(IJIRSE) International Journal of Innovative Research in Science & Engineering
ISSN (Online) 2347-3207

Analysis of Digital Watermarking Using DWT and PCA

Shikha Sharma, Anupam Agarwal
Electronics and Communication Department, Jagan Nath University,
NH-12, Chaksu Bypass, Tonk Road, Jaipur, Rajasthan, India

Abstract— **In recent years, the popularity of digital video based applications is accompanied by the need of copyright protection, to prevent illicit copying and distribution of digital information. Copyright protection inserts authentication data such as ownership information and logo in the digital media, without affecting its perceptual quality. In case of any dispute, authentication data is extracted from the media and can be used as an authoritative proof to prove the ownership. As a method of copyright protection, digital watermarking has recently emerged as a significant field of interest and a very active area of research. In this work the targeted semi-visible watermarking is approached using Haar Wavelet functions and Principal component analysis (PCA) based approach. The embedding is performed in PCA domain, by means of a secret key alpha.**

Keywords— Discrete Wavelet Transform, Principal Component Analysis, Digital Watermarking, Tuning Factor.

I. INTRODUCTION

The reproduction, manipulation and the distribution of digital multimedia (images, audio and image) via networks has becomes faster and easier. Digital multimedia has proven to be very powerful media for copying and distribution [1]. As the proprietors and creators of the digital products are aware of illegal copying of their products therefore, security copyright protections are important issues pertaining to multimedia applications and services. Earlier, the watermarking techniques were proposed for these aforesaid purposes in which the copyright information is embedded into multimedia data for protecting the ownership. The commonly utilised work of watermarking includes bank notes and stamps, it was this where the term watermark was first used [3]. Consequently, research is now being focused on watermarking schemes to protect multimedia information. In watermarking the data can be embedded in a manner that the contents of the message are hided [5]. The most suitable technology that can serve this purpose is none other than digital watermarking. Multifarious watermarking schemes have been proposed to camouflage copyright marks and other information in digital images, audio, image and other multimedia objects. One of the largest technological events of the last twenty years was the aggression of digital media within the entire vary of daily routine aspects. A watermarking is adding "ownership" information in multimedia contents to prove the authenticity along with embedding a data which is an unperceivable digital code, namely the watermark that carries information about the copyright status of the work to be protected [2]. The manner in which information is embedded into a cover image, watermarking techniques can be revealed either as spatial domain (wherein the watermarking embedding system directly alters the main data components, whiz to hide pixel in a cover image of that watermarked data) or transformed domain (wherein the watermarking scheme transforms the frequency i.e. LL, HL, HH, LH of data components to mask the watermarked data). The last technique has demonstrated to be further robustness than the spatial domain watermarking [1][4]. Several reversible transforms are taken to transform the cover image to its frequency domain representation like discrete fourier transform (DFT), discrete cosine transform (DCT) or discrete wavelet transform (DWT) [1]. Nevertheless the spatial techniques are not able to resist most of the common and predefined attacks like data manipulation, low-pass or high-pass filtering [1][4]. Since, monetary impact of some of the application areas like copyright protection, fingerprinting now a days are at risk and till today no effective algorithm has been designed to prevent illegal misuse of the interactive media contents like images and text. The ultimate goal of this dissertation work is to develop watermarking schemes for images that can bear the known possible and predetermined attacks, and varied image manoeuvre process.

II. THEORETICAL BACKGROUND

A. Digital Watermarking

Digital watermarking is the technique of hiding message in a carrier signal (information), such that it can be extracted later only by authenticated persons to provide copyright protection. Generally, a digital image watermark can be inserted into all forms of interactive media. Among them the most generally known interactive media are audio, image and pictures. It's easy to add a scene digital image watermark on a digital contents, it just requires inserting some content to the original cover image.

