Robust Digital Image-Adaptive Watermarking Using BSS Based Extraction Technique

Sangeeta Jadhav djsangeeta@rediffmail.com
Asst Prof./E & TC
AIT,Pune University
Pune (MS), 411015, India

Dr Anjali Bhalchandra asbhalchandra@yahoo.com
HOD & Prof/E & TC
BAM University
Aurangabad (MS), 431005, India

Abstract

In a digital watermarking scheme, it is not convenient to carry the original image all the time in order to detect the owner's signature from the watermarked image. Moreover, for those applications that require different watermarks for different copies, it is preferred to utilize some kind of watermark-independent algorithm for extraction process i.e. dewatermarking. Watermark embedding is performed in the blue channel, as it is less sensitive to human visual system .This paper proposes a new color image watermarking method ,which adopts Blind Source Separation (BSS) technique for watermark extraction. Single level Discrete Wavelet Transform (DWT) is used for embedding . The novelty of our scheme lies in determining the mixing matrix for BSS model during embedding. The determination of mixing matrix using Quasi-Newton's (BFGS) technique is based on texture analysis which uses energy content of the image. This makes our method image adaptive to embed the watermark into original image so as not to bring about a perceptible change in the marked image. An effort is also made to check feasibility of proposed method in device dependent color spaces viz. YIQ,YCbCr and HSI . BSS based on joint diagonalization of the time delayed covariance matrices algorithm is used for the extraction of watermark. The proposed method, undergoing different experiments, has shown its robustness against many attacks including rotation, low pass filtering, salt n paper noise addition and compression. The robustness evaluation is also carried out with respect to the spatial domain embedding.

Keywords: - DWT, BSS, BFGS, Mixing matrix, Attacks, Dewatermarking.

1. INTRODUCTION

With the development of network and multimedia techniques, data can now be distributed much faster and easier. Unfortunately, engineers still see immense technical challenges in discouraging unauthorized copying and distributing of electronic documents [1, 2]. Different kinds of

handwritten signatures, seals or watermarks have been used since ancient times as a way to identify the source or creator of document or picture. However, in digital world, digital technology for manipulating images has made it difficult to distinguish the visual truth. One potential solution for claiming the ownership is to use digital watermarks. A digital watermark is a transparent, invisible information pattern that is inserted into a suitable component of the data source by using a specific computer algorithm. In nature, the process of watermark embedding is the same as some special kind of patterns or under-written images are added into the host image, we can consider it as a mixture of host image and watermark , thus without host image, the watermark detection is equal to blind source separation in the receiver. Blind digital watermarking does not need the original images or video frames in the detection stage, thus it is the only feasible way to do watermarking in many multimedia applications, such as data monitoring or tracking on the internet, notification of copyright in playing DVD's. In particular some watermarking schemes require access to the 'published' watermarked signal that is the original signal just after adding the watermark. These schemes are referred as semi-blind watermarking schemes. Private watermarking [3] and non-blind-watermarking mean the same: the original cover signal is required during the detection process. The watermarked image is viewed as linear mixture of sources [4] i.e. original image and watermark and then we attempt to recover sources from their linear mixtures without resorting to any prior knowledge by using Blind Source Separation theory. Independent Component Analysis (ICA) is probably the most powerful and widely-used method for performing Blind Source Separation [15].
To present the basic principle of this new watermarking technique based on BSS, the paper is restricted to watermarking and dewatermarking with the simplest BSS model. The BSS model used to embed the watermark in the blue channel is shown below.

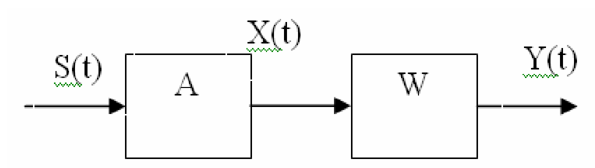


FIGURE 1: BSS model

.
The simplest BSS model assumes the existence of 'n' independent components i.e. the source signals S1,S2,S3……..Sn [S(t)], and the same number of linear and instantaneous mixtures X1,X2,……Xn [X(t)] of these sources . In vector matrix notation form the mixing matrix model can be represented as -

