ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 12, December 2017

Identify Protein Folding and Unfolding Structure using Digital Image processing

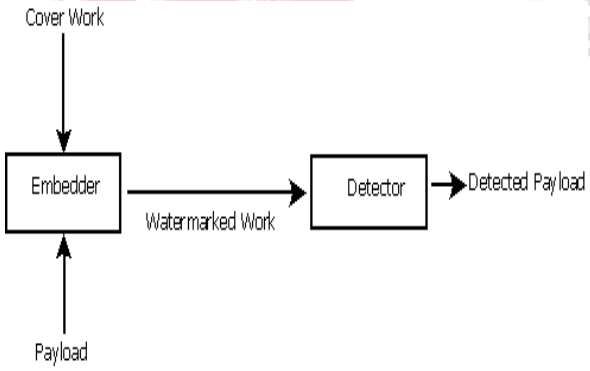
[1] Er. Sameer Dixit, [2] Shraddha Srivastava, [3] Er. Kamalesh Chandra Maurya
[1] Asst. Professor, Integral University, Lucknow. [2] HBTI, Kanpur.
[3]Asst. Professor, Integral University, Lucknow.

Abstract - **The unfolding of a protein can be described as a transition from a predominantly rigid, folded structure to an ensemble of denatured states. With the growing use and need of digital multimedia technology the security of multimedia information like audio, video, text and image have become a major concern. Integrity of image information is important especially when this type of data is used for authentication purpose e.g. medical diagnosis or court evidence. This is an interesting challenge and this is probably why so much attention has been drawn toward the development of digital information protection schemes. Many approaches are possible to protect visual data digital watermarking is probably the one that has received most interest. The idea of fragile watermarking of images is to embed information data within the image with an insensible form for human visual system but in a way that protects from attacks such as intentional or unintentional alteration in extensive content of the digital image. The goal is to produce an image that looks exactly the same to a human eye but still allows its positive identification in comparison with the owner's key if necessary. To cover the same objective this work pro- poses an efficient image authentication and recovery approach using block wise fragile watermarking.**

Keywords — watermarking. Watermark Embedding.

I. INTRODUCTION

The manner in which proteins, chains of amino acid residues, acquire their three dimensional conformation has long been a subject of inquiry and many different explanations of the underlying mechanism have been proposed is the practice of imperceptibly altering a work to embed a message about that work. This technique contains two high level elements Embedder and Detector as shown in figure .



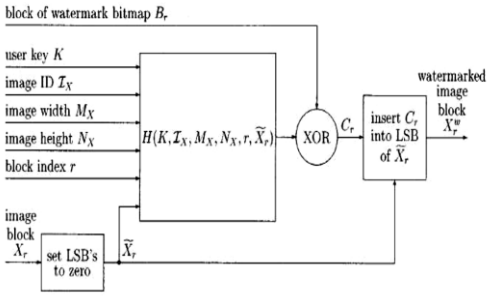
The embedder takes two inputs. One is the payload we want to embed (e.g. either the watermark or secret message), other is the cover work in which we want to embed the payload. The output of embedder is presented as an input to the detector.A generic watermarking system. Now a day's watermarking is used in various fields to provide security.

Some watermarking applications are

- ***Broadcast monitoring:*** Identifying when and where works are broadcast by recognizing watermarks embedded in them.
- ***Owner Identification***: Embedding the identity of the copy right owner as a watermark.
- Transaction tracking Using watermark to identify people who obtain content legally but illegally redistribute it.
- ***Content Authentication:*** Embedding signature information in content that can be later checked to verify it has not been tampered.
- ***Copy Control:*** Using watermark to tell recording equipment what content may not be recorded.
- ***Device Control:*** Using watermark to make devices, such as toys react to displayed content.
- ***Legacy Enhancement:*** Using the watermark to improve the functionality of existing system. Watermarks are broadly classified into three categories viz. fragile, semi fragile and robust watermark.
- ***Fragile watermark:*** A watermark that becomes undetectable after even minor modification of the work in which it is embedded.
- ***Robust watermark:*** A watermark designed to survive on legitimate and everyday usage of content.
- ***Semi fragile watermarking:*** A watermark that is unaffected by legitimate distortion but destroyed by illegitimate distortion.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 12, December 2017



