bases		BG : 0.54 ROI : 1.15	BG : 6.71 ROI : 14.36	51.55	31.01
SPIHT with Bior3.5 wavelet bases		BG : 0.59 ROI : 1.74	BG : 7.32 ROI : 21.74	43.99	31.70

Table 2 (Continued)

Compression Method	Image Output	Compression Performance		Perceptual Quality	
		BPP	CR (%)	MSE	PSNR (dB)
EZW with Haar wavelets		BG : 0.81	BG : 10.10	38.63	32.26
		ROI : 1.82	ROI : 22.73		
EZW with Bior3.5 wavelets		BG : 0.94	BG : 11.78	37.23	32.42
		ROI : 2.84	ROI : 35.50		

Though Global thresholding and Huffman encoding process shows very good compression performance, the perceptual quality of the compressed image is very bad. It can clearly be seen that the choice of progressive methods for compression such as the SPIHT and EZW are best suited for this application. SPIHT method with Haar wavelet bases is most suited for compression of images. This gives a final file size of about 21% of the original file size, while keeping the PSNR greater than 30 dB. Although there might be uses to Global thresholding and Huffman encoding in radar and infrared images; the use in medical images is limited.

5.2.2 Effect of Increasing Background Compression

Increasing the compression ratio of background will lead to lower file sizes. However, the quality of the image on the whole is lost. Figure 10 shows the effect of increasing the compression ratio in the background pixels, while maintaining the same compression ratio in the ROI pixels when using a SPIHT with Haar wavelet bases. The compressed file would be about 10% of the original file size with an overall peak signal-to-noise ratio of 22.68 dB while the PSNR of the ROI is 42.56 dB. Further increase in the compression ratio of the background leads to image shown in Figure 11. This has a file size of just 7.5% of the original file size while the PSNR of the ROI is 38.81 dB.

5.3 Study of the Processes Using Different Set of Images

Once the segmentation process, compression method and ratios are decided, it is important to study the effect of this procedure on various test images. A set of three other images were chosen for the comparison study. All gray scale images from different imaging fields – medical, military, everyday use. All files were passed through the same segmentation and compression algorithm and the output observed.

This effect of ROI based compression on different images is shown in the table 3 below. The compression ratio is that the combined compression ratio of both the ROI and background. The PSNR is calculated on the whole image. In all cases, the PSNR of the ROI was well above 40 dB. It can be clearly seen that the information in the ROI is not lost, however the file size reduction is achieved. The performance for medical images and military image are good. However, the process does not detect all the important regions of the image from the coins image, leaving out a coin from the ROI. Also the edges of the coins are lost to the background compression.

5.4 Comparison of Performance with Reference Papers

The results from the proposed method are compared with the results from the methods proposed in reference papers. Since different papers use different metrics for quantifying results and measuring performance, it is necessary to use similar metrics when comparing the results. Most papers use PSNR to measure the perceptual quality of the compressed image, while the compression performance is measured by a varied variety of metrics. The metrics and their meaning are described before the comparison of results is done in the following sections.

5.4.1 Comparison with [8] – MRI Image Data

A process where Haar wavelets are used to compress the background and ROI that are selected manually from the input medical (MRI) images is used in [8]. The compressed image is very blurry in the background losing most of the data due to lossy compression. The ROI is well preserved. The metric used for measurement of compression performance is file size reduction which can be translated to Compression Ratio as in Eqn. 11. The proposed method was implemented and tested for the same input data shown in Figure 3.

$$Compression\ Ratio\ \% = 100 - File\ Reduction\ \% \tag{11}$$

The compression method used in [8] yields a compression ratio of about 25% while the PSNR drops 24.58 dB. The details in the ROI are retained however the background is blurry due to the very lossy compression. The method proposed in this thesis achieves about 21% and the PSNR of about 31 dB. The background of the compressed method also well preserved. The comparison is shown in Table 4. When this data is averaged over multiple images, the method proposed in [8] provides a compression ratio of 26% and the average PSNR of about 25 dB. The method proposed in this thesis achieves a 20.62% compression while keeping the PSNR at about 31.8 dB.

5.4.2 Comparison with [2] – IR Image Data

An algorithm to automatically detect and segment the ROI and background parts of the image from battle field IR cameras was described in [2]. The compression of the images is achieved using a two-tier encoding scheme and a JPEG2000 compression. The compression performance is measured in bits-per-pixel and the perceptual quality is measured in PSNR of the ROI.

