Document Image Segmentation and Compression using Artificial Neural Networks and Evolutionary Methods

Thesis submitted to

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

By

AYSHA V

Reg No 2767

Under the guidance of

Dr B Kannan

DEPARTMENT OF COMPUTER APPLICATIONS

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY

COCHIN 682 022

Kerala

June 2013

Dr B Kannan (Supervising Guide)
Associate Professor
Dept of Computer Applications
Cochin University of Science and Technology

Certificate

Certified that the work presented in this thesis entitled ***"Document Image Segmentation and Compression using Artificial Neural Networks and Evolutionary Methods"*** *is based on the bonafide research work done by Aysha V under my guidance in the Department of Computer Applications, Cochin University of Science and Technology, Kochi -22 and has not been included in any other thesis submitted previously for the award of any degree.*

Kochi- 22

30-6-2013 ***Dr B Kannan***

Dr B Kannan (Supervising Guide)
Associate Professor
Dept of Computer Applications
Cochin University of Science and Technology

Certificate

This is to certify that all the relevant corrections and modifications suggested by the audience during the Pre-synopsis seminar and recommended by the Doctoral Committee of the candidate have been incorporated in the thesis.

Dr B Kannan

DECLARATION

I hereby declare that the present work entitled ***"Document Image Segmentation and Compression using Artificial Neural Networks and Evolutionary Methods"*** is based on the original work done by me under the guidance of Dr B Kannan, Associate Professor, Department of Computer Applications, Cochin University of Science and Technology, Kochi-22 and has not been included in any other thesis submitted previously for the award of any other degree.

Kochi -22 **AYSHA V**

29-6-2013

Acknowledgements

First of all I express my sincere gratitude to my guide Dr B Kannan, Associate Professor, Department of Computer Applications, CUSAT whose constructive criticism and analytical skills helped me to improve a lot throughout my work. Without his motivation I would not have published my papers at International Conferences.

I would like to extend my most sincere thanks to Dr S Babu Sundar, former Head, Department of Computer Applications, CUSAT for his guidance, encouragement and all that he had done for helping me to develop my professional and personal skills. He motivated me to present a paper in a National Conference held at Sri Saraswathi Thyagaraja College, Pollachi. I am certain that I will be benefited from his rigorous scientific approach and the critical thinking throughout my career.

I also want to thank Dr Pramod K V, Present Head of the Department of Computer Applications for all the favors that I received from his side. He has supported me always and motivated me to present three papers at ICMCM 2010. Also I thank all the office staff and faculty of Department of Computer Applications who helped me to complete my research work.

Most of all, my deepest thanks go to my husband B.Syamlal, my parents and my family. I cannot thank them enough for their love, support, sacrifice, motivation and their belief in me.

I would also like to thank Mr Prajith, Dept of Electronics, College of Engineering, Trivandrum, Mr Narayanan, Librarian of ER & DC, Trivandrum, Miss Geena, Research Scholar, Kannur University and Librarian, Central Library, CUSAT for supporting me with reference materials. I would also like to thank Dr AchuSankar S Nair for his informative website: www.achu.keralauniversity.edu on "How to do research". I also thank Mr Firoz Shah for informing me about conferences.

I also thank Mrs Babitha Raj, Akshaylal S, Renjini Krishna, Remya P.R, my friends and colleagues for helping me in typing, imaging and giving me constructive suggestions and comments.

Aysha V

Contents

Title	Page No

List of Tables

List of Figures

List of Algorithms

Chapter 1

INTRODUCTION

Contents

1.1 Motivation

Document image processing is relevant in maintaining documents as images for Digital Libraries, Engineering Drawings, communication through internet, facsimile etc. It also helps in maintaining the legacy of documents, the document archiving helps to maintain documents in websites, communicate through internet etc. The compressed documents help to reduce the bottle neck due to low band width.

Existing methods for segmentation and compression are suitable for Ordinary images and bi-level document images and not for all document images. A better method for document images is required.

1.2 Problem Statement

The work is intended to study the following important aspects of document image processing and develop new methods. (1) Segmentation of document images using adaptive interval valued neuro-fuzzy method. (2) Improving the segmentation procedure using Simulated Annealing technique. (3) Development of optimized compression algorithms using Genetic Algorithm and parallel Genetic Algorithm (4) Feature extraction of document images (5) Development of IV fuzzy rules.

This work also helps for feature extraction and foreground and background identification. The proposed work incorporates Evolutionary and hybrid methods for segmentation and compression of document images. A study of different neural networks used in image processing, the study of developments in the area of fuzzy logic etc is carried out in this work.

1.3 Document Image Processing

From the beginning of the scripting stage, human beings showed a tendency to preserve the documents for their successors. The writings or carvings in rocks, metal plates, leaves, papers etc were considered as documents. By the emergence of digital age, human beings started to scan the documents and store the replica as document images.

The aim of Document Image Processing is to retard the growth of paper bundles and to substitute for paper in storing and accessing information. It provides easy access to the electronic replicas of documents and cheap storage cost. Document Image Processing refers to the management of paper documents, records, forms by capturing, indexing, archiving, retrieving, and distributing them electronically. Document images are exact digitized replica of the original documents and allow document preservation. They are superior to paper documents because they can be economically stored, efficiently searched and browsed, quickly transmitted, and coherently linked together. The documents can be remotely retrieved by multiple users and manipulated using existing information technology. Moreover, unlimited number of hard copy printouts can be made for the convenience of the users [6].

1.3.1 Steps in Document Image Processing

Document image archiving is a typical application, in which incoming documents are digitized (unless they are initially digital), categorized, and archived in electronic form. The whole process may be set for fully automatic operation without human intervention. The digitization phase can be efficiently performed using scanners and facsimile, which is also relatively inexpensive technology. The *archiving* phase includes image enhancement, compression, recognition and indexing operations.

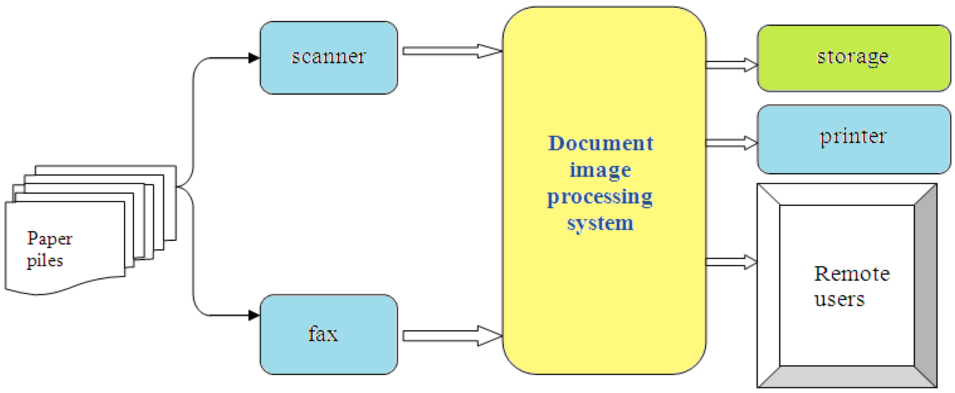


Fig 1.1 Document imaging

After digitization, document images are usually *enhanced* in order to correct deficiencies in the original and digitized documents. The enhancement is aimed to produce more readable version of the image (also known as *image restoration*) as well as to achieve better compression rates.

Enhancements methods include de-skew and de-speckle operations, line and border removal and *filtering*, and are useful for archival work where the quality of the original paper documents may have faded. Although the noise level can be low enough to affect the document readability, it can degrade the image enough to cause difficulties in the compression [4][5][6].

1.3.2. Present Scenario of Document Image Processing

Human Resources Management System is an example of Document Image Processing Application. It maintains records of employees for internal control purpose as well as to comply with legal requirements. Incoming resumes and recommendation letters are faxed or digitized paper documents should be categorized and distributed into appropriate archives for future examination without human intervention. OCR is performed for extracting the vital information from the documents (resumes) and for the indexing purposes. Original paper documents are disposed after digitization is completed [5].

In an image communication system, such as facsimile, image serves as a communication medium. The document is first digitized using an optical scanning device, and is then compressed and transmitted to the recipient, where it is re-printed or archived in an electronic form. The characteristic feature of this application is that the sender and the recipient are separated by a communication channel, usually a telephone line, which is the bottleneck of the system. The sender may not have sufficient memory to hold the entire image for the time between digitization and transfer. In this case, image scanning, compression and transmission are performed simultaneously [4].

As per the estimates of International Data Corporation (IDC) that about 8,000,000,000 line drawings exist in the world. Only about 13 % of them have been designed and originally stored in digital form. Nevertheless, there are still (and will continue to be) a large number of drawings that are stored as paper documents. A possible solution for engineering images is to perform a Raster-to-Vector conversion (RVC), where the bitmap image is segmented into vector primitives such as line segments, circles, arcs, etc. and stored with any CAD/CAM format. As per the opinion of Dr Eugene Ageenko, in Digital Spatial Libraries (DSL), raster map images are usually generated from a map database for digital publishing on CD-ROM or the Web. The main problem in digital spatial libraries is the huge storage size of the images. For example, a single map sheet of 5000×5000 pixels representing an area of 10×10 km^2 requires 12-25 Mb of memory and this is not the limit [6].

1.4 Objectives of the Present Study

- ✓ To develop a quad tree representation of the document image.
- ✓ To develop better segmentation techniques for document images.
- ✓ To extract features of document images
- ✓ To study and compare existing Document image compression techniques.
- ✓ To explore better algorithms for Document image compression.
- ✓ To develop methods to recognize foreground and back ground patterns from document images.
- ✓ To optimize segmentation and compression by Evolutionary methods like Genetic Algorithms, Simulated Annealing and Neural Networks techniques.
- ✓ To find methods to store the document image in a well compressed and reconstructable format after recognizing the patterns.

1.5 Scope of the Work

The work describes about a new document image segmentation method, a method of segmentation optimization using Simulated Annealing and a new method of compression optimization with Genetic Algorithms and parallel Genetic Algorithms of the document images for better storage and transmission through internet or facsimile. Since the work uses Evolutionary Algorithms we are able to give fast solutions for document image processing. Layered segmentation is possible with interval valued fuzzy sets. The IV fuzzy relational representation reduces storage space compared to other methods. It also helped in feature extraction. Optimization of document image processing algorithms will help in better performance. Comparative study of different algorithms in Different Document Databases will help to choose better algorithms. Literature survey of this work performed area wise study of Document image processing. In this work we have used Experimental Research methodology. And analysis of the algorithms are performed with the help of statistical and complexity analysis methods. More than two hundred document images are incorporated for study. In addition to that standard document image databases also explored.

1.6 System Framework

As a first step a data base of scanned document is created with variety of documents such as hand written, printed, with photographs, images, diagrams, tables etc. They are stored as JPEG images. The images are segmented into different regions using adaptive interval valued fuzzy set algorithm. The results are stored as IV-fuzzy relations. Then segmentation optimization is done with the help of Simulated Annealing algorithm.

Compression optimization is done with Genetic algorithms. Comparison of the work is done with contemporary works. Result and conclusions are listed at the end of each work.

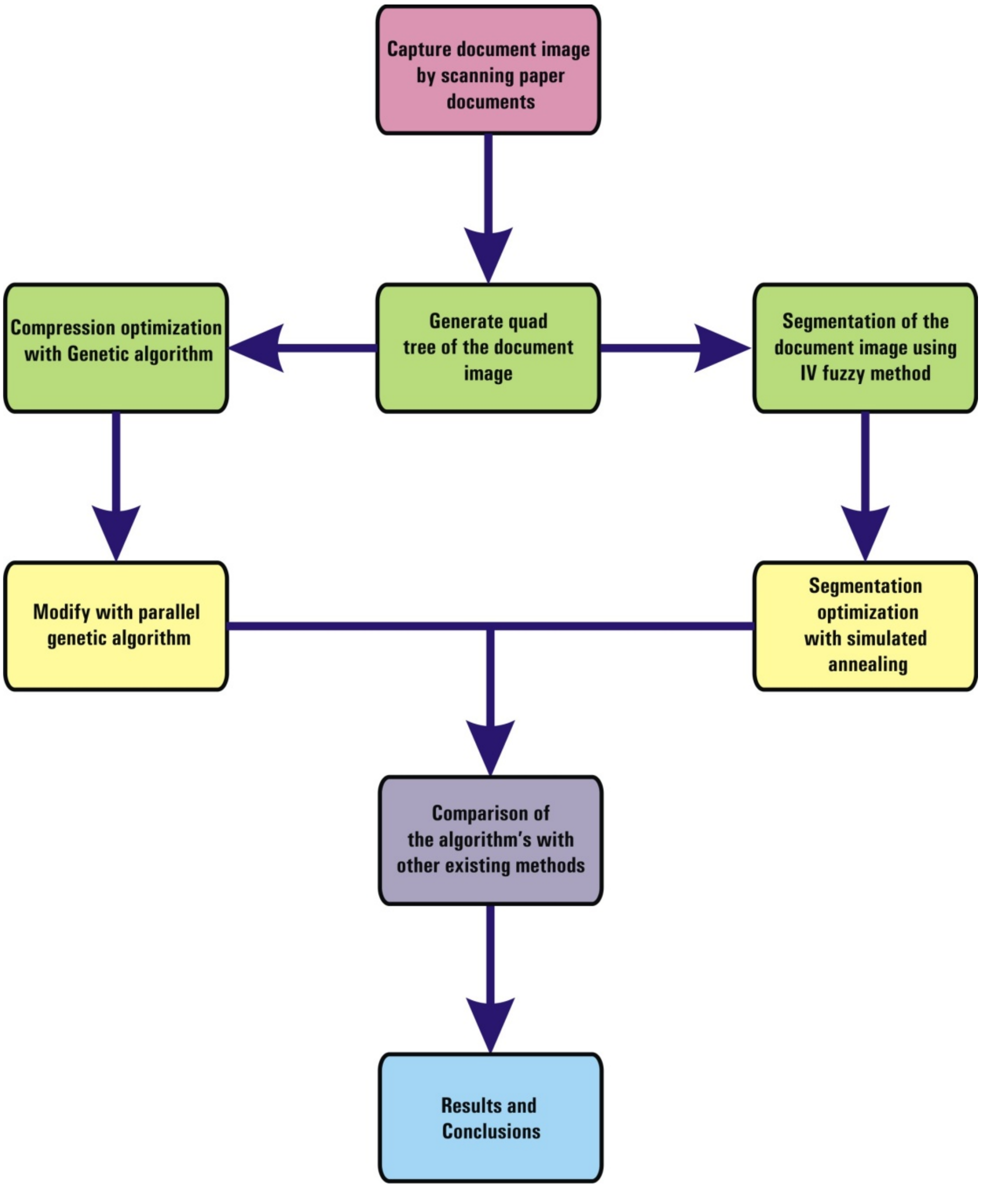


Fig: 1.2 Conceptual framework

1.7 Thesis Contributions

1. A new adaptive Interval Valued fuzzy based document image segmentation method.
2. IV Fuzzy Relational Representation to save storage space.
3. Optimization of segmentation with Simulated Annealing.
4. Compression optimization with Genetic Algorithm.
5. Compression optimization improvement with parallel Genetic Algorithm.
6. Comparative study of performance of different algorithms over standard document image databases is done.
7. Developed a new feature Extraction techniques and features useful for compression.

1.7.1 List of Research Papers

As a part of Research work various papers were presented and previously published in peer reviewed International Conference proceedings. They are listed below. Conferences were also attended on related topics.

1. Aysha V, Dr Kannan Balakrishnan, Dr S Babu Sundar, “ Image segmentation using Interval valued Adaptive Neuro Fuzzy Method”, International Conference on Mathematical Computing and Management 2010.
2. Aysha V, Dr Kannan Balakrishnan, Dr S Babu Sundar, “Optimization of Document image segmentation Using Simulated Annealing”,

International Conference on Mathematical Computing and Management 2010.

3. Aysha V, Dr Kannan Balakrishnan, Dr S Babu Sundar, Dr Pramod K V, “Document Image compression Optimization using Genetic algorithm”, International Conference on Mathematical Computing and Management 2010.

4. Aysha V, Dr Kannan Balakrishnan, Dr S Babu Sundar, “Parallel Genetic Algorithm for Document Image Compression Optimization” ICEIE 2010, Kyoto Japan

5. Aysha V, “Digital Image Processing and Artificial neural Network Approaches”, National Seminar on Image Processing and Criptography at Sree Saraswathy Thyagaraja College, Pollachi, 8th March 2008.

6. Aysha V, “Segmentation of Document Images for identifying or eliminating Background from text or pictures embedded in the text”, 21st Kerala Science Congress Kollam, 28-31 January 2009.

1.8 Organization of the thesis

The **first chapter** gives an Introduction to Document Image Processing, motivation, Relevance of Document image processing, explains the research problem, Objectives of the work, conceptual frame work of the proposed system, contributions in the thesis and organization of the thesis.

The **second chapter** deals with the literature survey of the related areas of Document image processing. The related areas are Digital Image Processing, Pattern Recognition, Artificial Neural Networks, Segmentation, Compression, Fuzzy Logic, IV Fuzzy Logic, Evolutionary Algorithms, and

Genetic Algorithms**.** This Chapter explains about the Present status in each of the areas and Limitations of the existing systems. A detailed study of fuzzy logic since its introduction in 1965 to current period is done and represented as a table. Then the need for soft computing and Evolutionary algorithms are listed. The Chapter deals with the image capturing techniques. Then it explains all the steps involved in Digital Image segmentation etc. It also covers a detailed study of the need for soft computing and evolutionary methods.

The **third chapter** describes a new method of Interval valued fuzzy sets in Document Image Segmentation. An adaptive Neuro Interval valued fuzzy model is designed with one hidden layer. Steps for creating quad tree are formulated. Its algorithm is also explained here. The construction of a fuzzy relation, and sample results in segmentation are explained in this chapter. This method suits different plain wise segmentation. Sample results of this method are shown in this Chapter.

The **fourth chapter** explains about a new method of optimizing segmentation with Simulated Annealing algorithm. During the annealing schedule, the intensity of the image is represented as temperature for Boltzmann's expression. Training sets, test sets, etc are decided and analysis of the work is done. For Simulated Annealing the work shows better results when the coefficient of the annealing schedule is set in between 0.8 and 0.9. The results show that the chances to get stuck at local minima are reduced when Simulated Annealing is used.

The **fifth chapter** describes about the new algorithm developed by us for document image compression optimization with Genetic Algorithm and Parallel Genetic Algorithm. The method applied over compressed document images with Huffman encoding and run Length encoding algorithms. A characteristic function with Hausdroff distance is used with the fitness function. The Optimization algorithms showed better performance over traditional methods.

The **sixth chapter** deals with Summary and Conclusion, major findings and suggestions. Conclusion, future scope and expansion of the work to other domains are explained here. The advantages of the Artificial Neural Network and evolutionary algorithms are listed here.

Sample outputs, comparative study with other methods etc are given at the end of each chapter. List of research papers, synopsis and references are given at the end.

1.9 Summary and Conclusion

The introductory chapter gave a gist about the thesis and thesis contributions. With Evolutionary methods we can digitize, store, transmit, process and retrieve document images easily and efficiently. In this thesis new approaches for document image segmentation, segmentation optimization and compression optimization are provided.

………………

Chapter 2

REVIEW OF LITERATURE

Contents

2.1 Introduction

Document Image Processing is a special area under Digital Image Processing, which have common steps but different methods. This attempt is for identifying the contributions in Document Image Processing and related areas. This research work needs an in depth study of the following areas for successfully completing segmentation and compression of document images. The allied areas involve Digital Image Processing, Data Compression Techniques, Pattern Recognition, Artificial Intelligence, Artificial Neural Networks, Fuzzy logic and Interval Valued fuzzy sets, Genetic Algorithms etc. Document Image Processing is important in today's world, because we need to handle bundles of paper documents. Even though computerization is done all over the world, the size of paper piles is also increasing. And most of the old documents, forms, palm leaf writings, hand written documents, records etc. are still in queue to enter into the electronic media. Computer Output Microfilm (COM) storage was one of the replica of documents in the early stages of computers. Now the scenario is changing and Document

Imaging is accepted as a fastest way to capture, store, index, retrieve, distribute/ share images of paper or historical documents. Nowadays various file formats like *.pdf (Portable Document Format), *.ps (Post Script) etc are available. For easy transfer of paper documents to Digital form Document Image Processing is relevant. Document Image Processing involves, image capturing, segmentation, enhancement, character recognition, compression etc. This study concentrates on segmentation and compression of paper document images. Document Image Processing involves the methods such as acquiring image, segmentation, enhancement, recognizing character patterns, background elimination, compression of document images, reconstructing document images etc. Various researches have been done on this area. And some of the research details relevant to this work are listed below.

The history of Digital Image processing begins with 1960s. Now, Digital Image processing has wide range of applications like Geographical Image processing, Space Image Processing, Medical Image processing, finger print analysis, video processing, Digital Steganography, Pattern Recognition etc. Digital Image processing is a most common method compared to analog signal (image) processing.

A digital image is represented as a two dimensional function $f(x,y)$ in which the amplitude of the function represents the intensity or gray level of the point or pixel (picture element or pel) at spatial coordinates (x,y) [84]. Digital image processing involves capturing of the image using scanner, barcode reader, digital camera, web cam or remote sensing probes or capturing of medical images like X-rays, MRI scan images, CT- scan images, Ultra Sound scan images or Doppler scan images and storing it in a computer, processing it digitally. The processing can be done for various purposes (like diagnosis with the help of medical images, analysis of surface of celestial bodies, analysis of spectrum, to clear unwanted data from an image, to compress the image for fast sending through the internet or to save storage space etc.) and that depends upon problem areas to be solved. Generally the processing involves image restoration, color image processing, wavelets processing, compression, enhancement, segmentation morphological processing, segmentation, various internal representations, object recognition etc.

Image enhancement means removing the noise from images, enlarging an image, removing unwanted portions to change focus to another object, selective colour change, orienting images or portion of the images, perspective correction or distortion, adding special effects, change background or merge with another image etc. Actually image enhancement is associated with image editing. When it is done with the help of image enhancement program it will become a digital image processing technique.

2.2 Previous Works in Image Segmentation

According to Dr Eugene Ageenko image capturing methods are OCR like barcode readers, scanners, digital cameras etc and Raster to Vector Conversion (RVC). The earlier works describes the application of wide range of methods for segmentation of images like, Hough transform(for line detection, peak detection and linking), Watershed segmentation, point, line, edge detection, Sobel edge detector(using mask), Prewitt Edge Detector(using mask), Roberts Edge Detector(using mask), Laplacian of a Gaussian (LoG) detector(this is a smoothing function), Zero – Crossings Detector (Same as LoG but convolution is carried out using specified filter function), Canny Edge Detector(Canny[1986])[83], Global and Local thresholding, Region based segmentation etc. And nowadays there is a trend to apply evolutionary methods to all areas of image processing. The main problem associated with Document Image Processing is that all the methods for Image processing cannot be directly applied to Document Image Processing. The decisions should vary from language to language, the amount of information in the document, variety of information present such as mathematical formulae, tables, pictures, plain text etc.

For segmentation purpose, sometimes, point, line and edge detection are necessary. The most common way to look for discontinuities is to run a mask through the image [20].

The Hough transform uses sparse matrices to transform the images [20]. In [67] the authors explain a method of semantic region growing with adaptive edge penalty for multivariate image segmentation. A method for analyzing ancient manuscripts is presented in [68].

For transforming a paper document into electronic format, geometric document analysis is required [4]. The analysis involves specifying the geometry of the maximal homogeneous regions and classifying them into text, image, table, drawing etc. In their paper they have presented a parameter free method for segmenting documents into maximal homogeneous regions and identifying the regions into text, image, table, drawing etc. A pyramidal quad tree structure is used for multi scale analysis and a periodicity measure is suggested to find a periodical attribute of text regions for page segmentation [4].

Furier-Mellin Transform is used for identifying and correcting Global affine transformations to tabular document images [5]. In [6], the authors used textual images which are compressed by constructing a library of symbols occurring in a document and the symbols in the original image then replaced with pointer into the code book to obtain a compressed representation of the image. The feature in wavelets, special domain based on angular distance span of shapes is used to extract the symbols.

In [2], the authors used Tree Structured Belief Networks (TSBNs) as prior models for scene segmentation. Their results show that for classification, a higher performance is obtained with the Multi-Layer-trained neural networks.

In [3] a text segmentation method using wavelet packet analysis and k-means clustering algorithm is presented. This approach assumes that the text and non text regions are considered as two different texture regions. The text segmentation is done by using wavelet packet analysis as a feature analysis method. The wavelet packet analysis is a method of wavelet decomposition that offers a richer range of possibilities for document image. From these multi scale features, they compute the local energy and intensify the features before adapting the k-means clustering algorithm based on the unsupervised learning rule. They claim that their text segmentation method is effective for document image scanned from newspapers and journals. A perspective rectification of document images using fuzzy set and morphological operations is mentioned in paper [29].

In another paper [15] the authors explain about segmentation of satellite images attained successfully with the help of feed forward Neural network with a kappa coefficient of 0.95. In [16], the authors try for automatic segmentation of displayed math zones from the document image, using only the spatial layout information of Math formulae and equations so as to help commercial OCR systems which cannot classify Math zones and also for identification and arrangement of math symbols by others. In this work the displayed math is classified into three categories DM I, DM II and DM III. That is ordinary expressions, Expressions with suffix or super script and expressions with Matrices or determinants.

2.3 Previous Works in Image Compression

There are various methods for data compression. Some of them are used for Image Compression.

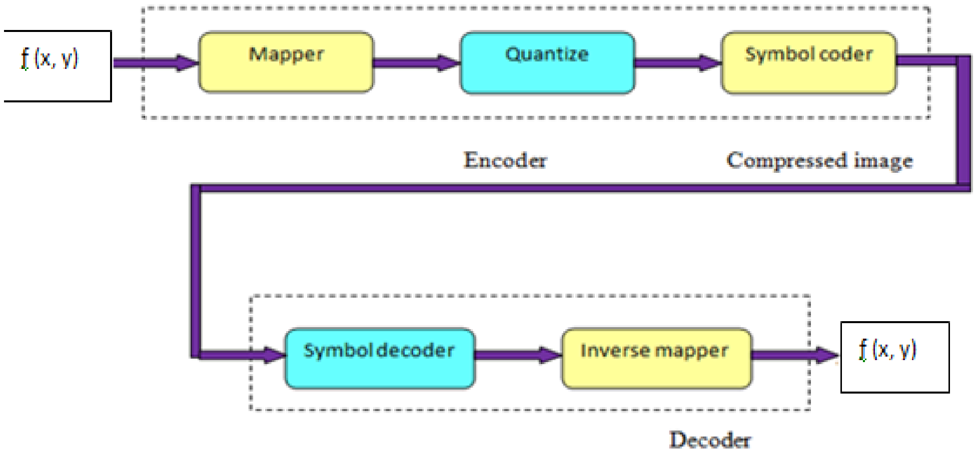


Fig: 2.1 Block Diagram for image compression [83]

Compression Ratio achieved is measured using the formula given below:

$$Compression\ Ratio = \frac{\text{length(output)}*100}{\text{length(input)}} \qquad (2.1)$$

Another measure is:

$$Compression\ rate = 1 - \text{compression} \ldots\ldots(2.2)$$

$$Compression = \frac{[\text{length(input)} - \text{length(output)}]}{\text{length(input)}]} * 100 \ldots\ldots(2.3)$$

This specifies the percentage of saved space. For example, 25% compression rate means that the uncompressed file was reduced by one fourth of its original size [84]. Peek Signal to Noise Ratio is computed using the following formula [22], which is used as a measure of similarity and dissimilarity.

$$\text{PSNR} = 10log_{10}(2^k - 1)^2/(\ e^2\,rms) \ldots\ldots(2.4)$$

Where the mean squared error is the quantity

$$e^2\ \text{rms} = \frac{1}{\text{mn}}\sum_{x=0}^{m-1}\sum_{y=0}^{n-1}[\ g(x,y) - f(x,y)]^2 \ldots\ldots(2.5)$$

There are different Compression Algorithms which uses Prefix-condition codes and the Kraft-McMillan Inequality [59], Huffman's coding algorithm [61], Variable length coding technique [60], Shannon Fano coding [62], Arithmetic Coding etc.

