ISSN: 0976-9102(ONLINE)
DOI: 10.21917/ijivp.2012.0053

ICA BASED DIGITAL IMAGE WATERMARKING BASED ON REDUNDANT DISCRETE WAVELET TRANSFORM

G. Thirugnanam[1] and S. Arulselvi[2]
Department of Instrumentation Engineering, Annamalai University, India
E-mail: [1]thirugnanam_me@yahoo.com and [2]ggtt_me@yahoo.com

Abstract

In recent years, access to multimedia data has become much easier due to rapid growth of the internet. While this is usually considered an improvement of everyday life, it also makes unauthorized copying and distributing of multimedia data much easier, therefore presenting a field of watermarking. Many literatures have reported about Discrete Wavelet Transform watermarking techniques for data security. However, DWT based watermarking schemes are found to be less robust against image processing attacks and the shift variance of Wavelet Packet Transform causes inaccurate extraction. In this paper, an attempt is made to develop a watermarking scheme based on RDWT and extraction using Independent Component Analysis (ICA). Various ICA like Fast ICA, Radical ICA, Pearson ICA, Jade ICA, Maximum Likelihood ICA, Mean Field ICA and Kernel ICA are proposed and implemented for extraction. Simulation results reveal that Fast ICA produces better Similarity Measure and Mutual Information over other ICA methods.

Keywords:

Authentication, Digital Watermarking, Independent Component Analysis, Redundant Discrete Wavelet Transform

1. INTRODUCTION

Recently, the tremendous growth of the internet has increased multimedia services, such as electronic commerce, pay-per-view, video–on-demand, electronic newspapers and peer to peer media sharing. As a result, multimedia data can be obtained quickly over high speed network connections. However, authors, publishers, owners and providers of multimedia data are reluctant to grant the distribution of their documents in a networked environment because the ease of intercepting, copying and redistributing electrical data in their exact original form encourages copyright violation. Therefore, it is crucial for the future development of networked multimedia systems that robust methods are developed to protect the intellectual property right of data owners against unauthorized copying and redistribution of the material made available on the network. Furthermore, it is an important issue to develop a robust watermarking scheme with a better tradeoff between robustness and imperceptibility.

Recently many literatures have reported the watermarking schemes based on Discrete Wavelet Transform (DWT) [1-3]. Among these schemes, the one which require the original information and secret keys for the watermarking extraction are called private watermark schemes. Schemes which require the watermark information and secret keys are called semi-private or semi-blind schemes. Schemes which need secret keys rather than the original information are called public or blind watermark schemes.

In general, the robustness of private watermark scheme is good to endure signal processing attacks. However, they are not feasible in real application, such as DVD copy protection where the original information may not be available for watermark detection. On the other hand, semi-blind and blind watermark schemes are more feasible in that situation. However, they have lower robustness than the private watermark schemes.

Watermarking techniques can be divided into two main groups: embedding watermarks in the spatial domain or in the frequency domain. Spatial domain watermarking directly embeds the watermark into the object while frequency domain watermarking embeds the watermark by changing frequency component values by an orthogonal transformation. DWT based watermarking is a widely used technique to embed a spread spectrum watermark into DWT coefficients. Moreover, the conventional watermark detection uses transform techniques to decompose the corrupted image, of which the ownership is determined, and from which the watermark is recovered. In general, toward the aim of watermark recovery, some detection systems require previous knowledge of the watermark such as its location, the strength, the threshold or the original image. Therefore, a watermarking algorithm must be able to embed watermarks to satisfy those requirements.

A new approach is proposed in this paper, where the watermarking method is based on RDWT for embedding the information. In general, RDWT produce an over complete, over sampled expansion system which is used to embed watermark to the transformed coefficients. The perceptual model is applied with stochastic approach for watermark embedding. This is based on computation of a noise visibility function (NVF) which has local image properties where the strength of watermarking is controlled. The result is that watermarks at texture and edge areas are stronger than flat areas. Besides, an intelligent detection technique based on ICA is implemented for extraction without the use of previous knowledge of the watermark and even the transformation process. In T.D.Hien[7] work, watermark is extracted using FastICA and multi- logo watermark is embedded. In this proposed work, various ICA are attempted to extract a single watermark. Robustness against various attacks like filtering, cropping, JPEG compression and transparencies of the proposed scheme are demonstrated with simulation results.