$$x=A*s \tag{1}$$

Where A is square (n x n) mixing matrix. W is separating matrix or demixing matrix and Y1,Y2,……Yn.[Y(t)] are estimated output sources which should be identical to sources represented by S(t).
Image watermarking techniques proposed so far can be categorized based on the domain used for watermarking embedding domain. The first class includes the spatial domain methods [9]. These embed the watermark by directly modifying the pixel values of the original image. The second contains Transform domain techniques, Discrete Fourier Transform (DFT), Discrete Wavelet Transform (DWT) [11], Discrete Cosine Transform (DCT) [10].The third class is the feature domain technique, where region, boundary and object characteristics are taken into account [16]. The first class includes the works of Adib et al [4] have proposed to use the blue channel as the embedding medium . In [5] the authors have benefited from a new decomposition of the color images by the use of hyper complex numbers, namely the Quaternion and they achieved their watermarking/data-hiding operation on the component of the quaternion Fourier Transform. In the recent past, significant attention has been drawn to Blind source Separation by Independent Component Analysis [7,8] and has received increasing care in different image data applications such as image data compression, recognition, analysis etc. The technique of BSS

has been extended to the field of watermarking images [6, 12]. In [4], several assumptions are made regarding values of the mixing coefficients, distribution of the watermark as well as the mixing process. The proposed BSS based method is more flexible in the sense that the system finds out best suited mixing matrix using Quasi-Newton (BFGS)algorithm to keep the watermark hidden in the selected image. Watermark embedding in wavelet domain with adaptive mixing matrix makes the method quite difficult to extract unknown watermark using the existing simple methods, for example using [12].

The objective of this paper is to introduce an efficient digital image watermarking scheme based on BSS theory adopting watermark embedding in wavelet domain, which is more robust to the dewatermarking attacks as compared to the methods [10] in spatial domain embedding.

In the present work the effort has been also put to check the feasibility of proposed method in device dependent color spaces namely HSI, YIQ and YCbCr. In device dependent color spaces, color produced depends on parameters used as well as the equipment used for the display.

The RGB color space is highly correlated except of the blue channel because of its low sensitivity to human perception. The same set of embedding and detecting procedure is applied to all the color spaces so as to achieve the best comparison among them. The simulation results are shown for blue channel of RGB color space.

A BSS/ICA algorithm based on Joint diagonalization of the time delayed covariance matrices [13] is used for the extraction of watermark.

The paper is structured as follows: section 2 describes the proposed watermarking method including the watermark embedding using DWT and estimation of mixing matrix. The watermark extraction using BSS/ICA algorithm is also discussed in this section. The simulation results are illustrated in section 3. The robustness testing w. r. t. spatial domain embedding is analysed in section 4 .Finally section 5 mentions conclusions and future work.

2. WATERMARKING SYSTEM

In the generic watermark embedding scheme, the inputs to the system are the original image and the watermark. To assure the identifiability of BSS model , it is required that the number of observed linear mixture inputs is at least equal to or larger than the number of independent sources [15] .

2.1 The Watermark Embedding Scheme

In this paper the effective watermark embedding consists of mainly three phases. In first phase the blue channels of the host image and watermark image are extracted. The size of host image selected is 512*512 (M x M) and size of watermark image is 64*64 (N x N) so that M>>N. In order to determine the sub-image of interest ,the host image is divided into 128*128 blocks and a sliding square window containing Nb number of such blocks in both the horizontal and vertical directions (a tentative sub-image) is considered. It has been shown that the energy of textured portion of image is high. Based on the energy content of the image, the two blue channel sub-images of size 128*128, one representing the smooth portion and other the textured one are taken out. In high textured area the visibility is low; therefore a textured sub-image is selected to embed the watermark. In second phase a single level DWT using haar wavelet function is applied to this textured sub-image and only the lowest frequency band (LL1 of size 64*64) is selected for embedding the watermark (size 64*64) . To have as many mixtures as sources, the mixing matrix A is selected to be a square matrix (order 2*2). The mixing operator 'A' has to be appropriately chosen such that the human vision can not determine that the message (watermark) is contained inside a host image. A Quasi-Newton (BFGS) algorithm [14] is used to estimate the mixing matrix A to keep the watermark hidden.