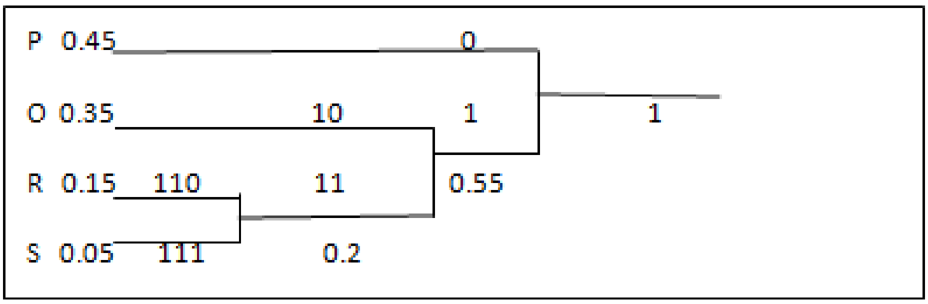


Fig 2.2: A Sample Huffman Tree

Huffman coding starts with a set of n symbols with known probabilities (or frequencies) of occurrence. A sample Huffman tree is given in Fig 2.2. The symbols are first arranged in descending order of their probabilities. The set of symbols is then divided into two subsets that have the same (or almost the same) probabilities. All symbols in one subset get assigned code that start with a 0, while the codes of the symbols in the other subset start with a 1. Each subset is then recursively divided into two and the second bit of all the codes is determined in a similar way. When a subset contains just two symbols, their codes are distinguished by adding one or more bit to each. The process continues until no more subsets remain. This coding technique gives good results only when an incoming symbol of probability of negative powers of 2 (i.e. its entropy is less). And Huffman coding is more efficient than Shannon Fano coding [62]. Arithmetic coding overcomes the difficulties faced by Huffman coding and Shannon Fano coding. Arithmetic coding assigns one code to the entire input stream. The code is a number in the range [0, 1). [64].

Dictionary methods use a list of phrases (the dictionary)[22][26][28], which hopefully includes many of the phrases in the source, to replace source fragments by pointers to the list. Comparison occurs if the pointers require less space than the corresponding fragments. v.42 bis uses dictionary method. The Papers published in the years 1977 and 1978 by Ziv and Lempel known as LZ77 and LZ78 respectively. Applications employing variations on LZ77 include LHare, PKZIP, GNU zip, Info ZIP and Portable Network Graphics (PNG). Lossless Image compression format designed as a GIF successor. LZ78 type schemes are used in modern communications (v.42 bis), the UNIX compress program, and in the GIF graphics format [63].The basic difference between LZ77 and LZ78 is in the management of the dictionary. In LZ77, the dictionary consists of fragments from a window (the sliding window) into recently seen text. LZ78 maintains a dictionary of phrases. In practice, LZ77 and LZ78 uses a more structural approach in managing a slow growing dictionary (possibly trading compression for speed at the coding stage.) and LZ77 has rapidly changing dictionary (which may offer better matches) and is faster for decoding. In applications, the choice of basic scheme may be complicated by various patent claims.

Document image Compression helps us to reduce storage space and to access data easily. There are two types of Compression techniques: Lossless and Lossy. Lossy Compression techniques are applicable for ordinary digital images, because of the limitation of our eyes. Even if certain pixel portions are lost, human eye can interpret the image. But this is not completely adoptable in the case of Document image compression. If some of the text portions are lost the meaning may change or the reader may be unable to deduce the meaning. Some of the conventional encoding technique for Lossy compression are Huffman coding, Inter pixel redundancy coding, Psycho Visual Redundancy coding [63], JPEG compression encoding, JPEG2000 compression encoding etc.

A method to compress the text plane using the pattern matching technique is called JB2 [11]. Wavelet transform and zero tree coding are used to compress the background and the text's color plane [11]. For two pass compression algorithm Huffman encoding [19] is used. The authors achieve a Signal to Noise Ratio (SNR) ratio of more than 20. This is a comparatively low value of SNR. The lossless methods like JPEG binary arithmetic coding, Lempel Ziv coding, Huffman coding etc achieved compression ratio from 2:1 to 3:1 [20].

Lossy methods like Vector Quantization (VQ), JPEG DCT, and Differential Pulse Code Modulation (DPCM) etc have achieved compression ratio of 10:1 [20]. Parallel implementation of the JPEG still image compression standard is done on the MasPar MP-1, a massively parallel SIMD computer [21]. They achieved a data rate of 0.742 bits/pixel for gray scale image and 1.178 bits/pixel for the color image. Execution time of the algorithm is 0.2205s and 0.7651s for grayscale and color images respectively. Discrete Cosine Transform (DCT) quantization and Huffman encoding are the core of the baseline JPEG compression algorithm, which execute in 0.0771s [gray scale] and 0.2097s [colour] and comprises less than 15 percent of the total execution time. In [1], the authors describe the compression of handwritten signatures and their reconstruction. They observed that, the lowest and highest reconstruction errors were 3.05 multiplied by 10^{-3}% and 0.01%, respectively. 28% improvement over JBIG is achieved [30]. The authors of the paper [31], proposes plain term interpretation of Culik's image compression, a very capable yet undeservingly under represented method. Another method [32] uses Mean Squared Error and Human Visual System. They could improve DCT. A method divides the image into Luminance and chrominance components for further operations [33]. Continuous wavelet transforms which is viewed as an extension of the Fourier Transform. It is a powerful tool for signal and image processing [34]. A compression ratio of 52:1 is achieved, through vector quantization algorithm [36]. Efficient entropy encoding of transform coefficients has been used to take advantage of different symbol rates. In [37], the authors propose a faster method, Ordered Web Process. They obtained a Peak Signal to Noise Ratio (PSNR) of 20.16 dB using 3436 bytes. In Comparison, the method of Shapiro attains 27.54dB with 2048 bytes and 30.23dB with 4090 bytes. In paper [38], provides CNN based mammogram analysis system for compression. Authors of another paper [39], have applied Global Preprocessing and level preprocessing for compression. In the paper[40] the authors have applied Adaptive image compression method based on wavelet Transform, bit image compression, fractal iteration and Huffman coding. In [41], the authors applied reduced dimension image compression, which is a lossy compression.

Authors of paper [42] have achieved a coding rate of 0.57 bits/pixel for Lena image, SNR=22.4dB and 22.7dB respectively for gray scale and colour images. The authors of the paper [43], claims that, in Magnetic Resonance Imaging (MRI) image, a compression ratio of 2.5:1 or coding rate of 3.1 bit per pixel (bpp) is obtained comparing favorably with other recently reported medical image compression. They used image size of 256x256 pixels. The embedded zero tree prediction wavelets (EZW) is explained in the work [44]. First order entropy with fast Huffman adoption codes used in paper [45]. The authors [46] have achieved approximate bit rate 0.623bits/pixel with a PSNR average 33dB. Code Division Filter (CDF) [47] wavelet filters have applied for JPEG 2000 images Bi4.4 and Bi1.1 showed better performance on de-correlation in the family of CDF filters. With Block Arithmetic Coding for Image Compression (BACIC) [48] the authors have achieved a bit rate of 7 bits per pixel. They mention that, Lossless JPEG uses Huffman coding.

The authors of [49], use an image with a resolution of 512×512 and the method used is adaptive zigzag reordering. In paper [50] the authors specify Vector Quantization- Kohonen's Self Organizing Feature Map (SOFM) as one of the well known method for VQ. It allows efficient codebooks design with interesting topological properties for Medical applications. The block size is restricted to small values (3×3, 4×4) which limit the compression rate. In another paper [51] the authors conclude that if Range block is small, compression ratio is small and PSNR is high.

Good multi filter properties [52] are used for the design and construction of multiwavelet filters. Simulated Annealing (SA) [3] shows highest performance in multifiltering. SLIm (Segmented layered Image) for separating text and line drawing from background images [7], in order to compress them both more effectively. This approach is different from DjVu, Tiff-FX, and MRC by being simple and fast. Fractal compression and statistical methods are also used for compression [9].

For Lossless Generalized Data Embedding, the authors [10], present a lossless recovery of the original image, achieved by compressing portions of the signal that are susceptible to embedding distortion and transmitting these compressed descriptions as a part of the embedded payload.

In [12] the authors explain an algorithm as follows: The input image is quantized to several quantized images with different number of quantized color. For each quantized image, it is put to 3D histogram analysis to find some specific colors, which are probable text candidates. Each bi-level image relative to its color candidate could be produced. By calculating some spatial features and relationships of characters, text candidates should be identified. Then combine all these to single quantization layer so that to localize text region accurately.

A method of integrated inter-character and inter-word spaces for water marking embedding is explained in [13]. An overlapping component which is of size 3 is utilized, whereby the relationship of the left and right spaces of the character is employed for the watermark embedding. The integrity of the document can be ensured by comparing the hash value of the character components of the document before and after watermark embedding, which can be applied to other line shifting and word shifting methods as well. While the authenticity of the document can be ensured by generating the gold-like sequence, which takes the secret key of the authorized user/owner as the seed value, and it is subsequently XORed with hash value of the character components of the document to generate the content-based watermark. The capacity of the water mark increases when compared to conventional line shifting and word shifting methods.

In Adaptive wavelet based image compression methods [23], when non stationary Multi Resolution analysis is performed compression ratio of 5, 10 and 20 are obtained with colour, gray scale and binary images. In [24] for each input image, 4 images are reconstructed from the compressed sub images corresponding to 2048,4096, 8192 and 16384 (largest magnitude) image sizes using non separable orthogonal linear phase reconstruction method. Paper [25] claims a compression ratio of 35% lossless compression. Data used were relevant in remote sensing. Here, Run length with Huffman, Quantization with Huffman etc is used. In the paper [27], mentions the compression ratio achieved as 12.5:1. In [32], the authors claim that during compression a PSNR ratio of 28.3 was achieved. A work on non linear projection scheme for data filling that matches the baseline JPEG coder and produces good compression results and improved image quality. Basic information on the image content, 3×3 pixels size, limits the compression rate [58]. The authors of the paper achieved a PSNR of 0.37dB to 0.45dB. Pixel smoothing is done with their algorithm. Four code books, with code words of different size are used. Image analysis for coding uses a quad tree scheme. Results are compared with those obtained using the standard JPEG image compression algorithms. A new arithmetic coding is used for compression which is superior to Huffman coding [55]. In paper [56], the authors explain the design of the Image Compression Laboratory (ICL) a visual environment supporting radiologists interactively compressing medical images still maintaining the diagnostic information. It uses Discrete Wavelet Transform (DWT), Zero Tree Quantizer and Arithmetic coder. The Reordering algorithm [57] uses $2N\log_2 N$ bits in which N means image resolution. In [59] proposes a method which consists of two phases: prediction phase and quantization phase. Multi valued threshold Quantization is used. The authors of paper [60], in their work, use ART2 Neural Network for compression. The authors [61], in their work, use three standard images with bit rates 0.25, 0.5 and 1.0 bpp respectively and having title sizes 64×64 and 128×128. The authors of the paper [62] mention that, in different sub bands different thresholds are used. A method of Parallel Virtual Machine (PVM) is also used for compression [63]. Regional Search is applied to reduce Compression time. Compression ratio achieved is 6.30:1 to 10.00:1. In the paper [64] the authors performed the following works: 1) Generated a given level dyadic wavelet transform of an input image. 2) Linearly quantize the wavelet coefficients with a proper dead zone. 3) Generate quad tree code and a 1-D integer sequence with the qualitative decomposition method. 4) Convert the integer sequence into an L sequence and R sequence with the data composition method and 5) compress the quad tree code the L sequence and the R sequence with an adaptive arithmetic code. Both the Lena image and Barbara image are compressed with the proposed algorithm.

In the paper [65] with their work achieved a Compression ratio of 8.81:1 to 10.50:1; PSNR 27.11 to 30.72. A method with Spiking Neuron Networks (SNNs) are often referred to as the 3rd generation of neural networks which have potential to solve problems related to biological stimuli [18]. They derive their strength and interest from an accurate modeling of synaptic interactions between neurons, taking into account the time of spike emission. SNNs overcome the computational power of neural networks made of threshold or sigmoid units. Based on dynamic event driven processing, they open up new horizons for developing models with an exponential capacity of memorizing and a strong ability to fast adaptation. Moreover, SNNs add a new dimension, the temporal axis, to the representation capacity and the processing abilities of neural networks.

In [17] the authors propose a hierarchical framework for document segmentation as an optimization problem. The model incorporates the dependencies between various levels of the hierarchy unlike traditional document segmentation algorithms. This framework is applied to learn the parameters of the document segmentation algorithm using optimization method like gradient descent and Q-learning. The novelty of their approach lies in learning the segmentation parameter in the absence of ground truth.

For fractal image and image sequence compression, Genetic Algorithm is used to compress an image [18]. Here Genetic Algorithms are used for randomly generated population of Local Iterated Function Systems (LIFS), the one whose attractor is the first frame in the sequence. ”Collage theorem” is considered as the foundation for the paper [18].

2.4 Soft Computing Techniques

Evolutionary algorithms or Genetic and Evolutionary Computations (GEC) like Genetic Algorithms, Genetic Programming, Classifier Systems, Evolution Strategies, artificial life, Hybrid of Neuro and Evolutionary computation methods are able to give fast solutions. And these algorithms can give fastest solutions for pattern classification, image and signal processing, forecasting and control. The concept of echo state networks can benefit from reservoirs, which are pre-trained on the input with the implicit plasticity rule [58]. Such methods are suitable for handling complex data. Fuzzy and IV fuzzy methods help to deal with imprecise data.

Automatic detection or classification of objects or events is known as Pattern Recognition. The individual items, objects or situations to be classified are referred as a pattern or sample. This involves edge detection or boundary detection of objects in images, hand written pattern recognition, recognition of audio video signals etc. Different successful methods of pattern recognition are available. Some of them are Probabilistic methods, Statistical methods, Syntactic methods, Non parametric decision making, Clustering, Artificial Neural Network methods etc. The application areas of pattern recognition are: Automated analysis of medical images obtained from microscopes and Computer Aided Tomography scanners, magnetic resonance images, nuclear medicine images, X-rays, photographs, machine assisted inspection, human speech recognition by computers, classification of seismic signals, selection of tax returns to audit, stocks to buy and people to insure etc, finger print identification, identification of persons from various characteristics like, finger print, hand shape and size, retinal scans, voice characteristics, typing patterns, hand writing etc, automatic inspection of printed circuits, printed character recognition , machine assisted analysis of satellite pictures, agriculture crops, snow and water reserves, mineral prospects etc, classification of ECG, EEG etc[22].

The design concepts for automatic pattern recognition are done by the ways in which pattern classes are characterized and defined.

1. Membership-roster concept- template matching
2. Common Property Concept- feature matching
3. Clustering concept-clustering properties eg: minimum distance classifier.

Principal Categories of methods for pattern recognition are:

1. Heuristic methods --- by trial and error
2. Mathematical methods--- deterministic and statistical methods
3. Linguistic (syntactic) methods--- context free and context sensitive

Patterns can be classified based on various measures of minimum distance between two patterns. One of such measure is Euclidian distance [22]. Another measure is the Mahalanobis distance [22]. Pattern Recognition means identifying and recognizing patterns or characters in the Document image. Pattern Recognition or Pattern classification is also relevant for segmentation. Syntactic Pattern Recognition and Semantic Pattern Recognition are the conventional methods. Nowadays AI methods or Neural Network methods like Multilayer Perceptrons, Spatio-temporal Pattern recognition are used. The editors [22] have collected and compiled many papers on pattern recognition, which describes a wide range of methods.

In their paper, the authors [8], have mentioned that for pattern learning they developed a weighted direction code histogram suitable for a gray scale character, which is divided into grids. The value of the feature vector is based on a summation of each grids edge power. By normalization the length of this feature vector luminal variation can be accepted. This is because such variation does not affect edge direction or edge power. This advanced feature is also able to absorb slight variation in edge position. Pattern Recognition process consists of feature extraction, length evaluation, screening and peak detection [22].

In the paper [3] authors introduce an interval fuzzy rule based method for the recognition of hand gestures acquired from a data glove, with an application to the recognition of hand gestures of the Brazilian Sign Language. To deal with the uncertainties in the data provided by the data glove, an approach based on interval fuzzy logic is used. The method uses the set of angles of finger joints for the classification of hand configurations and classifications of segments of hand gestures for recognizing gestures. The segmentation of gestures is based on the concept of monotonic gesture segments. The set of all lists of segment of a given set of gestures determine a set of finite automata able to recognize such gestures. In the International Conference on Image Processing 2004, Amir Said presented the paper in which the author proposes a method to identify those regions using a discrimination function that works with a nonlinear transform to reliably identify edges and at the same time avoid false positive detection on regions with complex patterns. It does so by exploiting the properties of histograms of coefficients of this block transform and their entropy function, which can be computed efficiently via table look up.

Simulated Annealing is analogous to physical annealing. In metallurgy, annealing is the process used to temper or harden metals and glass by heating them to a high temperature and gradually cooling them, thus allowing the material to coalesce into a low energy crystalline state [69].

Even with high speed computers searching certain pattern is very difficult. Artificial Neural Networks shows better performance over traditional search techniques. There are three different types of learning in artificial neural networks. They are Supervised, Unsupervised and Reinforced. Some of the Artificial Neural Networks used for image processing is consolidated in Table 2.1.

Table: 2.1 Different Neural networks used in the field of Image processing [25][26][35][38]

Sl No	Image Processing Area	Type of Neural Network Applied	Activation Function used
1	Pattern classification	Multi layer feed forward neural networks (perceptron)	$x = \sum_{i=1}^{m} wiai - \theta$ o/p s = f(x) $\Delta wi = \eta\, \delta\, ai$
2	Object Recognition , Feature extraction	Shared weight networks (Le Cun's)	double sigmoid function $f(a) = \left(\frac{2}{1+e^{-a}}\right) - 1$
3	Image Restoration	Regression Networks	Sigmoid function $f(a) = \frac{1}{1+e^{-a}}$ OR double sigmoid function $f(a) = \left(\frac{2}{1+e^{-a}}\right) - 1$ OR hyperbolic tangent function $\mathrm{h(a)} = \tanh(a)$
4	Image Segmentation	SOM	$\mathrm{Si} = \mathrm{f(xi)} = \mathrm{sgn(xi)}$ $Xi = \sum_{j=1}^{N} \mathrm{wijsj} - \theta\mathrm{i}$
5	Optimization and for identifying 3D objects in an image	HNN	$J(x) = \frac{1}{2} x^T W_x - x^T \theta$

6	Image segmentation and Compression	Kohonen's SOM	$S_i = f(x_i) = sgn(x_i)$ $Xi = \sum_{j=1}^{N} w_{ij} s_j - \theta_i$
7	Image recognition	Modified version of TDNN	$Y = \sum_{j=1}^{J} x_{jf_h} \left\{ \sum_{i=1}^{m+1} wjix[t - (i-1)\Delta x] \right\}$
8	Formatting pixel data	Adaptive or non adaptive NN	$Yi = \frac{w_i^T x}{\lVert wi \rVert^2}$
9	Enhancement	Auto associative ANN	$J(w) = 1/2 \sum \lVert x^k - wx^k \rVert^2$
10	Classification, hand written digit recognition, object recognition, face detection, text categorization[105]	SVM	$F(x, \alpha) = sign(\sum y_i \alpha_i k(x_i x) - b)$

The following steps should be used for designing ANN for specific applications [22] [38][35][26] [111] :

a) Choose the right ANN architecture

b) The use of prior knowledge about the problem in constructing both ANNs and Training sets

c) Thorough analysis of the black box character of ANNs.

d) Determine the input vector format and desired output vector.

e) Fix the learning rate parameter, step size

f) Converge the training algorithm at global minimum.

g) Analyze the convergence rate with Lyapunov Energy surface.

h) Choose the appropriate tools for practical implementation.

Traditional optimization algorithms like linear Programming, Transportation, Assignment, Nonlinear programming, Dynamic Programming, Inventory, queuing, Replacement, scheduling etc which existing since 1940, have not been tackled the following questions: Does an optimal solution exist? Is it unique? What is the procedure? How sensitive the optimal solution is? How the solution behaves for small changes in parameters? [78]

Nontraditional search and optimization methods have become popular in engineering optimization problems in recent past. These algorithms include: Simulated annealing (Kirkpatrik, et al, 1983), Ant colony optimization (Dorigo and Caro, 1999), Random cost (Kost and Baumann, 1990), Evolution strategy (Kost 1995), Genetic algorithms (Holland, 1975), Cellular automata (Wolfram, 1994). [78]

There are different search techniques available in mathematics. The general classification is given in figure 2.3. Genetic algorithm and evolutionary strategies mimic the principle of natural genetics and natural selection to construct search and optimization procedures. The idea of evolutionary computing was introduced in 1960 by I Rechenberg. Prof Holland of University of Michigan, Ann Arbor contributed much for the growth of this concept. The following figure illustrates different search methodologies [78].

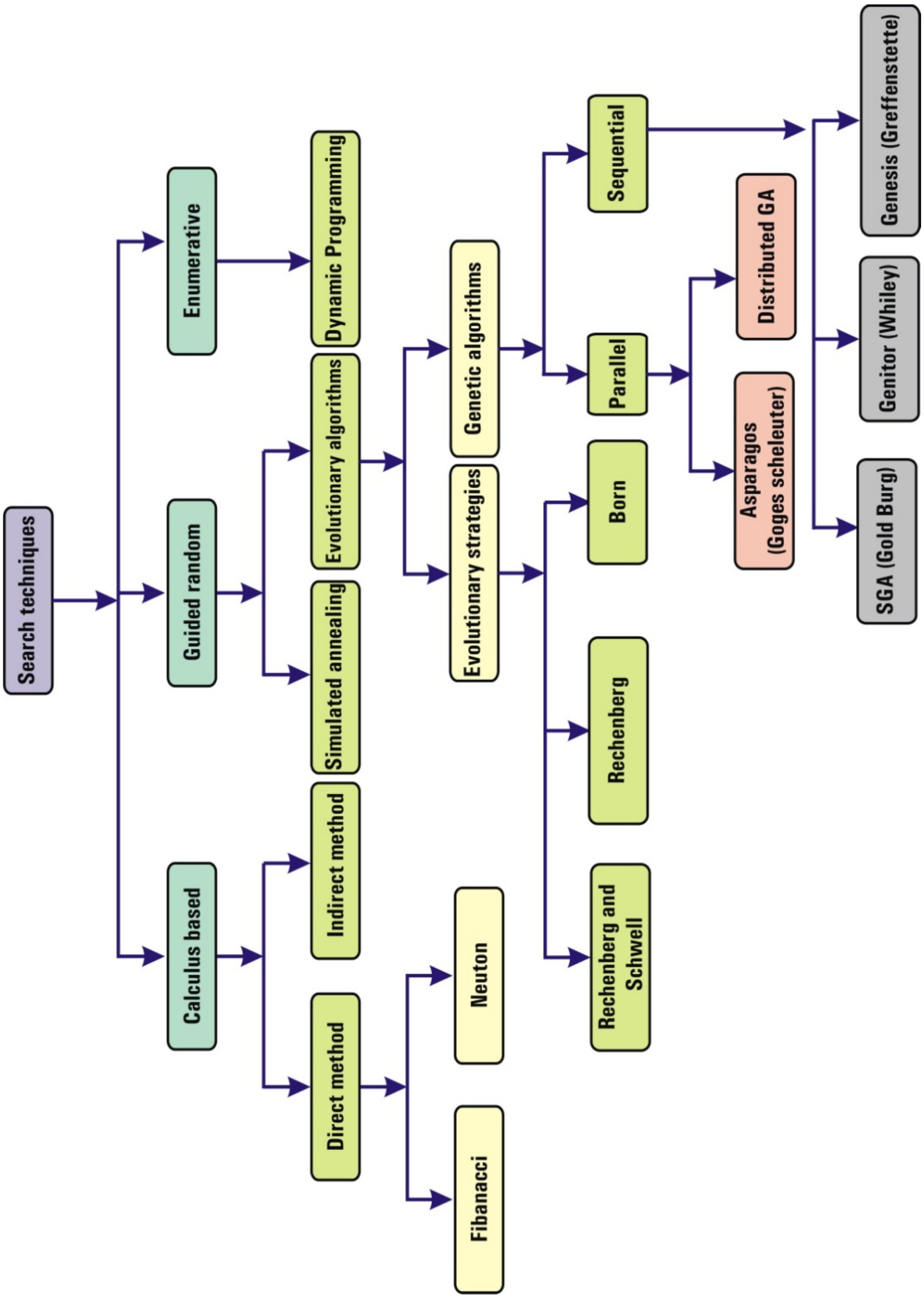


Fig: 2.3 Different Search techniques [78]

Genetic Algorithms (GA) are a particular class of evolutionary search algorithms for global optimization. A genetic algorithm is designed by two components:

1. Genetic representation of the solution domain
2. A fitness function to the solution domain

A standard representation of the solution is as an array of bits. Arrays of other types and structures also can be used. The main property that makes these genetic representations convenient is that their parts are easily aligned due to their fixed size that facilitates simple crossover operation. Variable length representations may also be used, but crossover implementation is more complex in this case. Tree-like representations are explored in Genetic Programming and graph-form representations are explored in Evolutionary Programming. The fitness function will vary from problem to problem. In some problems it is very difficult to define a fitness function. In such cases we can use an interactive Genetic Algorithm [70].

Once we have the genetic representation and the fitness function defined, GA proceeds to initialize a population of solutions randomly, and then improve it through repetitive application of mutation, crossover, inversion and selection operators.

Initially many individual solutions are randomly generated to form an initial population. The population size depends on the nature of the problem, but typically contains several hundreds or thousands of possible solutions. Traditionally, the population is generated randomly, covering the entire range of possible solutions (the *search space*). Occasionally, the solutions may be "seeded" in areas where optimal solutions are likely to be found [71] [72] [73].

During each successive generation, a proportion of the existing population is selected to breed a new generation. Individual solutions are selected through a *fitness-based* process, where fitter solutions (as measured by a fitness function) are typically more likely to be selected. Certain selection methods rate the fitness of each solution and preferentially select the best solutions [10] 107] [108] [110]. Different selection methods used in Genetic Algorithms are as follows: [74] [75] [76]

1. Roulette – wheel selection
2. Boltzmann Selection
3. Tournament Selection
4. Rank Selection
5. Steady State Selection.

Popular and well-studied selection methods include Roulette wheel selection and Tournament selection. Cross over and mutations are two genetic operations which will result in off springs which differ in their features from their parents [69]. This generational process is repeated until a termination condition has been reached [77] [78] [79]. Common terminating conditions are:

- A solution is found that satisfies minimum criteria
- Fixed number of generations reached
- Allocated budget (computation time/money) reached

- The highest ranking solution's fitness is reaching or has reached a plateau such that successive iterations no longer produce better Results
- Manual inspection
- Combinations of the above.

2.4.1 Fuzzy Sets:

The concept of Fuzzy logic was explained by Lofti A Zadeh in his paper "Fuzzy Sets" in 1965 at the University of California, Berkeley. Fuzzy sets are used to express vagueness or impreciseness. For fuzzy sets a membership function maps crisp values into an interval 0 to 1. And it is defined as: μ: X→ [0, 1]. Suppose X is a non-null set. A fuzzy set A of this set X is defined by the following set of pairs. In 1975 Zadeh made an extension of the concept of a fuzzy set by an Interval Valued fuzzy set (ie, fuzzy set with an Interval Valued membership function). The Interval Valued fuzzy set is referred to as IV fuzzy set. It is defined as follows:

$$A = (\mathrm{x}, [\,\mu A^{L}\,(\mathrm{x}), \mu A^{U}(\mathrm{x})]),\ x \in X \quad \text{..............................(2.6)}$$

Where, $\mu_A{}^L$,$\mu_A{}^U$ are fuzzy sets of X such that

$$\mu' A(\mathrm{x}) = [\mu AL\,(\mathrm{x}), \mu AU(\mathrm{x})],\ x \in X. \quad \text{................................(2.7)}$$

Let D[0,1] denotes the family of all closed intervals contained in the interval [0,1] .