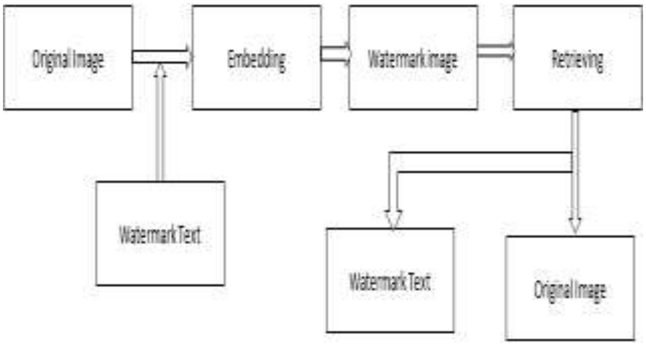


Fig.1: A general watermarking process

A challenge in watermarking is that, processing the watermarked image may remove or damage the embedded watermark, or make it more difficult to detect. The watermarked image may be treated for various inferences, containing the normal techniques that are used in a process; inadvertent bruise or impairment during repository, transference or retrieval over a net-work; or intentionally transforming by an owner for the principle of eradicating the inserted watermark. The main utilization of digital watermarking is marking media files to provide copyright information, as well as for source tracking. There are more Applications of Watermarking such as: copyright protection, fingerprinting, owner identification, labeling, tracking of material on web etc. For effective performance against watermark attacks the watermark must possess some properties such as: perceptual transparency, security, unobtrusive, Irremovable or irreplaceable by unauthorized interceptors, Robust and Lossless.

B. Classification of Digital Watermarking

Digital watermarking is a method that inserts an image known as a watermark into a cover image in such a manner that the watermark can be later on decrypted or extracted for the characterization of the cover image. Interactive media in goods are embedded by a watermark that is usually known as: the original cover signal, the host signal or bus work. A digital watermark is a distinguishing piece of information that is assigned to the data to be protected. This is an important requirement that the watermark cannot be easily extracted or removed from the watermarked object. Digital watermarking technique can be classified on different bases, such as Documentation, working domain and human perception. Watermarking techniques can be differentiated into the following four categories on documentation: Image Watermarking, Video Watermarking, Audio Watermarking and Text Watermarking

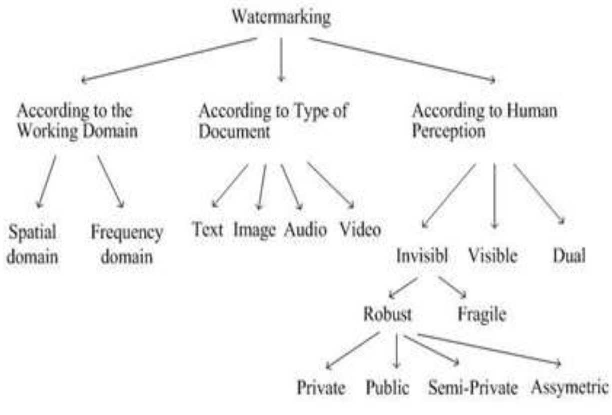


Fig 2: Classification of Watermarking

According to the human perception, the digital watermarks can be classified into different categories, as: Visible watermark, Invisible-Robust watermark, Invisible-Fragile watermark and Dual watermark. On the basis of Working field, watermark can be divided into spatial domain and Transform domain. When the watermark is embedded by modifying the pixel values of the host image or video directly by the pre-determined embedding scheme then this scheme is spatial domain techniques. Whereas when the coefficients of the transformed video frames are modified according to the pre-determined embedding scheme then it is considered as Transform domain. The watermark is embedded distributive on the whole domain of original information. It becomes hard to take away the embedded watermark proving Transform-domain technique to be more effectual, hardly noticeable, and more robust then spatial-domain technique. The predominantly used transform-domain techniques are Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT) [8]. Now a day's watermarking scheme mostly utilises Discrete Wavelet Transform (DWT) because of its multi resolution characteristics, its excellent capacity of localizing spatial-frequency and as it is very suitable to identify areas where watermark can be embedded imperceptibly. PCA technique has the inherent property of removing the correlation left; it then distributes the watermark bits over the sub-bands used for embedding, making the watermarking to be more robust.