2.2 Estimation Of Mixing Matrix (Statistical Model)

In the proposed method the sources namely original image and watermark are known. The concept of correlation cancellation is used to estimate the mixing matrix A [13].

Consider two zero mean vector signals $x(k) \in \mathbb{R}^m$ and $s(k) \in \mathbb{R}^n$ that are related by the linear transformation

x(k)=As(k) +e(k)
where A $\in \mathbb{R}^{m\times n}$ is unknown full rank mixing matrix and e(k) $\in \mathbb{R}^{m}$ is a vector of zero mean error , interference or noise depending on application. Generally vectors s(k) and x(k) are correlated i.e. $R_{xs} = E\{x\, s^T\} \neq 0$ but the error or noise e is uncorrelated with s, hence our objective is to find out the matrix A such that the new pair of vectors
e = x-As and s are no longer correlated with each other and can be expressed in terms of equation as-

$$R_{es}= E\{es^T\}=E\{(x-As)s^T\}= 0$$

The cross correlation matrix can be written as

$$R_{es} =E\{xs^T-Ass^T\}=R_{xs}-AR_{ss}$$

Hence the optimal mixing matrix can be expressed as

$$A_{opt}=R_{xs}\, R_{ss}^{-1} =E\{xs^T\}(E\{ss^T\})^{-1}$$

The same result is obtained by minimizing the mean square error cost function

$$J(e) = \frac{1}{2}E\{e^T e\}=E\{(x-As)^T(x-As)\}$$

$$= \frac{1}{2}(E\{x^Tx\}-E\{s^TA^Tx\}-E\{s^TA^Tx\}-E\{x^TAs\}+E\{s^TA^T As\})$$

By computing the gradient of the cost function J(e) w.r.t. A we obtain

$$\frac{\partial J(e)}{\partial A} = - E\{xs^T\}+AE\{ss^T\}$$

Hence applying the Quasi-Newton or BFGS approach, we obtain adaptive algorithm for the estimation of the mixing matrix to keep the watermark hidden in the host or original image.

$$\Delta A(k)= - V * \frac{\partial J(e)}{\partial A(k)} \qquad (2)$$

J(e) could be energy, entropy ,homogeneity or inertia of the original image. In this paper energy of the original image is used.V is a system constant. Δ A(k) = controlled rate of change of J w.r.t. mixing matrix A. Assuming the optimum value of A is achieved when gradient is zero.

$$A = \begin{bmatrix} a11 & a12 \\ a21 & a22 \end{bmatrix} \qquad (3)$$

Equation (2) becomes

$$\frac{\partial J(e)}{\partial A(k)} = \frac{\partial J}{\partial a11}+\frac{\partial J}{\partial a12}+\frac{\partial J}{\partial a21}+\frac{\partial J}{\partial a22}$$

But if we can simplify coefficients by certain assumptions

i.e. let a11 =a12=1
a12= 1-t
a22= 1+t

$$A=\begin{bmatrix} 1 & 1-t \\ 1 & 1+t \end{bmatrix}$$

Then instead of A we just have to check for ‘t’ so the equation gets reduce to

$$\frac{\partial J(e)}{\partial A(k)} = \frac{dJ(e)}{dt} \quad [\text{since } dA=dt]$$

In BFGS algorithm the value of system constant V =1 initially and as the process grows it gets updated and the value of 'A' gets converge easily.

In third phase, one of the compound sub-images (watermarked sub-image) is encrusted into the corresponding blocks of the earlier chosen region (by the BFGS algorithm) for the embedding ,in the original image called watermarked image and is open to the public. The remaining watermarked textured sub-image is kept secret by the copyright owner. It will constitute the secret key corresponding to the location at which the watermark is fused with the original (host) image.

