TAMPER LOCALIZATION FOR MEDICAL IMAGES USING CYCLIC REDUNDANCY CHECK (CRC)

WAN NUR SHAKIRA AIN BINTI WAN RANIZANG

TECHNICAL REPORT SUBMITTED IN FULFILMENT OF THE DEGREE OF COMPUTER SCIENCE

FACULTY OF COMPUTER SYSTEM AND SOFTWARE ENGINEERING

2013

ABSTRACT

The thesis focuses on the implementation of medical image watermarking in Digital Imaging and Communications in Medicine (DICOM). DICOM is an important part of information technology infrastructure in a health institution. Watermarking can be used to authenticate medical images and provide the additional security needed on top of the existing security measures that were already in place. Watermarking methods applied to medical images should be tampered or if not, an area known as Region Of Interest (ROI) needs to be defined on the image to retain the original information. The watermarked image produced should have visual quality similar to its original one. This thesis proposed a tamper localization watermarking for medical images using Cyclic Redundancy Check
(CRC). The confidentiality of medical images is increased by embedding CRC into image. I will use an ultrasound grayscale, CT-SCAN and also MRI. However, medical images require extreme care when embedding additional data within them because the additional information must not affect the image quality which leads to wrong diagnosis. So image is divided into ROI and Region Of Non Interest (RONI).

TABLE OF CONTENTS

LIST OF TABLES

Table Number **Page**

LIST OF FIGURES

Figure Number **Page**

LIST OF ABBREVIATIONS

DICOM	Digital Imaging and Communications in Medicine
LSB	Least Significant Bit
PSNR	Peak Signal-to-Noise Ratio
ROI	Region of Interest
RONI	Region of Non-Interest
MRI	Magnetic Resonance Imaging
CT- SCAN	Computerized Tomography
CRC	Cyclic Redundancy Check

INTRODUCTION

1.1 DIGITAL WATERMARKING

Digital watermarking started back in 1979 but it was not until 1990 that it gained popularity. A digital watermarking is a kind of marker convertly embedded in a noise tolerant signal such as audio or image data. It is typically used to identify ownership of the copyright of such signal. Watermarking is the process of computer- aided information hiding in a carrier signal, the hidden information should, but does not need to contain a relation to the carrier signal. It sometimes called “fingerprinting”, that allows copyright owners to incorporate into their work identifying information invisible to the human eye. When combined with new tracking services offered by some of the same companies that provide the watermarking technology, copyright owners can in theory find all illegal copies of their photos and music on the Internet and take appropriate legal action. It also be used to verify the authenticity or integrity of the carrier signal or to show the identity of its owners. Like traditional watermarks, digital watermarks are only perceptible under certain conditions, either after using some algorithm or imperceptible anytime else. Traditional watermarks may be applied to visible media , whereas in digital watermarking, the signal may be audio, pictures, video, texts or 3D models. A signal may carry several different watermarks at the same time. Unlike metadata that is added to the carrier signal, a digital watermark does not change the size of the carrier signal.

Both steganography and digital watermarking employ steganographic techniques to embed data convertly in noise signals. But whereas steganography aims for imperceptibility to human senses, digital watermarking tries to control the robustness as top priority. One application of digital watermarking is source tracking. A watermark is embedded into a digital signal at each point of distribution. If a copy of the work is found later, then the watermark may be retrieved from the copy and the source of the distribution is known.

1.2 MEDICAL IMAGING AND RECOVERY

Medical imaging is the technique and process used to create images of the human body or parts and function thereof for clinical purposes (medical procedures seeking to reveal, diagnose or examine disease) or medical science including the study of normal anatomy and physiology.

Medical images are produced by a wide variety of imaging equipment such as computed tomography (CT), magnetic resonance imaging, ultrasound and so on. Now, these medical images generally are stored in digital forms on different types of storage media such as compact discs (CDs) and digital versatile discs (DVDs). A patient can keep a copy of his medical image in a CD instead of a hard copy film. However, the digital images are very easy to modify by any image processing program in a computer. Hospitals, insurance companies, as well as patients might want to modify the image for various reasons. The tampered images may be used for illegal purposes. Therefore, how to protect a medical image, detect a tampered medical image and even recover the original image are important and urgent topics in the current digital age.

Image authentication can be achieved by embedding a message into the image and that embedded message is derived directly from the image itself. When an image will
be different from the original message. The authentication message can be called as a digital watermarking. The digital watermarking should be hidden into the image. However, the information hiding technique generally introduces some amount of noninvertible distortion in the image. The distortion might cause some legal problems for medical images. Recently, some lossless hiding techniques are proposed in some papers. In the lossless hiding techniques, the embedded images can be reversed into the original images without any distortion.

In general, only one authentication message is derived from the whole image. However, in this method, if an image is detected to be a tampered one, we do not have enough information to point out which region is modified. Hence, a block-based authentication technique is proposed for providing more detailed information. At first, a medical image is divided into several blocks and the authentication message of each

block is embedded into other blocks. Whether the block is tampered with can be checked by the message embedded in other blocks.

