بسم الله الرحمن الرحيم

Sudan University of Science and Technology

Collage of Graduates studies

Biomedical Engineering Department

MEDICAL IMAGE COMPRESSION USING HUFFMAN AND ARITHMETIC CODE ALOGARITHMS FOR TELEMEDICINE

Presented by:
Nidal Mohammed Abd Algader
Supervisor:
Dr. Mohammed Yagoub

Acknowledgments

After thanks Allah to give me power and patience to finishing this research I would like to thank my parents; Dr . Mohammed Yagoub for his help and support in this research; engineer Ahmed Medani for his time and help ; and thanks brothers , my friends for support for finishing this research .

المستخلص:

تعتبر الصور الطبية من اهم الوسائل التشخيصية في المجال الطبي ،وتستخدم هذه الصور في مجال التطبيب عن بعد كما يتم ارشفتها للاستفادة منها لاحقا . هنالك عدد من الانظمة القياسية المستخدمة في مجال الصور الطبية مثل ((PACS , DICOM تتميز الصور الطبية بضخامة حجمها حيث انها تحتاج الي تقليل حجمها عند ارسالها او تخزينها حيث انه عند تقليل الحجم نقلل بذلك حجم حزمة البيانات المستخدمة كما يقلل المساحة المستخدمة لتخزين هذه الصور . في هذا البحث تم انشاء خوارزميتين لضغط الصور الطبية وفك الضغط بطريقة فعالة بحيث لاتفقد اي بيانات من الصورة الاصلية تم استخدام طريقة (Huffman) للتشفير وايضا تم استخدام خوارزمية (arithmetic). تم تطبيق الخوارزميتين علي عدد من الصور الطبية (mammographic MRI and CT،) حيث تم قياس كفاءة كل خوارزمية علي حدى حيث وجد ان الصور المضغوطة وتم فك ضغطتها متطابقة تماما بنسبة 100% ولم تفقد اي معلومات ، تم الحصول ع نسبة ضغط اعلي عند تطبيق خوارزمية (arithmetic) حيث ان متوسط نسبة الضغط يساوي(3.2) اما خوارزمية(Huffman) فنسبة الضغط عند تطبيقا تساوي (3.14).

Abstract

Medical image play important role in medical investigation. Medical image use in telemedicine field too. When transferring or archiving was demanded, image needs to be compressed so as to reduce its size. Whereas reduction of image size contributes on reduce the bandwidth which is needed to transmit image and minimize the space of storage in archiving case. There are many standards used to deal with medical image like (PACS), (DICOM). In this research two algorithms were created to compress and decompress medical image without losing information compared with the original. Also performance of two algorithms would be measure. Huffman and Arithmetic were used and observe that the arithmetic algorithms work more best than Huffman algorithm. Whereas the compression ratio of Arithmetic algorithm equal to(3.2) and the compression ratio of Huffman algorithm is(3.14).

Chapter 1

1. Introduction

1.1 General overview

Medical imaging is the technique and process of creating visual representations of the interior of a body for clinical analysis and medical intervention, as well as visual representation of the function of some organs or tissues (physiology). Medical images give information of shape and function of organs of human body, this being one of the most important means for diagnosis. An expert physician uses images for diagnosis, together with other information. In most cases it is qualitative and subjective evaluation. The information conveyed by medical images is very difficult to exploit quantitatively and objectively. Increasingly, medical images are acquired or stored digitally. This is especially true of the images that are used in radiology applications [1].

Multimedia data, especially images have been increasing every day. Because of their large capacity, storing and transmitting are not easy and they need large storage devices and high bandwidth network systems [2].

As earlier, the medical images like X-ray, CT (computer Tomography), mammography, MRI , etc. are stored as matrix of binary digits in computer memory. To reduce memory storage required for archiving and transmitting digital images data is compressed before storage or transmitting and decompressed after retrieval or receive.

A large amount of image data is produced in field of medical image which can be stored in picture archiving and communication system (PACS) or hospital information system (HIS). Digital imaging and communications in medicine (DICOM) is the most comprehensive and accepted version of an imaging communications standard. In the digital image smallest element in the image is a pixel; the number of pixels per image is dependent on the required spatial resolution while the number of bits per pixel is determined by quantization accuracy needed for the application.

Image compression aims to reduce the size of image with or without loss of information from original image , it is aims to represent an image in fewer number of bits. We can achieve that by removing one or more of data redundancy.

There are three type of redundancy :

I) **Coding redundancy** is present when less than optimal code words are used. Examples of image coding schemes that explore coding redundancy are the Huffman coding and Arithmetic coding techniques [3]**.**

II) **Inter- pixel redundancy** This redundancy exists as image neighboring pixels are not statistically independent. It results from correlations between the neighboring pixels of an image [3]

III) **Phychovisual redundancy**: irrelevant information which is due to data that is ignored by the human visual system.

The inverse of this process called the decompression which is applied on compressed data/ image to get reconstructed image.

There are two techniques to compress image: first technique is lossy compression in which the original image lost some information when recovered from compressed image , It gives better compression ratio when compared with lossless compression. Lossy compression consist of many method : Vector Quantization, DCT based Transform, Wavelet Transform, and Fractal image compression [4]

Lossless compression: as their name implies, involve no loss of information. If data have been lossless compressed, the original data can be recovered exactly from the compressed data [5], used for data that can not tolerate any different between original image and recovered image. Lossless compression is suitable for medical image coding.

Compression of medical image is an important branch of biomedical engineering. PACS systems is used in Archiving, and telemedicine. In those branches lossless compression technique is used to save the characteristic of original data. DICOM (Digital Imaging and Communication in medicine) is standard of medical images communication; it developed to meet the needs of manufacturers and users of medical imaging equipment for interconnection of devices on standard networks. Its multiple parts provide a means for expansion and updating, and the design of the standard was aimed at allowing simplified development for all types of medical imaging. [4]

In DICOM standard JPEG2000 is added in new standard of DICOM image compression. With this addition in DICOM standard compress the original imaging data based on the JPEG and JPEG2000 image compressing standard. The JPEG is an image compression standard developed by the Joint Photographic Experts Group. It was formally accepted as an international in 1992 and designed for compressing either color or grayscale images of natural real-world scenes.

JPEG standard for image compression is a lossy and based on Discrete Coin Transform (DCT), the basic steps in encoder are: block division, Forward Discrete Cosine Transformation (FDCT), quantization, and entropy encoding. The decoder is the reverse all this process performed by encoder except quantization.

The JPEG-2000 standard has 6 parts. Part 1 is called the "core coding system", which describes the specifications of the decoder as a minimal requirement while the specifications of the encoding part are also included only as informative materials to allow further improvements in the encoder implementations. Part 2 is denoted as the extensions of Part 1, which adds more features (user defined wavelet transformation, etc) and sophistication to the core coding system for advanced users. Part 3 is for the motion JPEG-2000. Part 4 provides a set of compliance testing procedures for the implementations of the coding system in Part 1 as a tool for quality control. Part 5 introduces two free software implementations that perform the compression system for both the encoder and decoder in order to gain wide acceptance of the JPEG-2000. Part 6 defines a file format that stores compound images. [6]

In JPEG2000 standard embodied lossy and lossless compression at the same time, the steps of encoding process are: component transformation, tiling, wavelet transformation, quantization, coefficient bit modeling, arithmetic coding, and rate-distortion optimization The role of the decoder is to reverse the steps performed by the encoder, except the rate-distortion optimization step [6].

1.2 Statement of the Problem

When medical images are transmitted over networks, this requires broad band-width, time and cost to complete this operation. Moreover ,where the archiving for long term; for further use on researches or education. The problem of large space to store image is obvious.

Medical image contain important information about patients' health and must be transmitted or archived without loss of data from original image

1.3 Objectives

The objectives of this research are to:

1- Design an algorithm to compress medical image(MRI, Mammography) and increase compression ratio and decrease the loss of data in compression process.

2- Design an algorithm to decompress the compressed medical images with lossless

1.4 Research Method

In this research, MATLAB® program is used and applied two algorithms namely Huffman and arithmetic coding algorithms on MRI and mammographic images. Firstly medical image is loaded to the program, and then secondly, create a dictionary for the medical image. Then the third step is encoded (compress) image, fourth step is the decoding and reconstruction of the medical image. And final step is to assess the performance of algorithm.

1.5 Research Outline

The remainder of the research will be organized as follows:

Chapter 2: Literature review

Chapter 3: Theoretical background

Chapter 4: Methodology

Chapter 5: Results and discussion

Chapter 6: Conclusion and recommendation

Chapter2

3.Theoretical background

Image compression is a branch of computer science .It is the process of encoded image into small code with or without loss of information. The small size of code allows more images to be stored in small space of memory or drive desk and translating it over internet using small bandwidth and it requires less time to down_ load. There are two standers to deal with medical image (PACS) and (DICOM).

3.1 DICOM Standard

DICOM provides all the necessary tools for the diagnostically accurate representation and processing of medical imaging data. It is a transcriber of data ,storage and display protocol built and design to cover all functional aspects of digital medical imaging .

DICOM format has a header which contains information about the image, imaging modality and information about patient. To compress such a DICOM file, special attention should be given to header information. Distortion-limited wavelet image codec performs better in Case of medical images of large sizes [7] .

DICOM has truly shaped the landscape of contemporary medicine by providing [8]:

1. A universal standard of digital medicine. All current, digital image-acquisition devices produce DICOM images and communicate through DICOM networks. Current medical workflow is implicitly controlled by a multitude of DICOM rules.

2. Excellent image quality. For example, DICOM supports up to 65,536 (16 bits) shades of gray for monochrome image display, thus capturing the slightest nuances in medical imaging.

3. Full support for numerous image-acquisition parameters and different data types. Not only does DICOM store the images, but it also records a multitude of other image-related parameters such as patient 3D position, physical sizes of objects in the image, slice thickness, image exposure parameters, and so on.

4. Complete encoding of medical data. DICOM files and messages use more than 2000 standardized attributes (DICOM data dictionary) to convey various medical data from patient name to image color depth, to current patient diagnosis. These data are

often essential for accurate diagnostics, and capture all aspects of the current radiology.

5. Clarity in describing digital imaging devices and their functionality – the backbone of any medical imaging project. DICOM defines medical device functionality in very precise and device-independent terms. Working with medical devices through their DICOM interfaces becomes a very straightforward process, leaving little room for errors.

3.2 PACS

PACS are medical systems (consisting of necessary hardware and software) designed and used to run digital medical imaging. They comprise digital image acquisition devices (modalities – such as computed tomography (CT) scanners, or ultrasound), digital image archives (where the acquired images are stored), and workstations (where radiologists view the images).

PACS are directly related to DICOM. Their functionality is DICOM-driven, which ensures their interoperability. For that reason, any PACS device or software comes with its own DICOM Conformance Statement, which is a very important document explaining the extent to which the device supports the DICOM standard [8].