The method used in [2] achieves a BPP of 0.08 and a PSNR in the ROI of 31 dB. The method proposed in this thesis achieves a PSNR of about 40 dB with similar images while the BPP increases to 0.21. The details in the background are lost almost completely in both the methods while the information in the ROI is preserved. The reason for the increase in the BPP in the proposed method could be due to one or more of the following reasons - the presence of multiple targets / ROI segments in the image, irregular shape of the ROI as opposed to a rectangular ROI segment used in the method in [2] and due to the use of two-tier encoding scheme in [2].

The effects of two tier encoding and rectangular extensions of the Regions of interest need to be studied further to completely understand the reasons for the increased value of bit-per-pixel in the compression performance.

Figure 9 Effect of increasing compression of the background pixels – [CR : 10.2%]

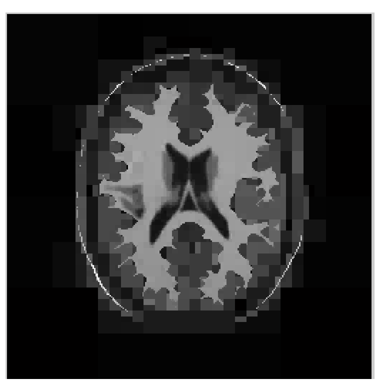
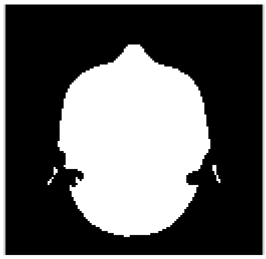
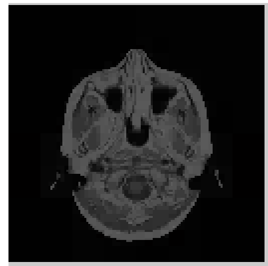
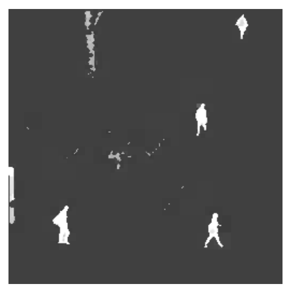
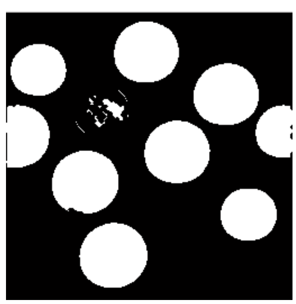


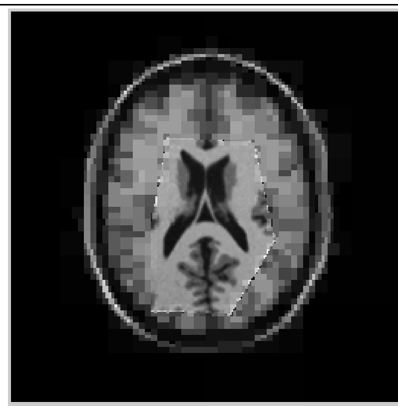
Figure 10 Effect of increasing compression of the background pixels – [CR: 7.5 %]

Table 3 Study of ROI based compression on different images

Input Image [2]	ROI Mask	Compressed Image	CR (%)	PSNR (dB)
			11.31	28.10
			3.31	17.33
			14.56	16.17

[2] Images 1 and 3 from the input images column are in public domain. Image 2 from the input images column is reproduced with permission from [16]. The permission is included in appendix A

Table 4 Comparing performance of [8] and the proposed scheme

Method	Output Image	CR (%)	PSNR (dB)
Proposed Method		20.96	31.70
Method from [2]		25.21	24.58

CHAPTER 6: CONCLUSION AND FURTHER WORK

6.1 Conclusion

A method for compressing gray scale images based on ROI was implemented. The ROI was detected and segmented automatically using Tsallis entropy approach. Improvements to this approach to decrease computation speed were also described and tested. Different methods of image compression with different compression parameters were implemented and tested.

The proposed method performs well for medical images such as MRI scans, IR images from battle field or surveillance cameras. The compressed images have a compressed ratio of at least 20% making this method good for transmission in small bandwidth connections. Since the details in the Regions of interest are preserved completely, the most important data is not damaged due to the compression process.