MOTIVATION

Almost all functions of a living organism at its cellular level are carried out through various classes of proteins. Our body consists of many kinds of cells, and each cell carries a copy of the genome– which is the program which a cell executes for its functioning. The genome is a polymer, or a chain; for us, it is a sequence of four kinds of molecules, known as nucleotides, and these four kinds are denoted as A, C, T, and G.Thus, a genome or DNA is a string over the alphabet {A, C, T, G} (the human genome has about 3 × 109 symbols). Certain subsequences of the genome sequence are genes. A gene is translated into a protein. Each triplet of nucleotides is a code of one of twenty amino acids. For example, TTT TAC TGC GGC is translated as the sequence of four amino acids: Phe Tyr Cys Gly. Thus, from a gene, the cell makes a sequence of amino acids such a sequence of amino acids is a protein. Each protein, which forms as a chain, folds up into a unique shape, unique for that sequence of amino acids. By structure of a protein, we mean this unique 3D shape of the protein. The functionality of a protein is due to its 3D shape, that is, its structure. By knowing the structure of a protein, we understand its function. It is relatively easy to find out the chain of amino acids that defines a protein. The protein folding problem is: given a sequence of amino acids, determine the 3D shape that this sequence will give rise to. What we are asking for is an algorithm, and as computer scientists we want an efficient, that is, a polynomial time algorithm. In a living cell, many chains of amino acids are getting formed all the time. Such a chain folds into its unique 3D shape usually within afew milliseconds, and then it functions as it should. Thus, nature seems to use an efficient algorithm to carry out protein folding. The rapid expansion of the internet in the past years has rapidly increased the availability of digital data such as audio, images and videos to the public. As we have witnessed in the past few years, the problem of protecting multimedia information becomes more and more important and a lot of copyright owners are concerned about protecting any illegal duplication or any intentional modification of their data or work. Some serious work needs to be done in order to maintain the availability of multimedia information but, in the meantime, the industry must come up with ways to protect intellectual property of creators, distributors or simple owners of such data. This is an interesting challenge and this is probably why so much attention has been drawn toward the development of digital information protection schemes. We can assume a case study, suppose a website administrator has created a website and uploaded some personal images on it, after some day he found that same images with some objectionable alteration are present in different website. In this situation there will be conflict that, which one is the correct image. This conflict may be resolved by fragile watermarking scheme by using third trusted party. If all images are watermarked using fragile watermarking by trusted third party and uploaded to website then at the time of conflict, using extraction algorithm we can know that which one is original image.

Goal

The goal of this thesis is to provide a fragile watermarking technique which is effective enough to detect the alteration in image as well as able to reconstruct the original extensive content of altered pixel.

First Proposal:

Pixel-wise generic fragile watermarking based
On ARA bits proposed algorithm is shown in figure is based on ARA bits where ARA stands for Authentication, Relational and Associative bits
Watermark Embedding Procedure
Consider a gray scale host image I, having dimension m xn. Then N represents the number of pixels, N= m x n. So the gray scale value at each pixel of the image is denoted by Pi

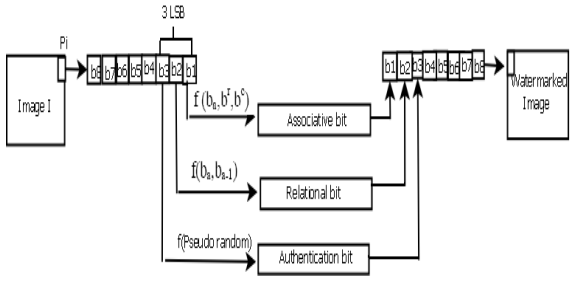


Figure: Block diagram of our first proposed approach

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 12, December 2017

Where, Pi ϵ (0….255), i = 1, 2, 3...N. Pi can be represented by 8 bits. So each individual bit of Pi is denoted by b (Pi, 7), b (Pi, 6), b(Pi, 5)...b(Pi, 0). Then the individual bit of any pixel Pi can be represented in binary form by following equation 1