This paper is organized as follows: Section 2 reviews the Wavelet Transform. Section 3 discusses the Redundant Discrete Wavelet Transform Section 4 explains Watermark embedding algorithm and in Section 5, Watermark extraction using various ICA are explained. Simulation results are presented in Section 6 and finally conclusions are drawn in Section 7.

2. WAVELET TRANSFORM

The main advantage of wavelets over Fourier analysis is that they allow both spatial and frequency resolution during decomposition. Wavelet transform allows the decomposition of the signal in narrow frequency bands while keeping the basis signals space limited. Fig.1 shows a two level DWT decomposition tree using low pass and high pass analysis filter banks $h(-m)$ and $g(-m)$ respectively. If the level of decomposition is increased, the approximate image will be more stable [3]. But the complexity increases and the amount of information that can be embedded will be decreased. As a compromised way, the original image is decomposed into two levels. In wavelet analysis, an original image can be decomposed into an approximate image LL1 and three detail images LH1, HL1 and HH1 as shown in Fig.1. Using wavelet analysis on the approximate image LL1 again, four lower-resolution sub-band images LL2 and three detail images LH2, HL2 and HH2 will be obtained and the approximate image holds the most information of the original image and others contain some high-frequency information such as the edge details. The main drawback of not embedding the watermark in the LL2 subband, which will lead to serious degradation of image quality. Hence, the watermark can be embedded either in HL2 or LH2. In this paper, HL2 sub-band is chosen to embed the watermark. $\downarrow$ and $\uparrow$ denotes downsampling and upsampling by a factor of two, respectively.

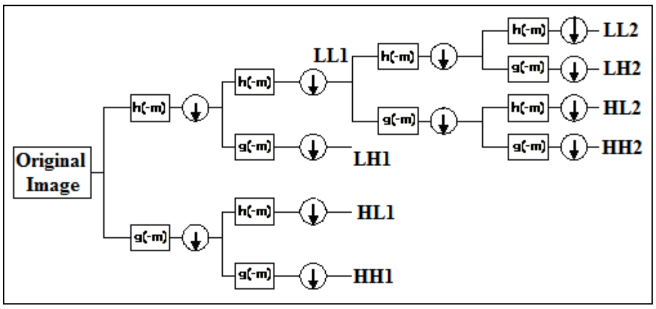

Fig.1. Two level 2D DWT Analysis Filter Banks

3. REDUNDANT DISCRETE WAVELET TRANSFORM

Unlike DWT and Wavelet Packets, RDWT gives an over complete representation of the input sequence and functions as a better approximation to the continuous wavelet transform. The RDWT is shift invariant, and its redundancy introduces an over complete frame expansion. It is known that frame expansion increases the robustness to additive noise, that is, addition of noise to transform coefficient results in less signal distortion for frame expansions than for orthogonal expansion [6]. RDWT has been proposed for signal detection and enhancement, since the RDWT maintains uniform sampling rate in time domain and in some respects, is a discrete approximation to the continuous wavelet transform. RDWT removes the down-sampling operation from the traditional critically sampled DWT and wavelet packets. The RDWT eliminates downsampling and upsampling of coefficients, and at each scale, the number of each sub-band size is doubles that of the input image as shown in Fig.3. The filters themselves are up-sampled to fit the growing data length. Lack of down-sampling in the RDWT analysis yields a redundant representation of the input sequence. Specifically, two valid descriptions of the coefficients exist after one stage of RDWT analysis. The advantage of this representation is that each RDWT coefficient is located within its sub-band in its spatially correct position.