Table: 2.2 Fuzzy logic development and its Applications in Image Processing- at a glance

Year	Contributions by	Area
1965	L Zadeh	Introduction of Fuzzy Sets[3]
1966	Zadeh et al	Pattern Recognition as interpolation of membership functions[29]
1969	Ruspini	Concept of Fuzzy Partitioning[29]
1970	Prewitt	First Approach toward Fuzzy Image Understanding[83]
1973	Dunn, Bezdek	First Fuzzy Clustering algorithm (FCM)[85]
1977	Pal	Fuzzy Approach to Speech Recognition[86]
1979	Rosenfeld	Fuzzy Geometry[3]
1980-86	Rosendfield et al., Pal et al.	Extension of Fuzzy Geometry, New methods for enhancement / segmentation[86]
1986 - 90	Dave/Krishnapuram/Bezdek	Different Fuzzy Clustering algorithms, [85] Rule-based Filters, Fuzzy Morphology[85]
2002	Lazhar Khiriji, Momet Gabborj	Adaptive fuzzy order Statistics- rational hybrid filters for color image Processing[86]
2003	Kulkarni and B Verma	Fuzzy logic based texture queries for CBIR[3]
2004	S Karitthakum, A Yaichareon et al	Classifying crispiness of freeze-dried Durian using fuzzy logic
2004	Lili Yun, keichi Uchimura, Shirji Workisake	Automatic Extraction of main road from Ikonos satellite imagery based on fuzzy reasoning
2005	Lior Shamir, Robert J Nemiroff	Algorithm for finding Astronomical objects in wide angle frames
2006	Kanchan Deshmukh and G N Shinde	Adaptive Neuro Fuzzy system for color image segmentation
2007	Gabriela Droj	Fuzzy theory in remote sensing, image classification
2007	Yu Wua Wong, Ligiong Tang, Donald Beily	Vision system for Robot guide system
2007	Sand Crastet	Micro structural imaging algorithms
2009	Harish Kundra, Er Aushima, Er Monika Verma	Image Enhancement Based on Fuzzy Logic
2009	Saikat Maity, Jaya sil	Colour image Biometric Approach for Segmentation using type -2 fuzzy sets
2010	Jarik Ozkul et al	Hierarchical Driver Aid for Parallel parking using Fuzzy Logic
2011	Rajini A , Bhavani R	Fuzzy C means clustering for Brain Images.[94]
2011	S.Yuan et al	Object shape detection –fuzzy generalized Hough transform.[95][96]
2011	Rameshwar rao R	Image Enhancement [99]
2012	Rattanalappalaon S et al	Feature matching[100]

2.5 Conclusion

Segmentation technique for Digital Image processing cannot be directly applied to Document Images. Segmentation method varies from language to language. Different types of connected component analysis should be performed for different languages. Lossy compression methods are not suitable for document images if it causes loss of vital information. For background elimination it is acceptable. In the existing methods the speed of processing is very low. Compression ratio achieved is very small. Even though JPEG provides highest compression ratio among the conventional methods, it is lossy. And conventional methods are much time consuming methods. Traditional methods also show difficulty in representing context information and setting prior knowledge. The work shows that better performance can be achieved with the help of evolutionary and hybrid methods. Studies of the above described areas are performed, since they are related to the proposed work and give ideas to carry out the proposed work. It can be concluded that there are numerous ANN models which are suitable for image processing. Tree Search and Graph search provisions in Genetic and Evolutionary Algorithms can be effectively utilized in Document Image Processing. If heuristic search in Artificial Intelligence is incorporated with search techniques in Document Image Processing will reduce the processing time.

………………

Chapter **3**

INTERVAL- VALUED FUZZY BASED DOCUMENT IMAGE SEGMENTATION

Contents

3.1 Introduction

Image Segmentation is an active research area in Pattern Recognition. Different regions of an image are identified for this method. Segmentation acts as the first step to provide description of the image. Segmentation divides the spatial domain into meaningful parts or regions. Most of the segmentation algorithms use psycho physical approaches for region identification. As in ordinary images a general algorithm is not suitable for document image segmentation. Some document image may be of pure text. Some document image may be a mixture of text and images. Some may consists of photographs, sketches etc. Certain document may be of handwritten text and pictures. Due to this problem the entire segmentation algorithms works on ad hoc basis. The new algorithm introduced here identifies the homogeneous property of lines of text like uniform spacing of lines, words etc. When a photograph occurs in the text, the boundary of the photograph is fixed by detecting edge and rectangular region in the text.

In order to segment the document image, it is necessary to classify the patterns in the document image. Different approaches of pattern classification are by using Polynomial classifier (PC), Multilayer Perceptron (MLP), Radial Basis Function, Nearest Neighbor classifier (NNC), Guassian Assumption Bayes Classifier (GBC)[22]. There are research results related to internal and external character segmentation methods for character segmentation. Soft computing is a term coined by Zadeh, which is basically a synergistic integration of three computing paradigms- neural networks, fuzzy logic and probabilistic reasoning to provide a frame work for flexible information processing, applications designed to operate in the real world. Bezdek called the synergizing of fuzzy logic, neural networks and genetic algorithms as computational intelligence. Soft computing technologies are robust by design and operated by trading off precision for tractability. Since they can handle uncertainty with ease, they conform better to real world situations and provide lower solution costs. Another approach known as Subset hood Product Fuzzy Neural Inference System (SuPFuNIS) is to integrate fuzzy systems with Neural Networks [22].

Table: 3.1 Different methods for Document Image Segmentation [22]:

Method proposed by	Method based on the metrics	Area of Application	Remarks
Kanai	Number of edit operations such as insertions deletions and moves	Commercial OCRs	To determine Zoning accuracy
Vincent et al	Bitmap level region based metric	Text based and non text based regions	Highly depends on pixel noise
Liang et al	Region area based metric	Mixed text and picture	Overlapping area of Ground truth zone and segmentation zone
Mao et al	Empirical evaluation of the segmentation algorithms in the presence of ground truth and performance	Document images	Automatically trains the algorithm with free parameters. Also improves performance of the segmentation algorithm.

Table: 3.2 Suitability of Segmentation Algorithms [22] [94] [97] [101] [109]

Segmentation Technique	Suitable for
Hough transform	Line detection, peak detection and linking
Watershed Segmentation	Point, line, edge detection
Sobel, Prewitt and Robert's method(uses mask)	Edge detection
Laplacian of the Gaussian(LoG)	Smoothing
Canny's method	Global and Local Thresholding
Region based segmentation	To classify into different regions.
O'Gorman	Page layout analysis of document images
Nagy et al	Structural segmentation and functional labeling
IV fuzzy based algorithm	To classify into different planes and segment images

Other Segmentation Algorithms [22]:

1. Recursive XY cut: This is a tree based Segmentation Algorithm. A document is recursively split horizontally and vertically until final criteria where a split is impossible to meet. The document forms the

root of the tree while each split expands the tree. The nodes form the complete segmentation of the document.

2. White Space Analysis: this algorithm concentrates on the background.
3. Constrained text-line detection: The constrained text line detection algorithms are another set of top down algorithm, which find the gutters of white spaces in the document. The white space rectangles are allowed to overlap within a particular threshold. They are sorted in order of decreasing quality which is further used to find the text lines within the document image.
4. Docstrum: It is a bottom up Segmentation Algorithm. The components are merged together found on a nearest neighbor based approach. This algorithm heavily depends on the thresholds that are set for a document. It clusters all the nearest neighbor components into particular regions which are further clustered into the regions of higher order text blocks.
5. Voronoi-diagram based algorithm: The Voronoi-diagram based Segmentation Algorithm extracts points on the boundaries of the connected components. The Voronoi cells are drawn surrounding the points. Superfluous edges formed by the Voronoi cells are removed using criteria functions which involve thresholds like distance and ratio of areas of adjacent cells to form the components.
6. Smearing: The Run Length Smearing Algorithm is a bottom up approach. Bitmaps are formed by grouping the pixels horizontally and vertically using thresholds of acceptable distance between the

components in each direction. The document is segmented by performing logical operation on the bitmaps. The performance of the algorithm highly depends on the thresholds used.

7. Grey level thresholding: iterative pixel classification, surface based separation, segmentation of colored images and edge detection are different segmentation algorithms among which our Artificial and IV fuzzy based methods supersedes other methods.

3.2 Image Representation

Document image is acquired with the help of a scanner and is stored as a JPEG image. JPEG format is used because to check whether compression is achieved in the JPEG image or not. The resolution of the image is retrieved as m × n. Using "imread" function in MATLAB, the image is read and represented as m × n × 3 matrix. Where the first, second and third matrix represents intensity values of red, green and blue respectively. The matrix contains gray scale values ranging from 0 – 255. For computational convenience the average value of the gray scale values of red, green and blue taken together. The following is the pixel intensity representation of a portion of a sample image. *The function f (x, y) =intensity of the pixel, where x and y indicates the pixel position of the image. For a document image the top left corner of the image is considered with coordinate (0, 0). A color image is just three functions of red, green and blue*. We can write this as a "vector-valued" function [82]:

$$f(x,y) = \begin{bmatrix} r(x,y) \\ g(x,y) \\ b(x,y) \end{bmatrix} \quad \text{...(3.1)}$$

A document image is scanned and stored as a "*.jpeg" file. The matrix representation of a gray scale (two dimensional) image will look as follows:

$$f(x,y) = \begin{bmatrix} 234 & 243 & 255 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 23 & 22 & 25 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 43 & 44 & 43 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 244 & 244 & 244 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

A document image with RGB Gray scale representation is given below:

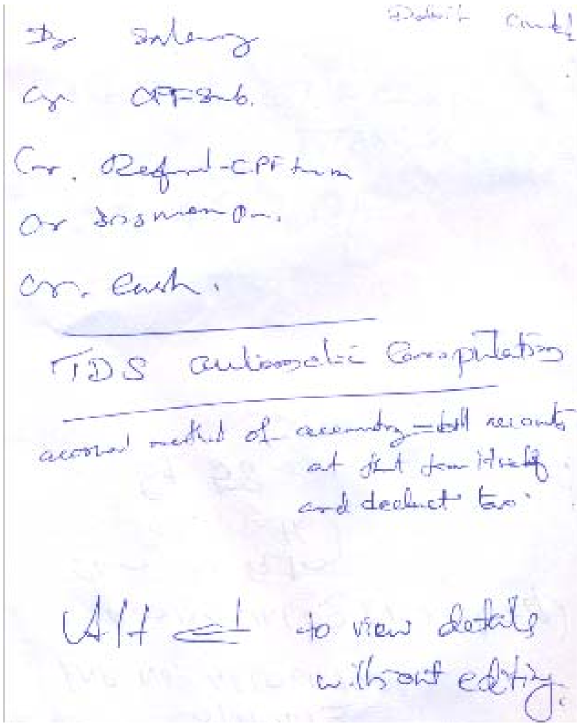


Fig 3.1: Original sample document image

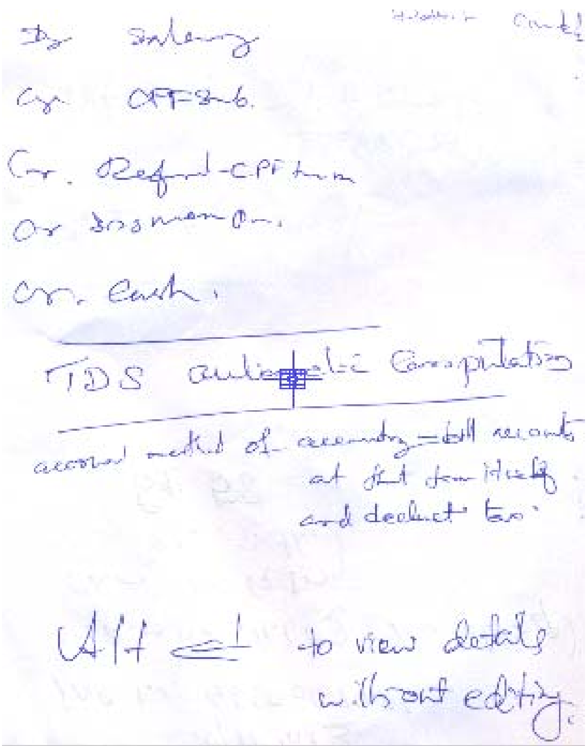


Fig 3.2 : Document image with marked portion

R:255 G:255 B:255	R:254 G:254 B:252	R:255 G:255 B:253	R:255 G:255 B:255	R:255 G:255 B:255	R:255 G:254 B:255	R:255 G:254 B:255	R:255 G:255 B:255	R:255 G:255 B:255	R:253 G:253 B:253	R:252 G:254 B:253	R:255 G:255 B:253
R:254 G:254 B:254	R:255 G:255 B:255	R:255 G:255 B:255	R:255 G:255 B:255	R:255 G:255 B:255	R:255 G:254 B:255	R:255 G:253 B:254	R:255 G:255 B:255	R:255 G:255 B:255	R:255 G:255 B:255	R:255 G:255 B:255	R:255 G:254 B:255
R:255 G:255 B:255	R:254 G:254 B:255	R:254 G:254 B:255	R:254 G:254 B:254	R:255 G:255 B:255	R:255 G:255 B:255	R:255 G:255 B:255	R:254 G:254 B:254	R:255 G:255 B:255	R:255 G:253 B:255	R:255 G:253 B:255	R:254 G:250 B:255
R:255 G:255 B:255	R:255 G:255 B:255	R:255 G:255 B:255	R:252 G:252 B:252	R:255 G:255 B:255	R:255 G:255 B:255	R:255 G:254 B:255	R:254 G:253 B:255	R:254 G:252 B:255	R:252 G:247 B:255	R:242 G:236 B:255	R:223 G:217 B:253
R:254 G:254 B:254	R:255 G:255 B:255	R:253 G:253 B:253	R:253 G:253 B:253	R:255 G:255 B:255	R:255 G:254 B:255	R:254 G:252 B:255	R:245 G:241 B:255	R:237 G:231 B:255	R:218 G:212 B:255	R:191 G:186 B:242	R:160 G:158 B:223
R:255 G:255 B:255	R:255 G:254 B:255	R:255 G:254 B:255	R:254 G:253 B:255	R:252 G:250 B:255	R:242 G:240 B:255	R:232 G:229 B:255	R:207 G:205 B:245	R:189 G:185 B:235	R:162 G:159 B:216	R:142 G:140 B:205	R:142 G:142 B:212
R:255 G:254 B:255	R:252 G:250 B:255	R:247 G:244 B:255	R:247 G:241 B:255	R:229 G:223 B:255	R:194 G:191 B:238	R:176 G:176 B:230	R:164 G:164 B:226	R:142 G:144 B:205	R:161 G:163 B:220	R:165 G:171 B:223	R:192 G:199 B:245
R:247 G:245	R:226 G:224	R:211 G:208	R:212 G:208	R:185 G:179	R:167 G:165	R:172 G:173	R:182 G:184	R:178 G:181	R:214 G:218	R:220 G:226	R:237 G:243

R:255 G:255 B:253	R:255 G:255 B:255	R:254 G:253 B:255	R:255 G:254 B:255	R:253 G:251 B:255	R:242 G:238 B:255	R:230 G:224 B:255	R:207 G:200 B:251	R:187 G:181 B:243	R:161 G:155 B:225	R:129 G:124 B:204	R:121 G:118 B:205
R:255 G:254 B:255	R:253 G:251 B:255	R:247 G:246 B:255	R:243 G:241 B:255	R:223 G:222 B:254	R:205 G:202 B:247	R:176 G:173 B:228	R:149 G:147 B:212	R:129 G:127 B:200	R:120 G:117 B:196	R:130 G:128 B:211	R:148 G:149 B:231
R:246 G:241 B:255	R:236 G:231 B:255	R:221 G:219 B:255	R:197 G:195 B:245	R:167 G:167 B:227	R:146 G:146 B:218	R:125 G:127 B:202	R:147 G:150 B:219	R:150 G:153 B:220	R:152 G:156 B:220	R:175 G:177 B:236	R:199 G:199 B:249
R:204 G:198 B:242	R:183 G:181 B:231	R:166 G:165 B:222	R:147 G:149 B:208	R:139 G:143 B:207	R:142 G:146 B:217	R:157 G:162 B:230	R:199 G:204 B:255	R:201 G:207 B:255	R:207 G:211 B:255	R:224 G:226 B:255	R:244 G:244 B:255
R:141 G:141 B:211	R:130 G:135 B:203	R:137 G:145 B:208	R:163 G:173 B:226	R:184 G:192 B:238	R:193 G:200 B:244	R:214 G:220 B:255	R:239 G:243 B:255	R:242 G:245 B:255	R:247 G:248 B:255	R:252 G:251 B:255	R:254 G:253 B:255
R:145 G:150 B:218	R:160 G:168 B:231	R:185 G:195 B:248	R:216 G:225 B:255	R:226 G:236 B:255	R:232 G:238 B:255	R:246 G:249 B:255	R:252 G:255 B:255	R:253 G:254 B:255	R:254 G:254 B:255	R:255 G:255 B:255	R:255 G:254 B:255
R:199 G:207 B:246	R:214 G:220 B:252	R:233 G:239 B:255	R:245 G:251 B:255	R:249 G:251 B:255	R:251 G:251 B:255	R:251 G:250 B:255	R:255 G:253 B:255	R:255 G:254 B:255	R:255 G:254 B:255	R:255 G:253 B:254	R:255 G:253 B:254
R:240 G:245	R:247 G:251	R:252 G:255	R:253 G:254	R:254 G:253	R:255 G:254	R:255 G:254	R:255 G:251	R:255 G:254	R:255 G:254	R:255 G:255	R:255 G:253

Fig 3.3: Gray scale values of the marked portion of fig (3.2)

3.2.1 Data Collection

A Document image Data base is created with 200 scanned document images of Different sizes both hand written and printed. In which 50 are hand written, 50 are with photographs and drawings 100 are of pure text. The Document image database is in *.jpeg format which is considered as our Data Base. The hand written document image is named as scan0001.jpeg to scan00050.jpeg. The printed document image is named as doc0001.jpeg to doc000150.jpeg.

3.3 Quad Tree Generation

The m × n × 1 matrix of average gray scale values of the document image is taken and the quad tree is generated. The document image is divided into quad tree by taking 'k' as number of levels in the tree. The first four quadrant of the tree is formed by dividing the image as follows:

0 to |(m-1)/2) | and 0 to |(n-1)/2| → first quadrant

0 to |(m-1)/2) | and |(n-1)/2|+1 to n → second quadrant

|(m-1)/2| + 1 to m and 0 to |(n-1)/2| → third quadrant

|(m-1)/2| + 1 to m and |(n-1)/2| +1 to n → fourth quadrant.

For every quadrant the procedure is repeated until k number of levels is reached or the region size become 64 × 64 matrices. So the number of levels in the quad tree is based on the size of the image.

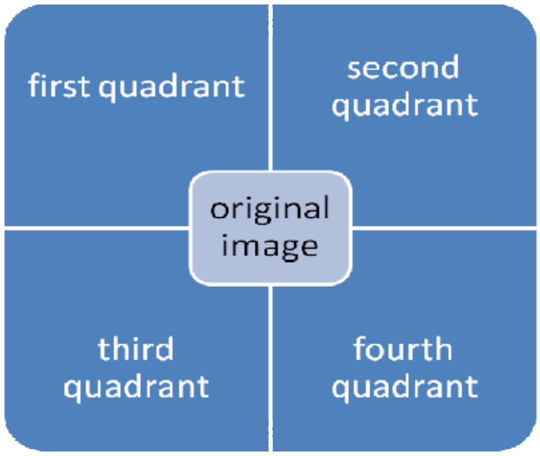


Fig 3.4: Different quadrants of the original image

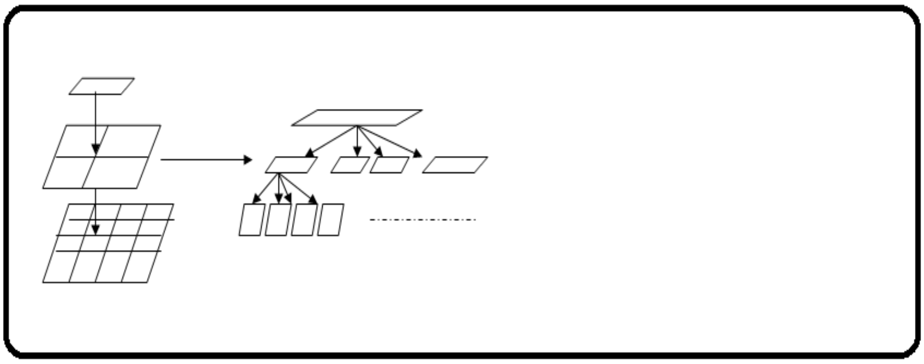


Fig 3.5: An Image is divided into different quadrants

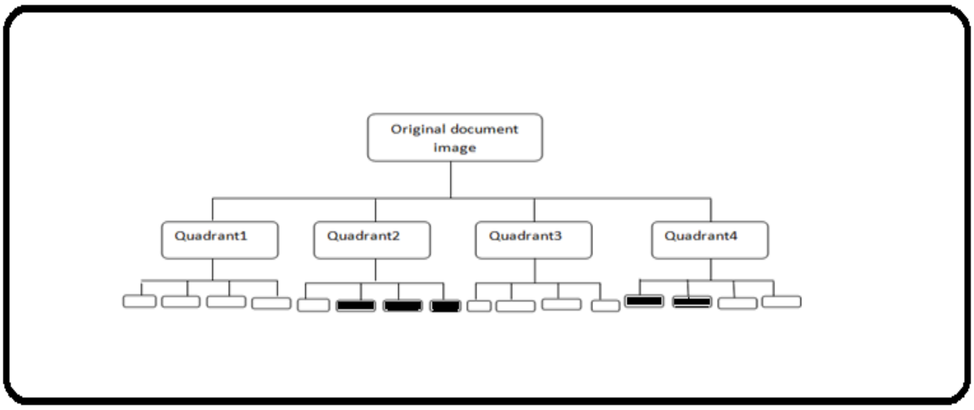


Fig 3.6: Quad tree representation of document image

In figure 3.6 the nodes marked in black colour are of uniform colour region, which will not be expanded further. But the other nodes will be explored further until our requirement satisfies. Each region of the quad tree is represented in the data structure as follows: Node (child (child …)). 1(1) is the first region in the sub division of quadrant 1. 1(2) is the second region in the sub division of quadrant 1. 2(1) is the first region in the sub division of quadrant 2. 2(2) is the second region in the sub division of quadrant 2. And so on. A region code which shows inner most nodes and outer nodes as 1 or 2 is in the same row. ie, 1(1), 1(2), 2(1),2(2) are in the same row. Similarly a region code which shows inner most nodes and outer nodes as 3 or 4 is in the same row. 3(3),3(4),4(3), 4(4) are in the same row. And so on. This is relevant in identifying a line of text (or a cluster).

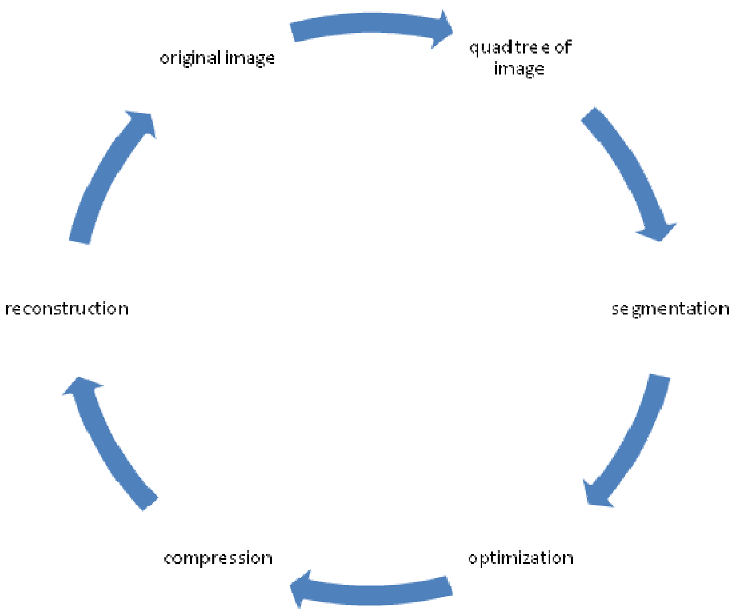


Fig 3.7: Utilization of quad tree for various applications

3.4 IV Fuzzification

3.4.1 Fuzzy operations

Let X is a set and A and B are two fuzzy sets of X. The membership function of A is denoted as μA(x) and membership function of b is denoted as μB(x). *Equality of two fuzzy sets:* The fuzzy sets A and B are equal then μA(x)= μB(x)$\forall x\ x\in X$. *Containment*: Let X is not null and A and B are two fuzzy sets of X with membership function μA(x)≤ μB(x))$\forall x\ x\in X$. *Normal fuzzy set*: A fuzzy set A of a set X is called a normal fuzzy set if and only if Max) ($x\in X$) is μA(x)=1. *Support of a fuzzy set A of the set X*: is the classical set {$x\in X$: μA(x)>0}, denoted by support (A). *α-cut or α-level set of a fuzzy set X* is the following crisp (ie, conventional) set given by A_α={$x\in X$: μA(x)≥ α}. *The complement of a fuzzy set A* which is a subset of X, written as A^c is defined as μA^c (x)=1- μA(x) $\forall x\ x\in X$. *Union*: The union of two fuzzy sets A and B is, μA(x) v μB(x), for $\forall x\ x\in X$, where "v" is the maximum operator. *Intersection* is μA(x) ^ μB(x), for)$\forall x\ x\in X$, where "^" is the minimum operator.

3.4.2 IV Fuzzy Sets and Operations:

In this work interval valued (IV) fuzzy sets are used (i.e. fuzzy set with an Interval Valued membership function). IV fuzzy sets are an extension of the concept of fuzzy sets made by L Zadeh in 1975. Gorzalczany used fuzzy sets for approximate reasoning. For this problem, fuzzy inference rules are set up based on the possible intensity values of pixels in the document image. IV fuzzy relations [Sanchez] are used to map region wise cluster information and this information is used for further compression of the document image.

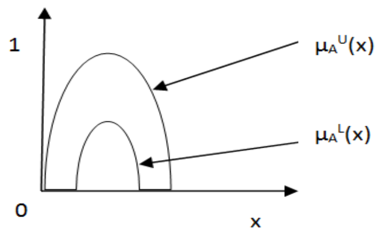


Fig 3.8: An Interval Valued fuzzy representation

The Interval Valued fuzzy set is referred to as IV fuzzy set. It is defined as, A=(x,[μ_A^L (x), μ_A^U(x)]), $x \in X$. Where, μ_A^U(x) and μ_A^L(x) are the upper and lower membership function values in the interval [0,1]. Where, μ_A^L, μ_A^U are fuzzy sets of X such that,

$$\mu'A(x) = [\mu A^L (x), \mu_A^U(x)], x \in X. \text{------------------------}(3.2)$$

Let D[0,1] denotes the family of all closed intervals contained in the interval [0,1] ie, $\forall a1, a2$. Let D[0,1] ∈ [0,1], $a1 \leq a2$, $[a1, a2]$ ∈ D[0,1]. Therefore, for each fixed x ∈X, μ'A(x) =D [0, 1].

Union of two IV fuzzy sets:

If X is a crisp set A and B are two IV fuzzy sets, then their union is an IV fuzzy set $A \cup B$ is given by, $A \cup B$ {(x, μ' $A \cup B$ (x))], where μ $A \cup B$ (x)= [$\mu A \cup B^L(x)$].

