WATERMARKING OF SINGLE AND MULTI FRAME ULTRASOUND MEDICAL IMAGES

GRAN BADSHAH

Thesis submitted in fulfillment of the requirements
for the award of the degree of
Doctor of Philosophy in Computer Science

Faculty of Computer Systems and Software Engineering
UNIVERSITI MALAYSIA PAHANG

FEBRUARY 2016

ABSTRACT

Digital communication of medical images may cause image manipulation for non-mandatory purposes. The originality of single and multi-frame medical images is strictly required to be used for patient health recovery. Watermarking is one of the latest security techniques, which can be used for protection of digital images. This dissertation has focused on watermarking of grayscale eight bits single and multi-frame ultrasound medical images. Watermarking of a medical image with heavy payload causes image degradation and directly affects patient diagnosis. Therefore, watermark must be lossless compressed to reduce its size without data loss. In the past Lempel-Ziv Welch (LZW) has been used for text, image pixel data lossless compression and in this research we extend this technique to binary watermark lossless compression to enhance the performance of region based watermarking techniques. In region based watermarking, an image is divided into region of interest (ROI) and region of non-interest (RONI). LZW compressed watermark based, a new region based watermarking scheme, tamper detection and recovery watermarking of medical images (TDARWMI) is developed. Produced results of the scheme are compared with those of other watermarking schemes and are found better in compression ratio, PSNR of watermarked and recovered images. In region based watermarking schemes, the tamper detection and recovery is restricted to ROI, while rest of the image remains under security threats. Possible wrong selection of ROI is another limitation of region based watermarking schemes. Non-proper selection of ROI may create problems during recovery at destination and can affect the required analysis. Therefore, another watermarking scheme, whole image authentication, tampers detection, localization and complete recovery (WIATDACR) is developed to provide security to a complete image. The LZW compressed watermark is embedded into image's ROI and RONI parts randomly. At destination, watermark is extracted and used for image authentication, tampers detection, localization and lossless recovery. This scheme can locate tampers occurred at any part of image and can perform recovery without data loss. The signature based security technique of DICOM multi-frame images has deficiency in tamper detection and recovery. Multi-frame image authentication, tamper detection, localization and recovery (MFATDLAR) watermarking scheme is formulated to facilitate a DICOM image's frame-wise authentication, tamper detection and lossless recovery. In this scheme the watermark of every frame is prepared, compressed and encapsulated into respective frame's RONI part. At destination the extracted watermark is used for the frame's authentication, tamper detection, localization and lossless recovery.

ABSTRAK

Komunikasi digital imej perubatan boleh menyebabkan manipulasi imej untuk tujuan bukan mandatori. Tunggal dan multi-frame imej perubatan keaslian adalah dilarang sama perlu digunakan untuk pemulihan kesihatan pesakit. Watermarking adalah salah satu teknik keselamatan yang terkini, boleh digunakan untuk perlindungan imej digital. Disertasi ini telah memberi tumpuan kepada watermarking skala kelabu lapan bit tunggal dan multi-frame ultrasound imej perubatan. Watermarking daripada imej perubatan dengan muatan berat menyebabkan degradasi imej dan secara langsung memberi kesan diagnosis pesakit. Oleh itu, watermark mesti lossless dimampatkan untuk mengurangkan saiznya tanpa kehilangan data. Lalu Lempel-Ziv Welch (LZW) telah digunakan untuk teks, imej piksel mampatan lossless data dan dalam kajian ini kita melanjutkan teknik ini kepada perduaan mampatan lossless watermark untuk meningkatkan prestasi teknik watermarking rantau berasaskan. Di rantau berdasarkan watermarking, imej dibahagikan kepada kawasan kepentingan (ROI) dan rantau bukan faedah (RONI). LZW watermark dimampatkan berasaskan, skim watermarking berdasarkan kawasan baru, pengesanan bega dan pemulihan watermarking imej perubatan (TDARWMI) dibangunkan. Keputusan yang dihasilkan skim ini dibandingkan dengan keputusan skim watermarking lain dan didapati lebih baik dalam nisbah mampatan, PSNR imej tera air dan didapatkan. Dalam skim watermarking rantau berdasarkan pengesanan bega dan pemulihan adalah terhad kepada ROI, manakala selebihnya imej kekal di bawah ancaman keselamatan. pemilihan salah kemungkinan ROI adalah satu lagi had skim watermarking rantau berasaskan. Bukan betul penandaan ROI boleh menimbulkan masalah di tempat tujuan dan boleh memberi kesan kepada analisis yang diperlukan. Oleh itu, satu lagi skim watermarking, pengesahan imej keseluruhan, mengganggu pengesanan tepat dan lengkap pemulihan (WIATDACR) dibangunkan untuk menyediakan keselamatan untuk imej yang lengkap. The LZW watermark dimampatkan tertanam ke dalam ROI imej dan bahagian-bahagian RONI rawak. Di destinasi tera air diekstrak dan digunakan untuk pengesahan imej, mengganggu pengesanan tepat dan pemulihan lossless. Skim ini boleh mengesan mengganggu berlaku di mana-mana bahagian imej dan boleh melakukan pemulihan tanpa kehilangan data. Keselamatan berdasarkan tandatangan DICOM imej pelbagai bingkai mempunyai pengesanan bega dan kekurangan pemulihan. DICOM skim pengesahan imej multi-frame, pengesanan bega, penyetempatan dan pemulihan watermarking (MFATDLAR) dibangunkan untuk membuat mungkin setiap pengesahan bingkai imej ini, pengesanan bega dan pemulihan lossless. A watermark setiap bingkai disediakan, dimampatkan dan dimuatkan ke dalam bingkai masing RONI bahagian. Di destinasi watermark yang diekstrak digunakan untuk pengesahan bingkai, pengesanan bega, penyetempatan dan pemulihan lossless.