III. PROPOSED TECHNIQUE

We have used two schemes of the watermarking technologies: DWT and PCA, to perform the embedding process.

A. Discrete Wavelet Transform (DWT)

Wavelets are functions defined over a finite interval and having an average value of zero. The basic idea of the wavelet transform is to represent any arbitrary occupation as a superposition of a set of such wavelets or basis functions. The 2-D discrete wavelet transforms (DWT) decompose the image in to sub-images, 3 detail and 1 approximation sub-image that resembles the original on ¼ scale of the original. The 2-D DWT is an application of the 1-D DWT in both the horizontal and also the vertical directions. The DWT mould an image into a lower resolution approximation image (LL) as well as horizontal (HL), vertical (LH) and diagonal (HH) detail. Frequency information and spatial information of the transformed data is exploited by wavelet based watermarking methods in multiple resolution to gain robustness [9]. Due to its excellent spatial-frequency localization properties DWT is very suitable to identify areas in the host video frame where a watermark can be embedded imperceptibly. The discrete wavelet transform stipulates great form of the human visual system [7]. Since the HVS is less sensitive to high frequencies, embedding the watermark in high frequency sub-bands makes the watermark more imperceptible while embedding in low frequencies makes it more robust against a variety of attacks. In Discrete Wavelet Transform Human Visual System (HVS) is more realistic than in Discrete Cosine Transform and Discrete Fourier Transform. DWT even has better multi resolution approach. DCT and DFT are full frame transform whereas DWT has spatial frequency localization. DWT provides both frequency and spatial description for an image. Here we have used Haar Wavelet Transform (DWT), which is in general a simple form of compression which involves averaging and differencing terms, storing detail coefficients, eliminating data, and reconstructing the matrix such that the resulting matrix is similar to the initial matrix [6-7].

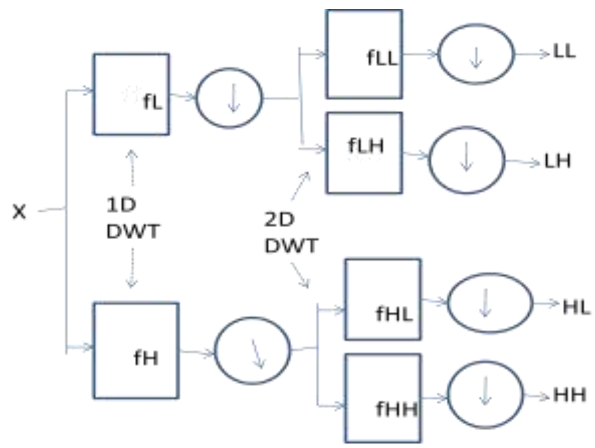


Fig. 3: 2D Discrete Wavelet Transform

B. Principle Component Analysis (PCA)

Principal component analysis (PCA) is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components. The number of principal Components is less than or equal to the number of original variables. PCA is a method of identifying patterns in data, and expressing the data in such a way so as to highlight their similarities and differences. Since patterns in data can be hard to find in data of high dimension, where the advantage of graphical representation is not available, PCA is a powerful tool for analyzing data. The other main advantage of PCA is that once these patterns in the data have been identified, the data can be compressed by reducing the number of dimensions, without much loss of information. It plots the data into a new coordinate system where the data with maximum covariance are plotted together and is known as the first principal component. Similarly, there are the second and third principal components and so on. The maximum energy concentration lies in the first principal component. The key ingredient is combined data set normally distributed, Therefore guaranteed to be independent. PCA Karl Pearson, was invented in 1901 by [10] as an analogue of the principal axes theorem in mechanics; it was later independently developed (and named) by Harold Hotellingin 1930s [11]. Principal component analysis can be achieved by disintegration of eigenvalues of an image correlation matrix or single value disintegration of an image matrix, basically after mean centering the image matrix for every variable [12]. Principal component analysis is the easiest form of the true eigenvector that works on multivariate synthesis.