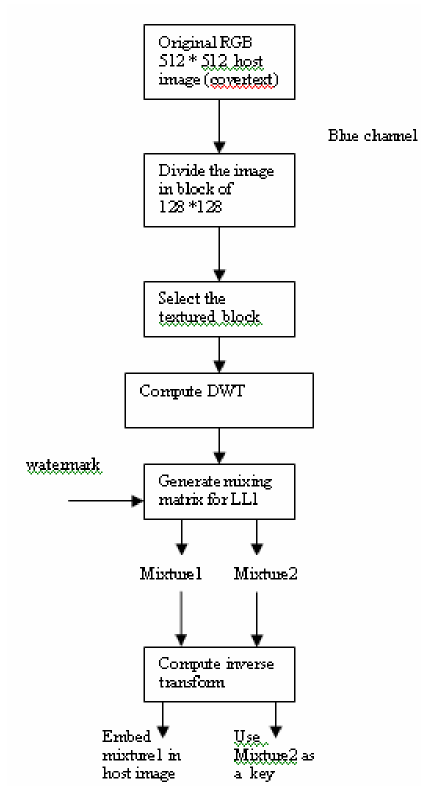


FIGURE 2: Flowchart showing the Watermark Embedding Process

As shown in Figure 2- Mixture1 and Mixture2 has following relationship-
Equation (1) i.e. x=A*s can be written in matrix form as

$$\begin{bmatrix} Mixture1 \\ Mixture2 \end{bmatrix} = A * \begin{bmatrix} Source1 \\ Source2 \end{bmatrix}$$

From Equation (3)

Mixture1=a11* Source1 + a12* Source2
Mixture2=a21* Source1 +a22* Source2

Thus a watermark embedding process is summarized in the following steps.

Step-1 Take the host and watermark color images, respectively of size (MxM) and (NxN) with M>>N. Select their blue channels.

Step-2 Select textured regions block based on energy metric. Take one level DWT and use LL1 for further processing.

Step-3 Obtain the mixing matrix using Quasi-Newton (BFGS) algorithm [14] in order to keep watermark hidden in textured sub-image to form the watermarked mixtures. Take inverse wavelet of watermarked mixtures.

Step-4 One of the compound sub-images (watermarked sub-image) is encrusted into the corresponding blocks of the earlier chosen region of high energy in the original image.The other secret watermarked mixture in blue channels must be kept for a prospective use in the watermark extraction process.

2.3 Watermark Extraction

A) PCA whitening –watermark detection

Standard Principal Component Analysis (PCA) is often used for whitening process [12], since it can compress information optimally in the mean-squared error sense, while filtering possible noise simultaneously.The PCA whitening matrix is given by

$V=D^{-1/2}U^{T}$

Where D is the diagonal matrix of data covariance matrix $E[X_j X_j^T]$ and U is its eigenmatrix,and E[.] denotes the expectation operator.

If the rank of D is equal to two for watermarked image ,meaning that there are totally two image sources. On the otherhand, if the image is unwatermarked image the rank D will be reduced to one.

After pre-whitening process, the sources are recovered by iteratively estimating the unmixing matrix W through a joint diagonalization of the time delayed covariance matrices algorithm[13]

As shown in Figure 3, the extraction process can be summarized in the following steps.

Step-1 Extract the marked block from the tampered watermarked image by using the first part of the key which is the position key.

Step- 2 Obtain the blue channels of the extracted blocks.

Step -3 Apply a PCA whitening process on associated blue channel .

Step-4 Post processing has to be done on the whitened blue channel. Joint diagonalization algorithm is used to recover both host image and watermark.

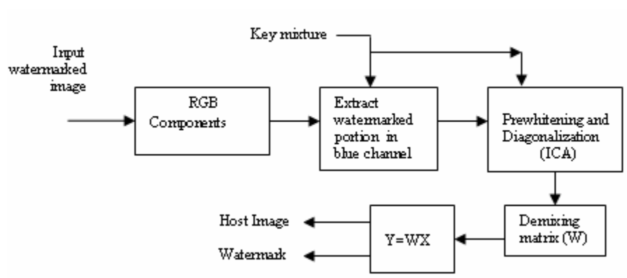
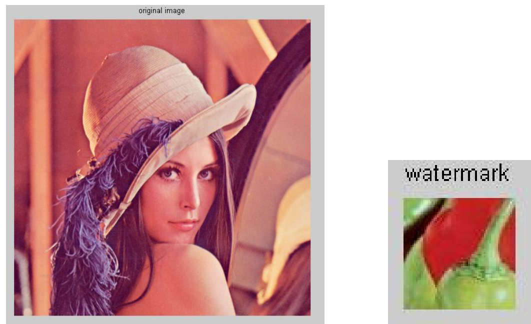