Where, $\mu A \cup B^L(x) = \max [\mu A^L (x), \mu B^L (x)]$

$\mu A \cup B^U (x) = \max [\mu_A^U(x), \mu_B^U(x)], \forall x$, x∈X.

Intersection of two IV fuzzy sets:

If A and B are two IV fuzzy sets, then their intersection $A \cap B$ is an IV fuzzy set, given by

$$A \cap B = \{(x, \mu' A \cap B_{(x)})\} \text{------------------- } (3.3)$$

$$\mu' A \cap B(x) = [\mu^L A \cap B(x)), \mu^U A \cap B(x)], x \in X. \text{----}(3.4)$$

$$\text{Where } \mu^L A \cap B(x) = \min [\mu^L A(x), \mu^L B(x)] \text{ ------------ } (3.5)$$

$$\mu^U A \cap B(x) = \min [\mu^U A(x), \mu^U B(x)] \text{ ---- --------}(3.6)$$

Let X, Y are two sets. A fuzzy relation R from x to y is a fuzzy subset of X×Y, characterized by its membership function

$\mu R: X \times Y \rightarrow [0, 1]$

A fuzzy relation R from X to Y is referred to as R(X→Y).

For example, X = {2, 7, 9}

Y = {1, 3, 8}

Table 3.3: A fuzzy relation

x-->y	1	3	8
2	.4	.2	.6
7	.3	.6	.7
9	.4	.2	.9

Here 9 is related to 3 with strength 0.2.

3.4.3. Fuzzification in Neuro Fuzzy Method

In Neuro Fuzzy evolutionary method fuzzy rule composition, fuzzification and defuzzification are done as follows:

- A fuzzification interface that fuzzifies numeric inputs by assigning grades of membership using fuzzy sets defined for that variable.
- A knowledge base that comprises of data derived or heuristic rule base which is used in the compositional rule of inference and defuzzification processes.
- An inference engine that infers fuzzy outputs by employing fuzzy implications and rules of inference of fuzzy logic.
- A defuzzification interface that yields a non-fuzzy control action from an inferred fuzzy control action.

A compositional rule for inference involves the following procedure:

b) Compute memberships of current inputs in the relevant antecedent fuzzy sets of a rule.

c) Since the antecedents are in conjunctive form, the AND operation is replaced by a minimum, that is we compute the minimum of the memberships found in step 1. In some models product is used to combine levels of antecedent firing.

d) Scale or clip the consequent fuzzy set of the rule by the minimum value found in step 2 since this gives the smallest degree to which the rule must fire.

e) Repeat steps a to c for each rule in the rule base

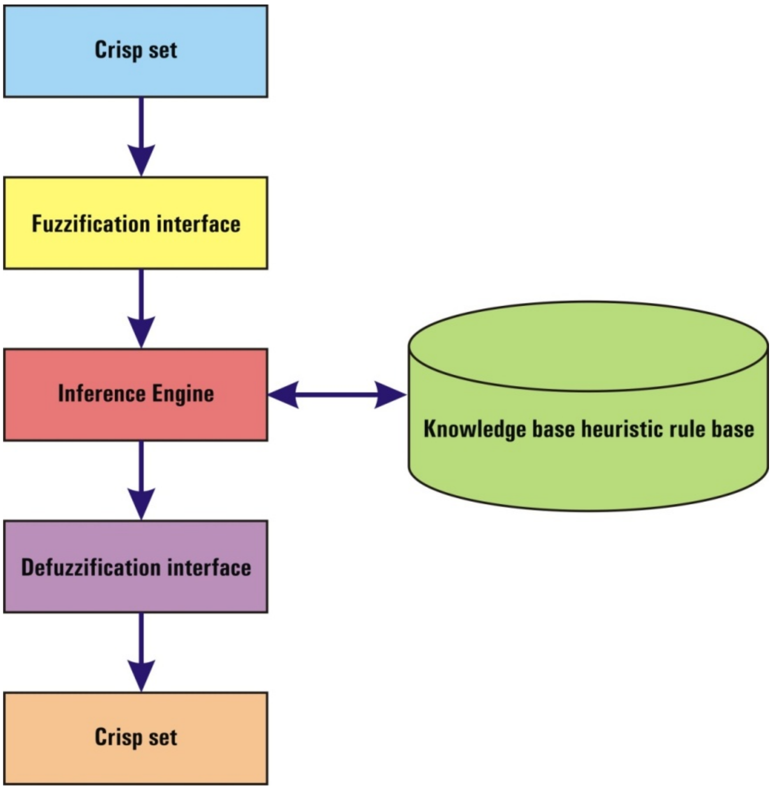


Fig 3.9 Fuzzification and defuzzifiation

In this work the property of a Document image is analyzed and zoning is done with interval valued fuzzy sets. Zoning means locating photographs, text regions etc. It is observed that human beings can skip the unwanted information and directly look at the information in which they are interested. And they are interested in the foreground features such as text, sketches, figures, formulae, images etc. They are not concentrating much on the background unless it is not comfortable. So while scanning the text page they directly locates the region of matter. In our work segmentation is done first because it should be done prior to compression. It is also observed that the document images are more structured than ordinary images (eg: uniform space in between lines, margins, uniform character size in a line etc) and also there are exceptions like, emphasis bold, italics, word art type text, multi column text , picture in the middle of the text etc. When hand written document is used there are again complexities like slant of the writing(slant angle are typically in the range of -45 degree to +45 degree, skew or tilt, underline, overlapped lines, discrete hand writing, imprecise punctuation, broken characters, similarly shaped words, ligatures (combined letters in natural languages), overlapping characters, cursive writing etc. So as a first stage the experiment is done with typed documents. Adaptive Neural Network method is adopted here in order to reduce noise. Noise reduction helps to locate the text lines easily. Neuro fuzzy method also helps to deal with impreciseness and it provides massive parallelism.

3.5 Representation of Segmented Images Using Interval Valued Fuzzy Relations

In our method the desired row and column intensity values represented as ordered pairs of minimum and maximum strength. In order to reduce storage space the index of the interval is also stored. The fuzzy rules are constituted for a document image based on an interval of intensity values. An interval is set for getting the desired intensity values.

Fuzzification is done with the following expression:

Pixel intensity (gray scale value)/255

Sample Fuzzy rules required for this problem is constituted as follows:

1. If the pixel intensity >=0 and pixel intensity <=10 then the IV fuzzy value will lie in the interval [0, 0.039216].

2. If the pixel intensity >=11 and pixel intensity <=20 then the IV fuzzy value will lie in the interval [0.043137, 0.078431]
3. If the pixel intensity >=21 and pixel intensity <=30 then the IV fuzzy value will lie in the interval [0.082353, 0.117647]
4. If the pixel intensity >=31 and pixel intensity <=40 then the IV fuzzy value will lie in the interval [0.121569, 0.156863].

The following figure shows the IV fuzzy value of the gray scale values used in this problem:

The inverse of an IV fuzzy relation R on X×Y is R^{-1}: $\mu R^{-1}(y,x)= \mu R(x,y)$.

Specifically, $\mu R^{L-1}(y, x)= \mu R^{L}(x,y)$ and $\mu R^{U-1}(y,x)= \mu R^{L}(x,y)$.

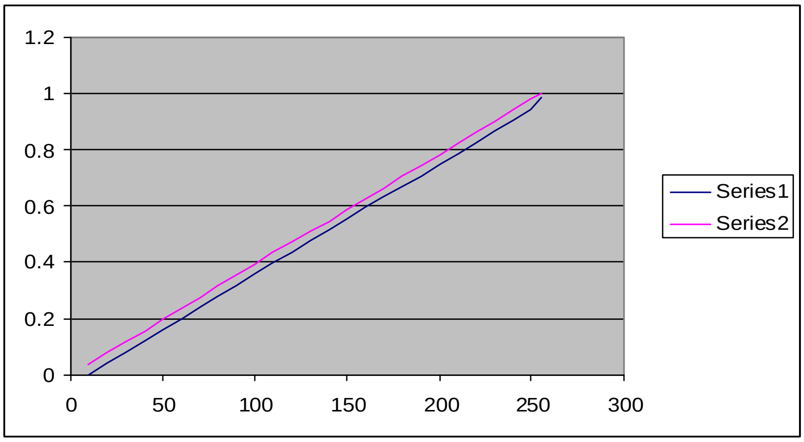


Fig 3.10: Fuzzification (intensities in X axis and fuzzy intervals in y axis)

In fig 3.10 the series 1 indicates fuzzy lower bound and series 2 indicates fuzzy upper bound.

Table 3.4: An Interval valued fuzzy relation

x→y	1	3	8
2	<0.4 0.5>	<0.12 0.13>	<0.6 0.7>
7	<0.6 0.7>	<0.8 0.9>	<0.6 0.7>
9	<0.5 0.6>	<0.7 0.8>	<0.4 0.5>

In Table 3.4 for the coordinate position (2, 3) the interval corresponds to rule 4 and it is referred as the 4^{th} interval. The IV fuzzy interval boundaries are given in Table 3.5:

Table 3.5: Fuzzy interval with boundaries

Intensity range	Fuzzy interval	Fuzzy Lower Boundary	Fuzzy Upper Boundary
0-10	1	0	0.039216
11-20	2	0.043137	0.078431
21-30	3	0.082353	0.117647
31-40	4	0.121569	0.156863
41-50	5	0.160784	0.196078
51-60	6	0.2	0.235294
61-70	7	0.239216	0.27451
71-80	8	0.278431	0.313725
81-90	9	0.317647	0.352941
91-100	10	0.356863	0.392157
101-110	11	0.396078	0.431373
111-120	12	0.435294	0.470588
121-130	13	0.47451	0.509804
131-140	14	0.513725	0.54902
141-150	15	0.552941	0.588235
151-160	16	0.592157	0.627451
161-170	17	0.631373	0.666667
171-180	18	0.670588	0.705882
181-190	19	0.709804	0.745098
191-200	20	0.74902	0.784314
201-210	21	0.788235	0.823529
211-220	22	0.827451	0.862745
221-230	23	0.866667	0.901961
231-240	24	0.905882	0.941176
241-250	25	0.945098	0.980392
251-255	26	0.984314	1

3.6 Feature Extraction for Segmentation

The feature of an image depends on the shape and intensity of pixels. The work helps to extract feature by classifying different layers of desired intensities. The user should submit the desired interval and retrieve the desired portions.

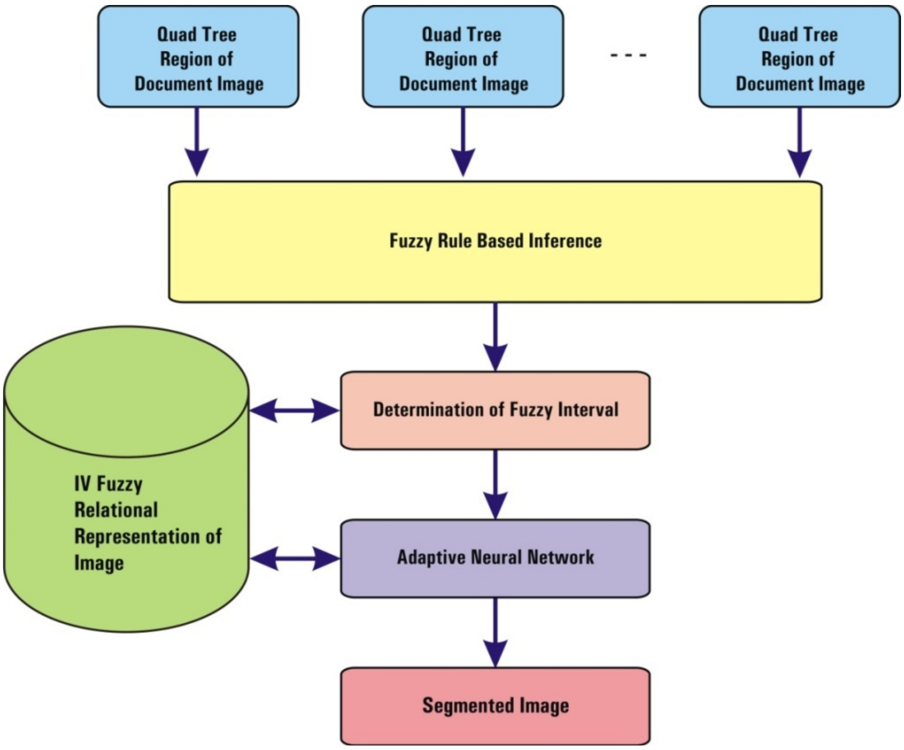


Fig: 3.11 IV fuzzy based Adaptive Neural Network inference system

3.7 Adaptive Neuro fuzzy based clustering Algorithm

The leaf nodes of the quad tree are analyzed for uniform regions. If any of the nodes are of uniform intensity values they are considered as homogeneous regions. Otherwise a threshold't' is applied to identify homogeneous regions. Fuzzy intervals help to identify clusters.

3.7.1 Design of the Adaptive Neural Network

In our work the Adaptive Neural Network consists of three layers: The input layer, hidden layer and output layer. There are n input layer neurons I_1, $I_2 \ldots I_n$; j hidden layer neurons H_1, $H_2 \ldots H_j$ and j output layer neurons $O_1, O_2 \ldots O_j$

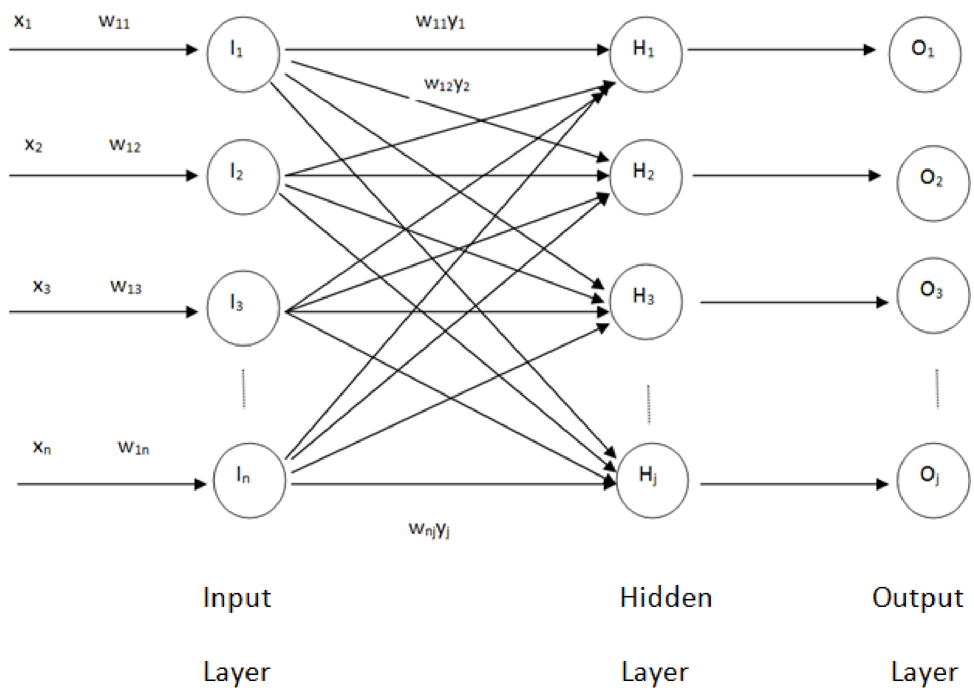


Fig 3.12: Adaptive IV fuzzy based Neural Network for document image segmentation

3.7.2 The Method

For the work 200 document images are scanned and stored as "*.jpeg" files. These files are considered as our document image database. 50

images are used for training the Adaptive Neural Network. Remaining images are used for testing the Neural Network. For training the Neural Network quad tree regions of images with input block sizes 64, 128, 256, 512 and 1024 are used. The input vector X is [X_1 X_2 …X_n] where n is the block size. There are 26 neurons in the hidden layer. [W_1 W_2 …W_j] is the weight vector. The adaptive activation function for the input layer is

$$Yi = \frac{\mathrm{wi}^{T} x}{\|\mathrm{w}\|^{2}} \quad \text{...(3.7)}$$

Where Y_i =[Y_1 Y_2 …Y_j] is the input vector for the hidden layer. In the hidden layer each neuron follows a winner takes all strategy (Kohonen SOM). It uses the activation function O_i=f (S_i) = sign (s_i)

$$S_i = \Sigma\, W_{ij}\, Y_i - \theta_i \quad \text{...(3.8)}$$

Where θ_i is the threshold value associated with hidden layer neuron

Each neuron in the hidden layer is fired only when the right interval is associated with the threshold θ_i. Region wise pixel information is accepted as the input vector for the input layer with block sizes 64, 128, 256, 512 and 1024. For each input block size the neural network is trained and the weight vectors converged. They are recorded for testing.

Algorithm 3.1 Document image segmentation

Step 1 Scan the document image store it in the database

Step 2 Using quad tree algorithm generate quad tree for the image.

Step 3 Select the quad tree regions one by one

Step 4 Apply IV fuzzy rules and fuzzify the image, store the IV fuzzy interval as an IV fuzzy Relation. Fix the threshold value θ_i

Step 5 Train the first layer with the adaptive activation function for the input layer

Step 6 Record the converged weight vectors

Step 7 Train the hidden layer with Kohonen's SOM

Step 8 Store the converged weight vectors

Step 9 Repeat the process for each training image

Step 10 Display the segmented image

Step 11 Stop

Algorithm 3.1.1 Training the first layer

Step 1 Input image regions

Step 2 Fix the input block size as 64 or 128 or 256 or 512 or 1024

Step 3 Input vector is $X = [X_1\ X_2 \ldots X_n]$ with block size given in step 2

Step 4 Initialize the weight vector as W= [1 1….1]

Step 5 Set the epochs and repeat the following iterations until the weight vector is Converged

Step 6 Compute the activation function

$$Yi = \frac{\mathrm{Wi}^{\mathrm{T}} X}{\|\mathrm{W}\|^2}$$

Step 7 W=W+ΔWi (Where ΔWi =ηδYi; η is the learning rate parameter usually in between 0.01 and 0.1; δ is in between 0 and 0.03; Yi is in between 0 and 1.

Step 8 Goto step 5

Step 9 Converged weight vector elements are in between 0 and 1

Step 10 Repeat the process for each training image

Step 11 stop

Algorithm 3.1.2 Training the hidden layer

Step 1 Output of first layer is the input for hidden layer

Step 2 Repeat the following for certain Epochs until the weight is converged

Step 3 Use the activation function $Si= \Sigma W_{ij} Y_i - \theta_i$

Step 4 If Si =0 then the i^{th} neuron is fired for the threshold

Step 5 Record the weight vector. Converged weight vector elements are in between 0 and 1.

Step 6 Go to step 2

Step 7 stop

Algorithm 3.2 Testing the adaptive neural network

Step 1 Input the quad tree regions of test image to trained Adaptive neural network

Step 2 Get the output of the neurons fired

Step 3 record it in the fuzzy relation

Step 4 Display the image from fuzzy relation

Step5 match it with original image

Step 6 Repeat this for all test images

Step 7 stop

Table 3.6: Relationship with input block size and epochs

Input block size	Epochs
64	10102
128	10258
256	10324
512	10397
1024	10460

3.8 Analysis of the Work

A comparative study of the Document Image Segmentation work with other methods is done. Quantitative comparison of segmentation algorithms are obtained by evaluating appropriate performance metrics. A large number of evaluation schemes are available to measure the capability of the segmentation algorithm of a document image. They are popular for different purposes in document understanding.

3.8.1 Comparison of the Work with Other Methods

Polynomial classifier(PC), Multilayer Perceptron(MLP), Radial Basis Function, Nearest Neighbor classifier(NNC), Guassian Assumption Bayes Classifier(GBC) Subset hood Product Fuzzy Neural Inference System (SuPFuNIS), O'Gorman(bottom up nearest neighbor clustering), Nagy et al top down approach for structural segmentation, Qingshneg Zhu, Zhongfu Wu, cian Wang and Zhangyu Zhu(Scale Fusion Approach of Document Image Segmentation), Syntactic Approach, mathematical formulae are available for document image segmentation. Each of these methods is suitable for layout analysis. Our method provides the facility of wide range of intervals. The following table compares the best classification rate attained by

Back Propagation Neural Network (BPNN), Self Organizing Feature Maps (SOM), Recursive Neural Network (RNN), Support Vector Machine (SVM) and IV fuzzy method

Table: 3.7 Best classification rates for each Neural Network classifier

Classifier	Best classification rate
BPNN	78
SOM	75
RNN	80.5
SVM	92.5
IV fuzzy method	95

Table 3.8: Confusion Matrix shows that ANN and IV Fuzzy methods supersedes ordinary Methods

Document Data bases →	UW III		ICDAR 2009		Our Document Database		
Items identified ↓	MLP	Ordinary methods	MLP	Ordinary methods	MLP/BPN	Ordinary methods	IV fuzzy methods
Non text classified as non text	98.9	95.36	96.70	95.44	98.92	95.56	98.98
Non text classified as text	1.1	4.64	3.3	4.56	1.08	4.44	1.02
Text classified as text	99.79	95.93	99.8	95.74	99.8	95.89	99.82
Text classified as non text	0.21	4.07	0.2	4.26	0.2	4.11	0.18
Photo graphs/images classified as images	98.9	95.36	96.69	95.43	98.85	95.84	99.07
Tables classified as tables	98.9	95.36	97.88	96.01	98.87	96	99.2
Segmentation accuracy	99.1225	95.50	97.77	95.66	99.11	95.82	99.26

3.9. Sample Outputs

Fig: 3.13 The original image of size (1104 ×842×3)

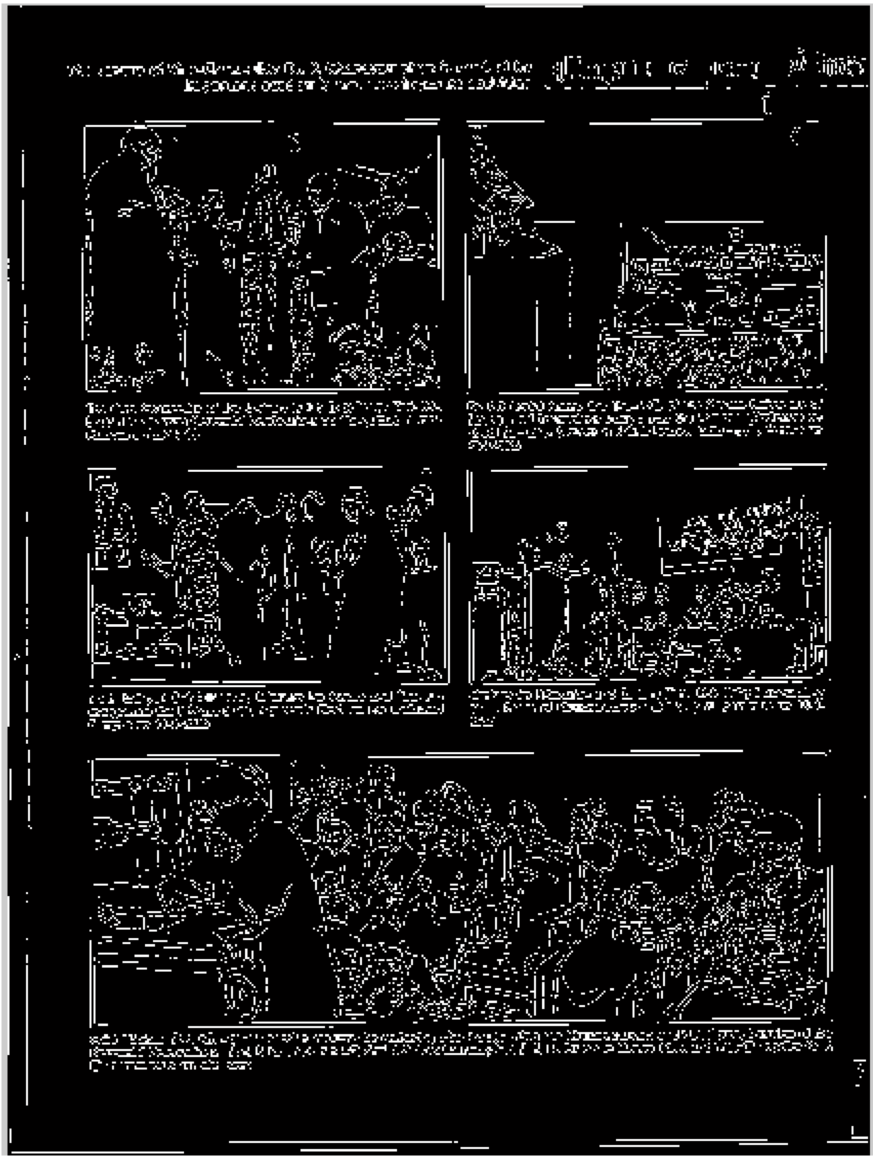


Fig: 3.14 Region wise features extracted of the original image

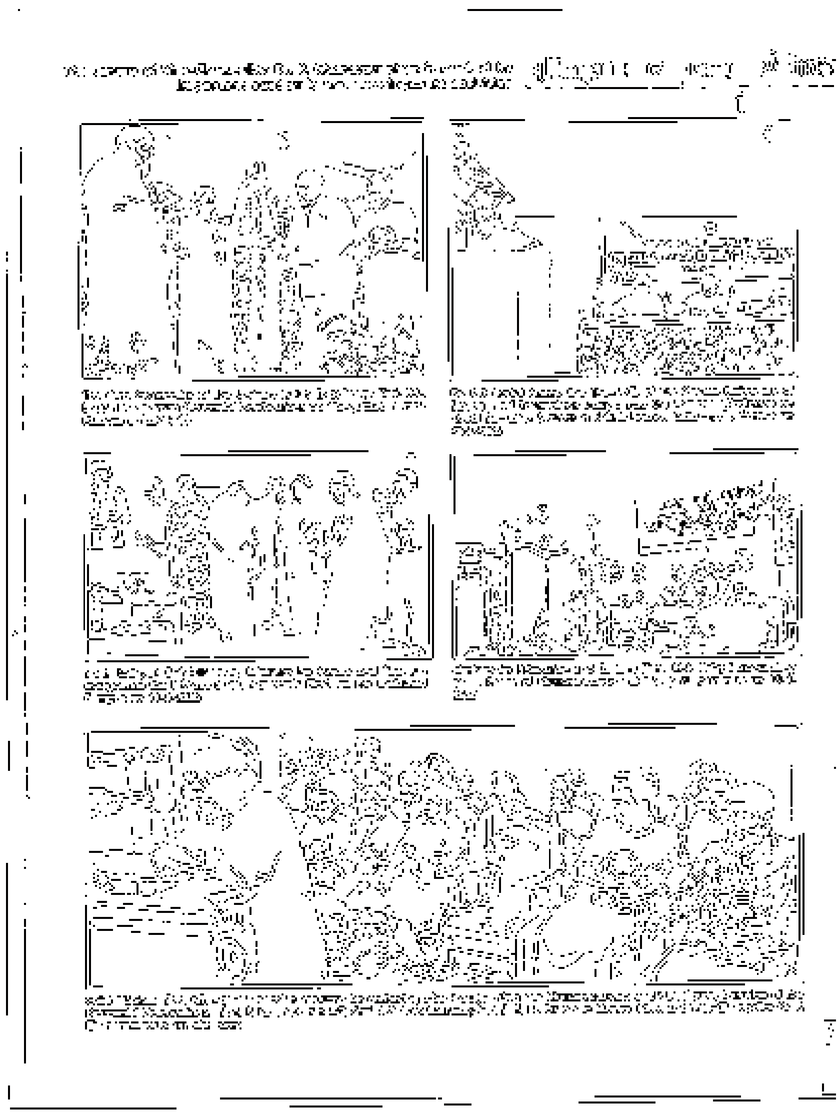


Fig: 3.15 feature extracted for the first fuzzy inerval

The tenure of Vice-Chancellor Dr. P. Chandramoh
important academic and develo

The then Chancellor of the University His Excellency Shri.
Bhatia inaugurates Decennial Celebrations on Mangattupara
Campus on 29-10-05