C. Methedology

In the proposed method we have to convert the watermark image into gray scale and is embedded into the cover image by decomposing it into DWT sub bands followed by the application of block based PCA on the sub-block of one sub-band.

- Algorithm for embedding :

Step 1: Click a picture through a webcam known as webcam image watermark.

Step2: Calculate the intensities of watermark image and webcam image watermark.

Step 3: Whose so ever intensity is greater will be embedded in the original image.

Step 4: Convert the n × n binary watermark logo into a vector W = { w1, w2, ……, wn × n } of '0's and '1's.

Step 5: Transform image from RGB to YUV color format.

Step 6: Apply 1-level DWT to the luminance (Y component) of image to obtain four sub-bands LL, LH, HL and HH of size N x N.

Step 7: Fragment the LL sub-band into k non-overlapping sub-blocks each of dimension n × n (of the same size as the watermark logo).

Step 8: Algorithm 2 is used for embedding with strength α into each sub-block by first obtaining the principal component scores for watermark bits. The general form for embedding is carried out as equation.

$$\text{Scorei} = \text{Scorei} + \alpha W \qquad \ldots\ldots\ldots\ldots (1)$$

Where, Scorei represents the principal component matrix of the ith sub-block.

Step 9: Obtain inverse PCA is applied on the modified PCA components of the sub-block of the frequency district of the LL sub-band to achieve the desired wavelet coefficient variables.

Step10: Inverse DWT is applied to obtain the watermarked luminance component of the image. Then convert the image back to its RGB components.

- Algorithm for extraction :

Step 1: Divide the watermarked (and possibly attacked) image into distinct frames and convert them from RGB to YUV format.

Step 2: Choose the luminance (Y) component of an image and DWT is applied to disintegrate the Y component into the four sub-bands LL, HL, LH, and HH of size N×N.

Step 3: Divide the LL sub-band into n × n non overlapping sub-blocks.

Step 4: Put PCA to every block that is selected in sub-band LL by using Algorithm (c).

Step 5: Provided from the LL sub-band of image, the watermark bits are drawn out from the principal components of every sub-block as in equation 2.

Wi' = (Scorei' – Scorei) /alpha ………… (2)

Where, Wi' is the watermark drawn out from the 'ith' sub-block.

- Algorithm for calculation of Principal

Component Analysis :

The LL sub-band coefficient indexes are changed into a new geometry set by manipulating the principal components of every sub-block (size n x n).

Step 1: Each sub-block is converted into a row vector Di with n2 elements (i=1, 2… k).

Step 2: Compute the mean μi and standard deviation σi (sigma) of the elements of vector *Di*.

Step 3: Compute Zi according to the following equation

Zi =(Di – □I)/sigma ………….. (3)

Here Zi represents a centered, scaled version of Di, of the same size as that of Di.

Step 4: Apply principal component analysis on Zi (size 1 x n2) to obtain the principal component coefficient matrix coeffi (size n2 × n2).

Step 5: Calculate vector Scorei as

Scorei = Zi X Coeffi ...………. (4)

Where, Scorei represents the principal component scores of the 'ith' sub-block.

D. Efficiency of watermarking

The efficiency of watermarking depends on two factors Peak Signal-To-Noise Ratio and Normalised Coefficients. The benchmark for identifying best watermarking scheme is that the NC value must be greater. NC is measure of the robustness of watermarking and its peak value is 1.It depends directly on the original and extracted watermark. PSNR is measure of the deviation of watermarked images (frames) from the original image (frame). PSNR depends on the value of MSE (mean square error).

If the noise free original image (size m x n) is denoted by I and K is its noisy approximation then MSE is given by:

$$MAX_{i=}\frac{1}{mn\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j)-K(i,j)]^2} \quad \ldots\ldots..(5)$$

PSNR**:** The Peak-Signal-To-Noise Ratio (PSNR) is used to calculate aberration of the watermarked as well as attacked frames from the original video frames and is denoted as:

$$PSNR = 10.\log_{10}(\frac{MAX_i^2}{MSE}) \quad \ldots\ldots\ldots.(6)$$

$$= 20.\log_{10}(\frac{MAX_i^2}{MSE})$$

$$= 20.\log_{10}(MAX_i) - 10.\log_{10}(MSE)$$

Where:

MAX_i represents the maximum possible pixel value of the image and when the pixels are represented using 8 bits per sample, this is 255.