FIGURE 3: Watermark Extraction Using BSS

3 SIMULATION RESULTS

3.1 Feasibility Of Proposed Method In Blue Channel Of RGB Color Space

Simulation experiments are conducted to demonstrate the feasibility and robustness of proposed BSS based watermark extraction method. Some results of DWT embedding are given below



(a) (b)

FIGURE 4: Original Image (a) and Watermark (b)

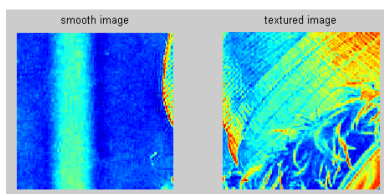


FIGURE 5: Smooth and Textured Portions of Original Image

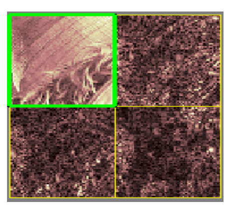


FIGURE 6: DWT of Textured Sub-image

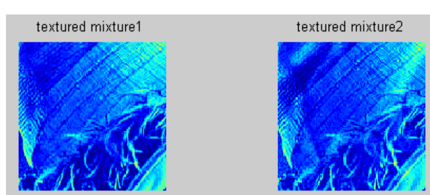
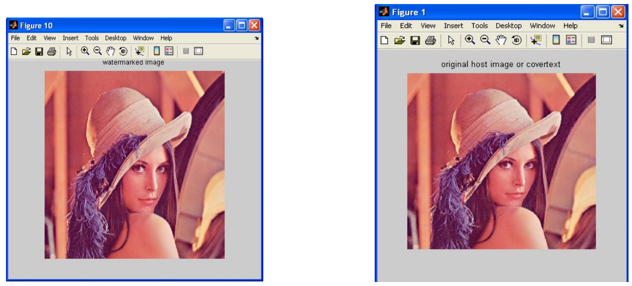


FIGURE 7 : Watermarked Mixture Sub-images



(a) (b)

FIGURE 8 : Watermarked Image (a) Original Host Image or Covertext (b)

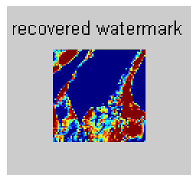


FIGURE 9: Recovered Watermark

3.2 Feasibility Of Proposed Method In Device Dependent Color Spaces

The proposed method of watermarking is tested over device dependent color spaces mentioned in Table 1 The Table 1 shows the value of PSNR and Correlation Coefficient computed by Equation (5) and (6) for recovered watermark using spatial domain embedding. The value of mixing matrix generated using BFGS method is also mentioned in the table.

Sr No.	Color Space	PSNR	Correlation Coefficient	Mixing Matrix
1	**YIQ** [:,:,1]	73.11	0.86	1 0.31 1 -0.1
2	**YCbCr** (I) Y	26.07	0.80	1 0.0023 1 -0.001
	(II) Cb	25.6	-0.84	1 0.0013 1 - 0.0013
	(III) Cr	25.85	0.96	1 0.001 1 -0.0022
3	**HSI/HSV** **(I) H**	Inf	0.53	1 0.864 1 -0.883
	(II) S	54.98	0.61	1 0.093 1 -0.24
	(III) I(Value)	53.54	0.285	1 0.474 1 -0.139

TABLE 1: Embedding in Spatial Domain Without Attack

In Table 2, the performance parameters PSNR and Correlation Coefficient computed for recovered watermark using BSS extraction technique; with DWT domain embedding is shown. The value of mixing matrix is also shown in the table.