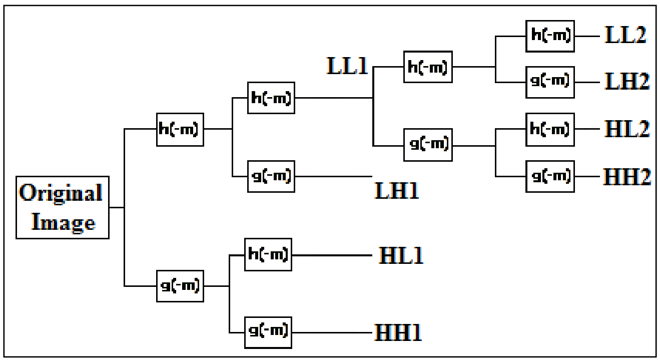

Fig.2. Two level 2D RDWT Analysis Filter Banks

4. WATERMARK EMBEDDING

The original image is decomposed into two levels using RDWT as shown in Fig.2. The approximation components at LL2 sub-band are not chosen to embed watermark because they will seriously degrade image quality. Similarly the diagonal detail coefficients HH2 are also not considered because security is poor when watermark is embedded. Hence, the middle sub-band HL2 is chosen to embed watermark based on the tradeoff between imperceptibility and robustness as shown in Table.1.

Table.1. Selection of sub-band to embed watermark

Sub-bands	PSNR(dB)	Similarity measure	Mutual Information
LL2	21.7943	0.9715	2.5254
HL2	47.7112	0.9779	2.5539
LH2	46.6801	0.9738	2.5412
HH2	47.7131	0.9536	1.8231

A stochastic model of the cover image is applied to an adaptive watermark by computing NVF with non-stationary Gaussian model [7]. In this case, NVF can be expressed by,

$$NVF(i,\ j) = \frac{1}{1+\sigma_X^2(i,\ j)} \tag{1}$$

where, $\sigma_X^2(i,\ j)$ denotes variance of the cover image in a window centered on the pixel with coordinates $(i,\ j)$. By applying NVF, the watermark in texture and edges becomes stronger than in flat areas. The watermark is embedded using the following equations,

$$I'HL_2(i,j) = HL_2(i,j) + E(HL_2)\alpha_1(1 - NVF(i,j))$$
$$W(i,j) + \frac{E(HL_2)}{10}.\alpha_1.NVF(i,j).W(i,j) \tag{2}$$

where, $I'HL_2(i, j)$ is watermarked transform coefficients, $E(HL_2)\alpha_1$ denote the watermark strengths of texture and $\frac{E(HL_2)}{10}.\alpha_1$ denote the watermark strengths of edge regions for HL subband. α_1 is the smoothing factor at the texture regions and flat regions and E denotes the mean and $W(i, j)$ is the watermark. To retrieve the watermarked image, inverse RDWT is performed.

5. INDEPENDENT COMPONENT ANALYSIS

ICA is a statistical technique for obtaining independent sources S from their linear mixtures X, when neither the original sources nor the actual mixing A are known. This is achieved by exploiting higher order signal statistics and optimization techniques[12,13]. The result of the separation process is a demixing matrix W, which can be used to obtain the estimated unknown sources, $\bar{S}$ from their mixtures[8,9]. This process is described by,

$$X = AS \rightarrow \bar{S} = WX \tag{3}$$

Various ICA algorithm used in this paper work for watermark extraction are discussed below:

5.1 FAST ICA ALGORITHM

Aapo Hyvarinen and Erkki Oja have proposed a Fast ICA algorithm and it is based on a fixed-point iteration scheme [8]. The operation of Fast ICA algorithm is outlined as follows:

i. The mean of the mixed signal X is subtracted so as to make X as a zero mean signal as

$$X = X - E[X],$$

where, $E[X]$ is the mean of the signal.

ii. Then covariance matrix is,

$$R = \mathrm{E}[XX^T]$$

obtained and eigen value decomposition is performed on it and is given by

$$R = \mathrm{E}\,D\,\mathrm{E}^T$$

where E is the orthonormal matrix of eigenvalues of R and D is the diagonal matrix of eigenvalues. Find the whitening matrix, P which transforms the covariance matrix into an identity matrix is given by,

$$P = Inv\left(sqrt(D) \times \mathrm{E}^T\right) \tag{4}$$

iii. Choose an initial weight vector W, such that the projection $W^T X$ maximizes non gaussianity as,

$$W^+ = E\left\{X * g\left(W^T X\right)\right\} - E\left\{g'\left(W^T\right)\right\}W \tag{5}$$

where, g is the derivative of the nonquadratric function. The variance of $W^{+T}X$ must be made unity. Since X is already whitened it is sufficient to constrain the norm of W^+ to be unity.