The extraction fidelity NC which is given by:

$$NC = \frac{\sum_i\sum_j RW(i,j).EW(i,j)}{\sqrt{\sum_i\sum_j RW(i,j)^2}\ \sqrt{\sum_i\sum_j EW(i,j)^2}} \ldots..(7)$$

Where:

RW is the reference watermark,
EW is the extracted watermark

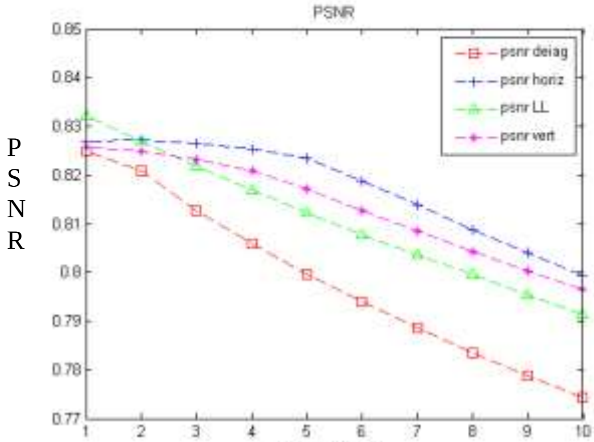
IV. RESULT AND ANALYSIS

This section presents the experimental results of embedding watermark image (Fig. 4) into the cover images (Fig. 5&6) according to the proposed scheme.