TABLE OF CONTENTS **Page**

LIST OF TABLES

LIST OF FIGURES

LIST OF ALGORITHMS

LIST OF ABBREVIATIONS

AES	Advance Encryption Standard
ANN	Artificial Neural Network
CRC	Cyclic Redundancy Check
CT	Computed Tomography
Db	Decibel
DCT	Direct Cosine Transformation
DFT	Direct Fourier Transformation
DICOM	Digital Imaging and Communication in Medicine
DS	Digital Signature
DWT	Direct Wavelet Transform
EPR	Electronic Patient Record
GIF	Graphic Interchange Format
HIS	Hospital Information System
IEEE	Institute of Electrical and Electronics Engineering
IRVL	Image Row Value Length
JPEG	Joint Photographic Expert Group
LBP	Local Binary Pattern
LSB	Least Significant Bit
LZW	Lempel-Ziv Welch
MAC	Machine Authentication Code
MDC	Manipulation Detection Code
MLP	Multilayer back Propagation
MFATDLR	Multi-frame authentication, tamper detection, localization and lossless recovery

MRI	Magnetic Resonance Imaging
MSE	Mean Square Error
NASA	National Aeronautics and Space Administration
NM	Nuclear Medicine
PACS	Picture Archive and Communication in Medicine System
PBM	Portable Bitmap
PET	Positron Emission Tomography
PNG	Portable Network Graphics
PSNR	Peak signal to noise ratio
RIS	Radiological Information System
ROI	Region of Interest
RONI	Region of Non-Interest
RSMA	ROI Segmentation and Multilevel Authentication
SHA	Secure Hash Algorithm
SNR	Signal to Noise Ratio
TALLOR	Tamper Localization and Lossless Recovery
TDARWMI	Tamper Detection and Recovery Watermarking of Medical Images
US	Ultrasound
VPN	Virtual Private Network
WIATDACR	Whole Image as Watermark for Authentication and Tamper Detection, Localization and Complete Recovery

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

History of the world is a witness of changes in technologies. Changes are made to upgrade the technologies but some upgradations have negative effects on their related applications. For example the invention and development of Internet technology has replaced formal, complex, costly and time consuming data communication system by an automatic, easy, cheaper and real-time responding one with data security threats. Internet technology has interconnected people around the globe to store, share, access and communicate data digitally with chances for hacker's manipulation. Digitization of data has facilitated data communication, effective storage, duplication, transportation with cost reduction. The use of digital technology has replaced data centralized records storage by decentralized one through globally accessible Internet with increase of security demands. Easy and open accessibility of Internet data may result into illegal alteration. Data alteration may happen intentionally and non-intentionally but affect the achievement of dedicated goals. Data alteration happens due to the unauthorized accesses or use of noisy communication channels. The illegal change in data originality is known as tampering. Data tampering causes problems of ownership, integrity, confidentiality and authentication of data. Digital medical image tampering is a problem that has been faced in electronic healthcare environment. Due to the lack of security unwanted outcomes may be produced and can loss the important digital image information. For more than last two decades, different security measures including digital watermarking have been tested to solve this problem (Abd-Eldayem, 2013). Watermarking is an efficient technique to be used for Internet based medical image's authentication, tamper detection, localization and lossless recovery.