Fig : 3 16 A portion of the original image with interval 3 and 25

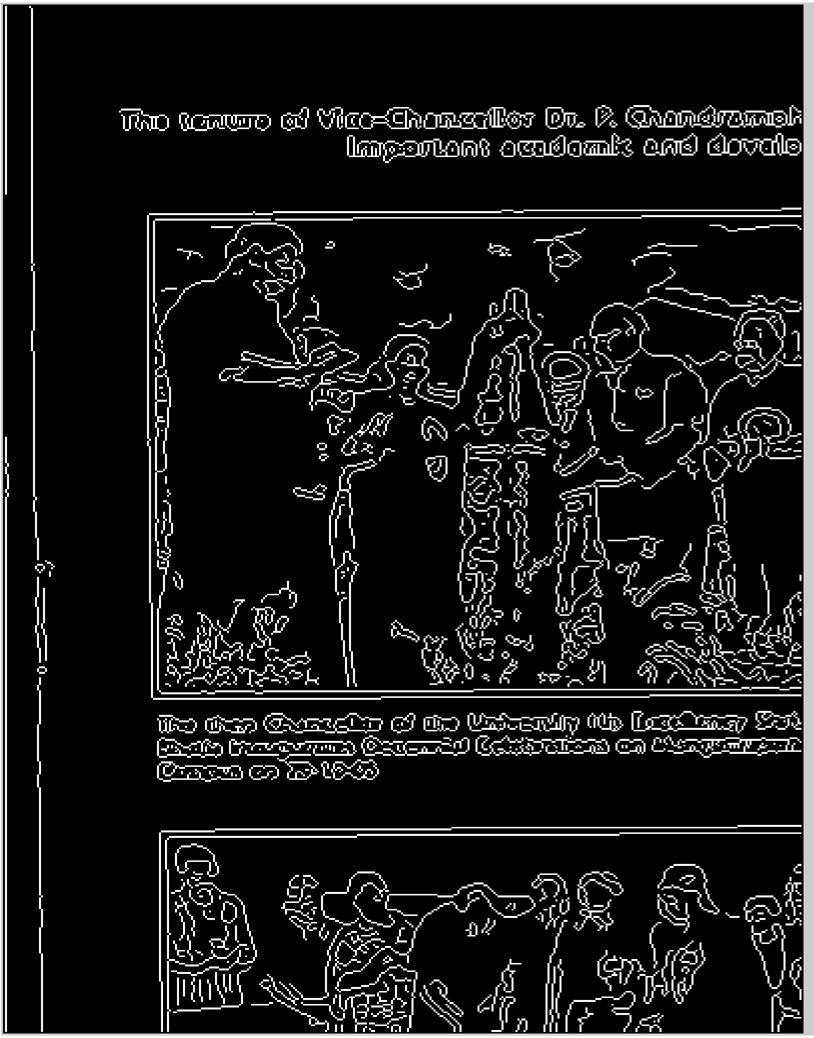


Fig: 3.17 Text and figure boundary identified

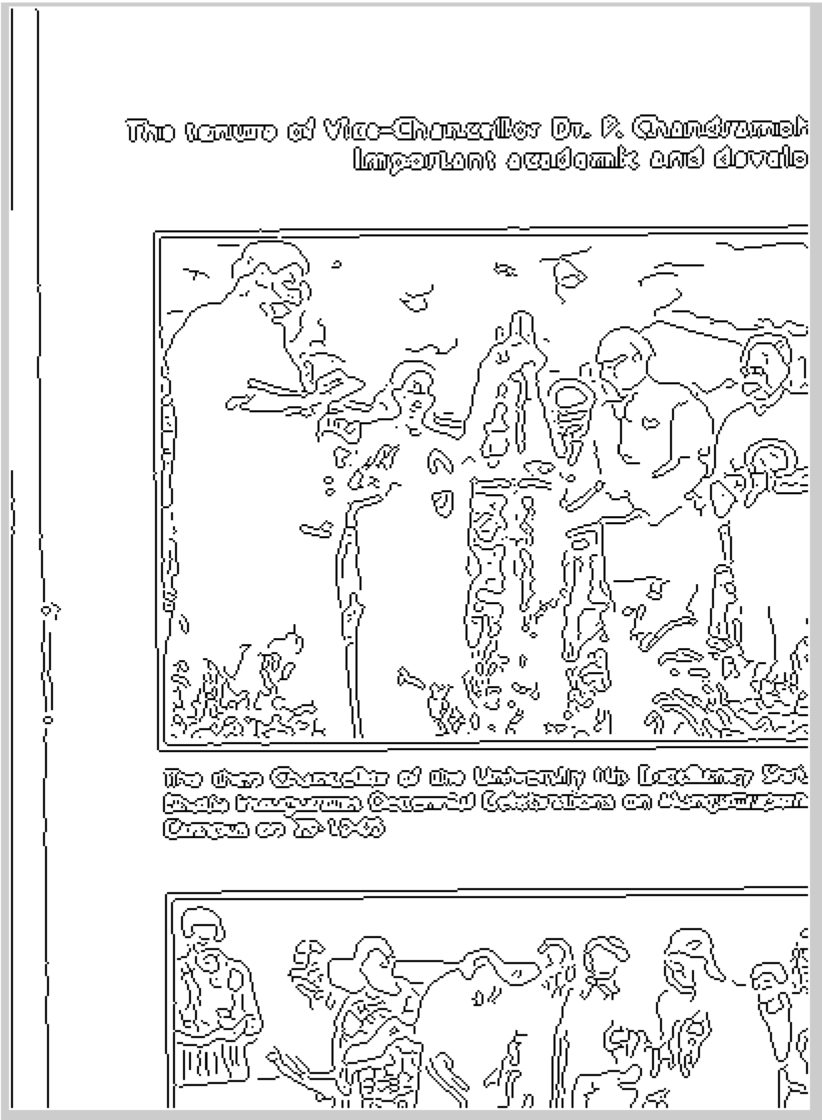


Fig: 3.18 Features and segmented regions

Fig: 3.19 Features of fuzzy interval 10 and 20

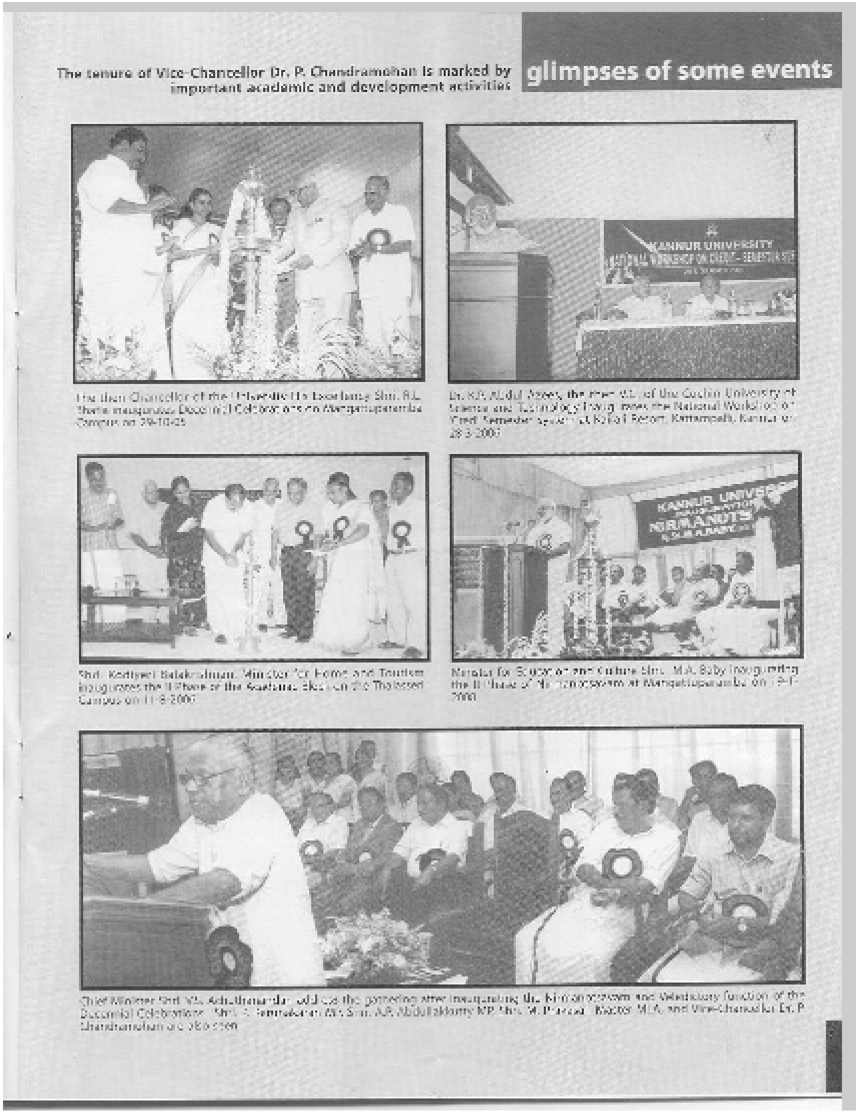


Fig: 3.20 Fuzzy intersection of intervals 3 and 18

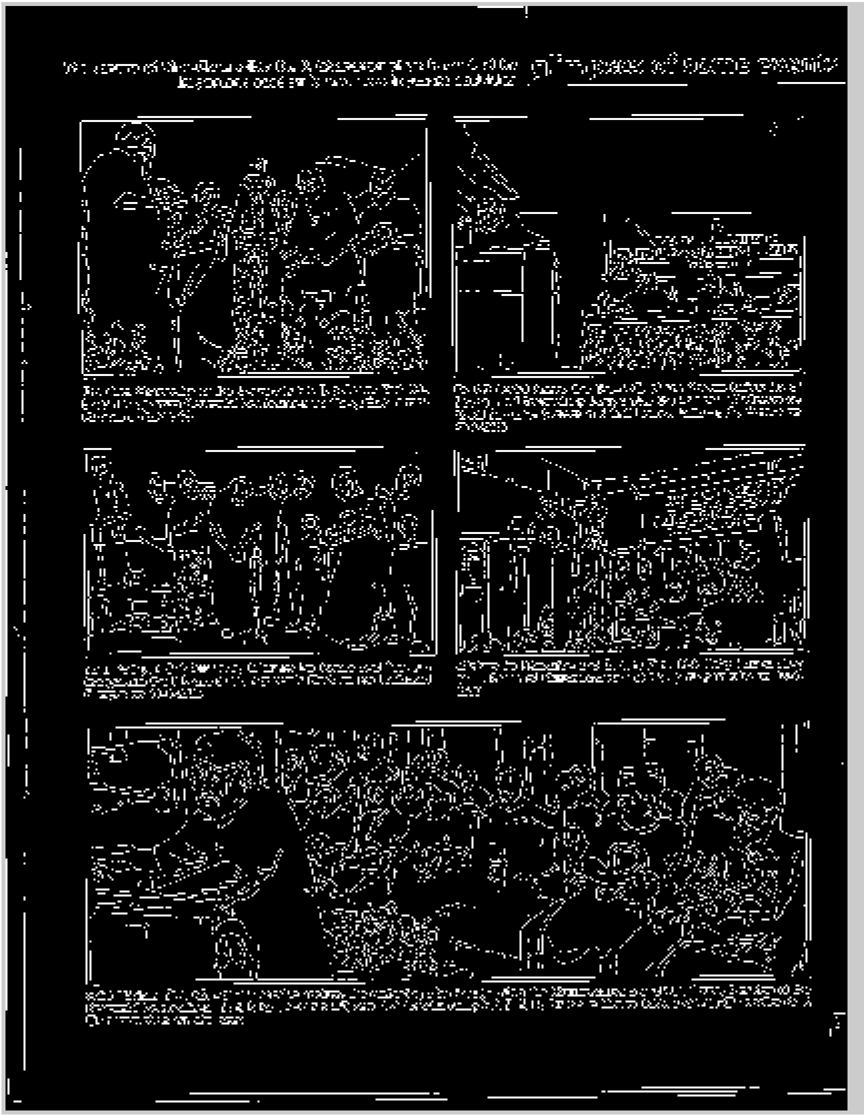


Fig: 3.21 Features analyzed after intersection of intensities

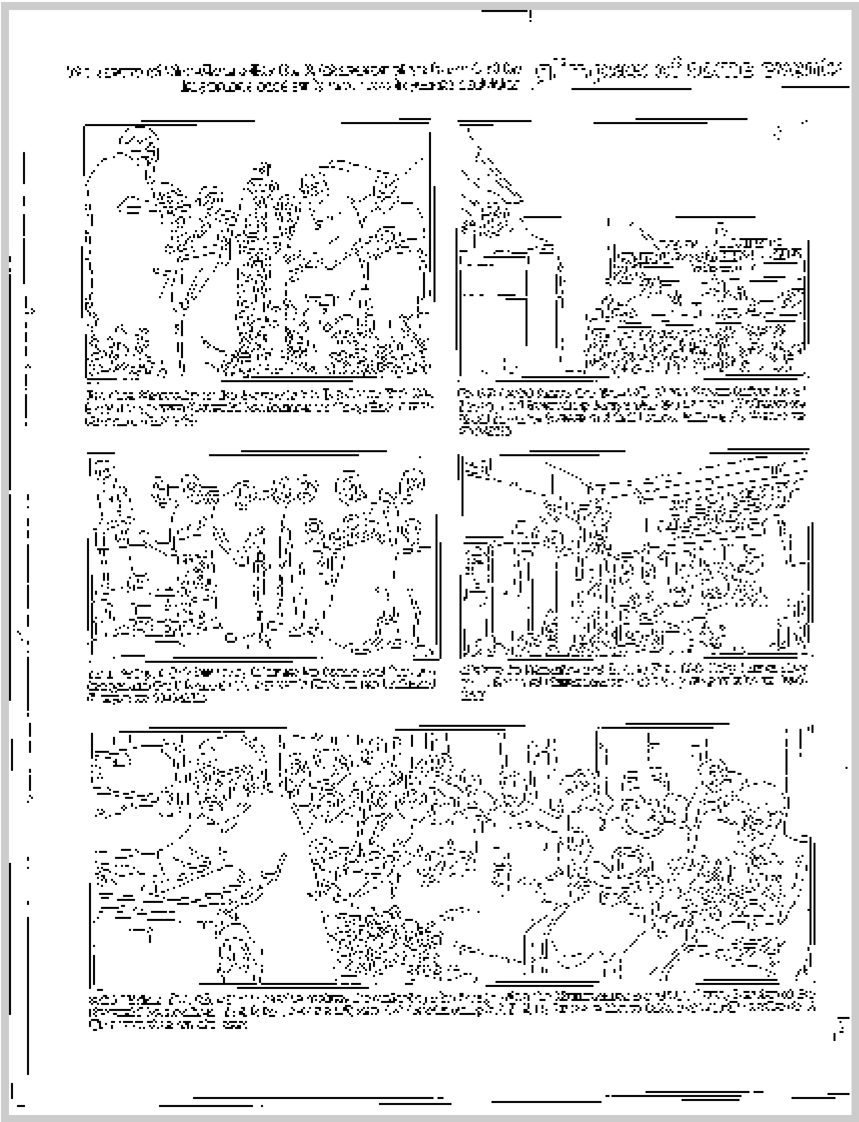
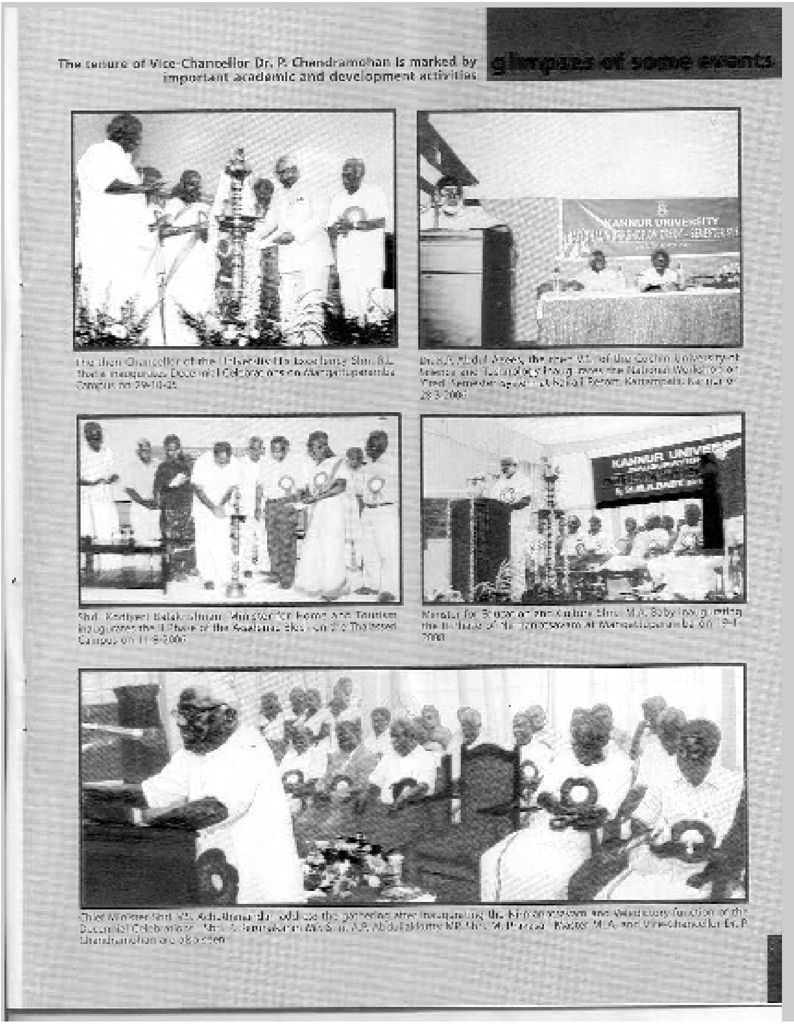


Fig : 3.22 Fuzzy union of the interval 3 and 18


The tenure of Vice-Chancellor Dr. P. Chandramohan is marked by important academic and development activities

glimpses of some events

Fig: 3.23 Desired portions are segmented and marked with strong rectangular regions

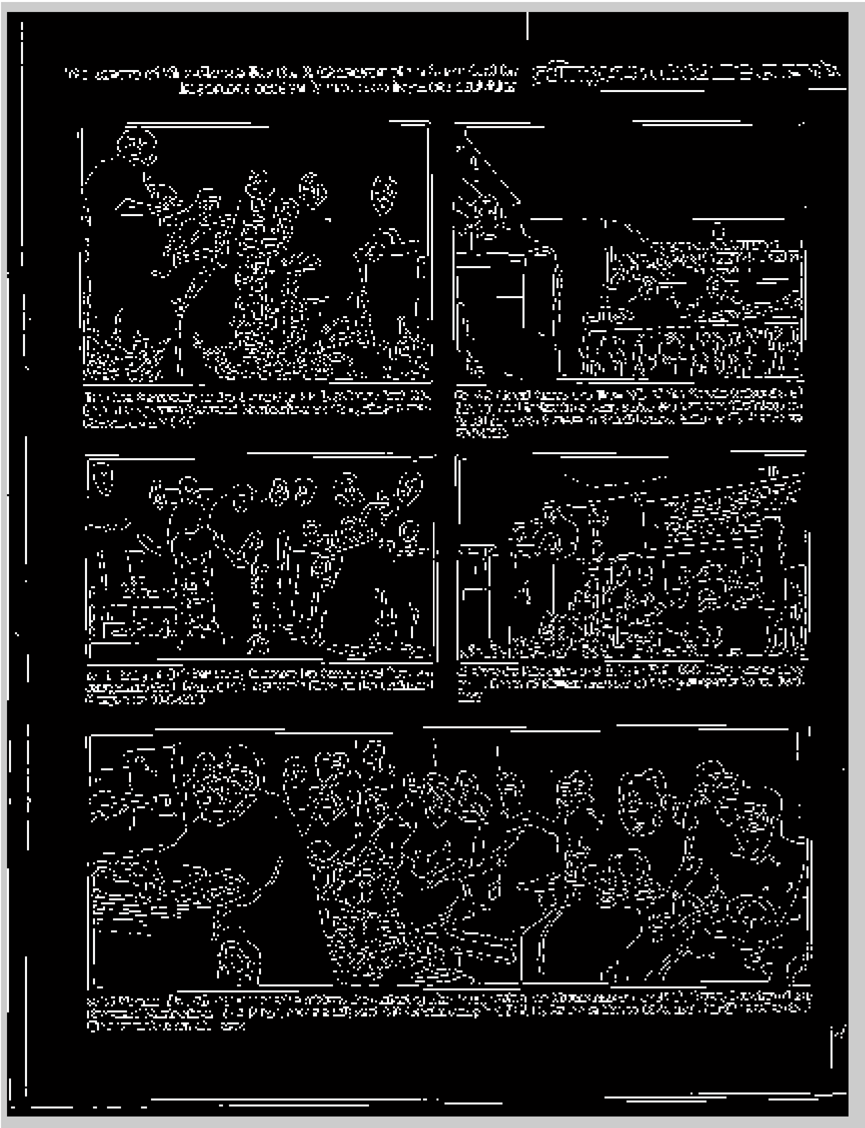


Fig: 3.24 Inverse image of interval 5

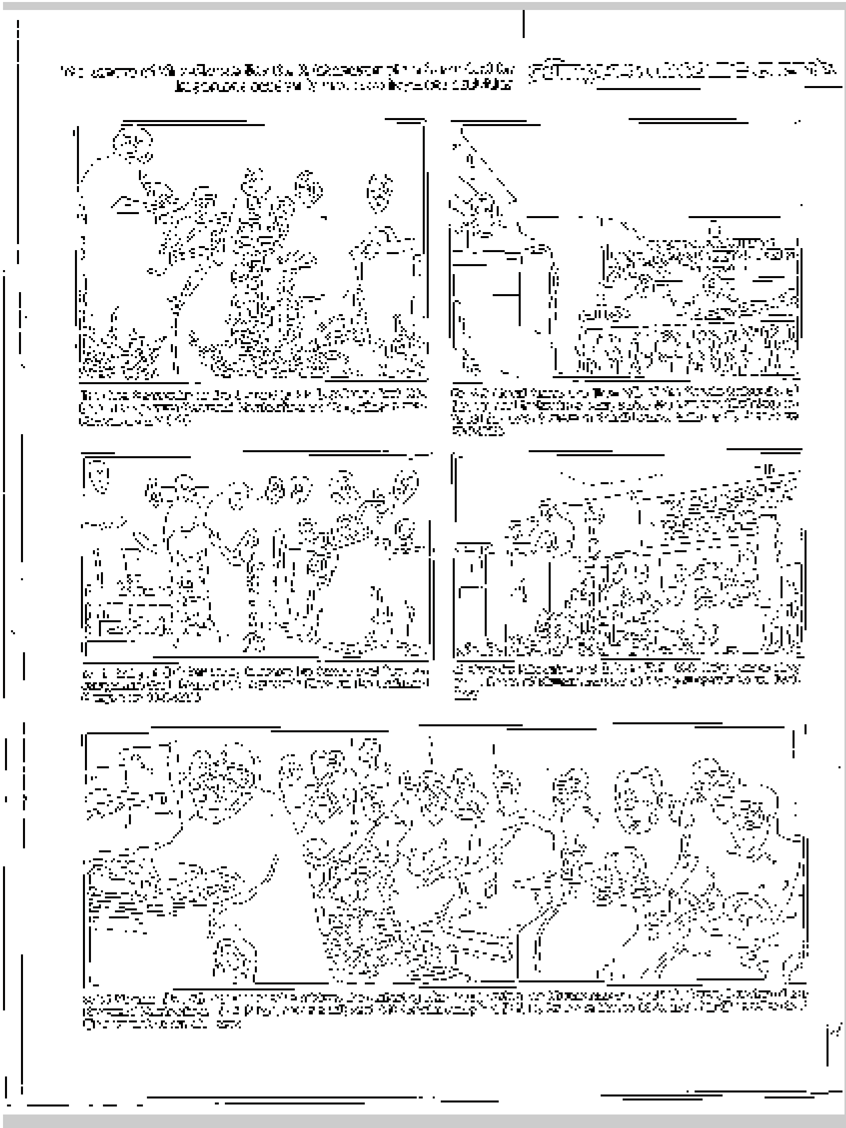
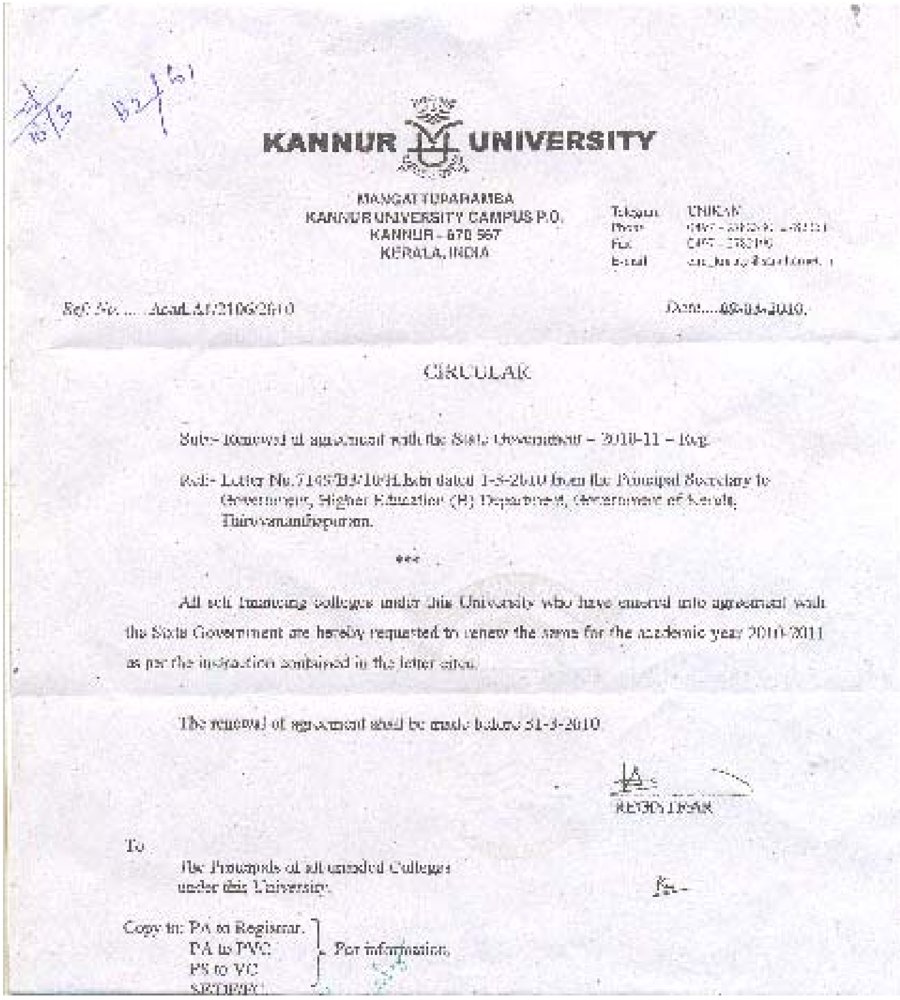


Fig: 3.25 Features extracted by the fuzzy interval 19



KANNUR UNIVERSITY

MANGATTUPARAMBA
KANNUR UNIVERSITY CAMPUS P.O.
KANNUR - 670 567
KERALA, INDIA

Ref. No. ..... Acad.A1/2106/2010 — Date.....09-03-2010.

CIRCULAR

Sub:- Renewal of agreement with the State Government – 2010-11 – Reg.

Ref:- Letter No.7145/B3/10/H.Edn dated 1-3-2010 from the Principal Secretary to Government, Higher Education (B) Department, Government of Kerala, Thiruvananthapuram.

***

All self financing colleges under this University who have entered into agreement with the State Government are hereby requested to renew the same for the academic year 2010-2011 as per the instruction contained in the letter cited.

The renewal of agreement shall be made before 31-3-2010.

REGISTRAR

To

The Principals of all unaided Colleges under this University.