Fig. 4: Watermark image

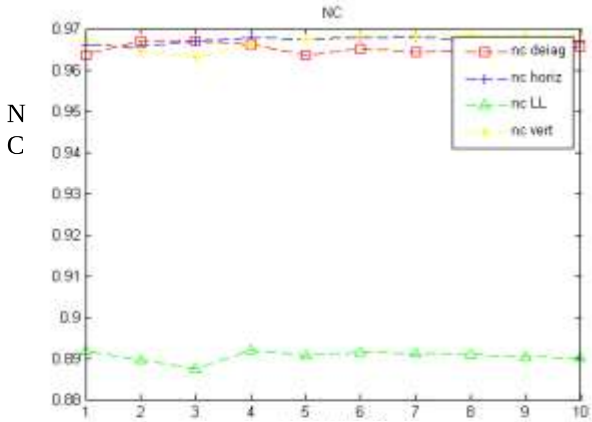
Fig. 5: Cover image 1

Fig. 6: Cover image 2



Visibility factor *10

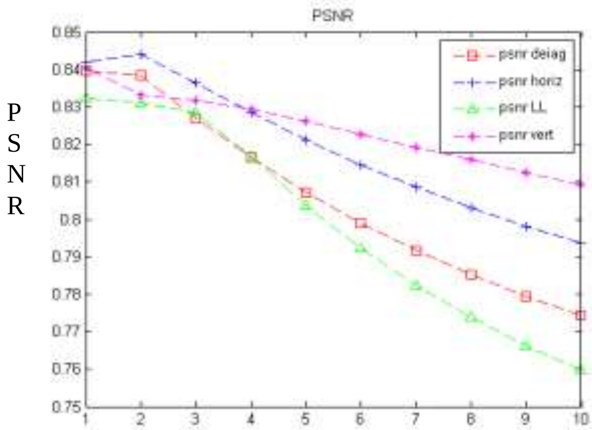
Fig. 7: Plot of PSNR



Visibility factor*10

Fig. 8: Plot of NC

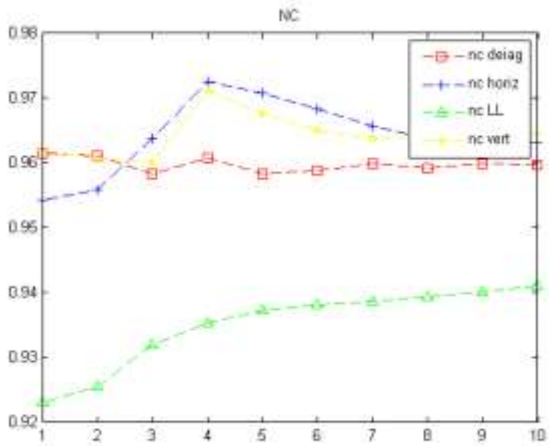
Fig. 7 & 8 shows the PSNR and NC plots for embedding watermark in different sub-bands of cover image 1 over tuning factor range 0.1 to 1.0



Visibility factor *10

Fig. 9: Plot of PSNR

N
C



Visibility factor*10

Fig. 10: Plot of NC

Fig. 9 & 10 shows the PSNR and NC plots for embedding watermark in different sub-bands of cover image 2 over tuning factor range 0.1 to 1.0 By interpretation of the graphs, we confirm that the PSNR and NC plots for watermarking in horizontal sub-band of image yields better results as compared with vertical, diagonal and lower resolution approximation (LL) sub-bands.

CONCLUSIONS

In this paper we have proposed a new scheme of watermarking in transform domain using DWT and PCA, where the visibility level of watermark is tuneable with a secrete key. In this scheme a range is shown from invisible watermarking to visible watermarking with different levels of visibility from 0.1 to 1.0. Here we have also compared the embedding in different sub-bands of cover image on the basis of PSNR and NC values. We obtained similar results, which show that watermarking in horizontal sub-band of cover image gives better PSNR and NC values.

REFERENCES

[1] Saraju P. Mohanti "Digital Watermarking: A Tutorial Review", Department Of Computer Science And Engineering, University Of South Florida, Tampa, 1999.

[2] W. Bender, D. Gruhl, N. Morimoto, and A. Lu. "Techniques for data hiding". IBM Systems Journal, Vol.35.(3/4), 1996, pp. 313-336.

[3] M.Arnold, M. Schmucker, and S.D. Wolthusen, "Techniques and application of Digital Watermarking and Content Protection", Eds. Northwood, Artech House, 2003.

[4] F.Bossen M. Kutter, F. Jorden, "digital signature of color images using amplitude demodulation," in Proc. Of SPIE storage and retrieval for image and video database, San lose, USA, vol. 3022-5, February 1997, pp. 518-526.

[5] B.M.Macq, J.J. Quisquater, "Cryptology for DigitalTV Broadcasting." Proc. of the IEEE, vol. 83, no. 6, pp. 944-957, Jun 1995.

[6] G. Caronni, "Assuring Ownership Rights for Digital Images", Proc. of Reliable IT Systems, VIS '95, Germany, pp. 251-263, 1995.

[7] N. Nikolaidis, I. Pitas, "Copyright Protection of Images using Robust Digital Signatures", in Proc. of, IEEE Int. Conf. On Acoustics, Speech, Signal Processing, vol. 4, pp. 168-2171, 1996.

[8] S. Sinha, P. Bardhan, S. Pramanick, A. Jagatramka, D. K. Kole and A. Chakraborty, "Digital Video Watermarking using Discrete Wavelet Transform and Principal Component Analysis", International Journal of Wisdom Based Computing, Vol. 1 (2), August 2011

[9] M. Chandra and S. Paney, "A DWT Domain Visible Watermarking Techniques for Digital Images", 2010 International Conference on Electronics and Information Engineering (ICEIE 2010).

[10] N. Terzija, "Digital Image Watermarking in The Wavelet Domain", Technical Report, Faculty of Engineering Sciences, Gerhard-Mercator-University at Duisburg, December 2002.

[11] H. Hotelling, Relations Between two sets of variates, Biometrika, 27, 321.

[12] H. Abdi. & L.J. Williams, "Principal component analysis", Wiley Interdisciplinary Reviews: Computational Statistics, 2: 433-459