$$W = W^+ / \|W^+\| \tag{6}$$

If W not converges means go back to step (iv).

iv. The demixing matrix is given by,

$$W = W^T \times P \tag{7}$$

and independent components are obtained by,

$$\bar{S} = W \times X \tag{8}$$

5.2 RADICAL ALGORITHM

The RADICAL (Robust, Accurate, Direct Independent Component Analysis Algorithm) estimates independent sources using differential entropy estimator based on 'm'-spacing estimator [10].

The data X is assumed to be whitened using whitening matrix P. The contrast function is to be minimized by RADICAL is almost equivalent to Vasicek estimator,

$$\hat{H}_{RADICAL}\left(Z^1,...,Z^N\right) = \frac{1}{N}\sum_{i=1}^{N-m_N} \log\left(\frac{N+1}{m_N}\left(Z^{(i+m_N)} - Z^{(i)}\right)\right) \tag{9}$$

where Z is the random variable and 'm' be the size of spacing. The value of 'm' is taken as $\sqrt{N}$ where N is the number of samples in each source. Let R be the number of replicated points per original data point to eliminate the local minima problem. The contrast function evaluated using Eq.(9) is given by,

$$H = R * Z^T. \tag{10}$$

The demixing matrix, W = RP and the estimated source is given by,

$$\hat{S} = WX. \tag{11}$$

5.3 PEARSON ICA

The Pearson-ICA algorithm is a mutual information-based method for blind separation of statistically independent source signals [10]. The data matrix X is considered to be a linear combination of statistically independent components as,

$$X = AS$$

where A is a linear mixing matrix and the columns of S contain the independent components of which at most one has Gaussian distribution. The goal of ICA is to find a matrix W such that the output,

$$\hat{S} = WX$$

is an estimate of possibly scaled and permutated source matrix S. In order to extract the independent component sources for a demixing matrix W, that minimizes the mutual information of the sources.

$$W_{k+1} = W_k + D\left(E\left\{\phi(y)y^T\right\} - diag(E\left\{\phi(y_i)y_i\right\})\right) \tag{12}$$

where, $D = diag\left(\frac{1}{(E\{\varphi(y_i)y_i\} - E\{\varphi'(y_i)\})}\right)$,

$\varphi(y) = \tanh(2y)$ and y_i is the current data.

5.4 JADE ICA

Another signal source separation technique is the JADE (Joint Approximate Diagonalization of Eigen Matrices) algorithm [9]. This exploits the fourth order moments in order to separate the source signals from mixed signals. The operation of JADE is as given below:

The whitened matrix U is calculated using whitening matrix P and data X is given by,

$$U=PX \tag{13}$$

The fourth cumulants of the whitened mixtures $\hat{Q}_i^Z$ are computed. Their 'm' most significant eigen values λ_i and their corresponding eigen matrices V_i are determined. An estimate of the unitary matrix R is obtained by maximizing the criteria $\lambda_i V_i$ by means of joint diagonalisation. An orthogonal contrast is optimized by finding the rotation matrix R such that the cumulant matrices are as diagonal as possible, that is, the Eq.(14),

$$R = \arg\min \sum_i Off(R^T \hat{Q}_i^Z R). \tag{14}$$

The demixing matrix W is estimated as, $W = RP$ and the components are estimated as, $\hat{S} = WX$.

5.5 MAXIMUM LIKELIHOOD ICA (ICA ML)

The Maximum Likelihood ICA (ICA ML) method for estimating the optimal unmixing matrix **W** [9]. ICA ML estimation is a standard statistical tool for finding parameter values (e.g., the unmixing matrix **W**) that provide the best fit of some data (e.g., the signals $\hat{S}$ extracted by **W**) to a given a model. The objective of ICA ML is to find an unmixing matrix **W** that yields extracted signals is given by,

$$\hat{S} = WX$$

where, X is the mixing matrix. This model incorporates the assumptions that source signals are non-Gaussian and independent. ICA ML is outlined as follows:

i. Centre the data to make its mean zero.

ii. Choose an initial separating matrix W, initial values γ_i, i=1,2,…n and learning rates μ and μ_γ randomly.