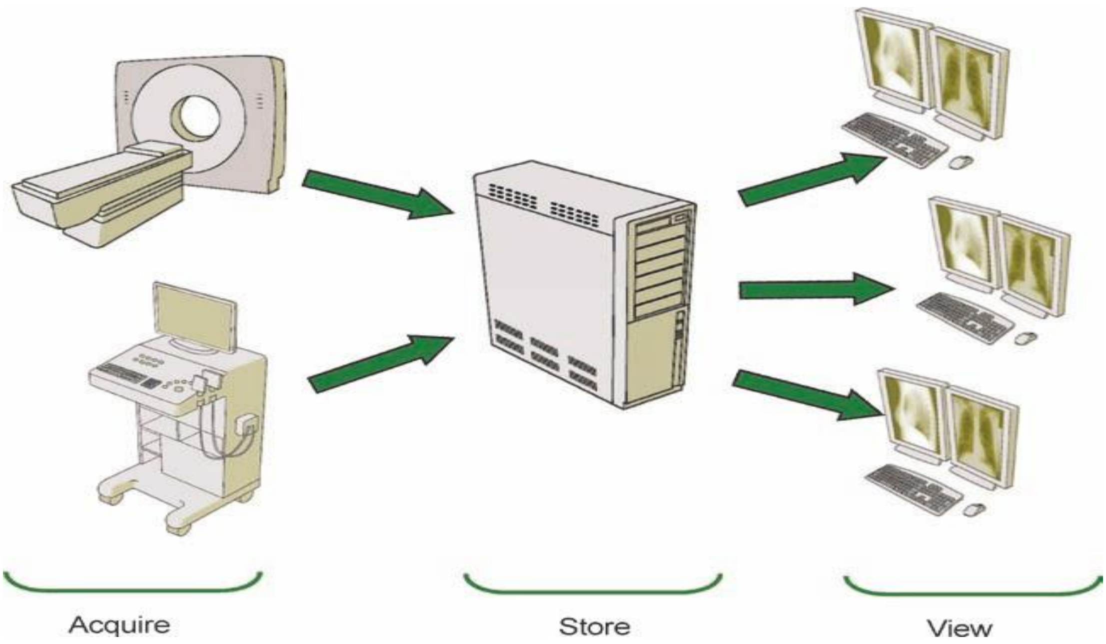


Fig 2.1: Major Picture Archiving and Communication System (PACS) components. Image acquisition devices (modalities) store images on a digital archive. [source]

The PAC system is the most visible component of a digital radiology department but is by no means the only one. A successful PACS requires a strong radiology information system (RIS) to feed it patient and exam information and to keep track of the life cycle of all exams from order placement to final result. The RIS ties together all the computer systems within the department and is typically the sole point of communication to the world outside the department, such as the hospital information system and the billing system [9].

The RIS directs information flow of exams from the ordering process, scheduling, and image acquisition through interpretation, communication of results, and billing. The PACS serves to receive and store the images from the modalities and to distribute them to radiologists for primary interpretation and throughout the healthcare enterprise for clinical review [9].

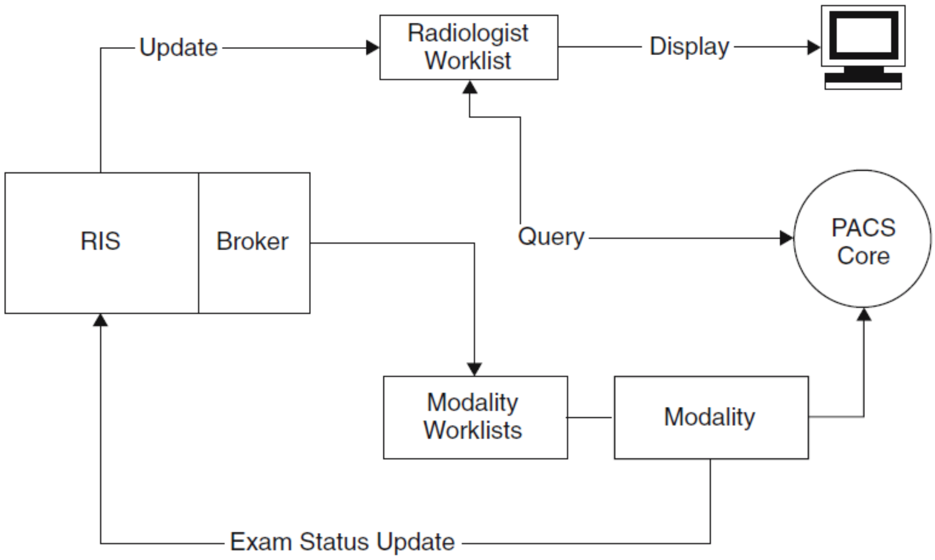


Fig (2.2): RIS-driven PACS workflow.

Following are the basic elements of a PACS:

- Image acquisition
- PACS core which consist of:
  - Database manager (e.g., Oracle, MS-SQL, Sybase)
  - Image archive (e.g., RAID, Jukebox)
  - Workflow/control software (image manager)
  - RIS interface.
- Interpretation workstations:

The workstation is where the physician and clinician see the results of the capture of the relevant exam information within the RIS and the images acquired and stored within the PACS [9].

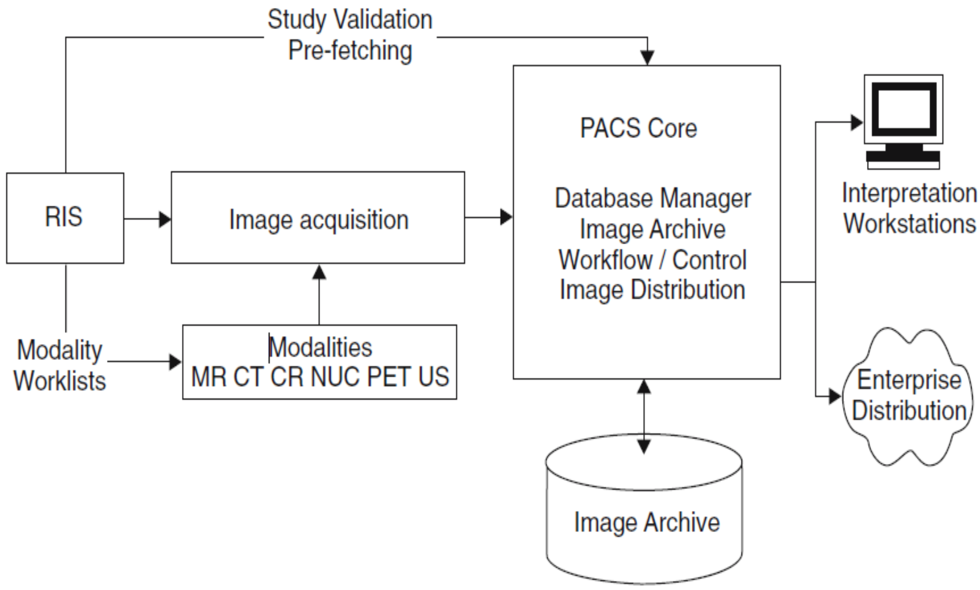


Fig (2.3): Data base management in PACS

3.3 Hospital information system(HIS)

A large computerized database management system that processes patient data in order to support patient care . The system is used by health care clinicians to access patient data and to plan, implement, and evaluate care. In other word HIS was defined as comprehensive software for patient's information integration for sending and exchange comprehensive patient's information between wards and other medical centers in order to expedite the process of patient care, improve quality, increase satisfaction and reduce costs [10].

The goal of Hospital Information System is supporting hospital activities in the levels of practical, tactical, and strategic .In other words, the goal of hospital information system is using of computers and communications equipment for collecting, storing, processing, readout, and communication between patients cares with administrative data on all hospital activities and comply needs of all consumers system [11].

3.4 Standard Health-Care Level 7 (HL7)

Standard HL7 is accepted standard (protocol) in the world that provides the common language for information exchange and electronic patient records in both domestic and abroad [11] .

HL7 is an international forum for health care aiming to work with professionals and health scientists to create standards for information exchange, management, and integration of electronic health information .HL7 strives for the usage of standards within and between health care organizations to increase efficiency and effectiveness of health care activities in a manner that is in favor of all [12].

In other words the goal of HL7 is facilitating communication in Health Care configuration [13].

3.5 The interface between (HIS) ,(PACS) and(DICOM)

An interface was created between the Department of Defense's hospital information system (HIS) and its two picture archiving and communication system (PACS)–based radiology information systems (RISs). The HIS is called the Composite Healthcare Computer System (CHCS), and the RISs are called the Medical Diagnostic Imaging System (MDIS) and the Digital Imaging Network (DIN)–PACS. Extensive mapping between dissimilar data protocols was required to translate data from the HIS into both RISs. The CHCS uses a Health Level 7 (HL7) protocol, whereas the MDIS uses the American College of Radiology–National Electrical Manufacturers Association 2.0 protocol and the DIN-PACS uses the Digital Imaging and Communications in Medicine (DICOM) 3.0 protocol. An interface engine was required to change some data formats, as well as to address some nonstandard HL7 data being output from the CHCS. In addition, there are differences in terminology between fields and segments in all three protocols. This interface is in use at 20 military facilities throughout the world. The interface reduces the amount of manual entry into more than one automated system to the smallest level possible. Data mapping during installation saved time, improved productivity, and increased user acceptance during PACS implementation. It also resulted in more standardized database entries in both the HIS (CHCS) and the RIS (PACS) [14] .

3.6 Telemedicine in health care field

Telemedicine is telecommunication technology integrated with the advancements in medical technology. The main purpose is to enhance health care delivery to a wider population. This telemedicine technology supports the transfer of pathological and imaging reports of patients across the telemedicine networks. The term telecommunications generally means electronic transmission of information over a distance. We can use modern telecommunication and information technologies for the provision of clinical care to individuals at a distance. This application is very efficient since patient records, stored electronically, can be made available through the Internet, resulting in the elimination of the need for physical storage and transfer of records. Furthermore, images and video can be included and transmitted as part of a computerized file.

Thus telemedicine technology offers the following benefits:

- Reduction in time and cost incurred in travel
- Easy and quick access to specialist
- Cost effective post treatment consultation
- Efficient use of medical resources.

The major areas of telemedicine technology are

- Tele-consultation
- Tele-diagnosis
- Tele-treatment
- Tele-education
- Tele-training
- Tele-monitoring
- Tele-support

3.7 Information theory and image compression

Data compression is sub-field of information theory because it deal with the redundancy. The information theory is a branch of mathematic that revolve various question about information including different way of storing and communicating data. Information theory uses entropy expression as a measure of information numbers which are encoded in a massage. The word entropy borrowed from thermodynamics, if the entropy of massage is high that means it contain more information. Some desirable properties of any compression method for medical images include: (i) high lossless compression ratios (ii) resolution scalability, which refers to the ability to decode the compressed image data at various resolutions and (iii) quality scalability, which refers to the ability to decode the compressed image at various qualities or signal-to-noise ratios (SNR) up to lossless reconstruction.

3.7.1 Entropy

Shannon borrowed the definition of entropy from statistical physics, where entropy represents the randomness or disorder of a system. In particular a system is assumed to have a set of possible states it can be in, and at a given time there is a probability distribution over those states. Entropy is then defined as [15]:

$$H(S) = \sum_{s \in S} P(s) \log_2 \frac{1}{P(s)}$$

Where S is the set of possible states, and p(s) is the probability of stat s ϵ S. This definition indicates that the more even the probabilities the higher the entropy (disorder) and the more biased the probabilities the lower the entropy.