By changing the q values during image segmentation process, it is possible to fine tune the ROI to better suit different applications. If the ROI is just a tumor or a lesion on the MRI, a higher q value would be able to achieve that. The selection of smaller ROI leads to a better compression ratio and hence a smaller file size. By changing the compression parameter such as making the background compression lossier will also improve the compression performance. The choice of a very lossy compression on the background on an MRI image (Figure 10) led to a compression ratio of about 10% while giving a clear ROI and still preserving background enough to identify the position of the ROI with respect to the background.

The proposed method performs very well on MRI images and sufficiently well on IR images. However, the use of the method on everyday images is limited because of the loss of information in the edge and background pixels may not be acceptable for all types of images.

6.2 Directions for Future Work

While encoding ROI into an image using Max-Shift algorithm or bit-by-bitplane encoding schemes, a rectangular or a circular ROI works best [9]. The effects of having regular regions of interest, such as a rectangular extension of the existing regions of interest needs to be studied further. Various other methods exist for compression that can be studied and applied to understand the effects of these on the compression of the images. The use of JPEG 2000 compression scheme and using a two-tier encoding scheme as used in [2], can be incorporated into the proposed scheme and effects of this investigated further. The encoding of the ROI using Max-Shift and bit-by-bitplane encoding schemes as discussed in [9] need to be studied and experimented with further. The use of fractal image compression and its effects on compression also need to be studied.

The use of this scheme in DICOM image transmission and other applications of ROI based compression such as extension of this process to video files are other avenues for future work.

REFERENCES

[1] N. Ostu, “A threshold selection method from gray-level histogram”, IEEE Trans. on SCM, Vol. 9, No. 1 pp: 62-66, 1979.

[2] S. Yuan; X. Xie, Q. Guo and R. Zhu, "Research on infrared image compression based on region of interest automatic detected," Proc. of the International Conference on Audio, Language and Image Processing (ICALIP), pp. 883-888, 2008.

[3] T. Pun "A new method for grey-level picture thresholding using the entropy of the histogram". Signal Processing, pp. 223-237, 1980.

[4] S. Kumar, M. Pant, and A.K. Ray, “2D Maximum Entropy Method for Image Thresholding Converge with Differential Evolution”, Advances in Mechanical Engineering and its Applications, vol. 2, no. 3, pp. 189-192, 2012.

[5] M. A. El-sayed , S. Abdel-khalek and E. Abdel-aziz, "Study of Efficient Technique Based On 2D Tsallis Entropy For Image Thresholding", International Journal on Computer Science and Engineering (IJCSE) , vol. 3, no. 9, 2011.

[6] J.N. Kapur, P.K. Sahoo and A. Wong: "A new method for gray-level picture thresholding using the entropy of the histogram. Computer Vision", Graphics and Image Processing, pp. 273-285, 1985.

[7] A. Khademi and S. Krishnan, "Comparison of JPEG 2000 and Other Lossless Compression Schemes for Digital Mammograms", Proc. of the 27th Annual International Conference of the Engineering in Medicine and Biology Society (EMBS), pp. 3771-3774, 2006

[8] R. Shah, P. Sharma and R. Shah, "Performance Analysis of Region of Interest Based Compression Method for Medical Images" Proc. of the Fourth International Conference on Advanced Computing & Communication Technologies (ACCT), pp. 53-58, 2014.

[9] C. Doukas and I. Maglogiannis, "Region of Interest Coding Techniques for Medical Image Compression" , IEEE Engineering in Medicine and Biology Magazine , vol. 26, no. 5, pp. 29-35, 2007.

[10] Y. Wu and S. Tai, "Medical image compression by discrete cosine transform spectral similarity strategy", IEEE Trans. on Information Technology in Biomedicine, vol. 5, no. 3, pp. 236-243, 2001.

[11] J. Jeng; C. Tseng and J. Hsieh, "Study on Huber Fractal Image Compression", IEEE Trans. on Image Processing, vol. 18, no. 5, pp. 995-1003, May 2009.

[12] L. Zheng, Q. Pan, G. Li and J. Liang, "Improvement of Grayscale Image Segmentation Based on PSO Algorithm" Proc. of the International Conference on Computer Sciences and Convergence Information Technology(ICCIT), pp. 442 - 446, 2009.

[13] C. Tsallis, "Statistical Physics", pp. 480-487, 1988.