 Compute $\hat{S} = WX$.

iii. If the nonlinearities are not fixed. Then,

 Update $\gamma_i^+ = (1-\mu_r)\gamma_i + \mu_r E\{-\tanh(y_i) y_i + (1-\tanh(y_i)^2\}$

iv. If $\gamma_i^+ > 0$ define g_i as

 $g^+(y) = -2\tanh(y)$, otherwise

 $g^-(y) = \tanh(y) - y$

v. Update the separating matrix is given by,

 a. $W^+ = W + \mu[I + g(y)y^T]B$

 b. where $g(y) = (g_1(y_1),..., g_n(y_n))^T$

vi. If not converged, go back to step (iii).

5.6 MEAN FIELD ICA(ICA MF)

The measured signal X is assumed to be an instantaneous linear mixing of the sources corrupted with additive white Gaussian noise η that is,

$$X = AS + \eta \tag{15}$$

where, A is a (time independent) mixing matrix, then the following likelihood for parameters and sources,

$$P\ (X \mid A, \Sigma, S) = \left(\det 2\pi\Sigma\right)^{-\frac{N}{2}} e^{-\frac{1}{2}T_r(X-AS)^T \Sigma^{-1}(X-AS)}.$$

The aim of this ICA is to recover the unknown quantities: the sources **S**, the mixing matrix **A** and the noise covariance Σ from the observed data. The posterior distribution of the sources is readily given by,

$$P^+(X \mid A, \Sigma, S) = \frac{P(X \mid A, \Sigma, S)}{P(X \mid A, \Sigma)} \tag{16}$$

where, $P(S)$ is a prior on the sources which might include temporal correlations [11].

5.7 KERNEL ICA

The kernel-ICA algorithm uses the contrast functions based on canonical correlation analysis in a reproducing hilbert kernel space [9]. The outline of the algorithm is given as follows:

i. Let X be the data and $K(X_i, K_j)$ be the kernel.

ii. Data is whitened using whitening matrix P.

iii. The contrast function $C(W)$ is minimized with respect to W.

iv. The contrast function is minimized in the following way:

 a. The centered Gram matrices K_1, K_2,…K_m of the estimated sources $\{y_1, y_2, \ldots y_m\}$, where $y_i = Wx_i$ are computed.

 b. The minimal eigen value of the generalized eigen vector equation, $\hat{\lambda}_F(K_1,..., K_m)$.

Then,

$$C(W) = \frac{1}{2} \log \hat{\lambda}_F \left(K_1,..., K_m\right) \tag{17}$$

The demixing matrix W is calculated using whitening matrix P. Then the independent components are estimated as,

$$\hat{S} = WX$$

To perform ICA, a linear mix of image with key is generated to demix the watermark signal from the mixtures. The novelty of this detector is that it does not require the transform process to separate *LH* and *HL* bands for watermark extraction and omits using original and embedding parameters such as watermark location and strength. In this paper, watermark is embedded in Wavelet packet and the location of embedding is the same in spatial domain. Therefore ICA is applied directly on the watermarked image. With the help of a random key *K* to create different mixtures, one can extract successfully the watermark to claim the ownership.

Mixtures are created by the following equations,

$$X_1 = a_{11}I + a_{12}W + a_{13}K \quad (18)$$

$$X_2 = a_{21}I + a_{22}W + a_{23}K \quad (19)$$

$$X_3 = a_{31}I + a_{32}W + a_{33}K \quad (20)$$

where, a is a mixing matrix, W is the watermark matrix and K is a random key in the embedding process. Applying the above mentioned ICA algorithms to those mixtures, matrix watermark W is extracted. Performance measures like similarity measure and mutual information between mixture signals and independent components are calculated using the formulae given below,

$$Sim(X, X') = \frac{X.X'}{\sqrt{X'.X'}} \quad (21)$$

where, X is the original watermark and X' is the extracted watermark. If X and Y are two random variables with joint distribution $\mu(x, y)$ and marginal distributions $\mu_x(x)$ and $\mu_y(y)$, then mutual information $I(X, Y)$ between X and Y is defined as,

$$I(X,Y) = \iint \mu(x,y) \times \log\left(\mu(x,y) / \mu_x(x)\mu_y(y)\right) dxdy \quad (22)$$