In the context of information theory Shannon simply replaced “state” with “message”, so S is a set of possible messages, and p(s) is the probability of message s ϵ S. Shannon also defined the notion of the self-information of a message as:

$$i(s) = \log_2 \frac{1}{p(s)}$$

This self-information represents the number of bits of information contained in it and, roughly speaking, the number of bits we should use to encode that message.

3.8 Modeling and coding

In general, data compression consists of taking a stream of symbols and transforming them into codes. If the compression is effective, the resulting stream of codes will be smaller than the original symbols [16].

The data compression algorithm can be divided into two phases. The first phase is a modeling ,in this phase we extract information about any redundancy that exists in the data and describe the redundancy in the form of a model . Establishing a certain code for a certain symbol or set of symbols is based on a model. Lossless data compression is generally implemented using one of two different types of modeling:

- statistical or dictionary-based. Statistical modeling reads in and encodes a single symbol at a time using the probability of that character's appearance. Dictionary-based modeling uses a single code to replace strings of symbols. In dictionary-based modeling, the coding problem is reduced in significance, leaving the model supremely important [16] .
- The second phase is coding, it is a description of the model and a description of how the data differ from the model are encoded.

In old way of coding ,when encoded symbols with exactly the number of bits of information the symbols contain for example if the character ' w ' give us five bits of information then it should be coded with exactly four bits.

Solving this coding problem in a reasonable manner was one of the first problems tackled by practitioners of Information Theory. Two approaches that worked well were Shannon-Fano coding and Huffman coding—two different ways of generating variable-length codes when given a probability table for a given set of symbols.

Huffman coding achieves the minimum amount of redundancy possible in a fixed set of variable-length codes. This doesn't mean that Huffman coding is an optimal coding method. It means that it provides the best approximation for coding symbols when using fixed-width codes [16].

3.9 Lossless compression

In lossless compression the Original data can be recovered exactly from the compressed data i.e the original image and the constructed image are identical or close to be identical. also lossless compression called noiseless since they do not add noise to signal and also known as entropy coding since it uses statistic / decomposition techniques to eliminate/minimize redundancy. This is an important requirement for medical imaging domains.

3.9.1 Run length encoding

Run-length encoding(RLE) is a very easy and simple technique of data compression, in which the count of occurrence of same data is stored as a single data value and single count. This is most useful for the data that contains many such runs for example, a simple colored image such as same color occur many time [5]. It is less useful with images that don't have many runs as it could greatly increase the files size The Run length encoding technique performs a lossless compression of input images that is based on sequences of identical values (runs).

Run Length Encoding provides efficient compression of data, whereas the data with large number of runs or large number pixel contains same intensity value.

But this encoding also has the drawback; RLE scheme does not always provide data compression. In some case where runs are smaller length or each pixel value has different intensity value from its adjacent pixel, this method performs very poorly and instead of compressing data.

3.9.2 Lempel–Ziv–Welch Coding (LZW)Compression

3.9.2.1 LZ77

The first compression algorithm described by Ziv and Lempel is commonly referred to as LZ77. It is relatively simple. The dictionary consists of all the strings in a window into the previously read input stream [16].

3.9.2.2 LZ78

The LZ78 program takes a different approach to building and maintaining the dictionary. Instead of having a limited-size window into the preceding text, LZ78 builds its dictionary out of all of the previously seen symbols in the input text. But instead of having carte blanche access to all the symbol strings in the preceding text, a dictionary of strings is built a single character at a time. This incremental procedure works very well at isolating frequently used strings and adding them to the table. Unlike LZ77 methods, strings in LZ78 can be extremely long, which allows for high compression ratios.

LZW is a universal lossless data compression algorithm created by Abraham Lempel, Jacob Ziv, and Terry Welch. It was published by Welch in 1984 as an improved implementation of the LZ78 algorithm published by Lempel and Ziv in 1978. LZW is a dictionary based coding. Dictionary based coding can be static or dynamic. In static dictionary coding, dictionary is fixed when the encoding and decoding processes. In dynamic dictionary coding, dictionary is updated on the fly [3].

In another way we can say that: LZW compression is the compression of files into smaller files using a table-based look up algorithm. The GIF and TIFF are the two commonly used file formats which are compress using the LZW compression.

The working principle of the LZW compression is as follows; an algorithm takes each input series of bits and created an entry into a table for that particular bit series. So when the next input is read any pattern that has the same result as the above is put back in the table and hence space is saved.

The compression algorithm used in the LZW compression is so powerful that you can compress to nearly half its original size.

3.9.3 Huffman coding

It is an entropy encoding algorithm used for lossless data compression in computer science and information theory [6].And it is a part from statistical model.

Basic principles of Huffman Coding

Huffman coding is a popular lossless Variable Length Coding (VLC) scheme, based on the following principles:

(a) Shorter code words are assigned to more probable symbols and longer code words are assigned to less probable symbols.

(b) No code word of a symbol is a prefix of another code word. This makes Huffman coding uniquely decodable.

(c) Every source symbol must have a unique code word assigned to it. In image compression systems), Huffman coding is performed on the quantized symbols.

Quite often, Huffman coding is used in conjunction with other lossless coding schemes, such as run-length coding,. Huffman coding uses a specific method for choosing the representation for each symbol, resulting in a prefix code.

$$H(S) \leq J < H(S)+1$$

Huffman's procedure creates the optimal code for asset of optimal code for a set of symbol and probabilities subject to the constraint that the symbols be coded one at a time.

The first step in Huffman's approach is to create a series of source Reductions by ordering the probabilities of the symbols under consideration and combining the lowest probability symbols into a signal symbol that replaces them in the next source reduction [17] table(3.1). Table (3.2) shows symbols are assigned to the two symbols on the right (the assignment to arbitrary ; reversing the order of the 0 and 1would work just as well).

Table (3.1): Huffman source reduction

Original source		Source reduction			
Symbol	probability	1	2	3	4
A2	0.4	0.4	0.4	0.4	0.6
A6	0.3	0.3	0.3	0.3	0.4
A1	0.1	0.1	0.2	0.3	
A4	0.1	0.1	0.1		
A3	0.06	0.1			
A5	0.04				

The second step in Huffman's procedure is to code each reduced source , starting with smallest source and working back to the original source the minimal length binary code for a two symbol source of course are the symbols 0 and 1.

Huffman's code are used in(CCITT, GBIG2; JPEG;MPEG1,2,4; H.261;H262; H263;H264) and other compression standards .

Table (3.2): Huffman code assignment procedure

Original source			Source reduction							
Symbol	probability	code	1		2		3		4	
A2	0.4	1	0.4	1	0.4	1	0.4	1	0.6	0
A6	0.3	00	0.3	00	0.3	00	0.3	00	0.4	1
A1	0.1	011	0.1	0011	0.2	01 0	0.3	01		
A4	0.1	0100	0.1	0100	0.1	01 1				
A3	0.06	01010	0.1	0.10101						
A5	0.04	01011								

In coding part , Huffman achieves the minimum amount of redundancy possible in fixed set of variable length code. But this doesn't mean the Huffman coding is an optimal coding method.

The problem with Huffman or shanon coding is that they use an integral number of bits in each code that means if the entropy is 2.5 bits the Huffman code must be either 2 or 3 bits that means the Huffman is inefficient due to using an integral number of bits per code but it easy to implement and very economical for both coding and decoding.

3.9.3.1 Optimality of Huffman codes

The necessary conditions for an optimal variable-length binary code are as follows [5]

Condition 1: Given any two letters (a_j) and (a_k), if $P[a_j] \geq P[a_k]$, then $l_j \leq l_k$, where l_j is the number of bits in the code word for a_j .

Condition 2: The two least probable letters have code words with the same maximum length l_m.

Condition 3: In the tree corresponding to the optimum code, there must be two branches stemming from each intermediate node.

Condition 4: Suppose we change an intermediate node into a leaf node by combining all the leaves descending from it into a composite word of a reduced alphabet. Then, if the original tree was optimal for the original alphabet, the reduced tree is optimal for the reduced alphabet If this condition were not satisfied, we could find a code with smaller average code length for the reduced alphabet and then simply expand the composite word again to get a new code tree that would have a shorter average length than our original "optimum" tree.

3.9.3.2 The length of Huffman code

Huffman coding procedure generate optimum code, the length of code depend on a number of things include the size of alphabet and probability of individual letters .

If we say that : Huffman code for the source (S) has average code length (J)then the Average code length should be between an entropy and entropy plus 1.

3.9.3.3 Adaptive Huffman Coding

In adaptive Huffman coding procedure, two component transmitter and receiver don't know evrything about statistics of the source sequence of the start transmission.

In the tree of both the transmitter and receiver consists of a single node that corresponding to all symbols not yet transmitted (NYT) and has weight zero.

As a transmission evolve ,nodes corresponding to symbols will be added to the tree and the tree is reshaped using update procedure .

First step before the beginning of transmission affixed code for each symbol is agreed upon between transmitter and receiver. A node for the symbols is then created and the symbol is taken out of the (NYT) list.

Both transmitter and receiver start with the same tree structure. The updating procedure used by both transmitter and receiver is identical there for the encoding and decoding process remain synchronized .

3.9.3.3. A update procedure

The function of the update procedure is to save the sibling property. In the update procedure the nodes must be in fixed order. This ordering is recorded by numbering the node, the largest node number is giving to the root of the tree and smallest number is point to the (NYT) node.

The number from the (NYT) node to the root of the tree are assigned in increasing order from the left to right and from lower level, the nodes with the same weight makes up block . Fig 3.4 shows the flowchart of the updating procedure. At the update procedure both transmitter and receiver are operate with the same information the tree at the transmitter is updated after each symbol encoded. At the receiver the tree is updated after each symbol is decoded.