Sr No.	Color Space	PSNR	Correlation Coefficient	Mixing Matrix
1	**YIQ** (I) [:, :, 1]	68.34	0.87	1 0.0017 1 -0.34
2	**YCbCr** (I) Y	26.28	0.89	1 0.001 1 -0.0014
	(II) Cb	26.03	0.79	1 0.0013 1 - 0.0013
	(III) Cr	25.80	0.96	1 0.0013 1 -0.0035
3	**HSI/HSV** **(I) H**	66.54	0.73	1 0.128 1 -0.907
	(II) S	55.25	0.612	1 0.001 1 -0.993
	(III) I(Value)	53.56	0.29	1 0.0012 1 -0.26

TABLE 2: Embedding in DWT Domain Without Attack

4 ROBUSTNESS TESTING WITH RESPECT TO SPATIAL DOMAIN EMBEDDING

The watermarking system should be robust against data distortions introduced through standard data processing and attacks. It should be virtually impossible for unauthorized users to remove it and practically the image quality must be degraded before the watermark is lost. There are many attacks against which image watermarking system could be judged . The attacks include average filtering ,rotation (+90^0),median filtering ,Salt n Paper noise and so on .These various attacks are applied to the watermarked images to evaluate whether the proposed dewatermarking system can recover the embedded watermark, thus measuring the robustness of the watermarking system to these types of attacks.

Mean Square Error (MSE), PSNR (Peak signal to Noise Ratio) and NC (Normalized Cross-Correlation) are used to estimate the quality of extracted watermark.

The equations used are defined as below-

$$MSE = \frac{1}{MN}\sum_{1}^{M}\sum_{1}^{N}\left[r(i,j) - r^{*}(i,j)\right]^2 \qquad (4)$$

Where r(i,j) represents pixel at location (i,j) of the original watermark and r^*(i ,j) represents the pixel at location (i ,j) of recovered watermark. M,N denotes the size of the pixel.

$$PSNR = 10\log_{10}\left[\frac{255^2}{MSE}\right] \qquad (5)$$

$$NC = \frac{\sum_{m=1}\sum_{n=1} W * W'}{\sqrt{\sum_{m=1}\sum_{n=1} W^2} \times \sqrt{\sum_{m=1}\sum_{n=1} W'^2}} \qquad (6)$$

Where W is original watermark and W' is recovered watermark with zero mean value each.

As shown in Figure 10 , the PSNR in dB is calculated by using Equation 5 and compared for both the types of embedding viz. Spatial domain and DWT domain embedding. It is observed that the PSNR obtained in DWT embedding is high under various attack conditions. In Figure 11, the correlation coefficients (NC) comparison is shown.