6. SIMULATION RESULTS

A gray scale of size 256 x 256 is considered as Original image (Flower image) as shown in Fig.3. The original image is decomposed using RDWT for two level. The watermark is embedded in the second level middle frequency subband (HL2) using the embedding Eq.(2). A binary image of size 256 x 256 is considered as the watermark image (Robut) as shown in Fig.4. The watermarked image is obtained using two level Inverse RDWT. The watermarked image is shown in Fig.5 and it is inferred that it resembles the original image as visual as well as statistical. The quality of the watermarked image is evaluated by calculating the Peak Signal to Noise Ratio (PSNR) between original and watermarked image using the formula,

$$PSNR = 10\log_{10}\frac{255^2}{MSE}(dB) \quad (23)$$

where, MSE is the Mean Square Error. The PSNR value calculated using RDWT is 47.7112 dB and for DWT is 41.6662dB. The robustness of the above watermarking scheme is validated against attacks like JPEG compression, Gaussian noise addition, Median filtering, Salt & Pepper noise addition and Cropping. Fig.6(a) shows the JPEG compressed watermarked image. Fig.6(b) shows the result of adding Gaussian noise with noise density of 0.5 and Fig.6(c) shows Median filtering, Fig.6(d) shows Salt & pepper noise and Fig.6(e) shows the cropped version of the watermarked image in which 50% of the image is removed and the X- ray image is included in that space. The JPEG compression attack is done on RDWT based watermarked image and watermark is extracted using various ICA. The performance measures are tabulated in Table.2 and it is plotted in Fig.8 and Fig.9. It is observed that the proposed scheme exhibits its robustness against the common image processing attacks. Various ICA like Fast ICA, Radical ICA, Pearson ICA, Jade ICA, Maximum Likelihood ICA, Mean Field ICA and Kernel ICA are attempted to implement for watermark extraction. The extracted watermarks from various ICA are shown in Fig.7(a) to Fig.7(g).