A typical watermark is a meaningful identification mark such as logo, fingerprint, sequence of text or numbers, a secret code, an image or part of image used for product identification, ownership proof, authentication and prevention of illegal distribution, duplication (Holliman and Memon, 2000). A watermark size is not restricted to a limit and can be long as to cover all the features of a product. Watermarks can be pasted onto or encapsulated into products for achievement of defined goals. The process of watermark pasting or encapsulation is called watermarking (Memon, 2010). Digital medical image security is very important in electronic healthcare system because pixels of a digital image can be easily manipulated due to easy access to the Internet. A watermarking technique should have the capability to detect, locate and restore such manipulations of medical images (Langelaar et al., 2000).

An efficient watermarking technique has the ability to keep the image data under protection even after image communication and watermark extraction. Watermarking technique is applicable to audio, video, text, image and other possible digital product. A product after watermark encapsulation is called watermarked product. For example, the watermark embedded image is called watermarked image. A watermark is embedded in a product for the aims of copyright protection, fingerprinting, integration, broadcast monitoring and authentication (Zhao et al., 2011). Copyright watermarks are used to prove the product ownership through backtracking, fingerprinting and to determine the source of illegal distribution. Copy protection watermarks prevent illegal copy of products while broadcast monitoring watermarks are used to control and ensure the agreed upon advertisements processing. In data hiding watermarking, the coded data is hided into a product for its safe sending to destination. In case of medical image watermarking the code may comprise of image metadata or diagnostic information or both. If image is tampered during communication then at destination tampers can be localized and recovered without losing any important data. Different watermarking techniques are available for image tamper detection, localization and recovery (Nyeem et al., 2012). These techniques are reviewed in the next chapter of literature review in details. The remaining chapters explore that how tamper detection, localization and recovery is performed in our proposed methodologies.