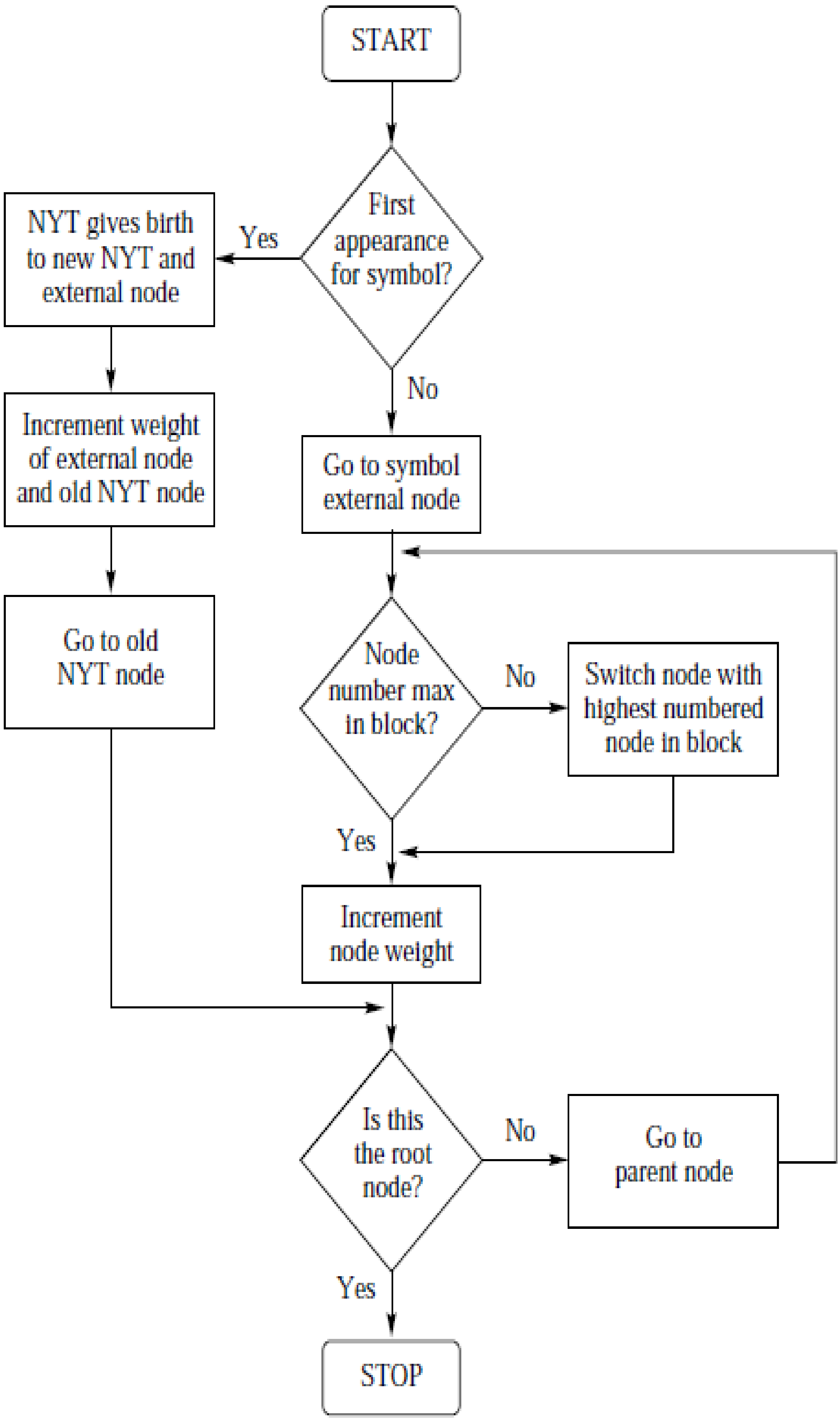


Fig (3.4) Update procedure for Huffman coding

3.9.3.3b Encoding procedure

In encoding procedure the tree at both at the encoder and decoder consist of a single node named (NYT) node in the first step in the encoded procedure is read the symbol if it first appears we send to the (NYT) followed by the previously agreed upon if isn't first appearance then the root of tree to the (NYT) node. After this step update procedure then check is this the last symbol stop or repeat all step above .

if the source of alphabet has a size m where

$$m = 2^e + r \text{ and } 0 < r < 2^e$$

The letter a_k is encoded as the (e+1)-bit binary representation of k-1, if $1 \leq k \leq 2r$; else, a_k is encoded as the e-bit binary representation of k-r -1.

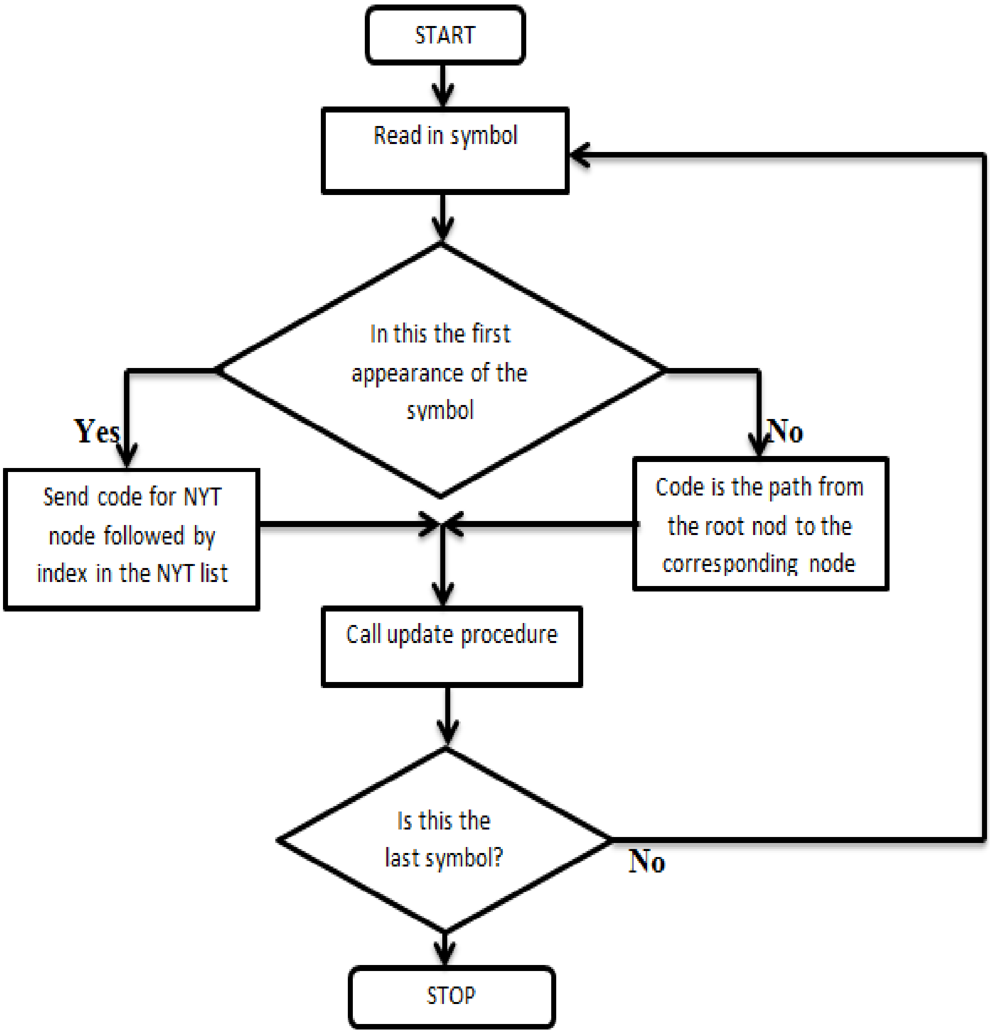


Fig (3.5) Encode procedure

3.9.3.3c Decoding procedure

we can decode the cod as showed in fig (3.6):

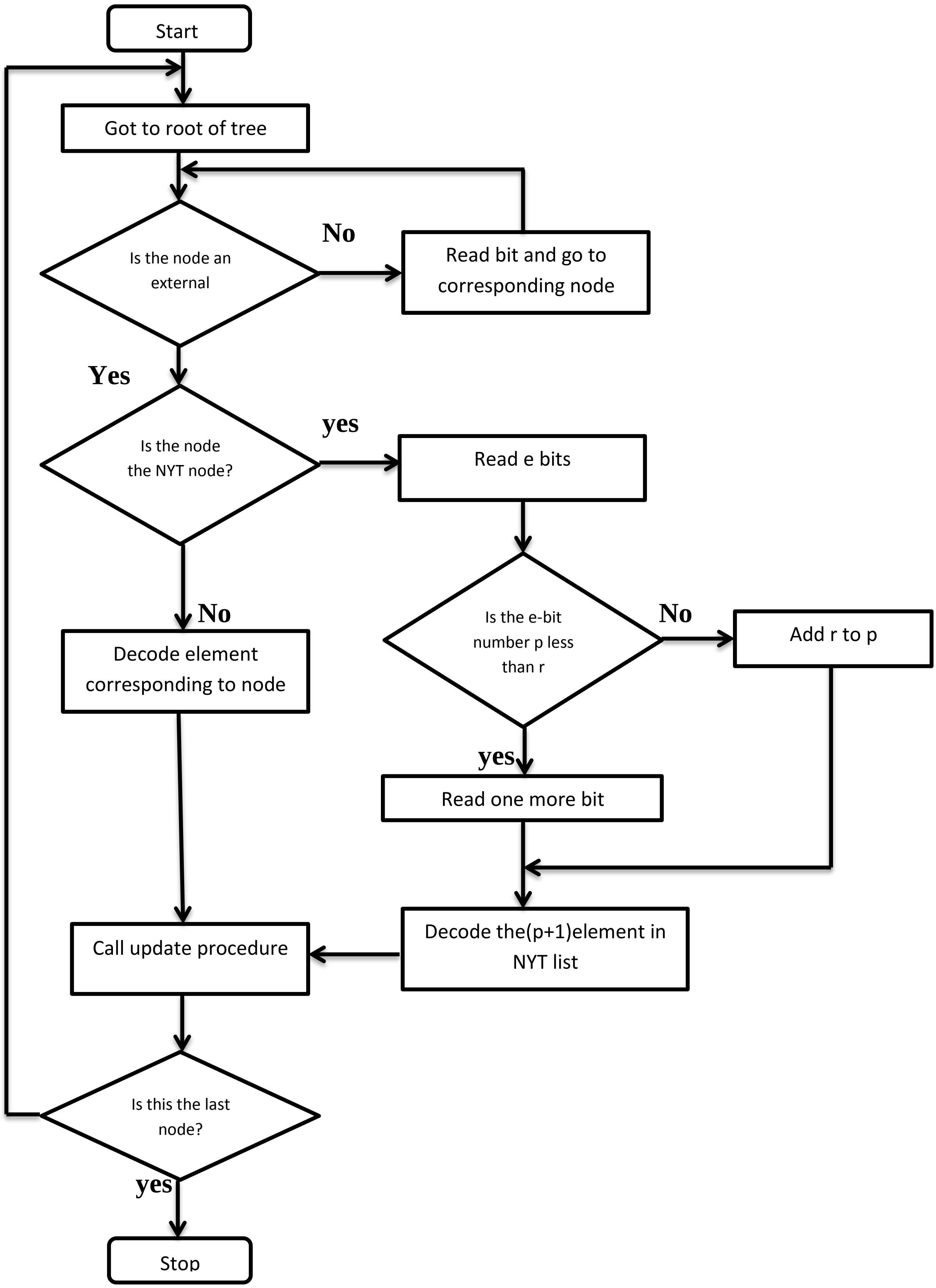


Fig (3.6) Decoding procedure

3.9.4arithmetic coding

Arithmetic coding is a form of entropy encoding used in lossless data coding. A massage is coded as a subinterval of the interval [0,1). Where [x,y) denotes a half open interval ,which includes x but excludes y

There are two basic concepts in arithmetic coding :

1- The probability of a symbol whereas the occurrence probabilities of source symbols determine the compression efficiency as well as the interval range of source symbols for the encoding process .
2- Encoding interval for a symbol , these interval ranges as are contained within the interval from zero to one and determine the compression output .

