International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:2, No:7, 2008

A Fragile Watermarking Scheme for Color Image Authentication

M. Hamad Hassan, and S.A.M. Gilani

***Abstract*—**In this paper, a fragile watermarking scheme is proposed for color image specified object's authentication. The color image is first transformed from *RGB* to *YST* color space, suitable for watermarking the color media. The *T* channel corresponds to the chrominance component of a color image and $YS \perp T$, therefore selected for embedding the watermark. The *T* channel is first divided into 2×2 non-overlapping blocks and the two *LSB*s are set to zero. The object that is to be authenticated is also divided into 2×2 non-overlapping blocks and each block's intensity mean is computed followed by eight bit encoding. The generated watermark is then embedded into *T* channel randomly selected 2×2 block's LSBs using *2D-Torus Automorphism.* Selection of block size is paramount for exact localization and recovery of work. The proposed scheme is blind, efficient and secure with ability to detect and locate even minor tampering applied to the image with full recovery of original work. The quality of watermarked media is quite high both subjectively and objectively. The technique is suitable for class of images with format such as gif, tif or bitmap.

***Keywords*—**Image Authentication, *LSBs, PSNR, 2D-Torus Automorphism, YST* Color Space.

I. Introduction

IN past decade, there has been exponential growth in the use of digital multimedia contents. The wideband networks made the exchange of multimedia contents, easy and fast. On the other hand, the availability of powerful image processing tools made it easy for user to do even imperceptible changes in the original work. As a result image authenticity has become greatly threatened. Generally image authentication verifies the integrity of a digital image. In past digital watermarking gave promising solutions for issues related to copyright protection and digital content authentication including images, audio, video and text and still in infancy. This area welcomed researchers from signal & image processing, information security, computer & electrical engineering and mathematics. Depending on the application, digital watermarking techniques can be classified into two main categories; *Robust and Fragile Watermarking Techniques.* The former is mainly used for copy right protection and fingerprinting applications, in which the goal of watermark is to sustain under all kinds of attacks that intend to remove the watermark while preserving the perceptual quality of the original work. The latter is used for data authentication which is sensitive against any kind of processing that is applied to the work.

In context of transmission and distribution of digital contents across open networks, the possibility of following two threats is always there:

Masquerade: Transformation of an original image into another with similar content. For instance, in a painting original author's signature is replaced by another signature. The tampered image still conveys some meanings, but creates confusion for the forensic applications.

Modification: The original image is transformed by swapping the contents, cropping, replacing portions of the image with content from another image or applying image transformations to change the original image structure. The image give some meanings but the actual information that user ever wanted is actually contained in the swapped area of the image or does not exits at all.

In this paper we have proposed a fragile watermarking scheme which is designed for color image particular object's authentication, for instance in the case of painting, the artist's signature might be the desired object to be authenticated against any kind of processing whatsoever. Similarly is the case for color images with company's monogram, institute logo or building name board. We have performed numerous experiments on different color images and some of them are selected for illustrations and discussed in the Section IV of this paper.

The given color image is first transformed from *RGB* to *YST* color space. This new color space is exclusively designed by Francesco et al. [4] for watermarking the color media. Details of *YST* color space is discussed in Section III of this paper. The *T* channel corresponds to the chrominance component of a color image and $YS \perp T$ therefore the *T* channel is explicitly selected for embedding the watermark information. After doing the color space transformation, the *T* channel is divided into 2×2 non-overlapping blocks and two LSBs of each block are set to zero. The object of an image that is to be authenticated is also divided into 2×2 non-

Manuscript received on April 30, 2006. This work was supported in part by the HEC, Pakistan under faculty development program.

M. Hamad Hassan is graduate student of Faculty of Computer Science & Engineering at GIK Institute, Pakistan (email: hamad_gikian@yahoo.com).

Dr. Asif Gilani is the Dean of Faculty of Computer Science & Engineering at GIK Institute, Pakistan (email: asif@giki.edu.pk).