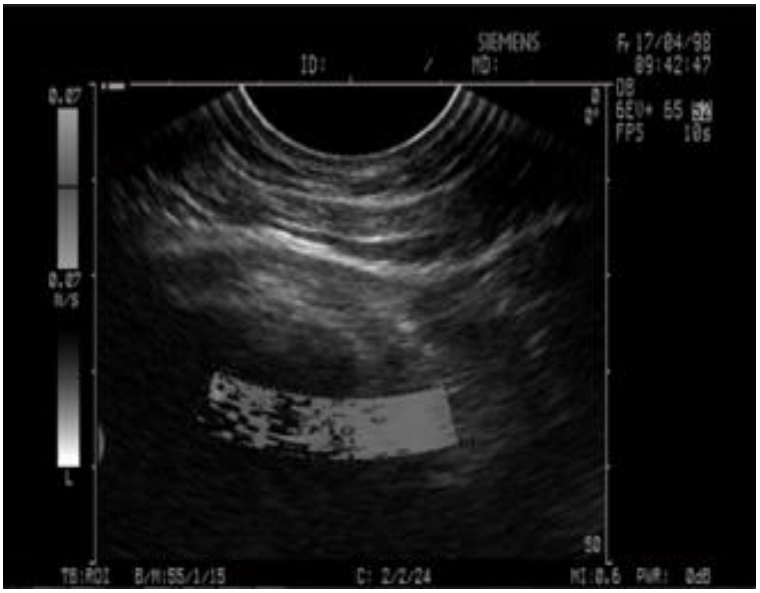

Figure 1.1: Ultrasound original image

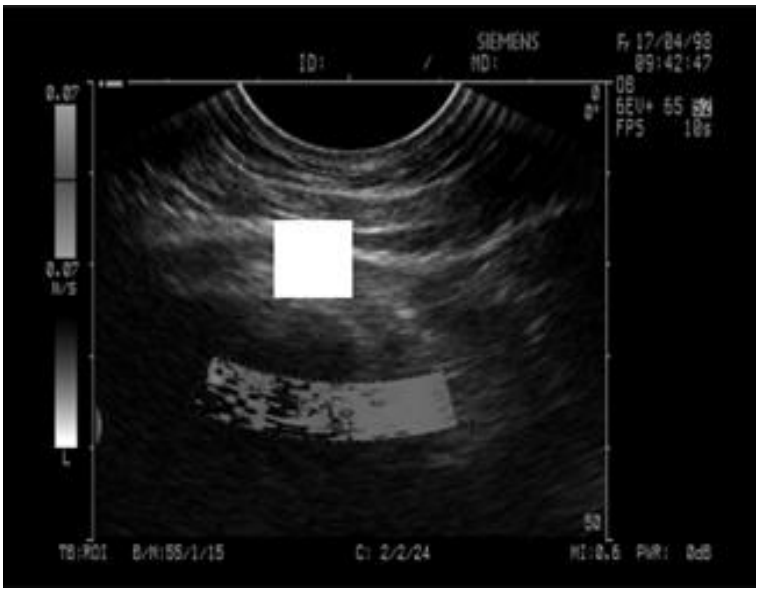

Figure 1.2: Ultrasound tampered image

Figures 1.1 and Figure 1.2, show the original and its tampered version of an ultrasound medical image respectively. The tampered image has been obtained from cropping the original image through the ImageJ software to show it as an example of image tampering in teleradiology environment. During patient's diagnosis, if the

doctors consider the cropped spot as an original part of image then it will surely mislead the final results and can put the patient life in danger.

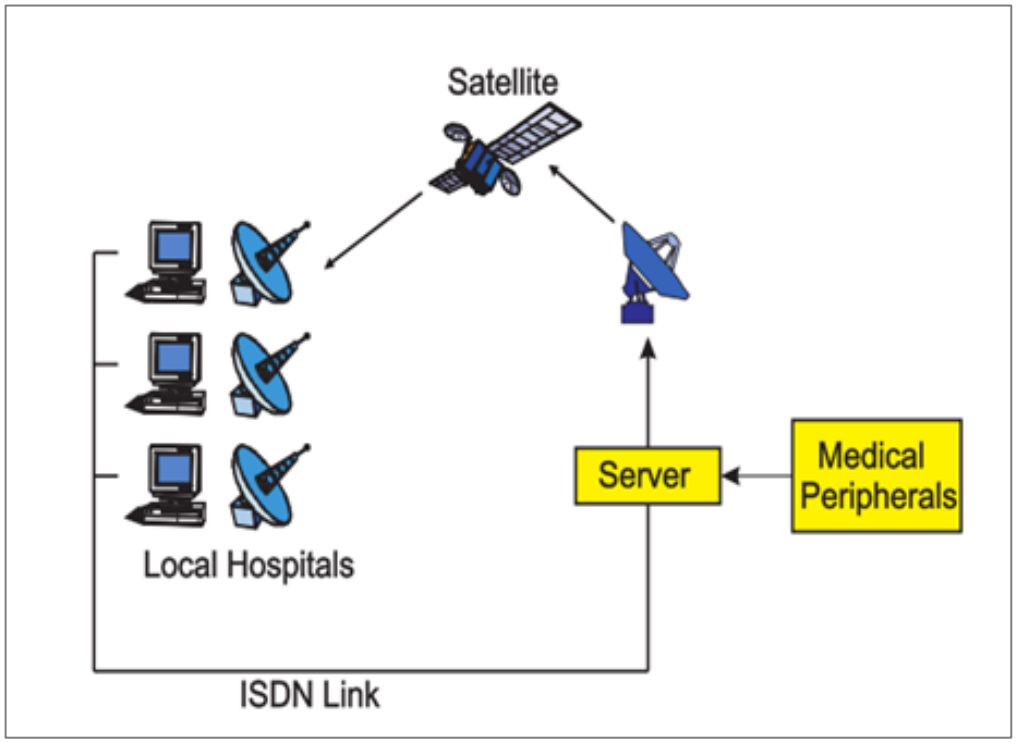

Figure 1.3: Different components involved with web-based medical image processing in distributed health care system

Figure 1.3 shows a diagram with visible main components used in research work related to distributed health care system located at Yonsei university of South Korea. The components links show the access mechanism of a medicle image specialist system at remote hospitals using satellite technology in telemedicine (Osborne, 2005). This is an example of picture archive and communication in medicine system (PACS) for the study of medical images. This system is basically derived from National Aeronautics and Space Administration (NASA), where the astronaut's medical information were collected on earth from space in audio and video formats.