Fig.3. Original Image

Fig.4. Watermark

Fig.5. Watermarked Image

Fig.6(a). Jpeg Compression

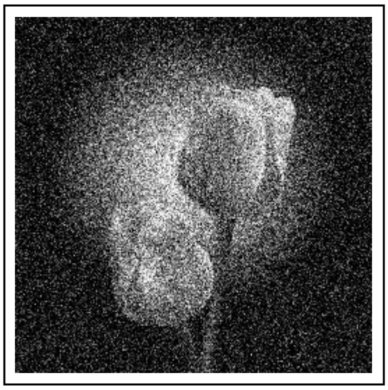

Fig.6(b). Gaussian Noise

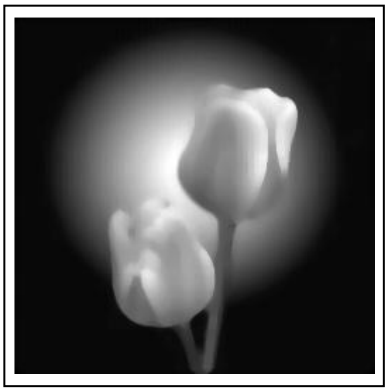

Fig.6(c). Median Filtering

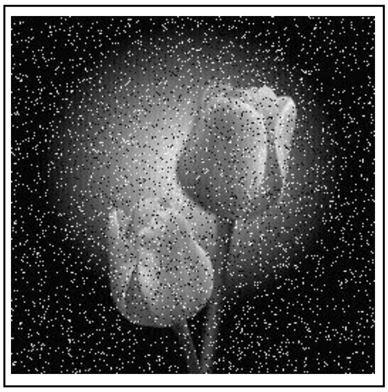

Fig.6(d). Salt & Pepper Noise

Fig.6(e). Cropping

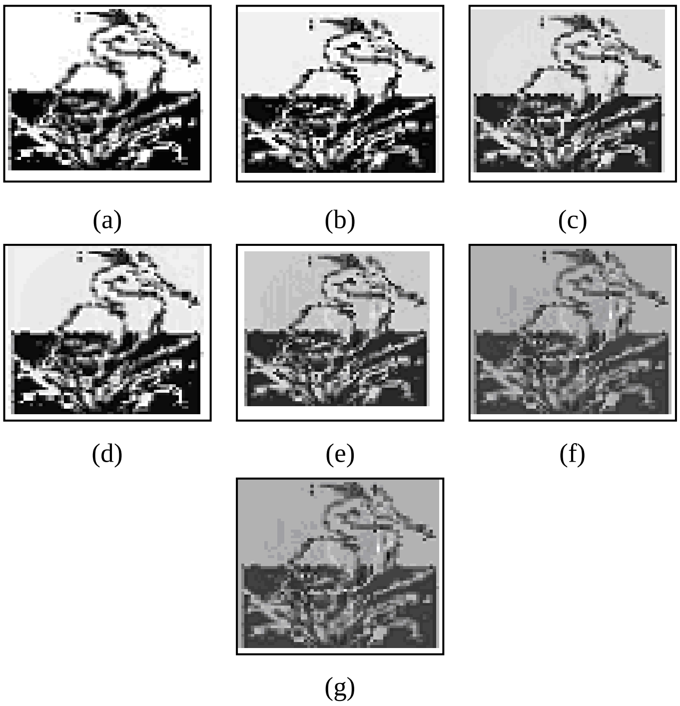

Fig.7(a) - 7(g). Extracted watermarks from various ICA

Table.2. Performance comparison of various ICA for RDWT

Various ICA	Similarity Measure	Mutual Information
Fast ICA	0.9797	2.5536
Pearson ICA	0.9791	1.5995
Radical ICA	0.9725	1.4232
Jade ICA	0.9787	1.9947
ICA ML	0.9661	1.3822
ICA MF	0.9648	1.3814
Kernel ICA	0.9670	1.3294