[14] "BrainWeb: Simulated Brain Database." <http://brainweb.bic.mni.mcgill.ca/brainweb/>

[15] J.S. Walker, "Wavelet-Based Image Compression", A Primer on Wavelets and Their Scientific Applications, 1999.

[16] J.W. Davis and M.A. Keck, "A Two-Stage Template Approach to Person Detection in Thermal Imagery," Proc. of the IEEE Workshops on Application of Computer Vision, (WACV/MOTIONS), vol. 1, pp. 364-369, 2005.

[17] G. Hu; M. Xiao and S. Yuan, "Detecting Automatically and Compression Algorithm for Infrared Image Based on Region of Interest," Proc. of the International Forum on Computer Science-Technology and Applications(IFCSTA) vol. 1, pp. 50-53, 2009

[18] R.K.-S. Kwan, A.C. Evans and G.B. Pike, "MRI simulation-based evaluation of image-processing and classification methods" IEEE Trans. on Medical Imaging, vol. 18 no.11 pp. 1085–97, 1999.

[19] D.L. Collins, A.P. Zijdenbos, V. Kollokian, J. Sled, N.J. Kabani, C.J. Holmes, and A.C. Evans, "Design and construction of a realistic digital brain phantom," IEEE Trans. on Medical Imaging, vol. 17, no. 3, pp. 463–468, 1998.

[20] C.A. Cocosco, V. Kollokian, R.K.-S. Kwan, A.C. Evans : "BrainWeb: Online Interface to a 3D MRI Simulated Brain Database" NeuroImage, vol. 5, no. 4, part 2/4, 1997 — Proceedings of 3-rd International Conference on Functional Mapping of the Human Brain, Copenhagen, May 1997.

APPENDICES

Appendix A Copyright Permissions

Below is permission for the use of Figure 3.

D. Louis Collins <louis.collins@mcgill.ca> 9:49 AM (14 hours ago)

to me, bert+bw

Adhokshaja,

you are more then welcome to use BrainWeb images in your thesis and in journal/conference publications. We ask only that you cite the papers listed on the BrainWeb site:

• R.K.-S. Kwan, A.C. Evans, G.B. Pike : "MRI simulation-based evaluation of image-processing and classification methods" IEEE Transactions on Medical Imaging. 18(11):1085–97, Nov 1999.

• D.L. Collins, A.P. Zijdenbos, V. Kollokian, J. Sled, N.J. Kabani, C.J. Holmes, and A.C. Evans, "Design and construction of a realistic digital brain phantom," IEEE Transactions on Medical Imaging, vol. 17, no. 3, pp. 463–468, 1998.

• C.A. Cocosco, V. Kollokian, R.K.-S. Kwan, A.C. Evans : "BrainWeb: Online Interface to a 3D MRI Simulated Brain Database" NeuroImage, vol.5, no.4, part 2/4, S425, 1997 — Proceedings of 3-rd International Conference on Functional Mapping of the Human Brain, Copenhagen, May 1997

Good luck with your masters thesis.

-Louis

D. Louis Collins, PhD McConnell Brain Imaging Center
Professor Montreal Neurological Institute
Neurology & Neurosurgery WB-314
Biomedical Engineering 3801 University St.
louis.collins@mcgill.ca Montreal, Quebec.
voice: (514)-398-4227
CANADA H3A 2B4
http://www.bic.mni.mcgill.ca/PeopleFaculty/CollinsDLouis

Below is permission for the use of Image 2 from Table 3.

Adhokshaja Achar Budihal Prasad <adhokshajaac@mail.usf.edu> 12:40 AM (7 hours ago)
to jwdavis

Sir,

I'm a Master's student in Electrical Engineering at the University of South Florida. I plan on using a couple of IR images from your "OTCBVS - Visual Computing and Image Processing Lab" database as a part of my master's thesis titled "Region of Interest based compression of gray scale images". I was hoping, I could get permission to use the images.
I'll be more than happy to provide with more information if you so desire. Hoping to hearing from you soon.

Thank you

Warm regards,

Adhokshaja Achar Budihal Prasad
Student, Dept. of Electrical Engineering
University of South Florida
Tampa, Fl, USA
Cell: +1 813 454 7165

Davis, James 8:09 AM (26 minutes ago)
to me

No problem. Just use the citation given on the webpage where the images are downloaded.

Jim Davis - Sent from my iPhone