1.2 WATERMARKING POSITION IN HIERARCHY OF DATA SECURITY TECHNIQUES

Data security remains a hot research topic for all the times because new security measures are needed with evaluation and development of communication technologies. Data security literature reports a number of techniques used for centralized and distributed data security. Figure 1.4 shows the hierarchical structure of data security techniques including watermarking (Cheddad et al., 2010). Data security techniques are basically divided into two types, cryptography and information hiding.

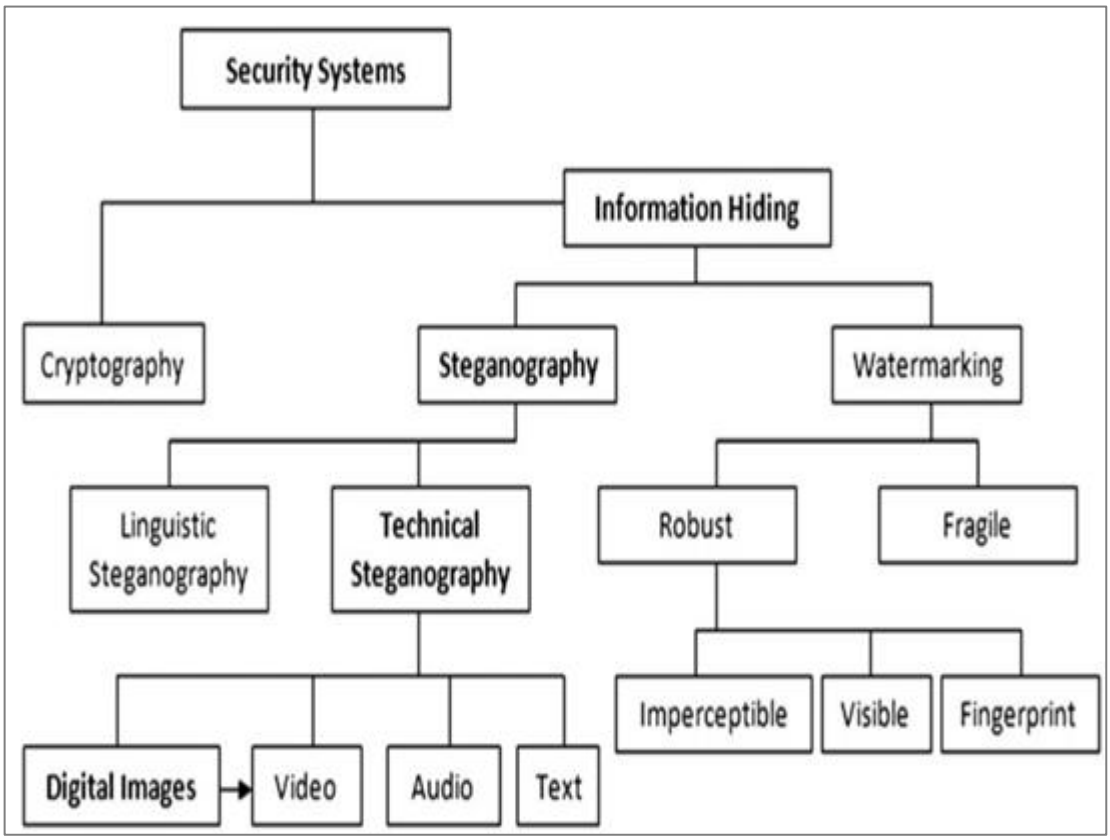

Figure 1.4: Hierarchical structure of data security techniques

Cryptographic techniques are used to get data in a reduced coded format as compared to the original data. The code is understandable only to the related people and remains secured from the outside people's understanding. Cryptography is a Greek word means secret writing; different functions are applied to data to get their secret codes. A MATLAB supported function SHA-256 is an example of such functions.

SHA-256 produces a 64-bytes hexadecimal code after application to data. The resulted code is used mostly for respective data authentication. The main limitation of a cryptographic technique is the only reporting of data authentication and has no capability to detect, locate and recover the tampers. Information hiding based security techniques are used to hide important information within a product. Hiding of an image data into image itself is an example of information hiding. The image serves as a carrier of hided data to be used at destination for intended purpose. Information hiding techniques are further divisible into steganography and watermarking. Steganography is also a Greek word means covered writing. Steganography data is the actual data hidden into a carrier without resulting into coded form like cryptography. Steganography aims to transfer secret data to destination with high security. A carrier medium is cared highly for the successful delivery of secret data and does not remain important after the hidden data extraction.