LSBs: Least Significant Bits, PSNR: Peak Signal to Noise Ratio

overlapping blocks after doing necessary resizing if desired. Then intensity mean of each block of object is computed and encoded upto eight binary bits to have the watermark information about each block of object. Followed by watermark generation, secure mapping of blocks of the *T* channel is performed based on *Torus Automorphism* presented by G. Voyatzis et al. [5] using a private key as discussed in Section IV of this paper. The desired object's each block information is then embedded into the mapped block's LSBs in a manner as shown in the Fig 2. The embedded watermark then helps not only in the authentication of work but with full recovery of original work. Our scheme is able to correctly localize the tampering in the object under consideration and recover it with probability of near one.

The rest of the paper is organized as: Section II discusses the related work, Section III briefs about the necessary background knowledge. Section IV explains the proposed scheme, Section V demonstrates the simulation results, and Section VI derives the concluding remarks.

II. Related Work

The survey of watermarking based authentication schemes is done in paper presented by T. Liu. et al. [6]. An early scheme for image authentication was presented by S. Walton [7] where checksums of image is computed and in combination with a seal, generates the watermark that votes for authentication later on. This work excited the idea of digital image authentication among the researchers working in the area of watermarking and many of them consider it in different and sophisticated ways. An efficient and easily computed method was proposed by Yeung and Mintzer [8] that embed a binary logo in an image in order to detect possible alterations in the image and at the same time provide some information about the image owner. Fridrich presented her schemes [9]-[11] where an approximation of the image is embedded in the *LSB*s of the original image for authentication and recovery of tampered work.

The present work belongs to the family of fragile watermarking that can detect any kind of processing whether legitimate or illegitimate performed on an image. Instead of binary logo of the company or organization, our scheme embeds the desired part of an image like artist's signature in the digital painting.

The main advantage of our technique is the ability to detect the slightest changes or tampering that might occur in the image and able to provide information about the location of attacks such as cropping or pixel modification, however, due its fragile nature rather semi-fragile it is not suitable for lossy compression like JPEG compression. The ability to detect the slightest tampering implies that the watermark is very sensitive to any change that might occur in any location in the image i.e. watermark acts as digital signature.

III. Background Knowledge

YST Color Space: The selection of color space is very important step in watermarking based applications. In this regard we considered *YST* color space, exclusively designed and recommended by Francesco et al. [4]. A color space is notation by which we specify colors i.e. human perception of the visible electromagnetic spectrum.

The *RGB* color space is good for image display but is not the best choice when analyzing images using the computer. The main disadvantage of the *RGB* color space is high correlation between its components. The value of cross-correlation between the *B* and *R* channel is numerically about 0.78, 0.98 between the *R* and *G* channel and 0.94 between the *G* and *B* channels respectively. Because of this high correlation between channels, the *RGB* domain is not suitable for image processing techniques, especially for watermarking the color media. The potential of these three channels can be exploited for the applications of watermarking, by decreasing the correlation among them.

Other colors spaces too exist which have the property of separating the luminance component from the chrominance component and with that at least the partial independence of chromaticity and luminance is achieved. Such color spaces includes *YCbCr, YUV* etc. where *Y* corresponds to the brightness portion of an image while *Cr, Cb* and *U, V* corresponds to the chrominance (color) components of an image.

In case of *YST* color space, *Y* corresponds to the brightness component as before while *S* and *T* channels correspond to the chrominance component of the color image. The new color space satisfies all the principal conditions that are:

i) The brightness must be the same to that of other two color spaces i.e. *YUV* and *YCbCr.*

ii) One component i.e. *S* must be ad hoc created to match the vector corresponding to the skin color.

iii) The transformation should be reversible.