Arithmetic coding is used in lossy and lossless compression and it is part of international standards like JPEG2000

In arithmetic coding a unique identifier or tag is generated for the sequence to be encoded. This tag corresponds to a binary fraction, which becomes the binary code for the sequence. In practice the generation of the tag and the binary code are the same process. However, the arithmetic coding approach is easier to understand if we conceptually divide the approach into two phases. In the first phase a unique identifier or tag is generated for a given sequence of symbols. This tag is then given a unique binary code. A unique arithmetic code can be generated for a sequence of length m without the need for generating code words for all sequences of length m. This is unlike the situation for Huffman codes.

1.Coding a Sequence: is the first step in arithmetic coding and have two part

i- Generating a Tag

The operation of generating the tag works by reducing the size of the interval in which the tag resides as more and more elements of the sequence are received.

ii- Deciphering the Tag

We have spent a considerable amount of time showing how a sequence can be assigned a unique tag, given a minimal amount of information. However, the tag is useless until we can also decipher it with minimal computational cost. Fortunately, deciphering the tag is a simple.

2. Generating a Binary Code:

If we want to find a binary code that will represent the sequence in unique and efficient manner.

$T_X(\mathrm{x})$ is a number in the interval [0,1). A binary code for $T_X(\mathrm{x})$ can be obtained by taking the binary representation of this number and truncating it to

$$I(x)=\left\lceil log \frac{1}{p(x)}\right\rceil + 1\text{bits.}$$

3.9.4.1 Adaptive Arithmetic Coding

In the adaptive case, we may not know ahead of time what the total number of symbols is going to be. In this case we have to pick the word length independent of the total count. However, given a word length m we know that we can only accommodate a total count of 2^{m-2} or less. Therefore, during the encoding and decoding processes when the total count approaches2^{m-2}, we have to go through a rescaling, or renormalization, operation. A simple rescaling operation is to divide all counts by 2 and rounding up the result so that no count gets rescaled to zero [5].

3.10 Comparison Arithmetic Coding versus Huffman

The Huffman coding always produces rounding errors, because its code length is restricted to multiples of a bit. This deviation from the theoretical optimum is much higher in comparison to the arithmetic coding's inaccuracies.

The efficiency of an arithmetic code is always better or at least identical to a Huffman code [18].

The problem with Huffman coding in the above message is that it can’t create codes with the exact information content required. In most cases it is a little above or a little below, leading to deviations from the optimum. But arithmetic coding gets to within a fraction of a percent of the actual information content, resulting in more accurate coding.

Chapter3

2.Literature review

In the following section some of the state of the art techniques are presented:

2.1 Dual Tree Complex Wavelet Transform and Arithmetic Coding Technique for compression:

Megha Vaishnav et al [19], compressed medical image using the dual tree wavelet transform and arithmetic coding technique. They compressed six medical images (MRI) using MATLAB version R2015a. In the dual tree wavelet transform calculates the complex transform of a signal using two separate DWT decompositions , (DTCWT) employs two real DWTs; the first DWT gives the real part of the transform while the second DWT gives the imaginary part. Arithmetic coding start by determining a model of the data-basically a prediction of what patterns will be found in the symbols of the message (used traditional way by divided it to interval , each interval is divided in several subinterval(decomposition) and produce new interval) . The obtained result as follow PSNR value is 65.30 with the low MSE 27.24 The average compression ratio is 2.32 and the structure similarity index is 0.915 for the tested images.

2.2 Using Huffman Coding technique

Nehal Markandeya et all [20], compressed images using new method of Huffman encoding, in their work investigates image compression using block truncation coding. Three algorithms were selected namely: the original block truncation coding (BTC), Absolute Moment block truncation coding (AMBTC) and Huffman coding. A comparison was performed between these algorithms . Block truncation coding (BTC) and Absolute Moment block truncation coding techniques rely on dividing image into non overlapping blocks. They differ in the way of selecting the quantization level in order to remove redundancy. In Huffman coding an input image is split into equal rows and columns and at final stage sum of all individual compressed images which not only provide better result but also the information content will be kept secure. It has been showed that the image compression using Huffman coding provides better image quality than image compression using BTC and AMBTC, and size image reduced to 40% . Moreover, the Huffman coding is quite faster compared to BTC.

2.3 lightweight Binary Arithmetic Coding for lossless compression

Joan Bartrina-Rapestaa et al [21], proposed method to compression called A contextual lightweight arithmetic coder for lossless compression of medical imagery. A binary arithmetic coder with fixed-length code words is adopted, thus avoiding the normalization procedure common in most implementations. The probability of each context is estimated through bitwise operations. In this work a set of image were collected from different medical device. Experimental results indicated that, on average, their proposal improved over the standard version of CCSDS-123 for lossless coding by nearly 0.1 bps. Compared with the coding techniques evaluated in this work their proposal provides improvements from 0.05 to 0.5 bps for JPEG2000, from 0.1 to 0.41 bps for JPEG-LS, from 0.1 to 0.68 bps for HEVC, and from 0.003 to 0.26 bps for CCSDS-123. In addition, the results indicated that the use of the four closest neighbors for the context formation is enough to properly exploit the contextual information of medical images in an arithmetic encoder when the data to be encoded are obtained from the predictor of CCSDS-123.

2.4 The Current Role of Image Compression Standards in Medical Imaging

Feng Liu et al [22] , Compared of the lossless performances of the image compression standards on the data sets described that JPEG-LS provided the best compression ratios on CT Colonography and MR Mammography data sets. The compression performance of JPEG-LS was among the top three on the remaining data sets. they also noted that the JPEG2000 yielded the highest average compression ratios on MR T2 Flair Axial and Digital Pathology data sets, and HEVC RA attained the highest average compression ratio on CT Lung data set. It was also observed that the availability of inter-frame and intra-frame prediction in HEVC RA resulted in higher average compression ratios compared to the intra-only prediction used in HEVC AI. For all data sets except MR Mammography, JPEG-XR achieved higher average compression ratios compared to HEVC-AR.

2.5 Comparative Study on Lossless Data Compression Techniques

Seema and Priyanka Anand [23] studied various technique of lossless compression include : Huffman encoding , Shannon fano coding , Run Length Coding (RLE) , Lempel -Ziv Welch (LZW), Embedded Zerotree Wavelet (EZW) , Set Partition in Hierarchical Trees (SPIHT) and found compression ratio , Peak Signal to Noise Ratio (PSNR), Mean Square Error (MSE) for each method. The result determined as follow:

Method	CR	PSNR	MSE
Huffman	1.60	21.21	23.29
Shannon-fano coding	1.5	38.51	8.9
Run Length encoding	0.29	39.38	7.49
LZW	3.34	47.79	5.23
SPIHT	8	61.53	0.54

2.6 Entropy Coding Techniques

Sandeep Kaur and Sukhjeet Singh [24] compared between various coding techniques. They compared between: Huffman technique , Unary coding, arithmetic coding , Shannon–Fano coding, Elias Gama coding, Tunstall coding , Golomb Coding, Rice coding , Universal Code (Data Compression), Shannon coding, Range encoding, Exponential-Golomb Coding and Fibonacci Coding.

From comparison among different coding techniques they found that Huffman coding is easier than arithmetic coding. Arithmetic algorithm yields much more compression ratio than Huffman algorithm while Huffman coding needs less execution time than the arithmetic coding. Also they found that Exponential-Golomb coding generalizes the gamma code to integers. They recommended universal codes are useful when Huffman codes are inconvenient.

2.7 An improved lossless image compression based arithmetic coding using mixture of non-parametric distributions

Masmoudi, et al [25]proposed an efficient scheme for block-based lossless image compression using adaptive arithmetic coding. The model estimates the statistics of each block and is designed by a finite mixture model of non-parametric distributions by exploiting the high correlation between neighboring blocks. The experimental results showed that the compression efficiency of arithmetic coding when using finite mixture model for data statistics estimation is increased, and when combined with the median edge detector, it outperformed JPEG-LS(The latest coding standard for lossless and near-lossless of continuous tone Images) by 2.6 %. An hybrid algorithm combining the results obtained when encoding image in pixel domain and prediction error domain did very well, beating JPEG-LS by 9.7 %.

2.8 A new arithmetic coding model for a block-based lossless image compression based on exploiting inter-block correlation

Atef Masmoudi [26] investigated a new approach for a block-based lossless image compression using arithmetic coding. The conventional arithmetic encoders encode and decode images pixel by pixel in raster scan order by using a statistical model which provides probabilities for the whole source symbols to be encoded. However, in the proposed scheme, the arithmetic encoders encode an image block by block from left to right, and block-row by block-row from top to bottom. The proposed model estimates the probability distribution of each block by exploiting the high correlation between neighboring image blocks. Therefore, the probability distribution of each block of pixels is estimated by minimizing the Kullback–Leibler distance between the exact probability distribution of that block and the probability distributions of its neighboring blocks in causal order. The results of comparative experiments showed significant improvements over conventional arithmetic encoders in both static and adaptive order-0 models, reducing the bitrate by an average of 15.5 and 16.4 % respectively.

2.9 An enhanced variable-length arithmetic coding and encryption scheme using chaotic maps

QiuzhenLin et al [27] , enhanced the simultaneous arithmetic coding and encryption scheme previously proposed. By encoding a block of variable number of symbols to a code word within the length of the computation register, the operating efficiency has been substantially improved. Moreover, the compressed sequence is processed by an additional diffusion operation which strengthens the security of the original scheme by having higher key and plaintext sensitivities. Simulation results show that the enhanced scheme runs faster than the original scheme and the traditional compress-then-encrypt approach at a comparable compression performance.

2.10 Evaluation of Huffman and Arithmetic Algorithms for Multimedia Compression Standards

Asadollah Shahbahrami et al [28], tested and implemented two algorithms to find which entropy coding Arithmetic or Huffman is more suitable by comparing the compression ratio performance. The algorithms were applied on different image contents and sizes.

They found that: the compression ratio of the arithmetic coding for different image sizes is higher than the Huffman coding. On the other hand, arithmetic coding needs more execution time than Huffman coding. This means that the high compression ratio of the arithmetic algorithm is not free. It needs more resources than Huffman algorithm. Another behavior that can be observed that by increasing image sizes the improvement of the compression ratio of the arithmetic coding increases more than the Huffman coding.

Chapter 4

4. Methodology and materials

Three kinds of medical image are used in this project : MRI ,mammography and CT

The MRI images are collected from Yastbshron's hospital. The collection comprised of 12 patints image: every image is 256×265 pixel 16 bit unsigned integers. Digital mammograms were obtained from the Mammographic Image Analysis Society (mini-MIAS) database. In this database, each image is 1024× 1024 pixels and 8-bit. And the CT images are 256×256 pixel 16 bit unsigned integers from DICOM library [29] .

The compression algorithm are implemented using MATLABE2015a

Huffman coding and arithmetic coding are used in this research and use measure performance function in matlab program.

4.1 Huffman Encoding

The method makes use of statistical occurrence frequencies (probabilities) to carry out the process. Each pixel intensity of the image is treated as a symbol.