Watermarking is also a data hiding technique, capable to hide and extract information with intensive care of medium because medium remains the main focus of watermarking process. For example a medical image watermark is used for image authentication, tamper detection, localization and recovery. The actual mean of a medical image watermarking is to ensure the image originality for patient health analysis. During watermarking, the important data is hided into images, audios, videos and texts as watermarks according to the requirements. A watermarking technique is robust or fragile in nature; robust watermarking focuses on watermark security while the fragile one is used for product contents security. Imperceptibility, visibility and fingerprinting describe different levels and types of image watermarking. Imperceptibility means how to make a watermark data non-influential to the originality of an image. Visibility means how to keep a watermark data invisible to the human sight to increase its security. Fingerprinting is used to back-track the process of data illegal distribution to reach the illegal distributor of important data. A secret key can be used for making a watermarking scheme more secured and efficient.

Table 1.1: Comparison of watermarking, cryptography and steganography based on their properties

S.no	Property	Watermarking	Cryptography	Steganography
1	History	Modern era, started after the creation of digital images	Modern era but earlier than digital watermarking	Ancient, invented before the creation of digital images
2	Secret data	Watermark	Plain text	Payload
3	Secret carrier	Image, audio and video files	Image and text	Any digital and non-digital media
4	Key	Optional	Necessary	Optional
5	Input file	At least one	Always one	At least one
6	Detection	Blind or non-blind	Blind	Blind
7	Result	Watermarked image or file	Cipher text or image	Stego image or file
8	Main purpose	Authentication, tamper detection, Recovery and copyrights protection	Data authentication	Data secret communication
9	Visibility	Visible and non-visible	Always visible	Always non-visible
10	Failure	If removed, demaged or altered	If decrypted	If detected
11	Choice of carrier selection	Restricted to focused image, video or audio file	Not available	Not restricted to a specific carrier
12	Carrier and secret data relationship	Carrier is more important while secret data remains its important attribute	No relation at all	Secret data is more important than carrier
13	Attack causes	Image processing operations	Cryptanalysis	Stego analysis
14	Main concern	Robustness	Robustness	Capacity

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

History of the world is a witness of changes in technologies. Changes are made to upgrade the technologies but some upgradations have negative effects on their related applications. For example the invention and development of Internet technology has replaced formal, complex, costly and time consuming data communication system by an automatic, easy, cheaper and real-time responding one with data security threats. Internet technology has interconnected people around the globe to store, share, access and communicate data digitally with chances for hacker's manipulation. Digitization of data has facilitated data communication, effective storage, duplication, transportation with cost reduction. The use of digital technology has replaced data centralized records storage by decentralized one through globally accessible Internet with increase of security demands. Easy and open accessibility of Internet data may result into illegal alteration. Data alteration may happen intentionally and non-intentionally but affect the achievement of dedicated goals. Data alteration happens due to the unauthorized accesses or use of noisy communication channels. The illegal change in data originality is known as tampering. Data tampering causes problems of ownership, integrity, confidentiality and authentication of data. Digital medical image tampering is a problem that has been faced in electronic healthcare environment. Due to the lack of security unwanted outcomes may be produced and can loss the important digital image information. For more than last two decades, different security measures including digital watermarking have been tested to solve this problem (Abd-Eldayem, 2013). Watermarking is an efficient technique to be used for Internet based medical image's authentication, tamper detection, localization and lossless recovery.

A typical watermark is a meaningful identification mark such as logo, fingerprint, sequence of text or numbers, a secret code, an image or part of image used for product identification, ownership proof, authentication and prevention of illegal distribution, duplication (Holliman and Memon, 2000). A watermark size is not restricted to a limit and can be long as to cover all the features of a product. Watermarks can be pasted onto or encapsulated into products for achievement of defined goals. The process of watermark pasting or encapsulation is called watermarking (Memon, 2010). Digital medical image security is very important in electronic healthcare system because pixels of a digital image can be easily manipulated due to easy access to the Internet. A watermarking technique should have the capability to detect, locate and restore such manipulations of medical images (Langelaar et al., 2000).