The two components *Y* and *S* form an angle of 52° because they were generated without imposing any orthogonal criterion. The *T* component is identified by the *Gram-Schmidt* procedure in order to have a component that is orthogonal to the plane containing *Y* and *S* components, in this way $YS \perp T$. Thus we have set of linear transformation matrix to convert color image from *RGB* to *YST* color space given by equation (1).

$$\begin{vmatrix} Y \\ S \\ T \end{vmatrix} = \begin{vmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{vmatrix} . \begin{vmatrix} R \\ G \\ B \end{vmatrix} \quad (1)$$

The diagrammatic representation of *YST* color space is given by:

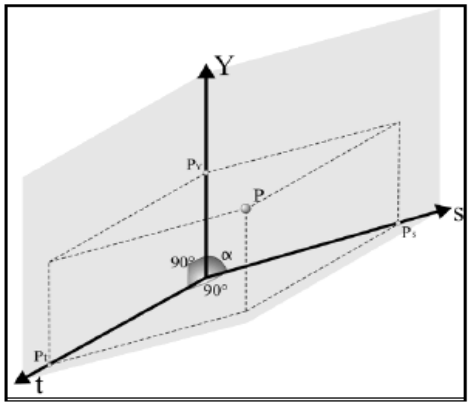


Fig. 1 YST Color Space Representation

2D-Torus Automorphism: 2D-Torus Automorphism can be considered as a permutation function or spatial transformation of a plane region. This transformation can be performed using the 2×2 matrix *A* with constant elements. A point (x, y) can be transformed to new point (x', y') using equation (2).

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ k & k+1 \end{pmatrix} . \begin{pmatrix} x \\ y \end{pmatrix} \bmod N \qquad (2)$$

Where $(x, y),\ (x', y') \in [0, N-1] x [0, N-1], N$, the number of blocks in each dimensions, and $k \in [0, N-1]$ is a private key. For in depth understanding of how *2D - Torus Automorphism* works, reader is recommended to follow the paper presented by G. Voyatzis et al. [5].

IV. Proposed Scheme

This section describes the principal phases that are required to implement our proposed authentication system. Each phase contains different steps in sequential order with self explanation. The following two steps are principally performed by the sending party of a network.

Watermark Generation

1. Read the color image.
2. Transform the given color image; say M, from *RGB* to *YST* color space equation (1).
3. Select the *T* channel, divide it into 2×2 non-overlapping blocks and set the two *LSBs* to zero.
4. Select the object or part of an image that is to be authenticated and divide it into 2×2 non-overlapping blocks after necessary resizing if desired.
5. Compute the intensity mean of each block of object and encode it upto eight binary bits to have the watermark information that is to be embedded in the *T* channel selected block's *LSBs* in the embedding phase.

Watermark Embedding

1. Select the *T* channel of image M, divide it into non-overlapping 2×2 blocks and generate blocks mapping sequence using the equation(2) with a private key based on *2D-Tours Authomorphism.*
2. Embed the desired object's each block information into the mapped block's two *LSBs* as shown in Fig 2.

						LSB#2	LSB#1
						b1	b2
						b3	b4
						b5	b6
						b7	b8

Fig. 2 *T* channel's 2×2 Block LSBs View

3. After completing the embedding process, concatenate the *Y, S* and *T* channels.
4. Transform the image from *YST* color space to *RGB* by taking inverse transform of equation (1) to have the watermarked image.

The following three steps are principally performed by the receiving party of a network.

Watermark Extraction & Tamper Detection

1. Transform the given watermarked image from *RGB* to *YST* color space using equation (1), and select the *T* channel.
2. Deploying the same private key as applied at the embedding phase; generate the 2×2 non-overlapping blocks mapping sequence using the equation (2).
3. Extract the watermark bits from each block *LSBs* in the manner as shown in the Fig 2.
4. After extraction of watermark, set the two *LSBs* of *T* channel to zero.
5. Select the desired object of an image that was intended to be authenticated by the proposed system; say m', and divide it into 2×2 non-overlapping blocks, after performing necessary resizing if desired.
6. Compute the intensity mean of each block and encode it upto eight bits to have watermark information about each block.
7. Now compare the extracted and generated watermark on bit/bit basis, if they are same and equal in number, then the object is authentic otherwise tampered.

8. If the object is found tampered, then identify that block and set its pixel value to zero for differentiation between authentic and tampered parts of a work.