Copy to: PA to Registrar, PA to PVC, PS to VC, SF/DF/FC — For information.

Fig : 3.26 Image of Size 1398×1256×3 classified into 3 regions based on the folding mark in the paper

3.10 Conclusion

A document image is classified into different planes and layers based on the requirement of the user. Once the text or image region with different thresholds is identified it helps to distinguish the foreground and background and in turn helps to eliminate unwanted information. Keeping the selected regions in fuzzy relation also reduces the storage space. Since the first row and column of a fuzzy relation are for the coordinate positions and each cell indicates the strength of the fuzzy interval the space is saved. In this work IV fuzzy rule base is constituted and an Adaptive Neural Network is designed with input layer, hidden layer and output layer for segmentation and extracting the features. The study conducted in different document databases shows that IV fuzzy method and Neural Network based method are better than ordinary segmentation algorithms.

………………

Chapter **4**

OPTIMIZATION OF DOCUMENT IMAGE SEGMENTATION WITH SIMULATED ANNEALING

Contents

4.1 Introduction

There are different image segmentation techniques such as Region growing, split and merge methods, fuzzy logic based segmentation, Fuzzy Hopfield Neural Network etc [89][90]. In the previous chapter we have described about the novel method to classify different regions, segment document images and economically save space. Document image segmentation with optimization helps us to correctly identify the clusters instead of using simple vector quantization techniques. When data space is large ordinary optimization techniques lead to retardation in speed. Stochastic Optimization technique such as Simulated Annealing works more efficiently in such situations. The technique Simulated Annealing (SA) [86] was first proposed in 1953 by Metropolis et al to simulate molecular processes. Krick Patricks, Gelatt and Vecchi used the idea as a method to resolve minimizing function of many variables such as hard problems [117] [118]. Simulated annealing is used for various applications like Seismic Applications, Image reconstruction algorithms, Optimizing Satellite image registration etc. [114]

[119] [120] [121] [122] [123] 124]. There are different Optimization Techniques related to Evolutionary methods which are given in table 4.1. When neural networks methods are simply used for classification purpose then there is a chance for getting converged at local minimum instead of global minimum. Simulated Annealing eliminates the chance to converge at local minima.

Table 4.1 Optimization Techniques [110][111][114][117][118][119][122][123][124]

Method	Description
Quantum Annealing	Uses quantum fluctuation to get through high but thin barriers in the target function
Stochastic Tunneling	A method to overcome the barriers of local minima through tunneling
Tabu Search	Able to provide uphill moves when it caught in local minima also avoids repeating in cycles by referring a 'taboo list' of solutions
Reactive Search optimization	Combining machine learning with optimization
Stochastic Gradient Descent	Able to perform many greedy searches from random initial states
Genetic Algorithms	It maintains a pool of solutions than one
Graduated Optimization	Smoothes the target function while optimizing
Ant colony Optimization(ACO)	Uses many agents to traverse the solution space
Cross entropy Method(CE)	Generates candidate solution with parameterized probability distribution and updates parameters and finds better samples in next iteration
Harmony Search	It mimics musicians as they improve their music
Stochastic optimization	It includes simulated annealing and other methods.
Particle swarm optimization	Modeled on swarm intelligence to get optimum solutions.
Intelligent Water Drops (IWD)	Mimics natural water drops for optimization
Parallel Tempering	Copies at different temperature or Hamiltonians to overcome potential barriers.
Simulated Annealing	Tries to get single best solution

Simulated annealing algorithm provides some heuristics to compare current state with neighboring states and provides more accurate results compared to other methods. Ordinary Neural Network methods like Back Propagation Network (BPN) have a disadvantage that during training, there is a chance to converge the weight vectors at local minima. BPN combining with Simulated Annealing eliminates the chance to converge at local minima and helps to converge at global minima. A document image is of complex nature but have certain common features within it. The different image regions in a Document Image are located and identified easily with the help of simulated annealing. An annealing schedule is required for the purpose. In this chapter we discuss about the method developed by us to optimize document image segmentation using Simulated Annealing (SA).

4.1.1 Simulated Annealing

Simulated Annealing is known by alternate names like Monte Carlo Annealing, Statistical Cooling, Probabilistic Hill Climbing, Stochastic Relaxation, and Probabilistic Exchange Algorithm. This algorithm is easy to implement and its simplicity is independent of the problem size. For implementing SA we have to consider the following matters:

- ✓ A method to generate initial configuration
- ✓ A transition or generation function to find a neighbor as next candidate.
- ✓ A cost function
- ✓ An evaluation criterion
- ✓ A stop criterion

Simulated annealing sometimes accepts candidates with higher cost to escape from local optimum. It adapts the parameters of the evaluation function during execution. It is functioning based upon the analogy with physical annealing. When heated solid is cooled down slowly, all particles arrange in the ground energy state. Probability of being in state with energy

$$E{:}P(E) = \frac{1*\exp\left(-\frac{E}{kb}*T\right)}{Z(T)} \quad \text{...(4.1)}$$

Where,

E - Energy

T - Temperature

kb - Boltzmann constant

Z(T) - Normalization factor – temperature dependent.

Simulation of cooling (Metropolis 1953) is as follows:-

a) At a fixed temperature T, Perturb(randomly) The current state to a new state
b) ΔE is the difference in energy between current and new state.
c) If $\Delta E<0$ (new state is lower) , accept new state as current state.
d) If $\Delta E>=0$, accept new state with probability
e) Pr(accepted)= exp(-ΔE/kB.T)
f) Eventually the system cools into thermal equilibrium at temperature T, then the formula mentioned before holds.
g) When equilibrium is reached, temperature T can be lowered and the process can be repeated.

Krickpatrick (1982) applied Simulated Annealing to optimization problems.

Simulated Annealing Cooling Schedule needs the following

- ✓ A starting temperature
- ✓ Final temperature
- ✓ Temperature decrement
- ✓ Iterations at each temperature

The starting temperature must be hot enough to allow moves to almost all neighborhood states. The starting temperature must not be so hot that we conduct a random search for a period of time. The problem is finding a suitable starting temperature. If we know the maximum change in the cost function we can use this to estimate the starting temperature. Start high, reduce quickly until about 60% of worse moves are accepted. Use this as the strategy for choosing starting temperature. Heat rapidly until a certain percentage is accepted then starts cooling. For choosing final temperature it is usual to let the temperature decrease until it reaches zero. However this can make the algorithm run for a lot longer, especially when a geometric cooling schedule is being used. In practice, it is not necessary to let the temperature reach zero because the chances of accepting a worse move are almost the same as the temperature being equal to zero. Therefore the stopping criterion can either be a suitably low temperature or when the system is "frozen" at the current temperature(ie, no better or worse moves are being accepted). Theory states that we should allow enough iteration at each temperature so that the system stabilizes at that temperature. In certain cases this may exponential to problem size. We need to compromise. We can either do this by doing a large number of iterations at a few temperatures, a small number of iterations at many Temperatures or a balance between the two. In linear expression temperature is adjusted as temp=temp-c. In Geometric expression temperature is adjusted as temp=temp*c. Experiences has shown that c should be between 0.8 and 0.9 with better results being found in the higher end of the range. Of course, the higher the value of c the longer it will take to decrement the temperature to the stopping criterion. One method is using a constant number of iterations at each temperature. Another method, first suggested by Lunchy in 1986 is to do one iteration at each temperature, but to decrease the temperature very slowly. The formula used by Lunchy is t=t/(1+ t) where   is a small value.

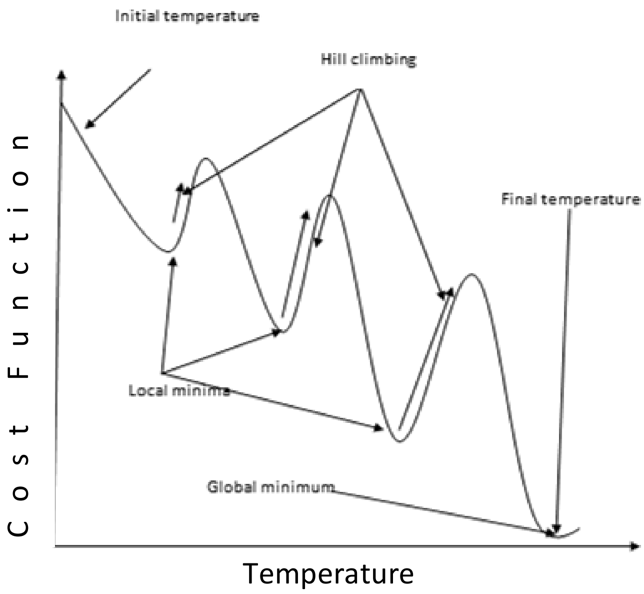


Fig 4.1 Convergence in simulated annealing

4.2 The Method

4.2.1 Quad Tree Generation

A quad tree of the original image is prepared first. The tree is explored such as the quadrant size becomes 1024×1024, 512×512, 256×256, 128×128, 64×64, 32×32, 16×16, 8×8 at each level, depending upon the size and nature of the image. When uniform leaf nodes are obtained that leaf nodes may not be explored further. A quad tree of document image of level 1 is given in the fig 4.2.

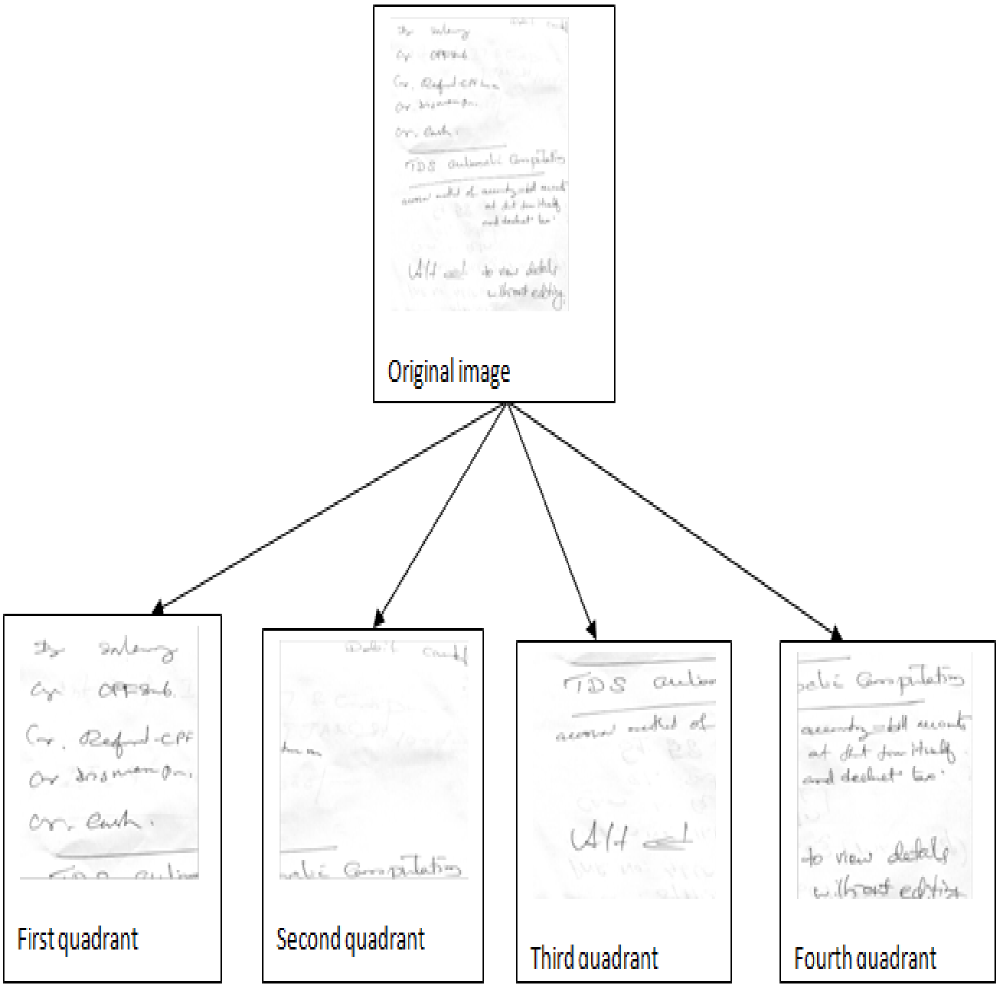


Fig 4.2: Quad tree of a document image of level 1

4.2.2 Cooling procedure

As mentioned in the previous chapter a neural network is designed for document image segmentation. The background is identified from the foreground text by change in intensities of the image. Hard limiting activation function ($y_i=w_ix_i-\varphi$, where φ is the threshold value) is used for the input and output layer and sigmoid function $\frac{1}{1+e^{-Ix}}$

ie, I=1 is used for the hidden layer. Input block sizes of 64, 128, 256, 512 and 1024 are used. The input vector X is $[X_1\ X_2\ \ldots X_n]$ where n is the block size. There are 26 neurons corresponding to 26 IV fuzzy intervals in the hidden layer. $[W_1\ W_2\ \ldots..W_n\ ]$ is the weight vector Weight vector is initialized as zeroes. For every previous state and current state, the intensity values are accounted for change in energy parameter. We started from the top left corner of each quad tree leaf node and the Energy level E is set as the intensity interval of the current pixel. Temperature is initialized as 1(ie, 255/255). And final temperature is kept as a value near to zero to achieve stability. A transition from one state to another state is determined by the Boltzmann probability distribution $p=1/(1+e^{-\Delta ET})$. Where ΔE is the change in intensity value compared current state to the new state. T is the temperature. The temperature of the annealing schedule is adjusted as follows: $T_{k+1}=cT_k$. Where c is a constant. This is the geometric expression for cooling. The value of 'c' is set in between 0.8 and 0.9 since this range acts as best values for the cooling schedule. Neighborhood of a pixel x_{ij} is considered as follows:

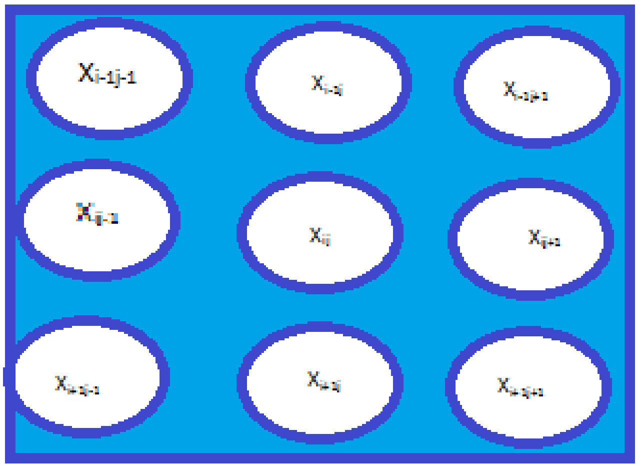


Fig: 4.3 Neighborhood of Pixel x_{ij}

In simulated annealing algorithm there is freedom to consider the entire neighborhood to find Optimal Solutions. If x_{ij} and x_{ij+1} are of same intensity values then ΔE will be equal to 0. This indicates that the pixels are continuous part of text or picture. Similarly this is considered for all Neighborhood pixels.

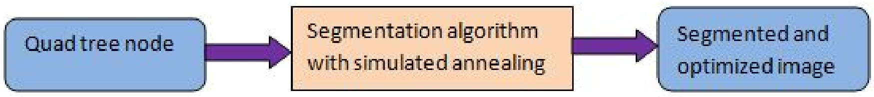


Fig 4.4: Segmentation Optimization using Simulated Annealing

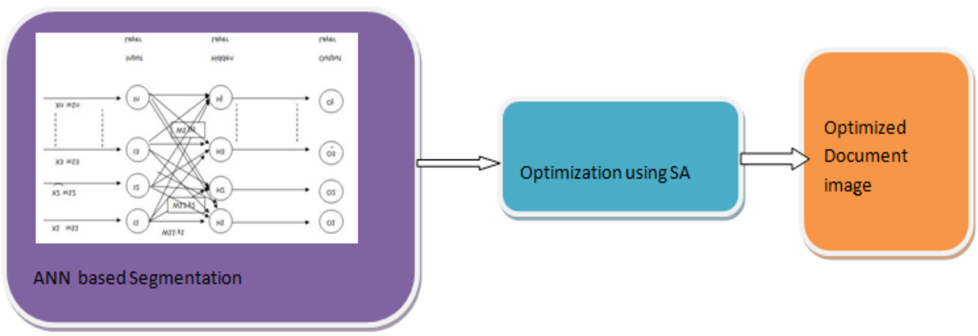


Fig 4.5 Detailed diagram of fig 4.4

Algorithm 4.1 Segmented image optimization using Simulated Annealing

Step1: Read the image

Step 2: Initialize the temperature ie, T=1

Step 3: initialize cooling= false

Step 4: While (! cooling) {

Step5: Anneal the system slightly (ie, move to the next pixel of the Image)

Step 6: Compute ΔE (Difference in intensity of the image pixels, change

in energy due to annealing schedule)

Step 7: If (ΔE <0) then accept this perturbation, this is the new system configuration. (This situation indicates that the nearby pixel intensities are almost same.)

Else accept, with probability p= $\frac{1}{1+\exp^{(-\Delta E/k * T)}}$

Where k=1 }

Step 8 : Do until(the system is in thermal equilibrium at this T=0)

Step 9: If (ΔE still decreasing over the last few temperatures) then

{T=0.9T cool the system

Else cooling=true}

Step 10: return (final configuration as low energy solution)

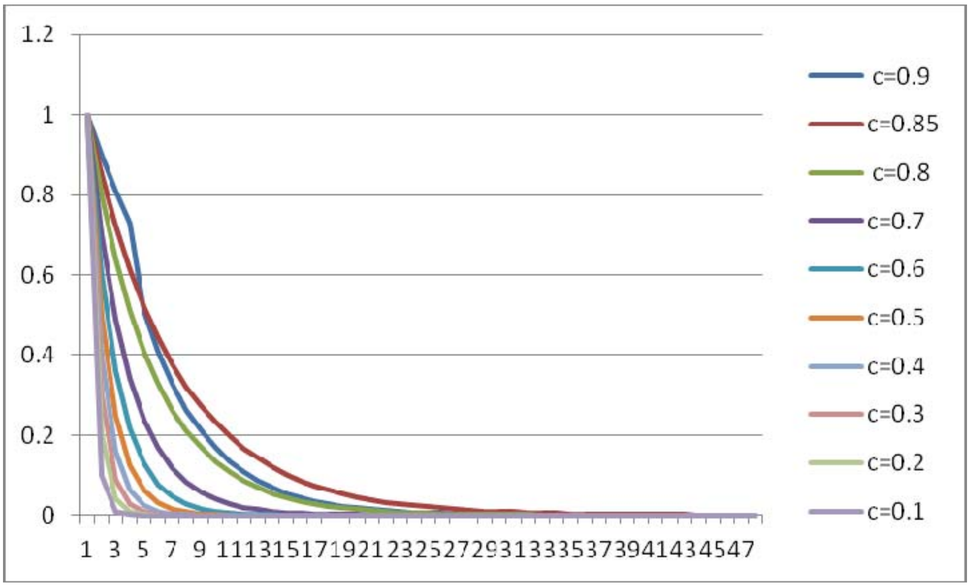
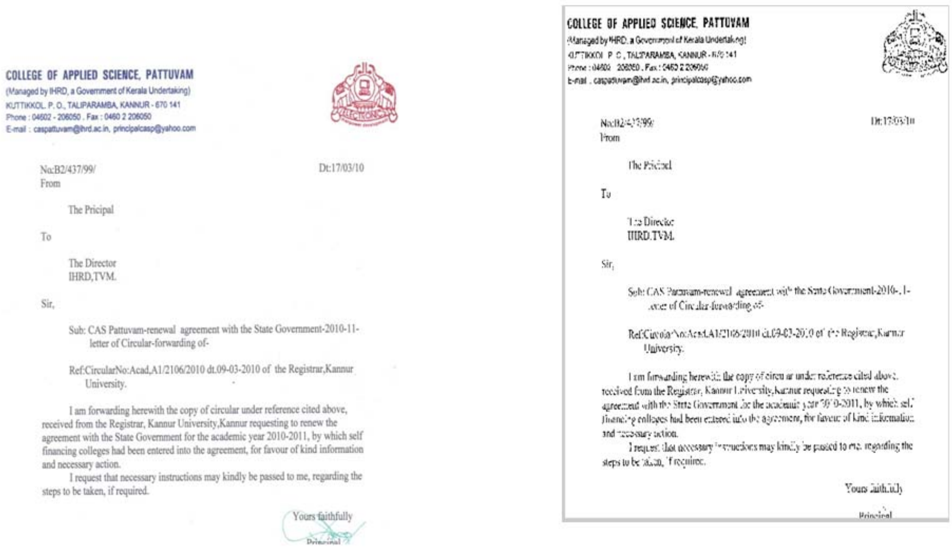


Fig: 4.6 Cooling schedule for different c values

Original image | After Optimized with SA



COLLEGE OF APPLIED SCIENCE, PATTUVAM
(Managed by IHRD, a Government of Kerala Undertaking)
KUTTIKKOL P. O., TALIPARAMBA, KANNUR - 670 141
Phone : 04602 - 206050 , Fax : 0460 2 206050
E-mail : caspattuvam@ihrd.ac.in, principalcasp@yahoo.com

No:B2/437/99/ Dt:17/03/10

From

The Pricipal

To

The Director
IHRD,TVM.

Sir,

Sub: CAS Pattuvam-renewal agreement with the State Government-2010-11-letter of Circular-forwarding of-

Ref:CircularNo:Acad,A1/2106/2010 dt.09-03-2010 of the Registrar,Kannur University.

I am forwarding herewith the copy of circular under reference cited above, received from the Registrar, Kannur University,Kannur requesting to renew the agreement with the State Government for the academic year 2010-2011, by which self financing colleges had been entered into the agreement, for favour of kind information and necessary action.

I request that necessary instructions may kindly be passed to me, regarding the steps to be taken, if required.

Yours faithfully

Principal

Fig 4:7 The first image is the original image, the second image shows pixels with prominent intensity values

4.3 Comparison of the Work with Other Methods

Different Optimization techniques are: Mathematical Programming, Network Analysis, Branch and Bound, gradient descent, Genetic Algorithm, Simulated Annealing, Tabu Search etc. Applicability of the optimization algorithms depends on the problems to be solved. The results are also depends on the type of the image we have used. All these algorithms are suitable for general optimization problems. Suitability of the algorithm for image segmentation depends on the nature of the document image. Gradient descent is suitable for problems without multiple local minima. If number of local minima is small simulated annealing shows poor performance. Simulated Annealing relies on a principle of locality. It performs best the number of local minima is very large.

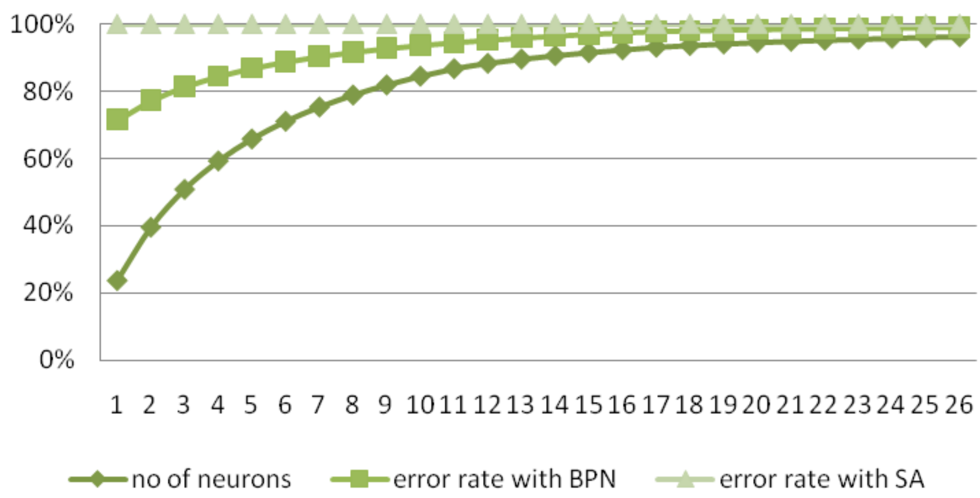


Fig: 4.8 Error rate comparison of BPN with Simulated Annealing

Table 4.2: Confusion Matrix shows that Neural Network methods with Simulated Annealing supersedes ordinary methods

Document Data bases →	Our documents			
Items identified ↓	**SA**	**MLP**	**Ordinary methods**	**IV fuzzy methods**
Non text classified as non text	98.985	98.92	95.56	98.98
Non text classified as text	1.0	1.08	4.44	1.02
Text classified as text	99.87	99.8	95.89	99.82
Text classified as non text	0.13	0.2	4.11	0.18
Photo graphs/images classified as images	99.86	98.85	95.84	99.07
Tables classified as tables	99.3	98.87	96	99.2
Segmentation accuracy	99.496	99.11	95.82	99.26

4.4 Conclusion

In this chapter we have discussed a new Document Image Segmentation Optimization method with Simulated Annealing algorithm. It avoids convergence at local minimum. In our work we used the Document image database prepared for our work. It consists of hand written as well as typed documents with different properties like text, image, photograph etc. The gradient factor for the sigmoid function is adjusted as 1 and the trial and error of various epochs are done and it converged for about 10000 epochs. For the hard limiting function the output is cut off below or above a range of threshold values. (It will depend on the background colour of the text values). The back propagation algorithm shows some anomalies when the paper is toned or when some stains present in the paper. So Simulated Annealing is used. For adjusting the Temperature the c is set as a value in between 0.8 and 0.9. The value of c in this range showed better performance in experiments. When a variation in intensity is occurred then the pixel position is checked. If the pixel is in the median text line and if it is a connected component it is left as it is. Otherwise the pixel intensity is changed as that of the background by checking the neighborhood pixel intensities. The experiment shows that segmentation based on Neural Network with Simulated Annealing is much better than that of traditional methods.

………………

Chapter **5**

DOCUMENT IMAGE COMPRESSION OPTIMIZATION USING GENETIC ALGORITHMS

Contents

5.1. Introduction

Compression of images and document images helps to manage office space, search time, to reduce number of office staff, transmission time etc [18][19][20][21][22]. Compression of Document images was very relevant for managing disk space for a long period. Due to the fastest technology it is easy to retrieve information and store files without much delay. But concerned with internet traffic, it is relevant to reduce the file size for transferring information with high speed [24][25][26][27][28]. And most of the internet service providers fixed their rate for certain GB of downloadable data [53][54][55][56][57][58][59][60]. In such situations, if there are efficient lossless compression algorithms available to compress document images significantly, then it will help to reduce congestion in the network [32][34][35][38]. And most of the available algorithms are for ordinary images, not for document images. So this work is an attempt to use

Evolutionary Algorithms (EA) for the optimization of lossless compression of document images [110][111][123].

The structure of the document image vary from language to language, context to context, content to content etc. Even though it contain certain common features like line spacing, word spacing, column spacing etc. there are certain difficulties and issues in identifying regions. ie certain difficulties in identifying boundaries, sometimes not exact boundaries, blurred edges, incomplete edges etc. In order to compress the data with ordinary algorithms it is difficult to deal with large two dimensional spaces of document images. The compression technique using evolutionary algorithms helps to improve the scanning procedure of document images and store the image in a compressed format at the scanner level itself [74][76][78][79][80][81]. There are a number of algorithms which comes under the broad category of Evolutionary Algorithms (EA) [82][91][92][108]. EA provides better solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, crossover, mutation and selection. Evolutionary techniques include Meta heuristic optimization algorithms such as Evolutionary algorithms and swarm intelligence. Evolutionary algorithms comprising of genetic algorithms, evolutionary programming (the term was first used by Lawrence J Fogel in US in 1960), evolution strategy and genetic programming. Swarm intelligence comprising of ant colony optimization and particle swarm optimization. Evolutionary techniques also extends to self organizing maps, differential evolution, artificial life, cultural algorithms, harmony search, artificial immune systems, learning classifier systems, learnable evolution models etc. Currently evolutionary programming is a wide evolutionary computing dialect with no fixed structure.