An efficient watermarking technique has the ability to keep the image data under protection even after image communication and watermark extraction. Watermarking technique is applicable to audio, video, text, image and other possible digital product. A product after watermark encapsulation is called watermarked product. For example, the watermark embedded image is called watermarked image. A watermark is embedded in a product for the aims of copyright protection, fingerprinting, integration, broadcast monitoring and authentication (Zhao et al., 2011). Copyright watermarks are used to prove the product ownership through backtracking, fingerprinting and to determine the source of illegal distribution. Copy protection watermarks prevent illegal copy of products while broadcast monitoring watermarks are used to control and ensure the agreed upon advertisements processing. In data hiding watermarking, the coded data is hided into a product for its safe sending to destination. In case of medical image watermarking the code may comprise of image metadata or diagnostic information or both. If image is tampered during communication then at destination tampers can be localized and recovered without losing any important data. Different watermarking techniques are available for image tamper detection, localization and recovery (Nyeem et al., 2012). These techniques are reviewed in the next chapter of literature review in details. The remaining chapters explore that how tamper detection, localization and recovery is performed in our proposed methodologies.

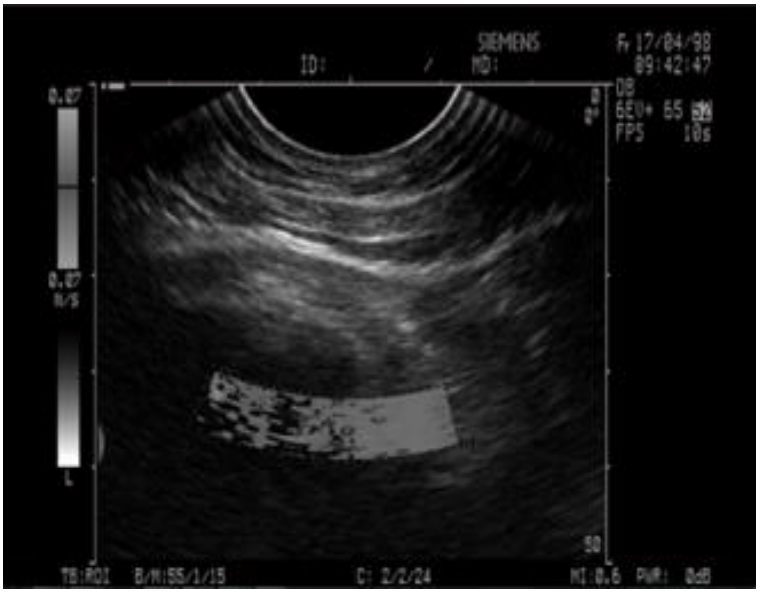

Figure 1.1: Ultrasound original image

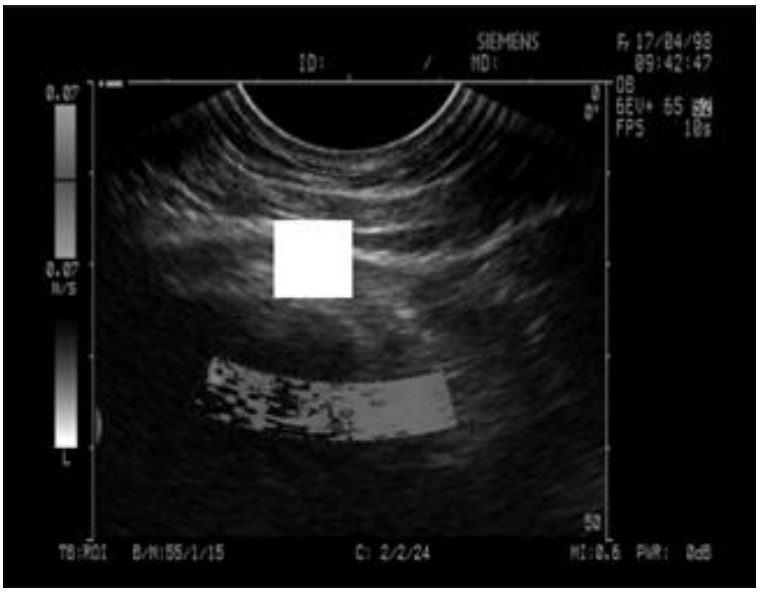

Figure 1.2: Ultrasound tampered image

Figures 1.1 and Figure 1.2, show the original and its tampered version of an ultrasound medical image respectively. The tampered image has been obtained from cropping the original image through the ImageJ software to show it as an example of image tampering in teleradiology environment. During patient's diagnosis, if the