b (Pi, u) = ⌞ Pi/2u ⌟ mod2, where u=0, 1,2...7

Step 1: The first LSB for a given pixel in host image I is called as Associative bit. As a first step we are taking the position of all pixels with respect to their row and column value then associate it with corresponding pixel by following way.
Consider any pixel of I as Pi. So we take only 5 MSBs of Pi represented as ba where a 2ϵ (3….7), for associative bit calculation. Similarly br and bc are binary representation of corresponding row and column value. Now calculate the following

As1 = Ex - OR (bra-3, ba), a = 7; 6…..3

As2 = Ex - OR (bca-3 ba), a = 7; 6…..3

Where, As1 represents bitwise Ex-OR operation between row value and pixel value whereas As2 represents bitwise Ex-OR operation between column value and pixel value. Hence Associative bit can be calculated as

Associative bit = (Xj=1, 2….5(As1j ^ As2j))mod 2

It can be observed that after 25 = 32th value of row and column, the first five LSBs of its binary value will be reinitialized from 0.This will happen till mth row and nth column.The first LSB of given pixel Pi will be replaced by this Associative bit.
Step 2: The second LSB of Pi is called as Relational bit. According to its name, it will show the relation among all MSBs of Pi. Calculation for this bit will be as follows

Relational bit = (Xv=7, 6…..4(bv ⊕ bv-1))mod 2

Due to Ex-OR operation, the MSB values will be tightly coupled with their sequence. The second LSB of given pixel Pi will be replaced by this Relational bit.
Step 3: The third LSB of Pi is called Authentication bit. For calculating it a binary matrix of size m x n is generated pseudo-randomly with the help of a secret key. The third LSB of each pixel Pi of the host image is replaced with the corresponding bit of the generated binary matrix. In this way, using ARA bits we replace all the three LSBs of host image. Let's understand this process with an example. If in image I, the pixel value of 12th row and 20th column is 139. So 139 = (10001011)2. Now we want to replace three LSBs i.e. 011 by ARA bits.

So first we calculate Associative bit, as 12 = (00001100)2 and 20 = (00010100)2.Now we calculate As1 and As2 as follows: third LSB of each pixel Pi of the host image is replaced with the corresponding bit of the generated binary matrix.

As1 = 1 ⊕ 0, 0 ⊕ 0, 0 ⊕ 1, 0 ⊕ 1, 1 ⊕ 0
= (1, 0, 1, 1, 1)

As2 = 1 ⊕ 0, 0 ⊕ 0, 0 ⊕ 1, 0 ⊕ 0, 1 ⊕ 1
= (1, 0, 1, 0, 0)

Associative bit = 1 x 1 + 0 x 0 + 1 x 1 + 1 x 0 + 1 x 0
= 2 mod 2 = 0

Hence the first LSB of 139 i.e. 1 is replaced by 0. Now we calculate the Relational bit by using equation (6) as:

Relational bit = (1 x 0 + 0 x 0 + 0 x 0 + 0 x 1) mod 2
= (1 + 0 + 0 + 1) mod 2
= 2 mod 2 = 0

Thus the second LSB i.e. 1 will be replaced by Relational bit 0. Now the value of authentication bit will depend on secret key. So in pseudo-random binary matrix whatever value will be at location 12th row and 20th column will be placed on position of third LSB.

Second Proposal:
Block-wise fragile watermarking with restoration capability
Proposed approach is based on k-Medoids clustering scheme which is representative object based technique. This algorithm is purely in spatial domain as illustrated in figure .