The main variation operator of Evolutionary programming is mutation. The members of the population are viewed as part of a specific species rather than members of the same species therefore each parent generates an offspring, using a (u + u) survivor selection. Evolution Strategy (ES) is an optimization technique based on ideas of adaptation and evolution. It was introduced in 1960s and later developed by Ingo Rechenberg, Hans-Paul Schwefel and his co workers in 1970s. Evolution strategies use natural problem-dependent representations, and primarily mutation and selection, as search operators. In common with EA the operators are applied in a loop. An iteration of the loop is called a generation. The sequence of generations is continued until a termination criterion is met. As far as real-valued search spaces are concerned, mutation is normally performed by adding a random value to each vector component. The step size or mutation strength (i.e. the standard deviation of the normal distribution) is often governed by self-adaptation. Individual step sizes for each coordinate or correlations between coordinates are either governed by self-adaptation or by covariance matrix adaptation [110][111][123].

The (environmental) selection in evolution strategies is deterministic and only based on the fitness rankings, not on the actual fitness values. The simplest evolution strategy operates on a population of size two: the current point (parent) and the result of its mutation. Only if the mutant's fitness is at least as good as the parent one, it becomes the parent of the next generation. Otherwise the mutant is disregarded. This is a *(1 + 1)-ES*. More generally, λ mutants can be generated and compete with the parent, called *(1 + λ)-ES*. In (1 + λ)-ES the best mutant becomes the parent of the next generation while the current parent is always disregarded. Contemporary derivatives of evolution strategy often use a population of μ parents and also recombination as an additional operator, called *(μ/ρ+ λ)-ES*. This is to avoid local optima.

Another optimization algorithm is Genetic Programming (GP), which is a specialization of Genetic Algorithms where each individual is a computer program. The optimization procedure is done with the help of a fitness landscape which determines a programs ability to perform a given computational task. In 1954, GP began with the EA first utilized by Nils Aall Baricelli applied to evolutionary simulations. In the 1960s and early 1970s, EA became widely recognized as optimization methods. Ingo Rechenberg and his group were able to solve complex engineering problems through ES as documented in his 1971 PhD thesis and the resulting 1973 book. In 1964 Lawrence J Fogel, one of the earliest practitioners of the GP methodology, applied evolutionary algorithms to the problem of discovering finite-state automata. Later GP-related work grew out of the learning classifier system community, which developed sets of sparse rules describing optimal policies for Markov decision processes. The first statement of modern "tree-based" Genetic Programming (that is, procedural languages organized in tree-based structures and operated on by suitably defined GA-operators) was given by Nichael L Cramer in 1985. This work was later greatly expanded by John R Kosa, a main proponent of GP who has pioneered the application of genetic programming in various complex optimization and search problems.

In the 1990s, GP was mainly used to solve relatively simple problems because it is very computationally intensive. Recently GP has produced many novel and outstanding results in areas such as quantum computing, electronic design, game playing, searching, and sorting due to improvements in GP technology and with the growth in CPU power. GP and GA are still in scenario due to its capability. Developing a theory for GP has been very difficult and so in the 1990s GP was considered a sort of outcast among search techniques. But after a series of breakthroughs in the early 2000s, the theory of GP has had a formidable and rapid development. Since GP is represented as tree structures in memory, any programming language which permits recursion can be used for writing programs.

Meta-Genetic Programming is the proposed Meta learning technique of evolving a genetic programming system using genetic programming itself. It suggests that chromosomes, crossover, and mutation were themselves evolved, therefore like their real life counterparts should be allowed to change on their own rather than being determined by a human programmer[70][76][81][108]. Meta-GP was formally proposed by Jorgen Schmidhuber in 1987, but some earlier efforts may be considered instances of the same technique, including Doug Lenats Eurisko. It is a recursive but terminating algorithm, allowing it to avoid infinite recursion. Critics of this idea often say this approach is overly broad in scope. However, it might be possible to constrain the fitness criterion onto a general class of results, and so obtain an evolved GP that would more efficiently produce results for sub-classes. This might take the form of a Meta evolved GP for producing human walking algorithms which is then used to evolve human running, jumping, etc. The fitness criterion applied to the Meta GP would simply be one of efficiency [110][111][123]

5.2 The Method

Document image compression methods are of two types lossless and lossy. There are different Algorithms for Data compression. Commonly used lossless data compression techniques are Data de duplication, Run length encoding, dictionary coders like LZ77, LZ78 and LZW[50][51][52], Burrows Wreder transform, Prediction by Partial Matching(PPM), context mixing, Dynamic Markov Compression(DMC)[42][43][44][46][47][48][49], Entropy encoding methods such as Huffman coding, Adaptive Huffman coding and Shannon-Fano coding, Arithmetic coding are more advanced compression techniques compared to the previous methods [37][38][39][40][41]. Range coding is same as arithmetic coding but treated in a slightly different way. Golomb coding is simple entropy coding for infinite input data with a geometric distribution. Universal codes are entropy coding for infinite input data with an arbitrary distribution. Two sub categories of Universal codes are Elias gamma coding and Fibonacci coding. Stepian-Wolf coding (SWC) is a lossless Distributed Source Coding (DSC). Lossy Data compression techniques are Discrete Cosine Transform, Fractal Compression and fractal transforms, Wavelet compression, Vector Quantization, linear Predictive coding, Modulo-N code for correlated data, A-law compander, Mu-Law compander, Wyner-Zino coding(WZC) a lossy Distributed Source Coding, etc. JPEG Implementations uses Discrete Cosine Transform, then Vector Quantization and then Huffman Coding. MPEG uses DCT and motion compensated prediction for video.

Best Lossless Compressors use Probabilistic methods. Jorma Rissanen invented a coding named Arithmetic coding. Witten, Neal and Cleary provided practical methods for the same. There is different Document Image compression standards available based on arithmetic coding. JBIG a bi-level image compression standard, DjVu a document compression standard, Dashu an inverse Arithmetic coder etc used Arithmetic coding as the base algorithm[42][43][47][48].

Various implementations follow lossless or lossy methods of mixed algorithms of above described categories. Human eye can recognize the images even if some of the pixels are lost in an image. A compression ratio around 6:1 is achieved with lossless techniques. And very high compression ratios were achieved with lossy methods. Mixed Content Compression (MCC) technique achieved 60:1, for JPEG ratio achieved is 35:1, for mono text documents it varies from 44:1 to 82:1, Raster Document Image Compression achieves 57:1 compression ratio with 0.420 bits per pixel bit rate [22]. For Document Images better compression ratios can be achieved by proper selection of compression algorithms, optimization techniques, background elimination etc.

We concentrate on Optimization techniques to achieve better speed and compression ratios by identifying and eliminating background, by reducing latency time for compression and reconstruction, by reducing the bits per pixel etc. We chose Genetic algorithm for optimization, because when search space is large Genetic Algorithm suits well [111]. After segmentation, sequential Genetic Algorithm and then parallel Genetic Algorithm are used and performance is analyzed.

As a first step we consider Genetic Algorithm. In 1960s John Holland proposed a method known as Genetic Algorithm (GA). This is useful and efficient when,

- The search space is large, complex or poorly understood
- The domain knowledge is scarce or expert knowledge is difficult to encode to the narrow search space.
- No mathematical analysis is available.
- Traditional search methods fail.

The advantage of the GA approach is the ease with which it can handle arbitrary kinds of constraints and objectives; all such things can be handled as weighted components of the fitness function, making it easy to adapt the GA scheduler to the particular requirements of a very wide range of possible overall objectives. The general form of Genetic Algorithm is given below:

Algorithm 5.1 General form of genetic Algorithm [108]

Step1 Randomly generate an initial population M (0)

Step 2 Compute and save the fitness u(m) for each individual m in the current Population M(t)

Step3 Define selection probabilities p (m) for each Individual m in M (t) so that p(m) is proportional to u(m)

Step 4 Generate M (t+1) by probabilistically selecting individuals From M(t) to produce offspring via genetic operators

Step5 Repeat step 2 until satisfying solution is obtained.

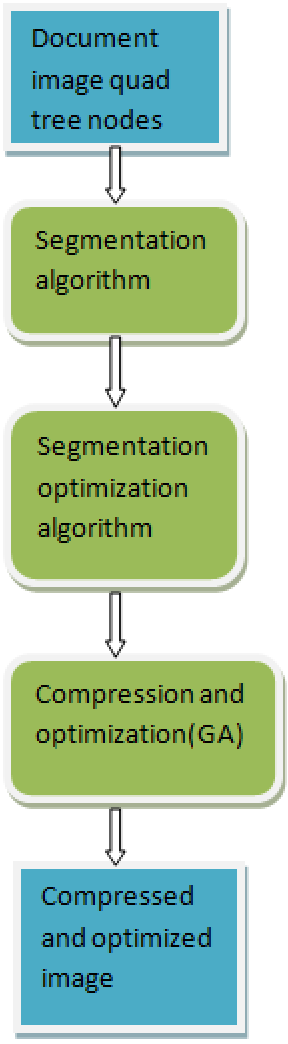


Fig 5.1: The Proposed Architecture.

Genetic Algorithms are adaptive search techniques that can learn high performance knowledge structures. The genetic algorithms' strength come from the implicitly parallel search of the solution space that it performs through a population of candidate solutions and this population is manipulated in the simulation. The candidate solutions represent every possible behavior of the robot and based on the overall performance of the candidates, each could be assigned a fitness value. Genetic operators could then be applied to improve the performance of the population of behaviors. One cycle of testing all of the competing behavior is defined as a generation, and is repeated until a good behavior is evolved. The good behavior is then applied to the real world. Also because of the nature of GA, the initial knowledge does not have to be very good. The usual steps involved in compressing an image are given below:

Algorithm 5.2 Image compression

Step1 Specify the Rate(intensities available) and Distortion (tolerable error) parameters for the target image.

Step 2 Divide the image data into classes, based on their importance.

Step 3 Divide the available intensities among these classes, such that the distortion is a minimum.

Step 4 Quantize each class separately using the intensity allocation information derived in step 3.

Step 5 Encode each class separately using an entropy coder and write to the file

Reconstructing the image from the compressed data is usually a faster process than compression. The steps involved are as follows:

Algorithm 5.3 Image reconstruction

Step 1 Read in the quantized data from the file, using entropy Encoder. (Reverse of step 5 in algorithm 5.2).

Step 2 Dequantize the data. (Reverse of step 4 in algorithm 5.2).

Step 3 Rebuild the image. (Reverse of step 2 in algorithm 5.2).

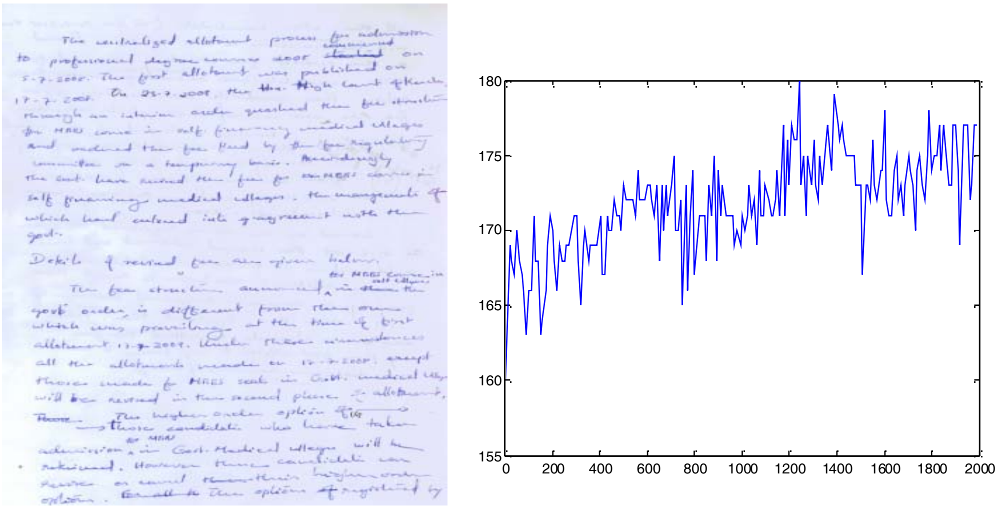
5.2.1. Compression Optimization using Genetic Algorithm and Parallel Genetic Algorithm

In order to compress the data with ordinary algorithms, it is difficult to deal with large two dimensional spaces of document images. The compression technique using genetic algorithms helps to improve the scanning procedure of document images and store the image in a compressed format at the scanner level itself. And most of the available algorithms are for ordinary images, not for document images. So this work is an attempt to use parallel genetic algorithm for optimization of the lossless compression of document images. Genetic Algorithms are a particular class of evolutionary search algorithms for global optimization. A genetic algorithm is designed by two components:

- Genetic representation of the problem domain
- A fitness function to the problem domain

5.2.1.1. Representation of the problem

A document is scanned and the image is stored as a "*.jpeg" file(fig 5.3). The image is represented as Quad Tree by recursively applying the method specified in Chapter 3 of this thesis for every quadrant:



(a) (b)

Fig 5.2 : (a) original hand written image and (b) The intensity plot of the image

Recursion stops when desired size of leaf node is obtained. The leaf node of the quad tree is a table with gray scale values. When a leaf node of 4×4, 8×8, 16×16 or 32×32 size is obtained, each matrix is considered as a ***population***. In a population a row is assigned as a ***chromosome***. A chromosome is formed with multiple or single genes. (Fig 5.4). When gene size increases the compression ratio also increases and the quality of reconstructed image decreases.

42	252	232	252	54	66	47	48
0	32	253	34	34	34	34	54
22	31	35	45	32	65	45	56
43	34	38	56	12	67	67	67
56	45	36	76	23	87	56	78
76	46	37	87	34	89	45	87
88	47	76	45	23	90	65	45
90	56	65	46	35	45	78	34

(a)

42	252	232	252	54	66	47	48

(b)

42	252	232	252

(c)

Fig 5.3 : a) A Sample population of size 8x8 b) A chromosome from the population of length 8 c) A gene of length 4

5.2.1.2. Designing the fitness function

In our method, the distance between two chromosomes X and Y are computed using Hausdorff Distance (d_H)

DISTANCE = $d_H(X, Y)$

$$d_H(X,Y) = \max\{\sup_{x\in X}\ {}_{y\in Y} \inf d(x,y),\ \sup_{y\in Y}\ {}_{x\in X} \inf d(x,y)\} \quad (5.1)$$

$$\text{FITNESS} = 1 - \text{DISTANCE}^{1/2} \quad (5.2)$$

Fitness is in the range 0 to 1.

Two chromosomes from a chosen population are compared and the fitness between the chromosomes is computed. The chromosome with highest fitness value is chosen.

5.2.1.3. The Selection Procedure

The population is selected based on the fitness of the individuals. The Tournament selection procedure is applied to select the best individual.

5.2.1.4. The Method and Operations

Mating and Crossover are applied with Constrained Run Length Encoding compression algorithm [1]. In this work multiple populations are considered simultaneously. Multiple chromosomes from each population are considered for achieving massive parallelism. A quantized index table is kept for deciding encoding index of the chromosomes. This index representation reduces the size of the encoded image.

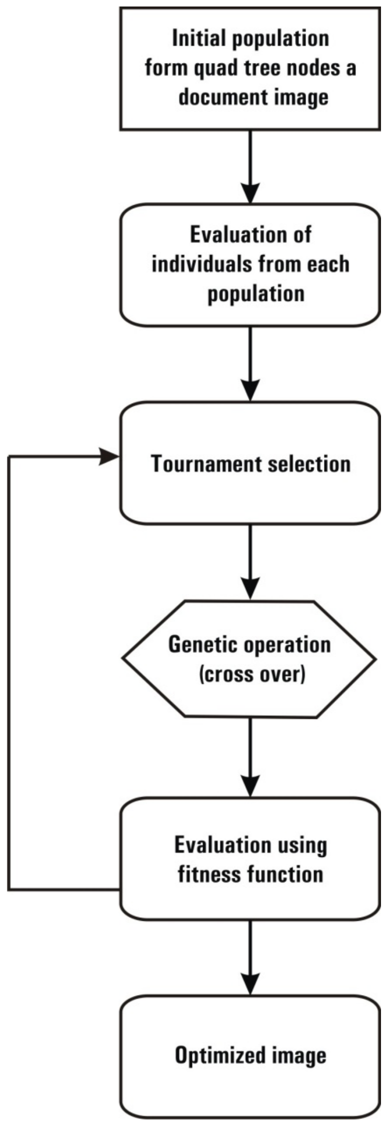


Fig 5.4 the sequential GA for compression optimization

The Algorithm 5.4 Compression Optimization

Step 1: Select n populations

Step 2: For each population select k chromosomes (repeat the following steps 3 through 6 until all chromosome and all populations are considered.)

Step 3: Apply a genetic operation

Step 4: Evaluate fitness using Hausdorff distance d_H (i.e. equation (5.1) and equation (5.2) as follows:

$$d_H(X,Y) = \max\{\sup_{x \in X} {}_{y \in Y} \inf d(x,y), \sup_{y \in Y} {}_{x \in X} \inf d(x,y)\}$$

$$\text{FITNESS} = 1 - \text{DISTANCE}^{1/2}$$

Step 5: Find out the chromosomes with highest fitness

Step 6: Get the index value of respected chromosome from index table. Store the index.

Step 7: stop;

Step 8: End

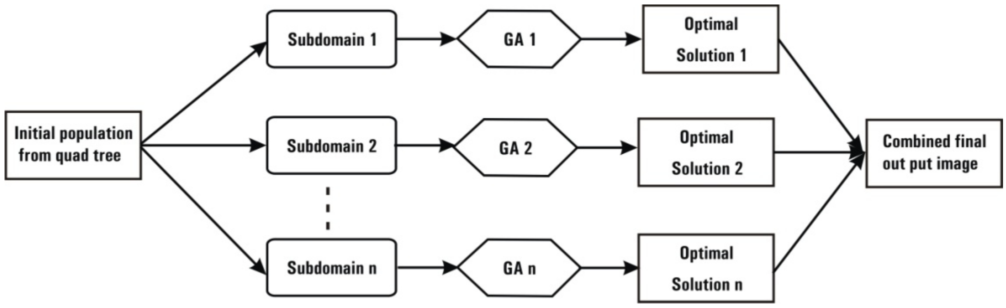


Fig 5.5 Parallel GA for Compression optimization

5.3 Analysis of the Work

The quality of the image is analyzed by the usefulness, naturalness of the image content. When one has to compare two images a fidelity value is used. The commonly used metrics are subjective measures and objective measures. One of the subjective measures is Mean Opinion Score (MOS). For large number of images it is not suitable. Objective measures are quantitative methods such as Peak Signal to Noise Ratio (PSNR), fidelity index Q, Mean Square error (MSE), etc.. PSNR in dB is calculated by using the following formula:

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{C^2}{\mathrm{MSE}}\right)$$

$$\mathrm{MSE} = \left(\frac{1}{mn}\right)\sum_{i=1}^{n}\sum_{j=1}^{n}\big(f(i,j) - g(i,j)\big)^2 \qquad \text{.......(5.3)}$$

Where mn is the size of the image and C represents the maximum color value for images. In our problem the maximum color value of the image is 255. ‘f" and ‘g’ are domain and range of images. Warry and Bovik in 2001 defined a fidelity index Q in terms of the linear correlation coefficient and the similarities between the mean and variance of two images [67]. This index is defined as

$$Q = \frac{4.\mu_x \mu_y .\sigma_{xy}}{\left(\mu_x^{\;2} + \mu_x^{\;2}\right)\left(\sigma_x^{\;2} + \sigma_y^{\;2}\right)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(5.4)$$

Where x and y are the original and tested images respectively, μ_x and μ_y are their means and σ_x and σ_y are their variances and σ_{xy} the Correlation. They defined the range Q as [-1, 1]. The value of 1 happens when the images are same (i.e. $y_i = x_i$ for all $_i$).

Table 5.1 Sample Compression Ratios achieved for Ordinary GA and Parallel GA

Sl no	Name of image	Size	PSNR dB	Compression ratio	Time in sec for ordinary GA	Time in sec for parallel GA
1	Doc001	1566x1102x3	27.28	5:1	273	162
2	Doc002	1346x1124x3	26.38	4:1	251	148
3	Doc003	1324x1096x3	26.34	4:1	247	146
4	Doc004	1545x1134x3	27.04	4:1	248	175
5	Doc005	1700x1044x3	26.06	5:1	245	142
6	Doc006	1376x1120x3	26.25	4:1	249	147
7	Doc007	1406x1224x3	27.53	4:1	273	169
8	Doc008	1488x1136x3	26.47	4:1	264	164

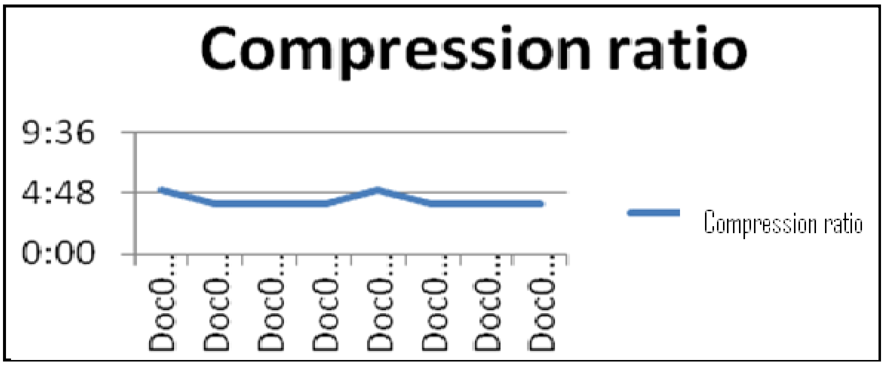


Fig: 5.6 Compression Ratio achieved for different test images

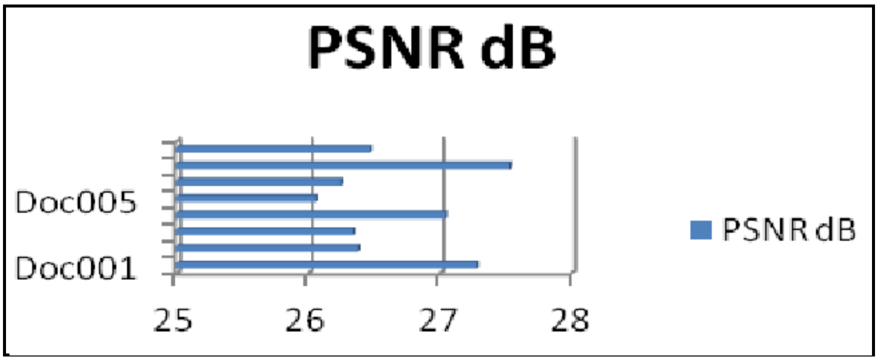


Fig 5.7 A Comparison of PSNR

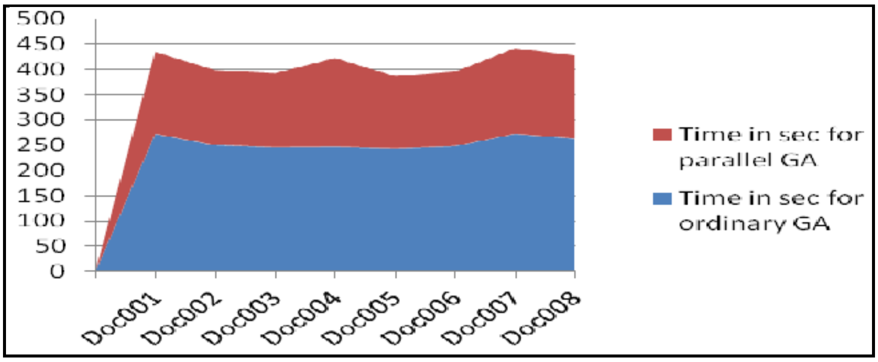


Fig 5.8 Area graph for Ordinary GA and Parallel GA

5.4 Conclusion

Compression of Document images is still relevant because transmission through internet suffers the bottle neck of bandwidth. There are various compression techniques such as run length encoding, Huffman and delta Huffman coding, fuzzy Hopfield Neural Net based coding etc. We have applied the algorithms to "*.jpeg" image. We have developed and used sequential Genetic Algorithm and parallel Genetic Algorithm for Compression Optimization. A fitness function is used with Hausdorff distance. For achieving the compression ratio both the algorithms performed well. In the case of execution time Parallel Genetic Algorithm Dominates Sequential Genetic Algorithm. Different methods for Document image compression are compared and found that optimization with Evolutionary Algorithms show better performance over other algorithms. Depending upon the nature of image the compression ratio achieved is varied from 4:1 to 5:1. The most important point to note is that this ratio is achieved over the already compressed JPEG image of documents. This is promising for document images. Consistently good PSNR of value around 27 indicates that the quality of the output is also good. When optimization with evolutionary algorithms is performed, background elimination of document images takes place effectively and it helps to overcome the bandwidth limitations of internet while transmitting huge amounts of document images.

………………

Chapter 6

SUMMARY AND CONCLUSION

Contents

6.1Introduction

Document Image Processing is very relevant in present scenario because large amount of paper documents needs digitization of data. Scanned images of documents are stored as different formats like *.jpeg, *.tiff etc. These images are processed in Document Image Processing. Document Image Segmentation corresponds to identification of different regions and objects in a document which humans can identify easily. Computers have no intelligent means for identifying objects and feature extraction. So we need to provide best algorithms for the same. Many such methods are already developed. Our attempt was plane wise segmentation of Document Images using Interval Valued fuzzy set methods. In the era of Internet Document Image Compression is relevant because poor band width is a constraint to transmit huge size of document images. So we have done the optimization of Document Image Segmentation and compression by properly layering the document Images, eliminating the back ground, by giving new algorithms etc.

6.2 Summary

For this work we have prepared a database of Document Images which consists of 200 documents which includes document images without any pictures or photo graphs, with images and photographs, hand written document images etc. A literature survey is done with more than 300 International papers on Document Imaging and selected about 125 papers which are closely related to Document Image Segmentation and Compression. Most of the papers have dealt with binary images, so we tried on colour images and gray scale images. Colour images are three dimensional images of red, green and blue intensities. As a first step we have done the plane wise segmentation of document images. Document image consists of various features. The planes refer the features of the document images. Layer wise isolation is done with the help of Interval Valued fuzzy sets. The optimization of segmentation is done with Simulated annealing which helped to avoid biasing towards a particular interval. Compression optimizations of Document Images were done with sequential and parallel genetic algorithm. A fitness function with Hausdorff Distance is used for the purpose.

6.2.1 Theoretical contributions of the work

In this work we have developed three different algorithms for segmentation, segmentation optimization and compression optimization. The work is done on *.jpeg files which are already lossy compressed files. So improvements over compressed files are achieved. Papers on related topic of our work are peer reviewed by experts and presented/ published in International conferences.

6.2.2 Social Contributions of the work

The work can be utilized for hardware level implementation in equipments like scanners, facsimile etc. and can be used for internet based document image transmission. The algorithms used can be extended to the fields like medical imaging, agriculture related imaging, satellite imaging etc.

6.3 Major Findings

A study of developments in the area of Document Image Processing and in fuzzy till the current year is performed and recorded in tabular format. Quad tree generation is done successfully. Fuzzy relation is used to store fuzzy intervals, which helps to store the required interval only and saves space even for JPEG images. Region wise segmentation of document images is done. This helps to identify the foreground from the background and in turn helps to eliminate the background for compression. Optimization of image segmentation is done with Simulated Annealing algorithm. Back propagation algorithm or other neural network algorithm shows convergence at local minimum. Simulated annealing rectifies the defects. The gradient factor for the sigmoid function is adjusted as 1 and the iteration converged about 10000 epochs. The probability determined the thresholds. When foreground and background changed, the threshold is also changed. The algorithm shows best performance when the coefficient of Temperature is adjusted in between 0.8 and 0.9. Median text line is identified by checking the intensities. Connected components of the text line are also identified. This algorithm shows better performance over traditional methods. As the next step Optimization of compression algorithm is done. We have used Hausdorff Distance based fitness function for the optimization algorithm. Depending upon the nature of image the compression technique ratio achieved is varied from 4:1 to 5:1. This is promising for document images because this ratio is achieved over already compressed JPEG document image. We have applied sequential and parallel genetic algorithm for compression optimization. The time conception was very less for parallel Genetic Algorithm. When optimization with evolutionary algorithms is performed, background elimination of document images takes place effectively and this helps to overcome the bandwidth limitations of internet while transmitting huge amounts of document images.