The lossless information hiding method cannot embed too many data into the images; it is too difficult to recover all regions in an image with accepted image quality. At the first proposed method, only an approximate small image when a block is detected to be tampered with. However, the recovered image is not good enough and the second method is proposed to embed the recovery information of only the most important region for diagnosis. The physician could indicate a region of interest (ROI) and the approximate image can be recovered from the information embedded in other blocks.

1.3 TAMPER LOCALIZATION WATERMARKING

Tamper localization is where the area of tampering can be identified. Once tampering is localized, tempered section can be recovered. Tampered image can be recovered by using embedded watermark that contain information of type original image. Approximate recovery is a concept to recover an image to approximately the original state.

Tamper-localization watermarking is useful in identifying corrupted image regions due to incidental data loss or finding the attacker temptation. A straightforward localizing method is to divide an image into blocks and embed integrity data into the blocks. However, this method is vulnerable to collage attack which assembles blocks of several authentic images or swaps blocks of the same image to forge a new authentic image. To defeat the collage attack, a naive method is to embed a unique image index, block number and block MAC into the block. Consequently, the challenge of localizing tampered regions is how to foil collage attack with smallest size of block watermarks or finest localization granularity.

Tamper localization is done by comparing the average value of each block in the ROI with the retrieved average value from the watermark. Tampered blocks can be recovered using the compressed ROI.

1.4 DIGITAL IMAGING AND COMMUNICATIONS IN MEDICINE (DICOM)

DICOM is the ineternational standard for medical images and related information (ISO 12052). It defines the formats for medical images that can be exchanged with the data and quality necessary for clinical use. DICOM is implemented in almost every radiology, cardiology imaging and radiotheraphy device (X-ray, CT, MRI, ultrasound and etc.) and increasingly in devices in other medical domains such as ophthalmology and dentistry. With tens of thousands of imaging devices in use, DICOM is one of the most widely deployed healthcare messaging standards in the world. There are literally billions of DICOM images currently in use for clinical care. Since its first publication in 1993, DICOM has revolutionized the practice of radiology, allowing the replacement of X-ray film with a fully digital workflow. Much as the Internet has become the platform for new consumer information applications, DICOM has enabled advanced medical imaging applications that have "changed the face of clinical medicine". From the emergency department, to cardiax stress testing, to breast cancer detection, DICOM is the standard that makes medical imaging work for doctors and for patients.

1.5 PROBLEM STATEMENT

Digital watermarking is the process of embedding information into digital multimedia content such that the information can later be extracted or detected for a variety of purposes including copy prevention and control. Digital watermarking has become an active and important area of research and development.

One of the requirements of an effective watermarking based authentication is tamper localization using Cyclic Redundancy Check (CRC) where the authentication watermark should be able to detect the localization of manipulated areas and verify other areas as authentic. Once tampering is localized, tampered section can be recovered.

Cyclic Redundancy Check (CRC) is a technique for detecting errors in digital data but not for making corrections when errors are detected. It is used primarily in data transmission. CRC also used to ensure that your data is fine when being transfered. It is

a checking procedure that quickly identifies when data has been damaged. But it does not mean all the data is lost forever. When data is transferred, it is usually in small blocks and each block is given a CRC value. If something goes wrong with the data between the time it leaves the source and arrives at its destination, the CRC sent at the source will no longer match the one that is calculated when data arrives. This is when the cyclic redundancy check error will appear.

Localization of a tampered image is useful for deducing the motive of the tampering or whether the modification is legitimate. Tampered image can be recovered by using embedded watermark that contain information of the original image. Approximate recovery is a concept to recover an image to approximately the original state.

1.6 RESEARCH OBJECTIVE

The objectives of the project are:

1. To develop a temper localization watermarking scheme using Cyclic Redundancy Check (CRC).

2. To test temper localization watermarking scheme using selected modality.

1.7 PROJECT SCOPE

There are two scope for this project:

1. Physicians or

    - Physicians diagnose their patients by relying on the provided electronic and digital data such as ultrasonic, computed tomography (CT), Magnetic Resonance Imaging (MRI) and X-Ray images.

2. Doctors

    - Hide invisible and robust medical information about a patient inside images.

1.8 THESIS ORGANIZATION

This thesis is divided into the following chapter:

Chapter 1: This chapter discuss on introduction to digital watermarking . Problem statement describes on the problem that faced by the previous research.

Chapter 2: The previous works on watermarking is reviewed in this chapter. It covers on watermarking scheme, tamper localization and also CRC.

Chapter 3: This chapter proposes on how to embed, retrieve and compare the CRC in medical images.

LITERATURE REVIEW

2.1 INTRODUCTION

This chapter introduces watermarking in details as well as its previous works. It consists of section 2.2 that introduces the components in a general watermarking scheme. Section 2.3 describes the classification of watermarking by domain. Section
2.4 introduces the concept of medical image watermarking. Section 2.5 introduces the tamper localization and recovery scheme and its previous works.