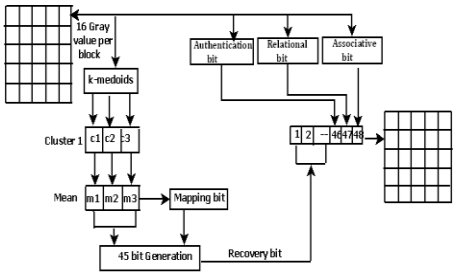


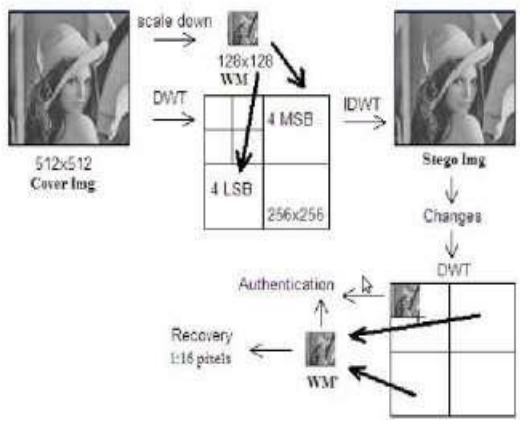
Figure : Block diagram of our Second proposed approach

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 12, December 2017

Watermark Embedding

Watermark embedding process can be classified into four different phases viz. clustering of pixels, Recovery bit generation, Authentication bit generation and Block mapping. Clustering of Pixels



Step 1: First of all remove first 3 LSBs of all pixels to reduce the gray scale value from [0,255] to [0, 31]. Now each will take 5 binary bits to represent it.

Step 2: Divide the image I into number of blocks having dimension 4 X 4. Hence each block will contain 16 gray scale values. So the total number of blocks will be N/16.

Step 3: Every block which contains 16 gray values will be inputted to k-medoids clustering algorithm to make three different clusters. Suppose we have n dataset and we have to make k clusters then k-medoids algorithm is as follows.

1. Randomly choose k gray values in a block as the initial representative seeds.
2. repeat
3. Assign every remaining gray value to the cluster with nearest representative seed.
4. Arbitrarily choose a non representative gray value.
5. Compute the total cost S of representative gray value with non representative gray value. Here cost is calculated by taking difference between both gray values.
6. If S < 0 then swap the representative gray value with non representative gray value.
7. until no change.
8. Finally we get 3 clusters for each block now calculates the mean for each cluster as a round integer. Let means are m1, m2, m3 then rearrange all mean m1, m2 and m3 in descending order. Here first phase of watermark embedding is completed and second is as follows Recovery bit generation For each block we generate 45 recovery bits. These bits are formed as a vector V.

Suppose three means for a block are m1,m2 and m3 where descending order of mean is m3 >m2 > m1

Step 4: Map the mean values with their corresponding two bit pattern as shown in table1.

Step 5: Convert the highest mean that is m3 here, into 5 bit binary form and put on first five indexes of vector V.

Step 6: Calculate D1, D2 and D3 by following equation

D1 = m3 - m2
D2 = m3 -m1
D3 = D2- D1

Step 7: Convert D1 and D3 into 4 bit binary form and put them into next eight indexes of vector V. Now we have occupied 13 position of vector V and we have to calculate32 more bits for recovery.

Step 8: Map the all 16 gray level values with the two bit mapping binary bits for their corresponding mean values of those clusters in which they belong. For example

B = Y11 Y12 Y13 Y14
Y21 Y22 Y23 Y24
Y31 Y32 Y33 Y34
Y41 Y42 Y43 Y44

Let B is one of the block having 16 gray level values and by using clustering algorithm we have calculated three clusters C1, C2 and C3 which has following values

C1 = Y11, Y12, Y33, Y34, Y31, Y32
C2 = Y13, Y14, Y41, Y42
C3 = Y43, Y44, Y21, Y22, Y23, Y24

Mean of C1, C2 and C3 are denoted by m1, m2 and m3 respectively and after arranging in descending order we get m3, m2, m1. According to table 1, m3 is mapped by 01 similarly m2 is mapped by 10 and m1 is mapped by 11. If we have to put two bit binary mapping bit for Y11, then it will be denoted by 11 because it belongs to the cluster C1 and mapping bit for C1 is 11. So by this way we will get 32 bit sequence for 16 gray values and it will be placed in next 32 indexes of vector V systematically. Now we have occupied 45 bits of V and we have 3 remaining bits.