6.4 Suggestions

The work can be extended with text pattern recognition for each language or object recognition from the document image may be performed. Historical documents can be digitized and compressed. If text in one language is extracted then translation can be done. The area is very vast. And sky is the limit.

6.5 Conclusion

Objectives of this work are satisfied. Analysis of the work is carried out at every stage and outputs are displayed. Document image processing is still relevant because many documents are to be digitized. And historical documents (manuscripts in thaliyola (palm leaf), thamra patra (copper plate etc)) should be preserved digitally for the coming generations. Compression of documents helps to reduce the bottleneck of internet bandwidth. It also helps for facsimile transmission. So the work is still relevant in today's scenario. The work shows considerable performance compared to traditional algorithms.

………………

REFERENCES

1. Flavio Bortolozzi, and Robert Sabourin. "Binary image compression using identity mapping backpropagation neural network." *Electronic Imaging '97*. International Society for Optics and Photonics, 1997.

2. Feng, Xiaojuan, Christopher K. I. Williams, and Stephen N. Felderhof. "Combining belief networks and neural networks for scene segmentation." *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 24.4 (2002): 467-483.

3. Bedregal, Benjamín Callejas, Graçaliz Pereira Dimuro, and Antônio Carlos da Rocha Costa. "Hand gesture recognition in an interval fuzzy approach.", Tendências em Matemática Aplicada e Computacional, v. 8, p. 21-31, 2007.

4. Lee, Seong-Whan, and Dae-Seok Ryu. "Parameter-free geometric document layout analysis." *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 23.11 (2001): 1240-1256.

5. Hutchison, Luke Alexander Daysh, and William A. Barrett. "Fast registration of tabular document images using the fourier-mellin transform." *Document Image Analysis for Libraries, 2004. Proceedings. First International Workshop on*. IEEE, 2004.

6. Saykol, Ediz, et al. "Content-based retrieval of historical Ottoman documents stored as textual images." *IEEE Transactions on Image Processing,* 13.3 (2004): 314-325.

7. Simard, Patrice Y., et al. "A foreground-background separation algorithm for image compression." *Data Compression Conference, 2004. Proceedings. DCC 2004*. IEEE, 2004.
8. Suzuki, A., N. Ito, and K. Arakawa. "Kanji recognition in scene images without detection of text fields-robust against variation of viewpoint, contrast, and background texture." *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*. Vol. 1. IEEE, 2004.
9. Séropian, A., M. Grimaldi, and Pr N. Vincent. "Differentiation of alphabets in handwritten texts." *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.* Vol. 2. IEEE, 2004.
10. Celik, Mehmet Utku, et al. "Lossless generalized-LSB data embedding." *IEEE Transactions on Image Processing,* 14.2 (2005): 253-266.
11. Wu, B-F., C-C. Chiu, and Y-L. Chen. "Algorithms for compressing compound document images with large text/background overlap." *IEE Proceedings-Vision, Image and Signal Processing* 151.6 (2004): 453-459.
12. Pei, Soo-Chang, and Yu-Ting Chuang. "Automatic text detection using multi-layer color quantization in complex color images." *IEEE International Conference on Multimedia and Expo, 2004. ICME'04. 2004*. Vol. 1. IEEE, 2004.

13. Yang, Huijuan, and Alex C. Kot. "Text document authentication by integrating inter character and word spaces watermarking." *IEEE International Conference on Multimedia and Expo, 2004. ICME'04. 2004*. Vol. 2. IEEE, 2004.

14. Pavlidis, George, Sofia Tsekeridou, and Christodoulos Chamzas. "JPEG-matched data filling of sparse images." *International Conference on Image Processing, 2004. ICIP'04. 2004* Vol. 1. IEEE, 2004.

15. Toshniwal, Mayank. "Satellite Image Classification Using Neural Networks." *Inter Conf*. 2005.

16. Chowdhury, S. P., S. Mandal, A. K. Das, and Bhabatosh Chanda. "Automated segmentation of math-zones from document images." In *7th International Conference on Document Analysis and Recognition*, vol. 2, pp. 755-759. 2003.

17. Kumar, KS Sesh, Anoop Namboodiri, and C. V. Jawahar. "Learning to segment document images." *Pattern Recognition and Machine Intelligence*. Springer Berlin Heidelberg, 2005. 471-476.

18. Vences, Lucıa, and Isaac Rudomin. "Genetic algorithms for fractal image and image sequence compression." *Proceedings Computacion Visual* (1997): 35-44.

19. Razavi, Abbas, Rutie Adar, Isaac Shenberg, Rafi Retter and Rami Friedlander, "VLSI implementation of an image compression algorithm with a new bit rate control capability." *IEEE International*

Conference on Acoustics, Speech, and Signal Processing, 1992. ICASSP-92., 1992. Vol. 5. IEEE, 1992.

20. Yeh, Pen-Shu, Gilles A. Moury, and Philippe Armbruster. "CCSDS data compression recommendation: development and status." *International Symposium on Optical Science and Technology*. International Society for Optics and Photonics, 2002.
21. Cook, Gregory W., and Edward J. Delp. "The use of high performance computing in JPEG image compression." *Signals, Systems and Computers, 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*. IEEE, 1993.
22. Gelenbe, Erol, and Mert Sungur. "Random network learning and image compression." *Neural Networks, 1994. IEEE World Congress on Computational Intelligence., 1994 IEEE International Conference on*. Vol. 6. IEEE, 1994.
23. Uhl, Andreas. "Digital image compression using wavelets and wavelet packets based on nonstationary and inhomogeneous multiresolution analyses." *SPIE's 1994 International Symposium on Optics, Imaging, and Instrumentation*. International Society for Optics and Photonics, 1994.
24. Venkataraman, Shankar, and Bernard C. Levy. "Nonseparable orthogonal linear phase perfect reconstruction filter banks and their application to image compression." *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*. Vol. 3. IEEE, 1994.

25. Hong G, G Hall and T J Tenell, “Joint entropy and multiband prediction for lossless image compression”, Image Processing and its Applications 1995, Fifth International Conference on IET, 1995.

26. Jiang, J. "Design of a LVQ neural network for compressed image indexing." *IEE conference publication*. Institution of Electrical Engineers, 1997.

27. Albanesi, M. G., and S. Bertoluzza. "Human vision model and wavelets for high-quality image compression." *Image Processing and its Applications, 1995., Fifth International Conference on*. IET, 1995.

28. Storer, James A., and Harald Helfgott. "Lossless image compression by block matching." *The Computer Journal* 40.2 and 3 (1997): 137-145.

29. Lu, Shijian, Ben M. Chen, and C. C. Ko. "Perspective rectification of document images using fuzzy set and morphological operations." *Image and Vision Computing* 23.5 (2005): 541-553.

30. Inglis, Stuart, and Ian H. Witten. *Bi-level document image compression using layout information*. Department of Computer Science, University of Waikato, 1996.

31. Kiselyov, Oleg, and Paul Fisher. "Image compression with iterated function systems, finite automata and zerotrees: Grand unification." *arXiv preprint cs/0003065* (2000).

32. Topiwala, Pankaj N. "Human-vision-based wavelet image coding." *Optical Science, Engineering and Instrumentation'97*. International Society for Optics and Photonics, 1997.

33. Bandy, C. P., and D. B. Koch. "A low bandwidth, still image transmission system." *Southeastcon'95.'Visualize the Future'., Proceedings., IEEE*. IEEE, 1995.

34. Moravie, P., et al. "Real-time image compression using SIMD architectures." *Computer Architectures for Machine Perception, 1995. Proceedings. CAMP'95*. IEEE, 1995.

35. Jackson, David Jeff, and Sidney Jwl Hannah. "Comparative analysis of image compression techniques." *System Theory, 1993. Proceedings SSST'93., Twenty-Fifth Southeastern Symposium on*. IEEE, 1993.

36. Kil, David H., and Frances Bongjoo Shin. "Reduced dimension image compression for remotely distributed underwater signal processing." *OCEANS'95. MTS/IEEE. Challenges of Our Changing Global Environment. Conference Proceedings.*. Vol. 2. IEEE, 1995.

37. Sankar Moni and R L Kashyap," A multi Resolution Stochastic Process for Image Compression", Acoustics, Speech and Signal Processing 1996 ICASSP-96. Conference Proceedings, 1996 IEEE International Conference1954-1957.

38. Venetianer, Peter L., and Tamas Roska. "Image compression by CNN." *Neural Networks, 1996., IEEE International Conference on*. Vol. 3. IEEE, 1996.

39. Saupe, Dietmar, and Hannes Hartenstein. "Lossless acceleration of fractal image compression by fast convolution." *Image Processing, 1996. Proceedings., International Conference on*. Vol. 1. IEEE, 1996.

40. Tian, Jinwen, Su Kang, Jian Liu, and Qian Gao. "A novel image compression encoding." *Signal Processing, 1996., 3rd International Conference on*, vol. 2, pp. 982-985. IEEE, 1996.

41. Shin, F. B., D. H. Kil, and G. J. Dobeck. "The impact of lossy image compression on automatic target recognition performance." *OCEANS'96. MTS/IEEE. Prospects for the 21st Century. Conference Proceedings*. Vol. 2. IEEE, 1996.

42. Li, Jiaming, and Jesse S. Jin. "An image coding method of high compression rate and clarity preserving." *Intelligent Information Systems, 1996., Australian and New Zealand Conference on*. IEEE, 1996.

43. Alirezaie, Javad, and M. E. Jernigan. "Medical image compression using B-spline wavelet transform." *Engineering in Medicine and Biology Society, 1995., IEEE 17th Annual Conference*. Vol. 1. IEEE, 1995.

44. Moni, Shankar, and S. Sista. "A framework for application specific image compression." *Data Compression Conference, 1997. DCC'97. Proceedings*. IEEE, 1997.

45. Said Amir, William A Pearl man, “Low Complexity wave form coding via Alphabet and Sample set Partitioning”, IEEE International Symposium on Information Theory, Institute of Electrical Engineers INC(IEE), 1997.

46. Jiang, Wen, Xiaolin Wu, and Wai Yin Ng. "VQ index coding for high-fidelity medical image compression." *Image Processing, 1997. Proceedings., International Conference on*. Vol. 3. IEEE, 1997.

47. Meng, Hongying, Zhihua Wang, and Li Zhang. "Performance analysis of the cdf wavelet filters in image compression." *Circuits and Systems, 2000. IEEE APCCAS 2000. The 2000 IEEE Asia-Pacific Conference on*. IEEE, 2000.

48. Reavy, Maire D., and Charles G. Boncelet. "BACIC: a new method for lossless bi-level and grayscale image compression." *Image Processing, 1997. Proceedings., International Conference on*. Vol. 2. IEEE, 1997.

49. Grosse, H-J., et al. "Sub-block classification using a neural network for adaptive zigzag reordering in JPEG-like image compression scheme." *Neural and Fuzzy Systems: Design, Hardware and Applications (Digest No: 1997/133), IEE Colloquium on*. IET, 1997.

50. Cazuguel, G., et al. "Medical image compression and characterization using vector quantization: an application of self-organizing maps and quadtree decomposition." *Information Technology Applications in Biomedicine, 1998. ITAB 98. Proceedings. 1998 IEEE International Conference on*. IEEE, 1998.

51. Penghui, Wu, and Zheng Baoyu. "A new image compression method based on HV fractal and DCT." *Communication Technology Proceedings, 1998. ICCT'98. 1998 International Conference on*. IEEE, 1998.

52. Tham, Jo Yew, et al. "A new multifilter design property for multiwavelet image compression." *Acoustics, Speech, and Signal Processing, 1999. Proceedings., 1999 IEEE International Conference on*. Vol. 3. IEEE, 1999.

53. Steil, Jochen J. "Online reservoir adaptation by intrinsic plasticity for backpropagation-decorrelation and echo state learning." *Neural Networks* 20.3 (2007): 353-364.

54. Cazuguel, Guy, et al. "Improving spatial vector quantization for image compression by use of a quadtree scheme. Application to echoendoscopic image compression." *Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE*. Vol. 2. IEEE.

55. Frydrych, M., and P. J. Toivanen. "Arithmetic coding with sliding window for control-point based image compression." Image processing and its applications- Conference Publication IEE(1999): 601-604.

56. Signoroni, A., et al. "Design of a visual environment for evaluating and customizing medical image compression techniques." *Image Analysis and Processing, 1999. Proceedings. International Conference on*. IEEE, 1999.

57. Phimoltares, S., K. Chamnongthai, and C. Lursinsap. "Hybrid binary image compression." *Signal Processing and Its Applications, 1999. ISSPA'99. Proceedings of the Fifth International Symposium on*. Vol. 2. IEEE, 1999.

58. Chang, Yung-Ching, et al. "Adaptive post-processing for region-based fractal image compression." *Data Compression Conference, 2000. Proceedings. DCC 2000*. IEEE, 2000.

59. Kuo, Chin-Hwa, Tzu-Chuan Chou, and Tay-Shen Wang. "An efficient spatial prediction-based image compression scheme." *Circuits and Systems for Video Technology, IEEE Transactions on* 12.10 (2002): 850-856.

60. Vlajic, Natalija, Howard C. Card, and Thomas Kunz. "Image-compression for wireless World Wide Web browsing: a neural network approach." *Neural Networks, 2000. IJCNN 2000, Proceedings of the IEEE-INNS-ENNS International Joint Conference on*. Vol. 1. IEEE, 2000.

61. Chung, Kuo-Liang, and Jhin-Sian Lin. "Faster and more robust point symmetry-based K-means algorithm." *Pattern Recognition* 40.2 (2007): 410-422.

62. Hsieh, Ching-Tang, et al. "Embedded image compression using the important-oriented-tree." *Multimedia Software Engineering, 2000. Proceedings. International Symposium on*. IEEE, 2000.

63. Wu, Pou-Yah. "Distributed fractal image compression on PVM for million-pixel images." *Information Networking, 2001. Proceedings. 15th International Conference on*. IEEE, 2001.

64. Yu, Tian-Hu, Zhihai He, and Sanjit K. Mitra. "Novel coding scheme for wavelet image compression." *Signals, Systems and Computers, 2000. Conference Record of the Thirty-Fourth Asilomar Conference on*. Vol. 2. IEEE, 2000.

65. Fisher, Yuval, E. W. Jacobs, and R. D. Boss. "Fractal image compression using iterated transforms." *Image and text compression*. Springer US, 1992. 35-61.

66. Mello, Carlos AB. "Synthesis of Images of Historical Documents for Web Visualization." *Multimedia Modelling Conference, 2004. Proceedings. 10th International*. IEEE, 2004.

67. Qin, A. K., and David A. Clausi. "Multivariate image segmentation using semantic region growing with adaptive edge penalty." *Image Processing, IEEE Transactions on* 19.8 (2010): 2157-2170.

68. Lettner, Martin, and Robert Sablatnig. "Multispectral imaging for analyzing ancient manuscripts." *Proceedings of the 17th European Signal Processing Conference (EURASIP, 2009)*. 2009.

69. Russell, Stuart Jonathan, et al. *Artificial intelligence: a modern approach*. Vol. 74. Englewood Cliffs: Prentice hall, 1995.

70. Eklund, Sven E. "A massively parallel architecture for distributed genetic algorithms." *Parallel Computing* 30.5 (2004): 647-676.

71. Seth Bacon, " A brief Overview of Parallel Genetic Algorithms" http://www.cs.iastate.edu/~cs425/reports/bacon-report.pdf

72. Fan, Yong, Tianzi Jiang, and David J. Evans. "Volumetric segmentation of brain images using parallel genetic algorithms." *Medical Imaging, IEEE Transactions on* 21.8 (2002): 904-909.

73. Yu, Qizhi, Chongcheng Chen, and Zhigeng Pan. "Parallel genetic algorithms on programmable graphics hardware." *Advances in Natural Computation*. Springer Berlin Heidelberg, 2005. 1051-1059.

74. Mantas Paulinas, Andrius Usinskas, " A Survey of Genetic Algorithm Application for Image Enhancement and Segmentation",ISSN 1392-124x *Information Technology and Control,* 2007, vol,36, No:3.

75. Barta, András, and István Vajk. "Document image analysis by probabilistic network and circuit diagram extraction." *INFORMATICA-LJUBLJANA-* 29.3 (2005): 291.

76. Otobe, Kazunori, Kei Tanaka, and M. A. S. A. Y. U. K. I. Hirafuji. "Knowledge acquisition on image processing based on genetic algorithms." *Proceedings of the IASTED International Conference on Signal and Image Processing*. 1998.

77. Alba, Enrique, Francisco Luna, and ANTONIO J. Nebro. "Advances in parallel heterogeneous genetic algorithms for continuous optimization." *International Journal of Applied Mathematics and Computer Science* 14.3 (2004): 317-334.

78. Ghosh, Susmita, Ashish Ghosh, and Sankar K. Pal. "Incorporating ancestors' influence in genetic algorithms." *Applied Intelligence* 18.1 (2003): 7-25.

79. Lim, Dudy, et al. "Efficient hierarchical parallel genetic algorithms using grid computing." *Future Generation Computer Systems* 23.4 (2007): 658-670.

80. Nedunchelian, Ramanujam, et al. "Dynamic task scheduling using parallel genetic algorithms for heterogeneous distributed

computing." *Proceedings of the 2006 International Conference on Grid, Las Vegas, Nevada, USA*. 2006.

81. Otoba, K., K. Tanaka, and M. Hitafuji. "Image processing and interactive selection with Java based on genetic algorithms." *Proceedings of the 3rd IFAC/CIGR Workshop on Artificial Intelligence in Agriculture*. 1998.
82. Cantú-Paz, Erick. "A survey of parallel genetic algorithms." *Calculateurs paralleles, reseaux et systems repartis* 10.2 (1998): 141-171.
83. Gonzalez, Rafael C., Richard E. Woods, and Steven L. Eddins. *Digital image processing using MATLAB*. Vol. 2. Tennessee: Gatesmark Publishing, 2009.
84. Rajasekaran, Sanguthevar, and GA Vijayalakshmi Pai. *Neural networks, fuzzy logic and genetic algorithms: synthesis and applications*. PHI Learning Pvt. Ltd., 2004.
85. Bunke, Horst, and Patrick Shen-pei Wang. *Handbook of character recognition and document image analysis*. World Scientific Publishing Company, 1997.
86. Choi, Hyeokho, and Richard G. Baraniuk. "Multiscale image segmentation using wavelet-domain hidden Markov models." *Image Processing, IEEE Transactions on* 10.9 (2001): 1309-1321.
87. Küçüktunç, Onur, Uğur Güdükbay, and Özgür Ulusoy. "Fuzzy color histogram-based video segmentation." *Computer Vision and Image Understanding* 114.1 (2010): 125-134.

88. Kaya, Metin. "Image Clustering and Compression Using An Annealed Fuzzy Hopfield Neural Network." *International Journal of Signal Processing* 1.2 (2005).

89. Khiyal, Malik Sikandar Hayat, Aihab Khan, and Amna Bibi. "Modified watershed algorithm for segmentation of 2D images." *Issues in Informing Science and Information Technology* 6 (2009): 877-886.

90. Cheng, Hui. *Document image segmentation and compression*. Diss. Purdue university, 1999.

91. Gordo, Oscar, et al. "Classification of Satellite Images by means of Fuzzy Rules generated by a Genetic Algorithm." *Latin America Transactions, IEEE (Revista IEEE America Latina)* 9.1 (2011): 743-748.

92. Fu, Dongmei, et al. "A novel immune image template set for fuzzy image segmentation and its application research." *Telecommunications and Signal Processing (TSP), 2011 34th International Conference on.* IEEE, 2011.

93. Rajini, N. Hema, and R. Bhavani. "Enhancing k-means and kernelized fuzzy c-means clustering with cluster center initialization in segmenting MRI brain images." *Electronics Computer Technology (ICECT), 2011 3rd International Conference on.* Vol. 2. IEEE, 2011.

94. Yuan, J. I., et al. "Research on Object Shape Detection from Image with High-Level Noise Based on Fuzzy Generalized Hough

Transform." *Multimedia and Signal Processing (CMSP), 2011 International Conference on*. Vol. 1. IEEE, 2011.

95. Ganta, Raghotham Reddy, et al. "Robust level set image segmentation based on modified fuzzy clustering." *Multimedia, Signal Processing and Communication Technologies (IMPACT), 2011 International Conference on*. IEEE, 2011.

96. Wang, Yanhong, Denghui Li, and Yaning Xu. "An improved image edge detection algorithm based on fuzzy sets." *Multimedia Technology (ICMT), 2011 International Conference on*. IEEE, 2011.

97. Xiao, Qinghan. "Using fuzzy adaptive fusion in face detection." *Computational Intelligence in Biometrics and Identity Management (CIBIM), 2011 IEEE Workshop on*. IEEE, 2011.

98. Guo Xian Jiu, Jiang Fen Jiao, Lixiang," Image enhancement method based on fuzzy set and sub division", ICAST 2011, page 174-176.

99. Rattanalappaiboon. S, Rtthipravat. P, "Fuzzy based feature matching for 3D re-construction from multiple waves", SICE Annual Conference 2012, pages 768-772

100. Askari. E, Cftekhari. A. M," Learning based objective evaluation of image segmentation Algorithm", IET Conference (IPR 2012 page 1-6)

101. Parvathi, Rangasamy, and Vassia Atanassova. "Intuitionistic fuzzy statistical tools for filters in image processing." *Intelligent Systems (IS), 2012 6th IEEE International Conference*. IEEE, 2012.

102. Miguel Pagola, Carloslopez-Molina, Javie Fernandez, Edurna. Barronechea and Humbertobustince," Interval type-2 fuzzy sets

constructed from several membership functions' application to the fuzzy thresholding" IEEE Transactions on fuzzy Algorithm, issue 99

103. Qiao, Jian, Naiding Yang, and Jie Gao. "Two-stage fuzzy logic controller for signalized intersection." *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on* 41.1 (2011): 178-184.

104. Reyna, Roberto A., et al. "Implementation of the SVM neural network generalization function for image processing." *Computer Architectures for Machine Perception, 2000. Proceedings. Fifth IEEE International Workshop on*. IEEE, 2000.

105. Gabbouj, Moncef. "Multidimensional particle swarm optimization and applications in data clustering and image retrieval." *Image Processing Theory Tools and Applications (IPTA), 2010 2nd International Conference on*. IEEE, 2010.

106. Veasicele. Z, Biddo. M," Evolutionary design of robust noise specific image filters", CEE IEEE congress 2011 pages 269-276.

107. Nakano, Yuta, et al. "Automatic construction of image transformation algorithms using feature based genetic image network." *Evolutionary Computation (CEC), 2010 IEEE Congress on*. IEEE, 2010.

108. Shirakawa, Shinichi, and Tomoharu Nagao. "Evolutionary image segmentation based on multiobjective clustering." *Evolutionary Computation, 2009. CEC'09. IEEE Congress on*. IEEE, 2009.

109. Gulum, Taylan Ozgur, Ahmet Yasin Erdogan, and Tulay Yildirim. "An edge detection technique using hybrid Ant Colony Optimization-

genetic algorithm."*Signal Processing and Communications Applications Conference (SIU), 2012 20th*. IEEE, 2012.

110. de Medeiros Napoles, S. H. L., and C. Zanchettin. "Offline handwritten signature verification through network radial basis functions optimized by Differential Evolution." *Neural Networks (IJCNN), The 2012 International Joint Conference on*. IEEE, 2012.

111. Khangar, Smita. V, Bharne, Dinesh. P, Malik, Latesh. G," Compression strategy for hand written gray level document images", ICCICT 2012 pages 1-5

112. Ritchey, Philip C., and Vernon J. Rego. "Hiding Secret Messages in Huffman Trees." *Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2012 Eighth International Conference on*. IEEE, 2012.

113. Zeng, Wenyi, and Yibin Zhao. "IV-similarity measure and IV-entropy of interval valued fuzzy sets." *Fuzzy Systems and Knowledge Discovery (FSKD), 2010 Seventh International Conference on*. Vol. 1. IEEE, 2010.

114. Jurio. A, Parnain. D, Lopez-Molina. C, Bustince. H, Mesian. R, Beliakev. G," A construction methord of interval valued fuzzy sets for image processing", advances in fuzzy logic systems(T2FVZZ), 2011, IEEE Symposium, page 16-22

115. Yingzhang," An improved T-S fuzzy neural network and its application in soft sensing for FCCU", Proceedings of the 8th world

congress on intelligent control and automation july 6-9, 2010 Jinan, China.

116. Fure-Ching Jeng, John W Woods, " Simulated Annealing in Compound Guassian Random fields", IEEE Transactions on Information Theory, vol 36, No:1, January 1990.
117. Abdelhakim Safir, Bertrand Zavidovique, " On the Synthesis of specific Image Processing Automata by a Simulated Annealing based design space Search", IEEE, ISCAS, 1989.
118. Grzeszczuk, Robert P., and David N. Levin. ""Brownian strings": Segmenting images with stochastically deformable contours." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 19.10 (1997): 1100-1114.
119. Alexander, S. K. "Image sampling by hierarchical annealing." *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*. Vol. 1. IEEE, 2003.
120. See, Eric SK. "A new novel PET/CT reconstruction algorithm by using prior image model with simulated annealing process." *Intelligent Signal Processing and Communication Systems, 2005. ISPACS 2005. Proceedings of 2005 International Symposium on*. IEEE, 2005.
121. Yagouni, Mohammed. "Using of meta heuristics for optimizing satellite image registration." *Computers & Industrial Engineering, 2009. CIE 2009. International Conference on*. IEEE, 2009.

122. Mohamadi, Hossein, Ali Nodehi, and Mohamad Tayarani. "A local search operator in Quantum Evolutionary Algorithm and its application in Fractal Image Compression." *Computer and Automation Engineering (ICCAE), 2010 The 2nd International Conference on*. Vol. 2. IEEE, 2010.

123. Huang, Kou-Yuan, and Yueh-Hsun Hsieh. "Very fast simulated annealing for pattern detection and seismic applications." *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*. IEEE, 2011.

124. Rafael C Gonzalez and Richard E Woods, "Digital Image Processing", Pearson Education. 2004

125. Bustince, H., E. Barrenechea, M. Pagola, and J. Fernandez. "Interval-valued fuzzy sets constructed from matrices: Application to edge detection." *Fuzzy Sets and Systems* 160, no. 13 (2009): 1819-1840.

………………

LIST OF PUBLICATIONS

1. Aysha V, Dr S Babu Sundar "Digital Image Processing and Artificial Neural Network Approaches", National Seminar on Image Processing and Cryptography at Sree Saraswathy Thyagaraja College, Pollachi, 8th March 2008.
2. Aysha V, " Segmentation of Document Images for Identifying or Eliminating Background from text or pictures embedded in the text", 21st Kerala Science Congress Kollam, 28-31 January 2009.
3. Aysha V, Dr Kannan Balakrishnan, Dr S Babu Sundar, " Image Segmentation using Interval Valued Adaptive Neuro Fuzzy Method", International Conference on Mathematical Computing and Management 2010.
4. Aysha V, Dr Kannan Balakrishnan, Dr S Babu Sundar, "Optimization of Document Image Segmentation Using Simulated Annealing", International Conference on Mathematical Computing and Management 2010.
5. Aysha V, Dr Kannan Balakrishnan, Dr S Babu Sundar, Dr Pramod K V, " Document Image Compression Optimization using Genetic algorithm", International Conference on Mathematical Computing and Management 2010.
6. Aysha V, Dr Kannan Balakrishnan, Dr S Babu Sundar, "Parallel Genetic Algorithm for Document Image Compression Optimization" ICEIE 2010, Kyoto Japan

………………