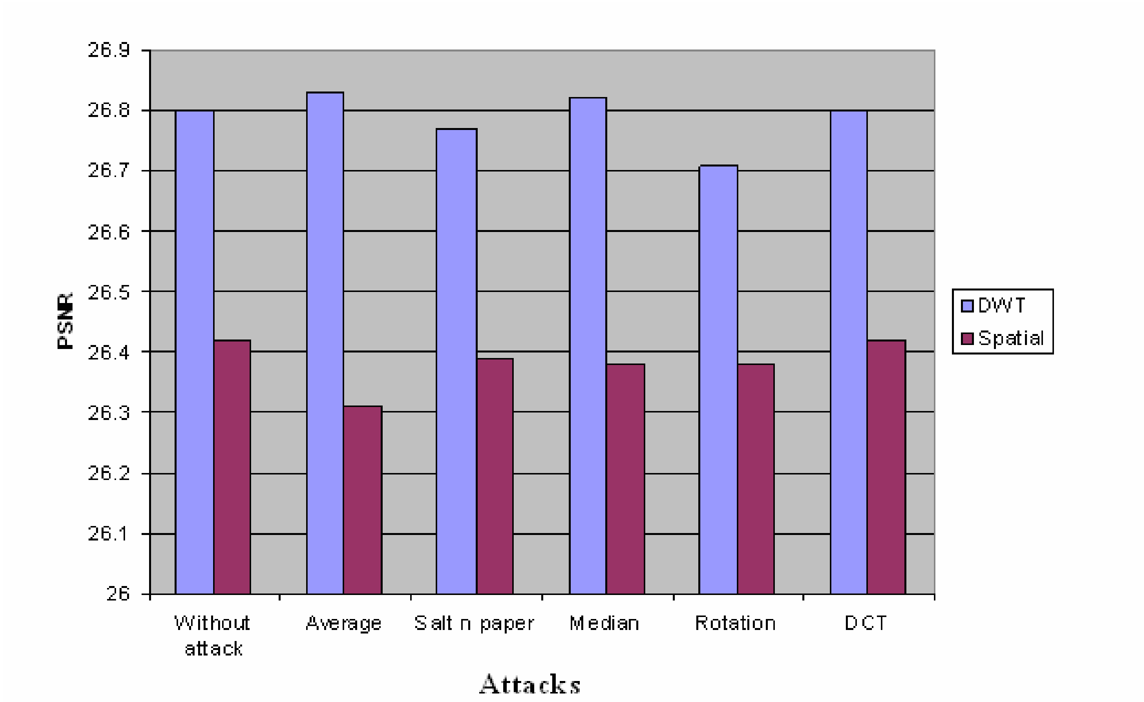


FIGURE 10: PSNR(dB) comparison for DWT based and spatial watermark embedding

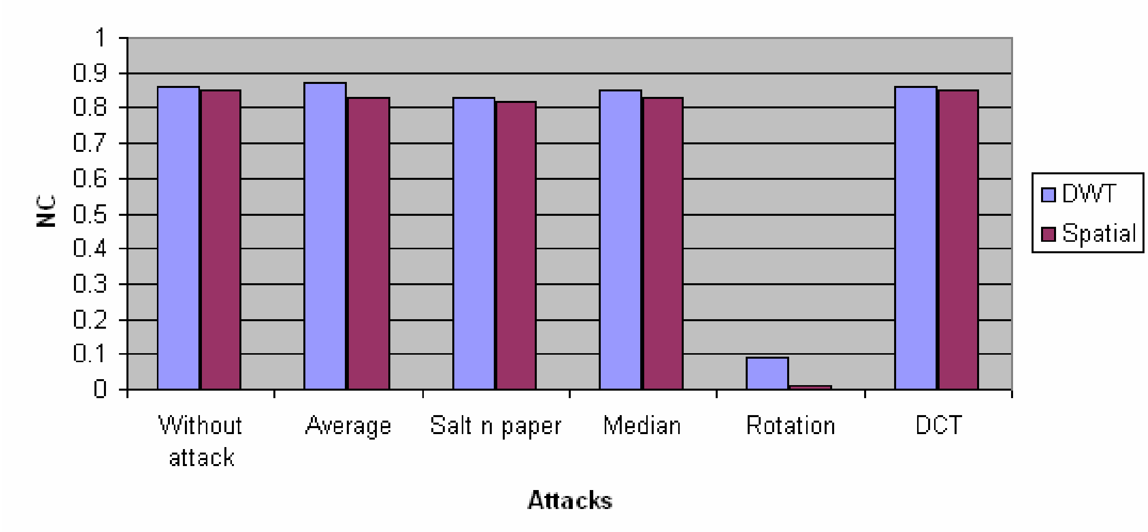


FIGURE 11: Normalized Correlation Coefficient (NC) Comparison

JPEG Quality variation is tested for DWT based and spatial based embedding and plotted against PSNR of recovered watermark.
As shown in graph ,it is observed that the PSNR values are more in case of DWT based embedding as compared to the Spatial based embedding proving that DWT based embedding is more robust against the
watermarking attacks .