2.2 GENERAL WATERMARKING SCHEME

General watermarking scheme is the process that embeds data called watermark or digital signature or tag or label into a multimedia object such that watermark can be detected or extracted later to make an assertion about the object. The object may be an image or audio or video.

A simple example of a digital watermark would be a visible “seal” placed

over an image to identify the copyright. However the watermark might contain additional information including the identity of the purchaser of a particular copy of the material.

In general, any watermarking scheme (algorithms) consists of three parts that is the watermark, the encoder (marking insertion algorithm) and the decoder and comparator (verification or extraction or detection algorithm).

The conceptual model of the watermarking system is explained in Figure2.1 (Podilchuk &Delp 2001). Original image depicts the carrier which needs protection. The watermark encoder embeds the watermark in to the cover image. The watermark can be a pseudo-random number or binary sequence. The optional key is used to enhance the security of the system. Decoder estimates the watermark from the received

image with the help of key and original image if required. Watermarked image is subjected to various forms of manipulations on communication channel.

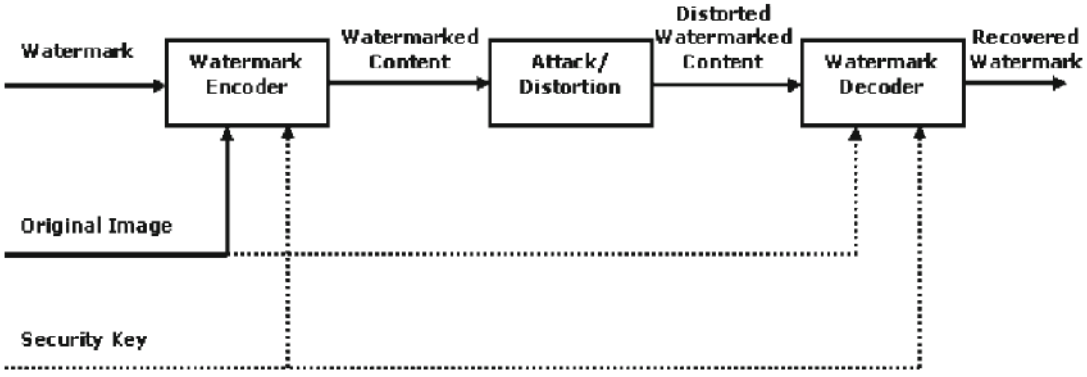

Figure 2.1 A typical watermarking system

Each owner has a unique watermark or an owner can also put different watermarks in different objects the marking algorithm incorporates the watermark into the object. The verification algorithm authenticates the object determining both the owner and the integrity of the object. Figure 2.2 illustrates the encoding process and
Figure 2.3 illustrates the decoding process.

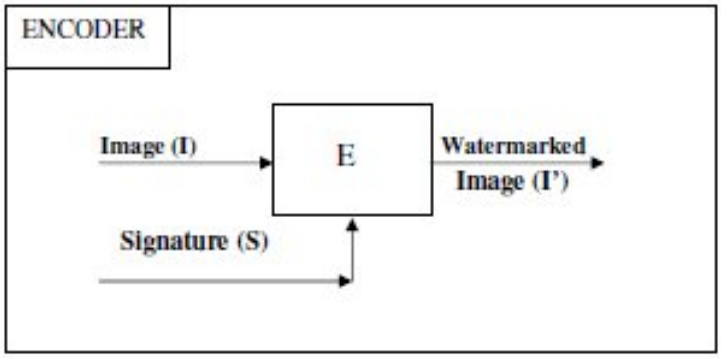

Figure 2.2 Encoding Process

Denote an image by I, a signature by S = { s1, s2, … } the watermarked image by I'. E is an encoder function, it takes an image I and a signature S and it generates a new image which is called watermarked image I', i.e.

$$E(I, S) = I' \quad (2.2)$$

It should be noted that the signature S may be dependent on image I. In such cases, the encoding process describes by (1.1) still holds.

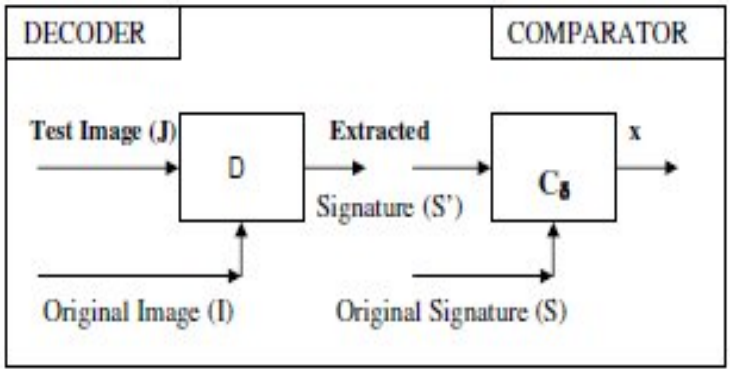

Figure 2.3 Decoding Process

A decoder function D takes an image J (J can be a watermarked or un-