CONCLUSION

This thesis proposes an image authentication and restoration approach using block-wise generic fragile watermarking which is based on k-medoids clustering technique. Many of the approaches for image restoration, proposed earlier, were in frequency domain but suggested

scheme is utterly in spatial domain. Experimental results and Table 5 show the efficiency of proposed scheme which is not only good enough toperceive the altered block with high accuracy but also able to restore those tampered blocks with good imperceptibility. K-medoids is a representative object based clustering technique which ensures that the gray level value which is used to replace other gray value within a block belongs to the same block. Hence we take the benefit of this property of kmedoids.Clustering for each block is done by using a common characteristic hence for each element within a single cluster has same characteristics.

FUTURE DIRECTION OF WORK

Proposed approach is capable to restore the extensive image content, only when forty eight bits of vector V for a block are preserved or at least first 13 bits from V are safe. But attack or any alteration in the image is unpredictable for the rate of change in the image hence there must be a very efficient approach to preserve forty eight bits for a block in the image itself. So that after alteration, first we recover the forty eight bits then using those we recover the extensive image content. It may require a hybrid approach which uses both spatial and frequency domain.

REFERENCES

[1] Anthony T. S. Ho, Xunzhan Zhu, Jun Shen,and Pina Marziliano. Fragile Watermarking Based on Encoding of the Zeros of the -Transform IEEE Transactions On Information Forensic and Security, VOL. 3, NO. 3, September 2008

[2] Eugene T. Lin and Edward J. Delp. A review of fragile watermarking. Center for Education and Research in Information Assurance and Security, Purdue University, West Lafayette, IN 47907-2086.

[3] Jiri Fridrich and Miroslav Goljan. Images with Self-Correcting Capabilities. IEEE, 0-7803-5467-2/99, 1999.

[4] Hongjie He, Jiashu Zhang, Fan Chen. Block-wise Fragile Watermarking Scheme Based on Scramble Encryption. IEEE 978-1-4244-4105-1/07, 2007.

[5] Hongjie He, JiaShu Zhang , and Heng-Ming Tai A Wavelet-Based Fragile Watermarking Scheme for Secure Image Authentication IWDW 2006 LNCS 4283, pp. 422432, 2006.

[6] Mi-Ae Kim and Won-Hyung Lee. A Content-Based Fragile Watermarking Scheme for Image Authentication Springer-Verlag Berlin Heidelberg, AWCC 2004, LNCS 3309, pp. 258-265, 2004.

[7] Ping Wah Wong and Nasir Memon Secret and Public Key Image Watermarking Schemes for Image Authentication and Ownership Verification IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 10, NO. 10, OCTOBER 2001

[8] Shengbing CHE, Bin MA, Zuguo CHE. An Adaptive and Fragile Image Watermarking Algorithm Based on Composite Chaotic Iterative Dynamic System IEEE DOI 10.1109/IIH-MSP.2008.24,2008

[9] Shivendra Shivani, Anoop Patel, Sushila Kamble, Suneeta Agarwal. An Efficient Pixel wise Fragile Watermarking Scheme based on ARA bits ACM ICCCS, Rourkela, February 2011 ISBN: 978-1-4503-0464-1.

[10] Shivendra Shivani, Anoop Patel, Akhilendra, Suneeta Agarwal. Self Recoverable Block wise Fragile Watermarking Scheme based on Histogram Segmentation ACM ICWET, February 2011 ISBN:978-1-4503-0449-8.

[11] Shivendra Shivani, Anoop Patel, Sushila Kamble, Suneeta Agarwal. Image Authentication and Recovery using Block wise Fragile Watermarking based on K-medoids Clustering Approach. IJCA Vol-15 , Issue 1 ISBN:978-93-80747-77-5.