Recovery of Tampered Work

1. Once the tamper detection is performed correctly, identify the source block for the tampered block using the equation (2) with the same private key as used in embedding and tamper detection phase.

2. Generate the pixel value from the eight bit watermark that was extracted and set all the four pixels of tampered block to this restored value.

V. Results

The simulations were conducted on Intel machine with 2.4 GHz processor and 512 MB of RAM. MATLAB 7.0 and Photoshop 7.0 were used for the implementation of proposed scheme and image processing operations.

PSNR Measurement:

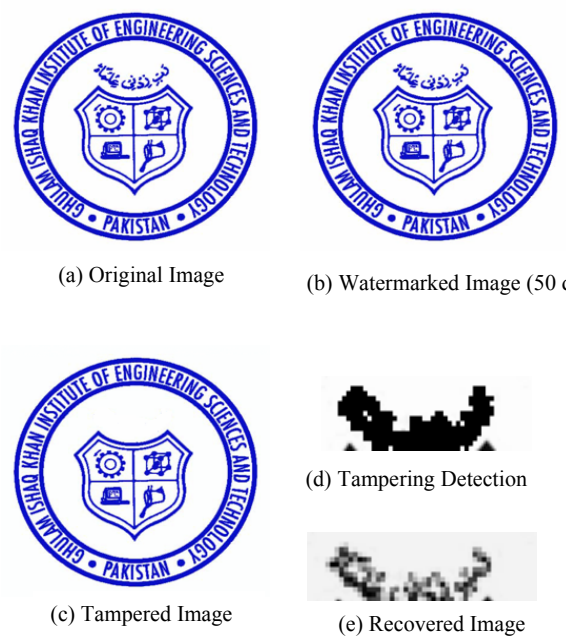
One commonly used measure to evaluate the imperceptibility of the watermarked image is the peak signal to noise ratio *(PSNR)* which is given by equation (3).

$$PSNR = 10.\log_{10}\left(\frac{255}{MSE}\right)(dB) \quad (3)$$

TABLE I
Quality Matrix (*PSNR*)

Test Image	Format	Size	PSNR (dB)
GIKI Pakistan Logo	tiff	256x256	50.01
GIKI FEE & FCSE	tiff	256x256	49.66

Table I shows the *PSNR* values computed for images used in our experiment for the implementation and verification of the proposed scheme.



(a) Original Image (b) Watermarked Image (50 dB)

(c) Tampered Image (d) Tampering Detection (e) Recovered Image

Fig. 3 Simulation Results: GIKI Logo Authentication

Fig. 3(a), (b), (c), (d) and (e) shows the original image, watermarked image with PSNR of 52 dB, tampered image, detected image and recovered image respectively

(a) Original Image (b) Watermarked Image (49 dB)

(c) Tampered Image (d) Tampering Detection (e) Recovered Image

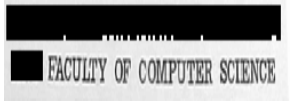
Fig. 4 Simulation Results: GIKI FEE & FCSE Building Name Authentication

Fig. 4(a), (b), (c), (d) and (e) shows the original image, watermarked image with PSNR of 49 dB, tampered image, detected image and recovered image respectively

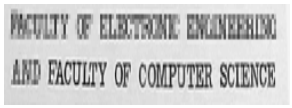
(a) Original Image

(b) Watermarked Image (49 dB)

(c) Tampered Image



(d) Tampering Detection



(e) Recovered Image

Fig. 5 Simulation Results: GIKI FEE & FCSE Building Name Authentication

Fig. 5(a), (b), (c), (d) and (e) shows the original image, watermarked image with PSNR of 49 dB, tampered image, detected image and recovered image respectively