CHAPTER 3

WATERMARK COMPRESSION IN DIGITAL IMAGE WATERMARKING

3.1 INTRODUCTION

Digital imaging techniques produce different size images, suitable for particular applications. The images out of proper size make problems for some secondary applications. For example large size medical images require more storage space, need more transmission time and bandwidth in teleradiology. Besides these obstacles, during image communication an image contents may be changed due to noisy communication channels or hackers manipulation. Medical image data is very sensitive and cannot afford changes to its original dataset. Digital watermarking techniques have been used to detect and recover illegal changes made to images. The watermarking of a medical image with heavy payload causes image perceptual degradation and affects patient diagnosis. To reduce watermark payload and maintain the image perceptual, diagnostic qualities standard, the watermark needs to be compressed. The compression should be in such a way to ensure the compressed data originality. For sensitive data like medical image watermarks, lossless compression is used. Watermark lossless compression ensures no loss of data while reducing the size of dataset. Watermarking of medical images with reduced number of bits preserves the image perceptual and diagnostic qualities. This chapter focuses on watermark lossless compression using different techniques including LZW. LZW compression technique is also applicable to binary watermarks. The binary version watermarks contain of large number of repeating sequences of binaries and most suitable for lossless compression using LZW technique. LZW is a better approach to compress data having repeating sequences of data elements like binaries (Nelson et al., 1989).

3.2 METHODOLOGY

This chapter methodology is watermark lossless compression using LZW technique. For this purpose, ten different samples of ultrasound medical images are chosen for watermark generation and compression experiments. Each sample is divided into ROI and RONI; ROI is the central and rich in information part, while RONI is the remaining image. Watermark is generated from combination of image selected pixels as ROI and secret key. The LZW selection for watermark compression is made after its comparison with other conventional compression techniques such as PNG, GIF, PBM, JPEG and JPEG 2000 of lossless versions. This selection is made on the basis of more bits reduction and good compression ratio.

3.2.1 Watermark Generation and Compression

A 100 * 100 size pixels segment of each image sample is selected as ROI. ROI size is allowed to vary but here we keep it fixed for results comparison of different compression techniques. A watermarking secret key is generated and applied to watermark to get a secured watermark. Watermark security is necessary to hide the originality of this important data from attackers. Here watermark is the combination of ROI and ROI hash code, while secret key is the binary of ROI hash repeated equal to the watermark size. A XOR operation is performed between watermark and secret key to obtain the secured watermark. LZW is a dictionary based compression technique applied to data for elimination of repeating sequences (Nelson et al., 1989). PNG, GIF, PBM and JPEG are conventional compression methods have been tested on the same size watermark compression to compare the results with LZW. For LZW compression, first the watermark is converted to binary and stored in a single row binary array. This conversion gives binaries unique sequences for LZW effective compression. During compression, every unique sequence is replaced by a decimal number called string code. More number of bits reductions makes LZW as the best choice of binary watermark lossless compression in medical image watermarking.

Before starting LZW compression of a binary watermark, an array as a dictionary is initialized with two stings, ‘0’ and ‘1’. These are only two unique values in watermark binary stream otherwise the dictionary is initialized to the distinct characters in an alphabetic string. The combinations of binaries of binary watermark give different repeating sequences. During watermark compression process unique strings are formed based on binary sequences and inserted into dictionary. Every unique string is allotted a decimal code and inserted to another array called codes table. LZW Algorithm 3.1 checks the availability of newly constructed string in the dictionary. In case of its uniqueness, the string and its allotted code are inserted into dictionary codes table as shown in Figure 3.2. This process of unique strings formation, codes allocation and insertion continue till the whole watermark is compressed. At the completion, codes table values give the LZW compressed watermark. The following Algorithm 3.1 explains step by step process of watermark lossless compression. The final dictionary and codes table are used for watermark lossless decompression after the watermark extraction at destination (Alarabeyyat et al., 2012).

Algorithm 3.1: LZW algorithm for binary watermark compression

1. Convert watermark to binary
2. Dictionary={‘0’,’1’}
3. String = get the first binary from binary watermark
4. WHILE get the next available binary and continue
5. Next-binary = Get the next binary
6. IF String + Next-binary exist in dictionary then
7. String = String + Next-binary
8. ELSE
9. Assign decimal code to String and insert into code table
10. Add String + Next-binary to the dictionary
11. String=Next-binary
12. END of IF
13. END of WHILE
14. Output dictionary and code table