4.2 Proposed Huffman encoding logarithm

This section explains the algorithmic details of the proposed Huffman coding. After the image is read into the development environment (MATLAB ®),we found the predictive code for image. The image histogram is calculated and then, pixel intensity probabilities were inferred. A dictionary is required by the Huffman algorithm was also developed and fed into the encoding algorithm as input.

In the decoding step the code that was created in the Huffman encoding is passed into the decoding algorithm with dictionary also as an input. The image was reconstructed , the last step is compression statistics and visualization

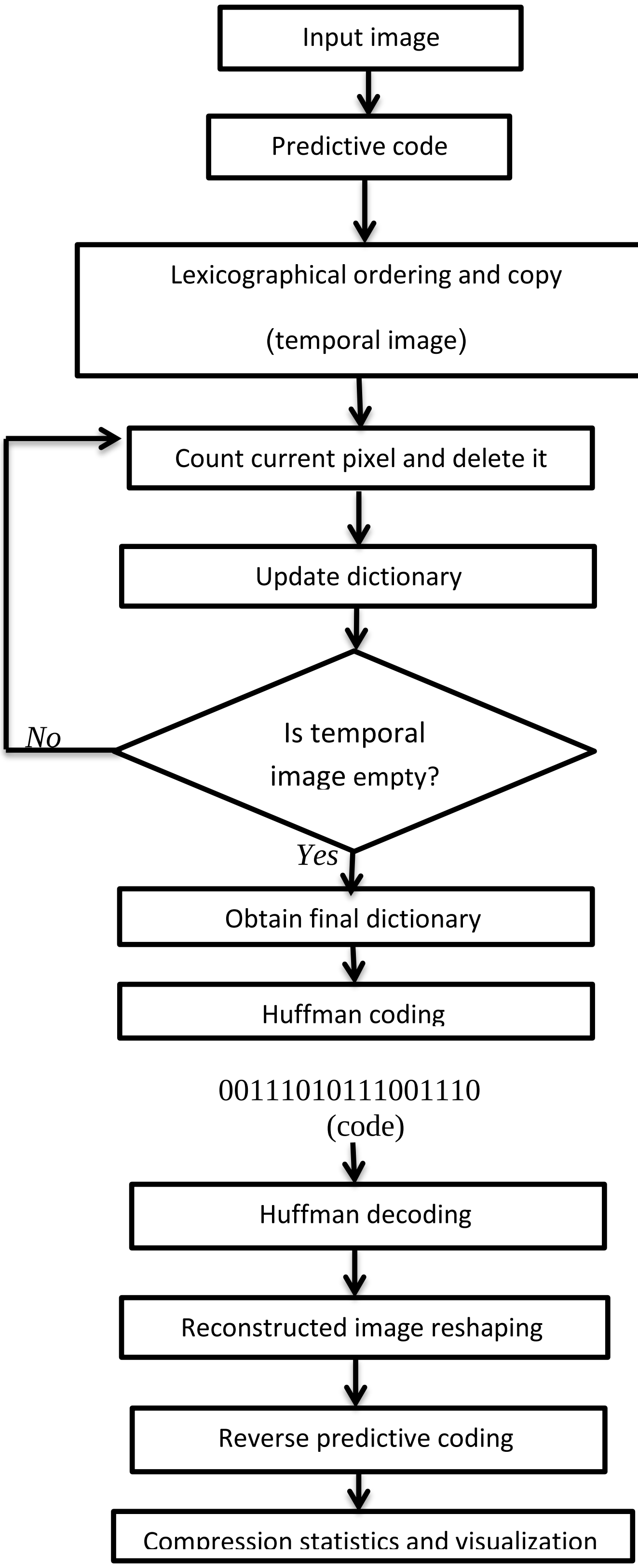


Fig (4.2):proposed Huffman coding Scheme(code and decode phases)

4.3 Arithmetic Encoding

Arithmetic coding bypasses the idea of replacing an input symbol with a specific code. It replaces a stream of input symbols with a single floating-point out .The output of an arithmetic coding process is a single number less than 1 and greater than or equal to 0. This single number can be uniquely decoded to create the exact stream of symbols that went into its construction. To construct the output number, the symbols are assigned a set probabilities put number.

By another way unique identifier or a tag is generate for particular sequences of symbols without the need to generate all possible code words for sequences at the same length as where the case of Huffman encoding .

Tag generated by mapping it symbol to real line segment from 0 to 1 and keep zooming into the segment as more symbol observed arithmetic coding is also suitable for dialing with short dictionaries .

4.4 Proposed Arithmetic Coding logarithm

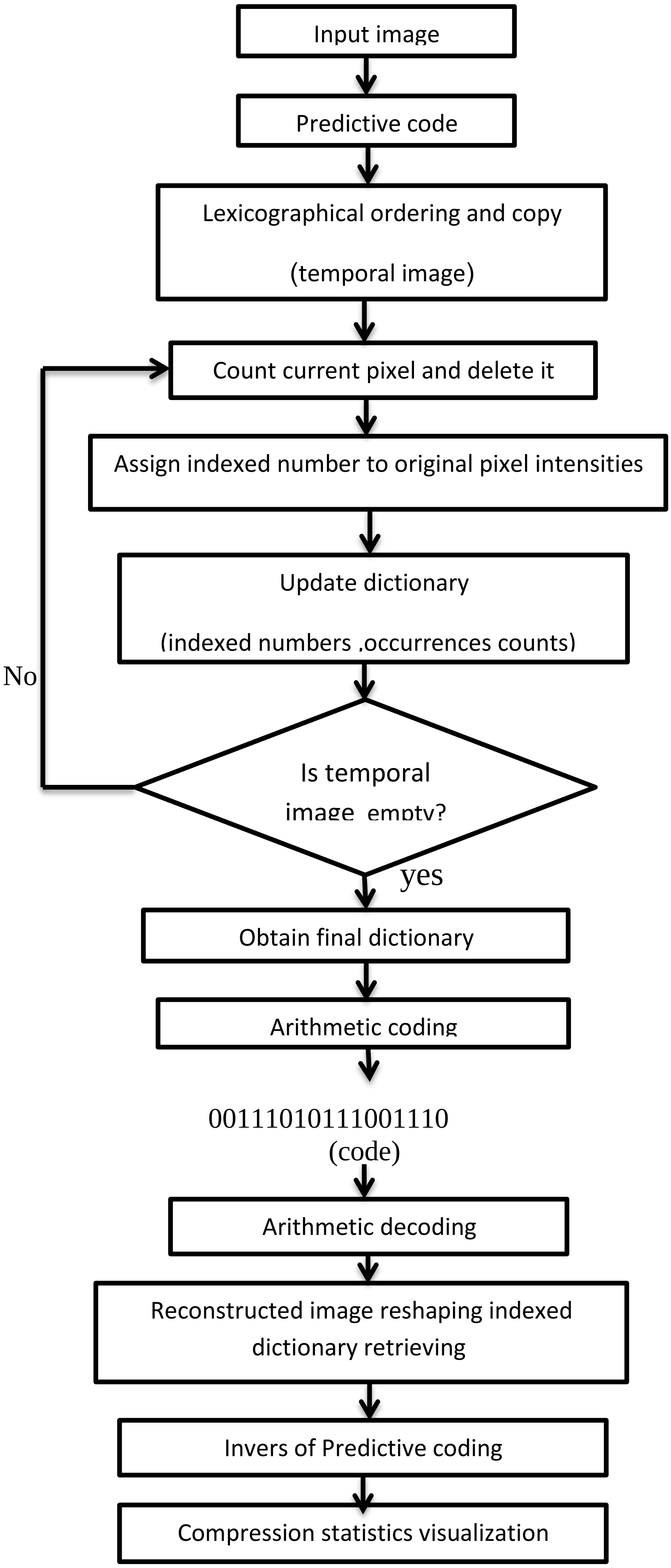
As known earlier the arithmetic coding algorithm does not require a dictionary to operate. But there is some technical requirement to the arithmetic coding function available to work. This technical requirement is that MATLAB® requires entries of the function to be in integer incrementing fashion; e.g. the first element in the entire array must have index of one then second entry must have index two etc. To meet this requirement a 'Special dictionary' was designed.

The design of the dictionary works exactly as successive counting and deleting descried in the Huffman coding above but with some changes will be described hereafter.

After the algorithm firstly count the very first pixel in the image and then deleting this pixel intensity in the entire image. The first element of the dictionary is created by recording the not the pixel intensity value as in Huffman above, but a value with is the index that is fed to algorithm. In other words this goes like saying the pixel gets counted the index gets recorded. This of course complicates the decoding as will describe next. The probability in the other hand is not stored as probability but the original pixel counts is now stored. i.e. the pixel actual occurrences.

The second pixel intensity value then enters and gets counted in the copy image and then gets deleted in the temporal image, and the dictionary is then updated by incrementing the counts by one with its correspondent counts. This continues until the image gets empty with no entries and then the loop exits leaving the dictionary of the entire image having integer indexes ordered from one to total number of intensities.

This dictionary is then passed as input to the arithmetic coding algorithm. In the decoding step the code that was created earlier in the arithmetic encoding is passed into the decoding algorithm with dictionary also as an input. But the recovered image is now a lexicographically ordered single row and does not contain the original pixel intensities but contains the ordered intensity values the dictionary, to recover the image this row gets reshaped to the original image dimensions and then the dictionary is replaced after the decoding with original pixel intestines.



Fig(4.4): proposed Arithmetic coding Scheme(code and decode phases)

4.5 Measure performance

A compression algorithm could be evaluated in a number of different ways. There are many tradeoffs when comparing a compression algorithm. More specifically, we could measure the relative complexity of the algorithm, the memory required to implement the algorithm, how fast the algorithm performs on a given machine, the amount of maximum compression that could be achieved, and how closely the reconstruction resembles the original**.** In this research we used compression ratio and data rate

4.5.1 Compression ratio

Compression ratio (CR) is a quantitative metric of the amount compression achieved and is defined as, the total number of bits in the original image divided by the number of bits used to represent the compressed image.

$$\boldsymbol{Compression\ ratio} = \frac{\boldsymbol{Uncompreesed\ Data\ size}}{\boldsymbol{Compressed\ data\ size}}$$

4.5.2 Data Rate

The data rate is a term to denote the average number of bits required to represent a single sample, or the number of bits per second transferred when talking in communication literature.

In telecommunications, it is common use to express the data rate in bits per seconds (bit/s). In data communication, the data rate is often expressed in bytes per second (B/s).
Data rate is given by following equation:

$$\boldsymbol{Data\ Rate} = \mathbf{1} - \frac{\boldsymbol{compressed\ data\ rate}}{\boldsymbol{Uncompressed\ data\ rate}}$$

Chapter 5

5.Result and discussion

In this work, a total number of 50 images was tested, 20 mammographic images from MIAS database with spatial resolution of 1024x1024 pixels with pixel intensity 8bits per pixel in portable gray map format (.pgm), have been compressed by arithmetic and Huffman coding algorithms. Other 20 MRI images were used and with spatial resolution of 512x512 pixels with pixel intensity 16bits per pixels in DICOM format (.dcm) and arithmetic and Huffman algorithms was applied too. Last type of image was used CT images, 10 images with spatial resolution of 512x512 pixels with pixel intensity 16bits per pixels in DICOM format (.dcm) and arithmetic and Huffman algorithms was applied .