VI. Concluding Remarks

In this paper we presented a fragile watermarking scheme, designed for color image particular object's authentication. For example in the case of painting, the artist's signature would be the paramount for us to verify against any kind of processing whatsoever. Similarly is the case for color images with company monogram, institute logo or building name board. The given color image is first transformed from *RGB* to *YST* color space. This new color space is exclusively designed by Francesco et al. [4] for watermarking the color media. The *T* channel corresponds to the chrominance component of a color image and $YS \perp T$, therefore the *T* channel is explicitly selected for embedding the watermark information. After doing the color space transformation, the *T* channel is divided into 2×2 non-overlapping blocks and two *LSBs* of each block are set to zero. The object of an image that is to be authenticated is also divided into 2×2 non-overlapping blocks after doing necessary resizing if desired. Then intensity mean of each block of object is computed and encoded upto eight bits to have the watermark information about each block of object. Followed by watermark generation, secure mapping of blocks of the *T* channel is generated based on *2D-Torus Automorphism* presented by G. Voyatzis et al. [5] using a private key. The desired object's each block information is then embedded into the mapped block's *LSBs*. The embedded watermark then helps not only in the authentication of work but with full recovery of original work. Our scheme is able to correctly localize the tampering in the object under consideration and recover it with probability of nearly one.

References

[1] I. Kostopoulos., A.N. Skodras, and D. Christodou-lakis, "Self Authentication of Colour Images", Proc. Of the European Conf. on Electronic Imaging & Visual Arts, Florence, Italy, March 26-30, 2001.
[2] Phen Lan Lin, Chung-Kai Hsieh, Po-Whei Huang,"A Hierarchical Digital Watermarking Method for Image Tamper Detection and Recovery, The Journal of Pattern Recognition, Elsevier, 2005.
[3] Jagdish C. Patra, Kah K. Ang and Ee-Luang Ang, "Hierarchical Multiple Image Watermarking for Image Authentication and Ownership Verification", ICIP, 2004.
[4] Francesco Benedetto, Gaetano Giunta, Alessandro Neri, "A New Color Space Domain for Digital Watermarking in Multimedia Applications", ICIP, 2005.
[5] G. Voyatzis, I. Pitas, "Applications of Toral Automorphism in Image Watermarking," ICIP 1996, Vol II, pp.237-240, 1996.
[6] T. Liu and Z.D. Qiu, "The Survey of Digital Watermarking Based Image Authentication Techniques", 6th International Conference, pp. 1566-1559, 2002.
[7] S. Walton, "Image Authentication for a Slippery New Age", Dr. Dobb's Journal of Software Tools for Professional Programmers, Vol. 20, Apr. 1995.
[8] M. Yeung and F. Mintzer, "An Invisible Watermarking Technique for Image Verification", Proc. ICIP 1997, Santa Barbara, California, Oct. 1997.
[9] J. Fridrich, "Image Watermarking for Tamper Detection", Proc. ICIP Chicago, Oct 1998.
[10] J.Fridrich and M.Goljan, "Protection of Digital Images using Self Embedding", Symposium on Content Security and Data Hiding in Digital Media, Newark, NJ, USA, May 1999.
[11] J. Fridrich, "Methods for Tamper Detection in Digital Images", Multimedia and Security Workshop at ACM Multimedia 1999, Orlando, Florida, USA, Oct, 1999.

M. Hamad Hassan did his BS(CS) and MIT from Peshawar and Iqra University respectively. At present, he is HEC Scholar at Faculty of Computer Science & Engineering, Ghulam Ishaq Khan Institute of Engineering Sciences & Technology, Pakistan for his MS in Computer System Engineering. He is also faculty member at the Institute of Information Technology, Kohat University of Science & Technology, Pakistan. His research interests include Digital Image Watermarking & Cryptography for Information Security.

Dr. Asif Gilani did his M.Sc from Islamia University Pakistan and Ph.D in Copyright Protection from University of Patras, Greece. He is Dean of Faculty of Computer Science & Engineering at Ghulam Ishaq Khan Institute of Engineering Sciences & Technology, Pakistan. His research interests include Digital Image Watermarking, Steganography and Image Authentication. He has published number of research papers internationally. At present he is supervising many MS/Ph.D students at GIK Institute. He is also at the list of HEC and PCST approved Ph.D supervisors.