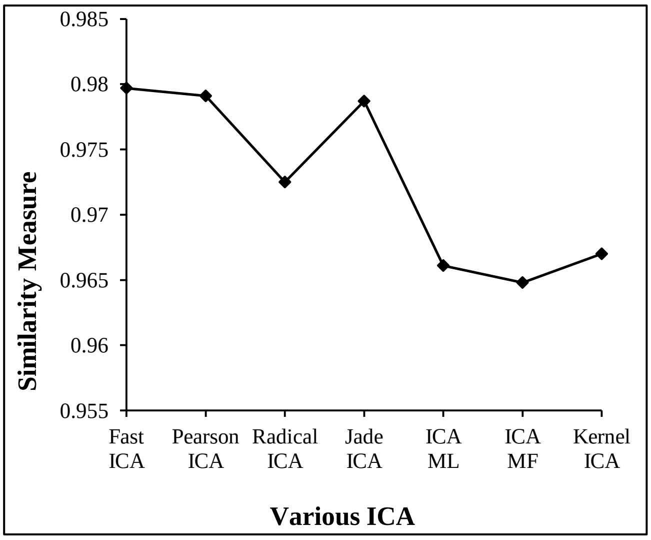

Fig.8. Similarity Measure for various ICA

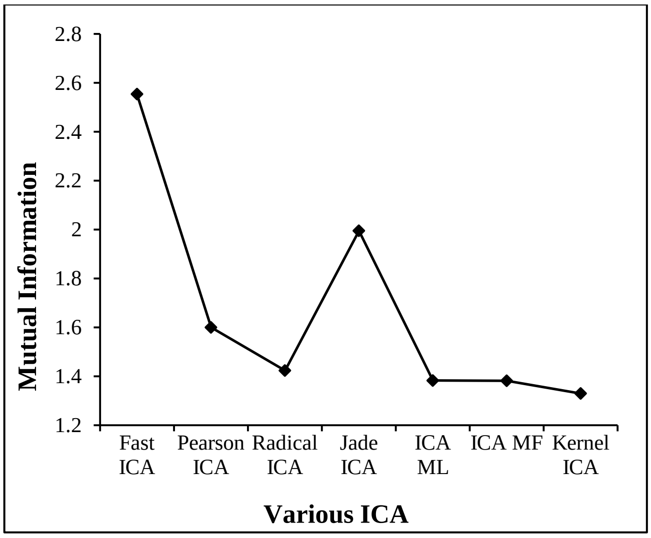

Fig.9. Mutual Information for various ICA

Table.3. Performance comparison of DWT and RDWT

Attacks	Similarity Measure		Mutual Information	
	DWT	RDWT	DWT	RDWT
JPEG compression	0.9599	0.9797	1.1546	2.5536
Gaussian Noise	0.9578	0.9762	1.1316	2.9915
Median Filtering	0.9593	0.9719	1.1521	2.5591
Salt & pepper noise	0.9541	0.9715	1.1368	2.9101
Cropping	0.9581	0.9777	1.1108	2.2728

7. CONCLUSION

In this paper, an attempt is made to implement RDWT for Digital Image Watermarking and Extraction using various ICA techniques. From the results, it is proved that RDWT posses a high PSNR value and better Similarity Measure as well as Mutual Information values. Among various ICA, the performance of Fast ICA is superior to other ICAs. The Robustness of the Proposed Scheme is also evaluated and compared with DWT against various Image Processing attacks.

REFERENCES

[1] M.D. Swanson, M. Kobayashi, and A.H. Tewfik, "Multimedia data embedding and watermarking technologies," *Proceedings of the IEEE*, Vol. 86, No.6, pp.1064-1087, 1998.

[2] H.J.M. Wang, P.C. Su and C.C.J. Kuo, "Wavelet-based digital image watermarking", *Journal of Optics Express*, Vol. 3, No. 12, pp. 491-496, 1998.

[3] D. Kundur and D. Hatzinakos, "Digital watermarking using multiresolution wavelet decomposition," in *Proceedings of the IEEE International Conference on Acoustics, Speech, Signal Processing(ICASSP)*, Vol. 5, pp. 2969-2972, 1998.

[4] Alexander H.Paquet, Saif Zahir and Rabab K. Ward, "Wavelet Packets based Image retrieval" *IEEE International Conference on Image Processing*, pp. 455-460, 2000.

[5] Alexander H.Paquet, Rabab K.Ward and Ioannis Pitas, "Wavelet Packets based Image digital watermarking for image verification and authentication" *Signal Processing*, Vol.83, pp. 2117-2132, 2003.

[6] Mayank Vasta, Richa Singh and Afzel Noore, "Feature based RDWT watermarking for multimodal biometric system," *Journal of Image and Vision Computing*, Vol. 27, pp. 293- 304, 2009.

[7] Thai Duy Hien., Zensho Nakao and Yen-Wei Chen, "Robust multi-logo watermarking by RDWT and ICA," *Signal Processing*, pp. 2981-2993, 2006.

[8] Hyvarinen, J. Karhunen and E. Oja, "*Independent component analysis*", First Edition, John Wiley & Sons, Inc., New York, 2002.

[9] V. Krishnaveni, S. Jayaraman, P.M. Manojkumar, K. Shivakumar and Ramadoss, "Comparison of independent component analysis algorithms for removal of Ocular Artifacts from Electroencephalo gram" *Journal of Measurement science review*, Vol. 5, 2005.

[10] Juha Karvanen., Jan Eriksson and Visa Koivunen, "Pearson system based method for blind separation", *Proceedings of Second International Workshop on Independent Component Analysis and Blind Signal Separation*, pp. 585 - 590, 2000.

[11] A.D.F.R. Pedro Hojen-Sorensen, Ole Winther and Lars Kai Hansen, "Mean Field Approaches to Independent Component Analysis", *Journal of Neural Computation*, Vol. 14, No. 4, pp. 889-918, 2002.

[12] S. Bounkong, B.Toch, D.Saad and D. Lowe, "ICA for watermarking digital images", *Journal of Machine Learning Research*, Vol. 4, pp. 1471- 1498, 2003.

[13] M. Shen, X. Zhang, L. Sun, P J Beadle and F H Y Chan, "A method for digital image watermarking using ICA", *4th International Symposium on Independent Component Analysis and Blind Source Separation*, pp. 209-214, 2003.