After applying compression algorithm a compression ratio and data rate were founded.

5.1 Huffman algorithm

The results of Huffman algorithm had shown greater effect in compression process of mammographic images than MRI and CT images. In three types of medical image the image before compress and decompress is identical. CT and MRI images shown in results tables below.

Table(5.1a)Huffman for MIAS data base mammographic images

No	Name	CR	Data rate
1	mdb013.pgm	3.7007	0.7298
2	mdb014.pgm	3.8890	0.7429
3	mdb015.pgm	3.4101	0.7068
4	mdb016.pgm	3.3468	0.7012
5	mdb017.pgm	3.3464	0.7012
6	mdb018.pgm	3.7301	0.7319
7	mdb019.pgm	3.9151	0.7446
8	mdb020.pgm	3.5612	0.7192
9	mdb021.pgm	3.4111	0.7068
10	mdb022.pgm	4.2597	0.7652
11	mdb023.pgm	4.3991	0.7727
12	mdb024.pgm	3.3884	0.7049

13	mdb025.pgm	2.7092	0.6309
14	mdb026.pgm	2.4616	0.5938
15	mdb027.pgm	2.4809	0.5969
16	mdb028.pgm	2.6927	0.6286
17	mdb029.pgm	2.8164	0.6449
18	mdb030.pgm	3.5929	0.7217
19	mdb031.pgm	3.6127	0.7232
20	mdb032.pgm	2.7749	0.6396
	Mean	3.37495	0.69534
	Standard deviation	0.546682	0.053705

As shown in table(5.1a) it's found that the algorithm has mean compression ratio of3.37495 with standard deviation 0.546682, whereas the mean value of data rate 0.69534with standard deviation 0.053705.

Table(5.1b) huffman for MRI images:

No	Name	CR	Data rate
1	IM1	2.9498	0.6610
2	IM2	2.9158	0.6570
3	IM3	2.9913	0.6657
4	IM4	2.9988	0.6665
5	IM5	3.0584	0.6730
6	IM6	3.0425	0.6713
7	IM7	2.9894	0.6655
8	IM8	2.9956	0.6662
9	IM9	2.9798	0.6644
10	IM10	3.0451	0.6716
11	IM11	3.0339	0.6704
12	IM12	3.0238	0.6693
13	IM13	3.0364	0.6707
14	IM14	3.0822	0.6756
15	IM15	3.1636	0.6839
16	IM16	3.1398	0.6815
17	IM17	3.3068	0.6976
18	IM18	3.4225	0.7078
19	IM19	3.4730	0.7121
20	IM20	3.5151	0.7155
	Mean	3.10818	0.67733
	Standard deviation	0.177885	0.017285

When we applied Huffman algorithm on MRI image its found mean of compression ratio for MRI image equals to 3.10818 with standard deviation 0.177885, and mean value of data rate equal to 0.67733with standard deviation 0.017285 ,table(5.1b).

Table(5.1c)Huffman for CT images

No	Name	CR	Data rate
1	CT1	2.9639	0.6626
2	CT2	2.9383	0.6597
3	CT3	2.9368	0.6595
4	CT4	2.9434	0.6603
5	CT5	2.9510	0.6611
6	CT6	2.9596	0.6621
7	CT7	2.9693	0.6632
8	CT8	2.9748	0.6638
9	CT9	2.9742	0.6638
10	CT10	2.9719	0.6635
	Mean	2.95832	0.66196
	Stander deviation	0.014913	0.001692

In implement of Huffman algorithm on CT images found mean of compression ratio for CT equal to 2.95832 with standard deviation 0.014913, and mean value of data rate equal to0.66196 with stander deviation 0.001692, table (5.1c).

5.2Arithmetic algorithm

The Arithmetic algorithm had shown greater effect in compression process of mammographic images than MRI and CT images that showed below in tables (5.2a), (5.2b) and (5.2c).

Table(5.2a) Arithmetic for MIAS data base mammographic images

No	Name	CR	Data Rate
1	mdb013.pgm	3.8624	0.7411
2	mdb014.pgm	4.1146	0.7570
3	mdb015.pgm	3.4359	0.7090
4	mdb016.pgm	3.3629	0.7026
5	mdb017.pgm	3.3637	0.7027
6	mdb018.pgm	3.8395	0.7396
7	mdb019.pgm	4.0586	0.7536
8	mdb020.pgm	3.6730	0.7277
9	mdb021.pgm	3.4716	0.7120
10	mdb022.pgm	4.6716	0.7859
11	mdb023.pgm	4.9087	0.7963
12	mdb024.pgm	3.4411	0.7094
13	mdb025.pgm	2.7395	0.6350
14	mdb026.pgm	2.4897	0.5983
15	mdb027.pgm	2.5123	0.6020
16	mdb028.pgm	2.7299	0.6337
17	mdb029.pgm	2.8397	0.6479
18	mdb030.pgm	3.7039	0.7300
19	mdb031.pgm	3.7255	0.7316
20	mdb032.pgm	2.7996	0.6428
	mean	3.474575	0.70291
	Standard deviation	0.660799	0.05777

Table (5.2b) arithmetic coding algorithm for MRI images

No	Name	CR	Data Rate
1	IM1	3.0339	0.6704
2	IM2	3.0069	0.6674
3	IM3	3.0112	0.6679
4	IM4	3.0188	0.6687
5	IM5	3.0809	0.6754

6	IM6	3.0645	0.6737
7	IM7	3.0104	0.6678
8	IM8	3.0174	0.6686
9	IM9	3.0014	0.6668
10	IM10	3.0673	0.6740
11	IM11	3.0558	0.6728
12	IM12	3.0461	0.6717
13	IM13	3.0585	0.6730
14	IM14	3.1053	0.6780
15	IM15	3.1910	0.6866
16	IM16	3.2852	0.6956
17	IM17	3.3400	0.7006
18	IM18	3.4567	0.7107
19	IM19	3.5080	0.7149
20	IM20	3.5498	0.7183
	Mean	3.145455	0.681145
	Standard deviation	0.180098	0.017108

Table(5.2c) Arithmetic for CT images

No	Name	CT	Data rate
1	CT1	2.9832	0.6648
2	CT2	2.9542	0.6615
3	CT3	2.9520	0.6612
4	CT4	2.9598	0.6621
5	CT5	2.9675	0.6630
6	CT6	2.9778	0.6642
7	CT7	2.9886	0.6654
8	CT8	2.9943	0.6660
9	CT9	2.9938	0.6660
10	CT10	2.9916	0.6657
	mean	2.97628	0.66399
	Standard deviation	0.016647	0.001892

From above tables (5.2a) , (5.2b) and (5.2c) found that , the mean of compression ratio of MRI was equal to 3.145455 with standard deviation 0.180098, And mean of data rate is 0.681145with standard deviation 0.017108.

When arithmetic algorithm was applied on of mammogram images it's found that the mean of compression ratio was equal to 3.474575 with standard deviation 0.660799, the mean of data rate was equal to 0.70291with standard deviation 0.05777. In implementation of arithmetic algorithm on CT images the result was founded whereas the mean of compression ratio is equal 2.97628with standard deviation 0.016647, mean of data rate is 0.66399 with standard deviation 0.001892.

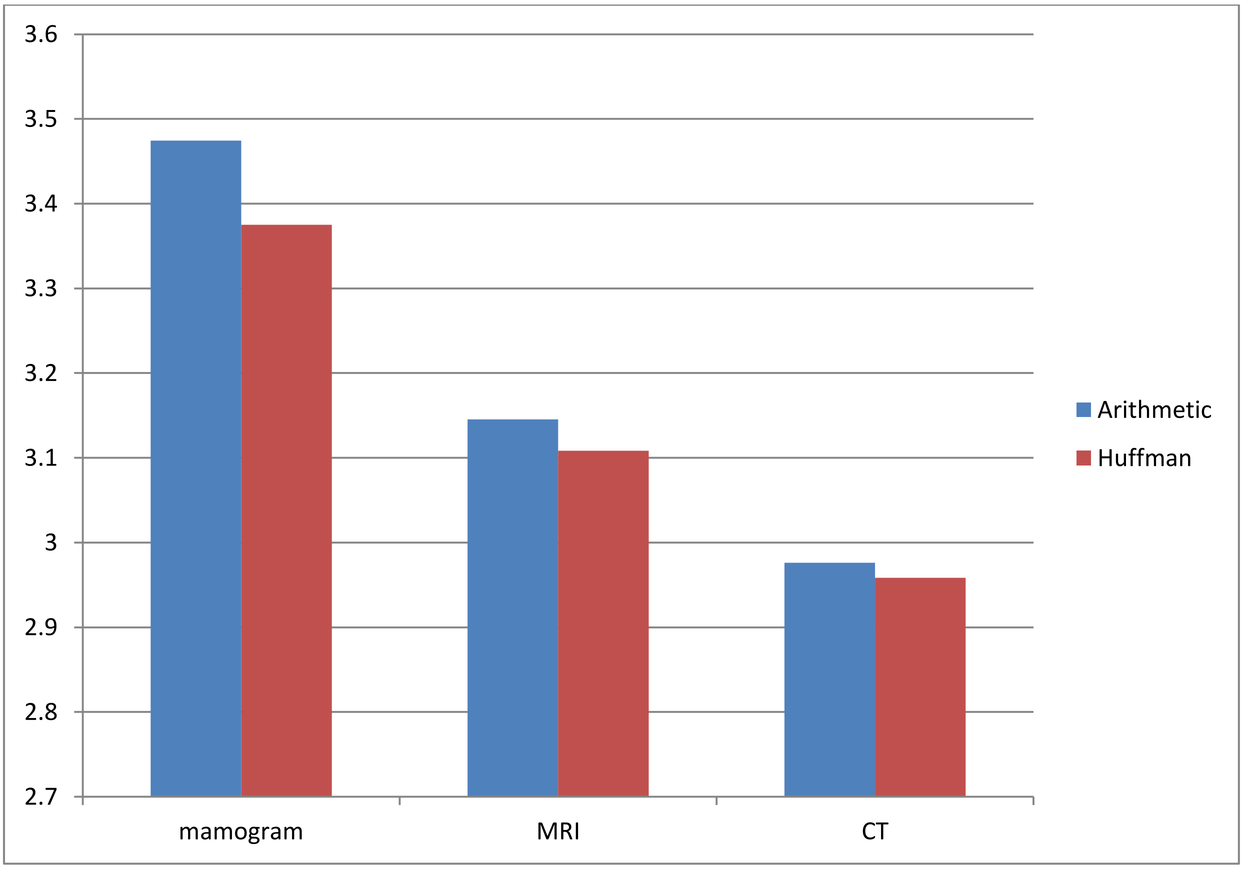
5.3 Compression between result of Arithmetic and Huffman algorithm

From implement arithmetic and Huffman algorithms on mammographic, MRI and CT images found that:

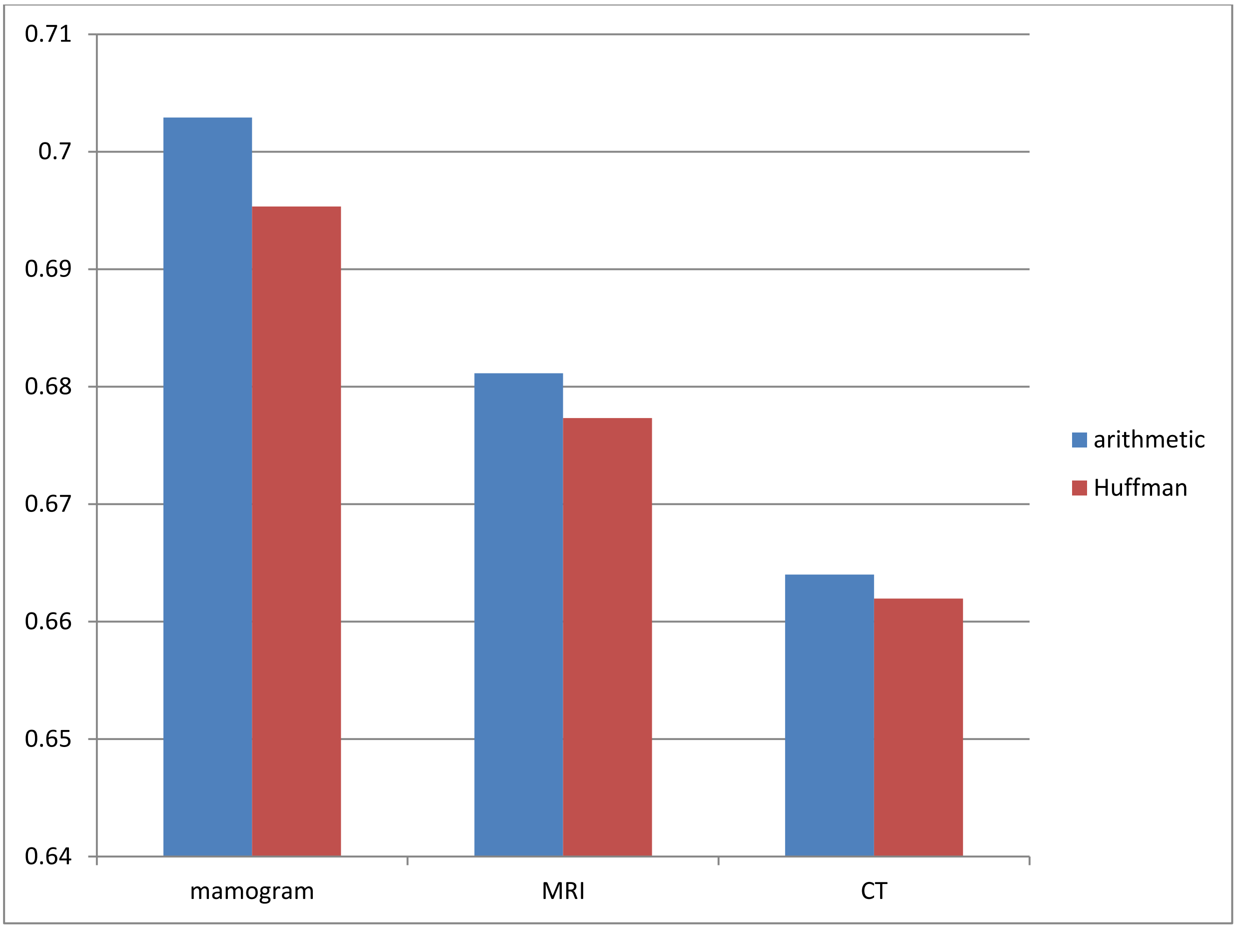
The compression ratio and data rate in Arithmetic algorithms is better than Huffman

Two algorithms give clear image after reconstructed from encoded image; and constructed image is identical to original image.

.



Fig(5.3a) compression ratio of (mammogram, MRI and CT images) using Huffman and arithmetic algorithms.



Fig(5.3b) data rate for mammogram ,MRI and CT images using arithmetic and Huffman algorithms .

Chapter 6

6.Conclusions and Recommendations

6.1 Conclusions

The main objective in this research was to decrease the size of medical image to be used in telemedicine and archiving systems.

In this research Huffman algorithm and arithmetic algorithm were applied on medical images.

MATLAB programming environment was used to apply this algorithm on 20 mammograms images, 10 CT images and 20 MRI images.

In two algorithms images were encoded by generate a dictionary which gave every pixel specific entry; the process is called encoding process and aims to reduce the size of image.

Then, dictionary was used along with size of original images to reconstruct back the image and return it to its original status (decoding process), and measure the performance of each algorithm by calculated compression ratio, data rate, peak signal to noise ratio and mean squire error

In two algorithms mean of peak signal to noise ratio was equal to infinity and mean of mean square error was equal to zero which means that the reconstructed image was identical to the original image.

The two algorithms were performing with percentage compression ratio mean of (31.1 %) , (29.2%) , (33.7%) for MRI , mammographic and CT images respectively.

6.2 Recommendations

Following are a few recommendations that could help in the extension of this work

- Try to send encoded image over internet and calculate bit error
- Use it for long term archiving of medical images and make data base.
- Find a new method to save the encoded image and dictionary in the same file.

References

[1] Janet, J., Mohandass, D. and Meenalosini, S., 2011. Lossless compression techniques for Medical images in Telemedicine. In Advances in Telemedicine: Technologies, Enabling ..Factors and Scenarios. InTech

[2] Weiser, M., 1991. The Computer for the 21 st Century. Scientific american, 265(3), .pp.94-105

[3] Vijayvargiya, G., Silakari, S. and Pandey, R., 2013. A survey: various techniques of image " .".compression. arXiv preprint arXiv:1311.6877

[4] KOTAK, H.B. and VALAKI, S.A., A STUDY OF IMAGE COMPRESSION TECHNIQUES AND ITS " .".APPLICATION IN TELEMEDICINE AND TELECONSULTATION

[5] ..Sayood, K., 2017. Introduction to data compression. Morgan Kaufmann

[6] Sharma, M., 2010. Compression using Huffman coding. IJCSNS International Journal of " .".Computer Science and Network Security, 10(5), pp.133-141

[7] Thomas, D.S., Moorthi, M. and Muthalagu, R., 2014. Medical image compression based " on automated ROI selection for telemedicine application. International journal of .".engineering and computer science, 3, pp.3638-3642

[8] Pianykh, O.S., 2009. Digital imaging and communications in medicine (DICOM): a ..Business Media & practical introduction and survival guide. Springer Science

[9] Dreyer, K.J., Hirschhorn, D.S., Thrall, J.H. and PACS, M., 2006. A guide to the digital ..revolution. USA: Springer Science+ Business Media

[10] .systems information hospital of study comparative and Analytical .2007 ،.M ،Aghajani .:(4)11 ،Tazkie Tebo

[11] Aghazadeh, S., Aliyev, A. and Ebrahimnezhad, M., 2012. Review the role of hospital " information systems in medical services development. International Journal of .".Computer Theory and Engineering, 4(6), p.866

[12] Qazisaeed, M., Safdari, R. and Davarpanah, A., 2006. Management of Medical and ..Health information. Tehran: Tehran University of Medical Sciences, 215

[13] The role and importance of hospital information systems in hospital “R. B. Azar, 2002

in Proceedings of National Seminar on Hospital Management, Tehran, ،managment .Emam Hosein University

[14] Kinsey, T.V., Horton, M.C. and Lewis, T.E., 2000. Interfacing the PACS and the HIS: " .".results of a 5-year implementation. Radiographics, 20(3), pp.883-891

[15] Masood, S., Sharif, M., Yasmin, M., Raza, M. and Mohsin, S., 2013. Brain image " .".Compression: A brief survey. Research Journal of Applied Sciences, 5(1), pp.49-59

[16] T Books, & Nelson, M. and Gailly, J.L., 1995. The data compression book 2nd edition. M ..New York, NY

[17] Gonzalez, R.C., Woods, R.E. and Eddins, S.L., 2009. Digital image processing using ..MATLAB, Gatesmark, USA

[18] Heng-Jian, L. and Jia-Shu, Z., 2010. Embedding adaptive arithmetic coder in chaos-" .".based cryptography. Chinese Physics B, 19(5), p.050508

[19] Vaishnav, M., Kamargaonkar, C. and Sharma, M., 2017. Medical Image Compression " .".Using Dual Tree Complex Wavelet Transform and Arithmetic Coding Technique

[20] Markandeya, N. and Patil, S., 2017. Image Compression Using Huffman Coding. " .".(International Journal Of Engineering And Computer Science, 6(1

[21] Llinàs, F. and -W., Aulí.J., Marcellin, M ،Bartrina-Rapesta, J., Sanchez, V., Serra-Sagristà" Blanes, I., 2017, September. Lossless medical image compression through lightweight *In Applications of Digital Image Processing XL (Vol. 10396, p.* "،.binary arithmetic coding *..103960S). International Society for Optics and Photonics*

[22] Liu, F., Hernandez-Cabronero, M., Sanchez, V., Marcellin, M.W. and Bilgin, A., 2017. " The Current Role of Image Compression Standards in Medical Imaging. Information, .".8(4), p.131

[23] Anand, S.P., 2017. Comparative Study on Lossless Data Compression Techniques. " .".International Journal of Scientific Research and Management, 5(8), pp.6630-6637

[24] Kaur, S. and Singh, S., 2016. Entropy Coding and Different Coding Techniques. Journal " of Network Communications and Emerging Technologies (JNCET) www. jncet. org, .".(6(5

[25] Masmoudi, A., Puech, W. and Masmoudi, A., 2015. An improved lossless image " compression based arithmetic coding using mixture of non-parametric distributions. .".Multimedia Tools and Applications, 74(23), pp.10605-10619

[26] Masmoudi, A. and Masmoudi, A., 2015. A new arithmetic coding model for a block-" based lossless image compression based on exploiting inter-block correlation. Signal,

."Image and Video Processing, 9(5), pp.1021-1027

[27] Lin, Q., Wong, K.W. and Chen, J., 2013. An enhanced variable-length arithmetic coding " and encryption scheme using chaotic maps. Journal of Systems and Software, 86(5), .".pp.1384-1389

[28] Shahbahrami, A., Bahrampour, R., Rostami, M.S. and Mobarhan, M.A., 2011. Evaluation " of huffman and arithmetic algorithms for multimedia compression standards. arXiv .".preprint arXiv:1109.0216

[29] .https://www.dicomlibrary.com/meddream/md5/?study

[30] Michael W. Marcellin, ،Bartrina-Rapesta, Joan, Victor Sanchez, Joan Serra-Sagristà" Llinàs, and Ian Blanes. "Lossless medical image compression through -Francesc Aulí lightweight binary arithmetic coding." In Applications of Digital Image Processing XL, ."vol

[31] A STUDY OF IMAGE COMPRESSION TECHNIQUESAND ITS APPLICATION IN TELEMEDICINE ANDTELECONSULTATION1 HIMALI B. KOTAK, 2 SANJAY A. VALAKI1, 2 Department of Computer Engineering, Government Polytechnic, Bhuj, .GujaratTecnological University, Gujarat. India.san