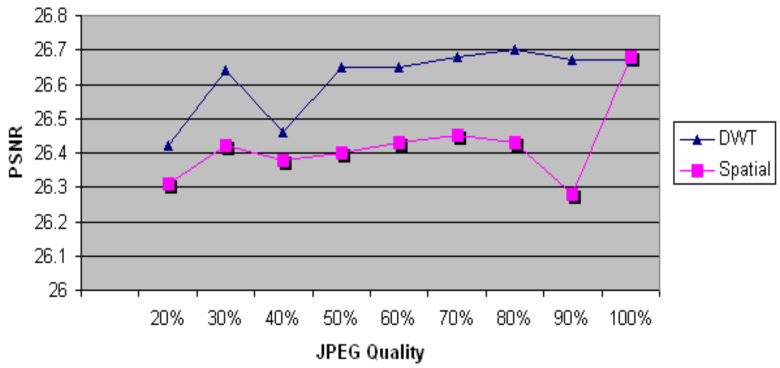


FIGURE 12: PSNR (dB) comparison for JPEG Quality variation

5 CONSLUSION & FUTURE WORK

In this paper, we proposed a digital color image watermarking system using wavelet(DWT) domain embedding and adopting Blind Source Separation theory along with RGB decomposition to extract watermark. The novelty of our scheme lies in determining the mixing matrix for BSS model , based on energy content of the image using Quasi Newton (BFGS) method. This makes our method image adaptive to embed any image watermark into original host image. The effort has been also put to check the feasibility of proposed method in device dependent color spaces for the application of image watermarking. The watermark is readily detected by Principal Component Analysis(PCA) whitening process .The watermark is further separated by using BSS/ICA algorithm based on Joint diagonalization of the time delayed covariance matrices. The performance of the proposed method can be evaluated in terms of normalized correlation coefficient and PSNR with respect to spatial domain watermark embedding. Experimental results demonstrate the proposed watermarking scheme is more robust to various attacks as compared to spatial domain watermark embedding.

In future research work, it is proposed to implement the watermark extraction process in time-frequency domain using DWT in order to improve the performance for different types of images as well as to make the proposed watermarking scheme more robust against various attacks.

6 REFERNCES

[1] Peter H W Wong, Oscar C Au and YM Yeung, “A novel blind multiple watermarking technique for Images”, in IEEE transactions on circuits and ystems for video technology, vol.13, No. 8,pp 813 -830,yr 2003

[2] Chiou-Ting Hsu, Ja-Ling Wu, “Hidden Digital Watermarks in Images,” in IEEE transaction on image processing, vol.8 No.1, pp 58-68,Jan1999.

[3] Ibrahim Nasir et al. “ a New Robust Watermarking Scheme for Color Image in spatial Domain”, ieeexplore.ieee.org/iel5/4618740/4618741/04618875.pdf.

[4] Moilim Amir, Abdellah Adib et al , “Color Images Watermarking by Means of Independent Component Analysis”,in IEEE transaction pp 347-350,yr-2007.

[5] Tsz Kin Tsui et al , “ Quaternion Image Watermarking using the Spatio-Chromatic Fourier Coefficients Analysis”, in ACM 1-59593-447-2/06/0010, yr-2006.

[6] Thang V. Nguyen et al., “A Novel Digital Image Watermarking Scheme Using Blind Source Separation”, Proc of IEEE, ICIP,pp2649-2652,Yr 2004.

[7] Yang Su –min,Wang Jia-zhen,ZHANG Zheng-bao Ding Guo-liang, “A digital watermark based on blind signal separation algorithm,” in ICIP 2006 proceedings.

[8] Thang Viet et al. “A Simple ICA –based Digital Image Watermarking Scheme”,DSP archive ,Volume 18,issue 5,yr-2008.

[9] Peter H W Wong, Oscar C Au and YM Yeung, "A novel blind multiple watermarking technique for Images", in IEEE transactions on circuits and ystems for video technology, vol.13, No. 8,pp 813 -830,yr 2003
[10] Chiou-Ting Hsu, Ja-Ling Wu, "Hidden Digital Watermarks in Images," in IEEE transaction on image processing, vol.8 No.1, pp 58-68,Jan1999.
[11] Shang-Lin Hsieh et al , "a Color Image Watermarking Scheme Based on Secret Sharing and Wavelet Transform" ,IEEE,pp-2143-2147,Yr 2007.
[12] Dan Yu ,Farook Sattar et al , "Watermark Detection and Extraction using Independent component analysis Method",EURASIP Journal on Applied Signal Processing 2002:1,92-104
[13] A.Cichoki,Shun-ichi Amari , "Adaptive Blind Signal nd Image Processing –Learning Algorihms and pplications", Oct-2005,pp 0-175.
[14] L F Escudero, "On Diagonally preconditioning the 2-Steps BFGS Method With Accumulated Steps for Supra Scale Linearly Constrained Non Linear Programming"published in QUESTIIO-1982,vol-06 no-04, Yr 1982.
[15] A Hyvarinen and E. Oja , " Independent Analysis a Tutorial", http://www.cis.hut.fi/projects/ica/April99.
[16] SUN Jiande et al. "A Blind Video Watermarking Scheme based on ICA and Shot Segmentation",Science in China Series F Information Sciences,vol-49, No-3,pp302-312 Yr 2006.